CS 3331 Numerical Methods

Lecture 2: Functions of One Variable

Cherung Lee

Outline

- Introduction
- Solving nonlinear equations: find x^* such that $f(x^*) = 0$.
  - Binary search methods: (Bisection, regula falsi)
  - Newton-typed methods: (Newton's method, secant method)
  - Higher order methods: (Muller's method)
- Accelerating convergence: Aitken's Δ^2 method

Introduction

Motivating problem

- How to estimate compound interest rate?
  - Example: Suppose a bank loans you 200,000 with compound interest rate. After 10 year, you need to repay 400,000 (principal+interest). Suppose the frequency of compounding is yearly. How much is the annual percentage rate (APR)?
- Equation of the compound interest: $20,000(1+r)^{10} = 40,000$.
  - How to solve $f(r) = (1+r)^{10} - 2 = 0$?
  - $r = \sqrt[10]{2} - 1 \approx 7.1773\%$

Amortized Loan

- Loan repaid in a series of payments for principal and interest.
- Formula: (r: interest-rate, a: payment, n: period)
  - Suppose x_k is the debt in the k's period.

$$\begin{aligned} x_k &= (1+r)x_{k-1} - a = (1+r)^2 x_{k-2} - (1+r)a - a = \ldots \\ &= x_0(1+r)^k - a\frac{(1+r)^k - 1}{r} \end{aligned}$$

  - x_0 is the principal and $x_n = 0 \Rightarrow x_0(1+r)^n - a\frac{(1+r)^n - 1}{r} = 0$.
- How to solve $f(r) = 20(1+r)^{10} - 4\frac{(1+r)^{10} - 1}{r} = 0$?

Useful tools from calculus LVF pp.10

- Intermediate value theorem

  If $f(x)$ is a *continuous function* on the interval $[a,b]$, and $f(a) < 0 < f(b)$ or $f(b) < 0 < f(a)$, then there is a number $c \in [a,b]$ such that $f(c) = 0$.

- Taylor's theorem

  If $f(x)$ and all its kth derivatives are continuous on $[a,b]$, $k = 1 \cdots n$, and $f^{(n+1)}$ exists on (a,b), then for any $c \in (a,b)$ and $x \in [a,b]$, (ξ is between c and x.)

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x-c)^k + \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-c)^{n+1}.$$

Solving Nonlinear Equations

Bisection method LVF pp.52-55

- Binary search on the given interval $[a, b]$.
  - Suppose $f(a)$ and $f(b)$ have opposite signs.
  - Let $m = (a + b)/2$. Three things could happen for $f(m)$.
    - $f(m) = 0 \Rightarrow m$ is the solution.
    - $f(m)$ has the same sigh as $