Lecture 6: Last Week ...

• Solution of Single Nonlinear Equation in One Unknown
  ➢ Two Examples
    ➢ Projectile Motion
    ➢ Mortgages/Loans
  ➢ Bisection
  ➢ Newton’s Method
Rate of Convergence:
Since the interval size is reduced by a factor of 2 at each step, the interval size after $k$ steps is $(b-a)/2^k$, which converges to 0 as $k \to \infty$.

To obtain an interval of size $2\delta$ we need

$$\frac{|b-a|}{2^k} \leq 2\delta \iff 2^{k+1} \geq \frac{|b-a|}{\delta} \iff k \geq \log_2 \left( \frac{|b-a|}{\delta} \right) - 1$$
Theorem: If $f \in C^2$, if $x_0$ is sufficiently close to a root $x^*$ of $f$, and if $f'(x^*) \neq 0$, then Newton’s Method converges to $x^*$ and ultimately the convergence rate is quadratic; that is there exists a constant $C^* = \left| \frac{f''(x^*)}{2f'(x^*)} \right|$ such that:

$$\lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^2} = C^*$$

Note: This means that for any $C \geq C^*$ there exists $K$ such that for all $k \geq K$, $|x_{k+1} - x^*| \leq C \cdot |x_k - x^*|^2$

Proof: Page 86 of textbook.
Theorem: If \( f \in C^{p+1} \) for \( p \geq 1 \), if \( x_0 \) is sufficiently close to a root \( x^* \) of \( f \), and if \( f'(x^*) = ... = f^{(p)}(x^*) = 0 \) but \( f^{(p+1)}(x^*) \neq 0 \), then Newton’s Method converges linearly to \( x^* \) with the error ultimately being reduced by about the factor \( p/p+1 \) at each step, that is:

\[
\lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|} = \frac{p}{p+1}
\]

Proof: Page 88 of textbook.

Examples: \( f(x) = x^2 \) (multiplicity \( p = 1 \)), \( f(x) = x^3 \) (\( p = 2 \))
Lecture 6: Today ...

• Solution of Single Nonlinear Equation in One Unknown (end)
  ➢ Quasi-Newton Methods
  ➢ Fixed-point Methods

• Conditioning of Problems
• Stability of Algorithms
Avoiding Derivatives:

- Iterations of the form:

  \[ x_{k+1} = x_k - \frac{f(x_k)}{g_k}, \quad \text{where} \quad g_k \approx f'(x_k) \]

  are called quasi-Newton methods.
Lecture 6: Quasi-Newton’s Methods

Constant Slope Method

\[ x_{k+1} = x_k - \frac{f(x_k)}{g}, \quad \text{where} \quad g = f'(x_0) \]

- If the slope of \( f \) does not change much as one iterates, then one might expect this method to mimic the behavior of Newton’s Method.
Convergence of Constant Slope Method

\[ f(x_k) = (x_k - x^*) f'(x^*) + O((x_k - x^*)^2) \]

\[ e_{k+1} = e_k - \frac{f(x_k)}{g} = e_k \left(1 - \frac{f'(x^*)}{g}\right) + O(e_k^2) \]

- In general, we cannot expect better than linear convergence.
- Variations on this idea ...
Secant Method

Iterations of the form:

\[ x_{k+1} = x_k - \frac{f(x_k)}{g_k}, \quad \text{where} \quad g_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \]

hence

\[ x_{k+1} = x_k - \frac{f(x_k) \cdot (x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad \text{for } k = 1, 2, \ldots \]
Secant Method

- To begin one needs two starting points.
- $g_k$ is the slope of the secant line through the points $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$.
- $g_k \rightarrow f'(x_k)$ as $x_{k-1} \rightarrow x_k$. 
Secant Method - Example

Let \( f(x) \equiv x^2 - 2 = 0 \). Taking \( x_0 = 1 \) & \( x_1 = 2 \), the secant method generates the following approximations and errors:

\[
\begin{align*}
  x_2 &= 1.3333 & e_2 &= -0.0809 \\
  x_3 &= 1.4 & e_3 &= -0.0142 \\
  x_4 &= 1.4146 & e_4 &= 4.2 \times 10^{-4} \\
  x_5 &= 1.4142114 & e_5 &= -2.1 \times 10^{-6}
\end{align*}
\]

Faster than linear, but slower than quadratic
Theorem: If \( f \in C^2 \), if \( x_0 \) is sufficiently close to a root \( x^* \) of \( f \), and if \( f'(x^*) \neq 0 \), then the error \( e_k \) in the Secant Method satisfies:

\[
\lim_{k \to \infty} \frac{e_{k+1}}{e_k \cdot e_{k-1}} = \lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*| \cdot |x_{k-1} - x^*|} = C^*
\]

where \( C^* = |f''(x^*)/2f'(x^*)| \). This means that for \( k \) large enough

\[ e_{k+1} \approx C^* \cdot e_k \cdot e_{k-1} \]

Note: The order of convergence of the secant method is

\[
\frac{1+\sqrt{5}}{2} \approx 1.62
\]
Proof: Subtracting $x^*$ from the Secant Method iteration:

$$e_{k+1} = e_k - \frac{f(x_k) \cdot (e_k - e_{k-1})}{f(x_k) - f(x_{k-1})}$$

$$e_{k+1} = e_k \cdot \left( \frac{f(x_k) - f(x_{k-1})}{f(x_k) - f(x_{k-1})} \right) - f(x_k) \cdot \left( \frac{e_k - e_{k-1}}{f(x_k) - f(x_{k-1})} \right)$$

$$e_{k+1} = e_k e_{k-1} \left( \frac{f(x_k) / e_k - f(x_{k-1}) / e_{k-1}}{f(x_k) - f(x_{k-1})} \right)$$

$$\frac{e_{k+1}}{e_k e_{k-1}} = \left( \frac{f(x_k) / e_k - f(x_{k-1}) / e_{k-1}}{x_k - x_{k-1}} \cdot \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right)$$
Proof: (continued)

\[
\frac{e_{k+1}}{e_k e_{k-1}} = \left( \frac{f(x_k)/e_k - f(x_{k-1})/e_{k-1}}{x_k - x_{k-1}} \cdot \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right)
\]

\[
\frac{e_{k+1}}{e_k e_{k-1}} = \left( \frac{f(x_k) - f(x^*)}{x_k - x^*} - \frac{f(x_{k-1}) - f(x^*)}{x_{k-1} - x^*} \right)
\]

\[
\frac{f''(x^*)}{2}
\]

\[
\frac{1}{f'(x^*)}
\]
Equation $f(x)=0$ can be expressed as fixed point problem $x = \varphi(x)$ for some $\varphi$, and vice versa.

\[
f(x) = 0 \quad \Rightarrow \quad \varphi(x) \equiv x \pm f(x) = x
\]

\[
\varphi(x) = x \quad \Rightarrow \quad f(x) \equiv x - \varphi(x) = 0
\]

\[
\varphi(x) = x \quad \Rightarrow \quad f(x) \equiv e^{(x-\varphi(x))} - 1 = 0
\]
A natural approach to solve a fixed point problem is to start with an initial guess $x_0$ and then iterate according to:

$$x_{k+1} = \varphi(x_k)$$

**Newton:** $\varphi(x) \equiv x - f(x)/f'(x)$

**Constant Slope:** $\varphi(x) \equiv x - f(x)/g$

**Secant Method:** does not fit this pattern.
Lecture 6: Fixed Point Methods

Geometrically fixed points lie at the intersection of the line $y=x$ and the graph of $y=\varphi(x)$.

**Example:** Find the fixed points of $\varphi(x)=x^2-6$. 

![Graph showing the intersection of the line $y=x$ and the graph of $y=x^2-6$.]
Lecture 6: Fixed Point Methods

The iteration may display monotonic convergence (upper left) oscillatory convergence (upper right), monotonic divergence (lower left), or oscillatory divergence (lower right).
Lecture 6: Fixed Point Methods

What determines whether a fixed point iteration will converge?

**Theorem:** Assume that \( \varphi \in C^2 \) and that \( |\varphi'(x)| < 1 \) in some interval \([x^* - \delta, x^* + \delta]\) centered about a fixed point \( x^* \) of \( \varphi \). If \( x_0 \) is in this interval, then the fixed point iteration \( x_{k+1} = \varphi(x_k) \) converges to \( x^* \).

**Proof:**

\[
x_{k+1} = \varphi(x_k) = \varphi(x^*) + (x_k - x^*) \cdot \varphi' (\xi) = x^* + (x_k - x^*) \cdot \varphi' (\xi)
\]

\[
e_{k+1} = e_k \cdot \varphi' (\xi)
\]
Example: Let us consider three fixed point iteration schemes that can be derived from the equation \( f(x) = x^3 + 6x^2 - 8 = 0 \), which has a solution in the interval \([1, 2] \), since \( f(1) = -1 < 0 \) and \( f(2) = 24 > 0 \).

\[
\varphi_1(x) = x^3 + 6x^2 + x - 8
\]

\[
\varphi_2(x) = \sqrt{\frac{8}{x + 6}}
\]

\[
\varphi_3(x) = \sqrt{\frac{8 - x^3}{6}}
\]

Study behavior of the sequences when \( x_0 = 1.5 \)
Actually \( \varphi'(x) \) need not exist in order to have convergence.

**Theorem:** If \( \varphi \) is a *contraction* (on all of \( \mathbb{R} \)), then it has a unique fixed point \( x^* \) and the iteration \( x_{k+1} = \varphi(x_k) \) converges to \( x^* \) for any \( x_0 \).

**Def’n:** \( \varphi \) is a *contraction* if there exists a constant \( L < 1 \) such that for all \( x \) and \( y \)

\[
| \varphi(x) - \varphi(y) | \leq L |x-y|
\]

**Proof:** \( \{x_k\} \) is a Cauchy sequence, hence converges.

**Note:** Theorem can be limited to interval \([a,b]\).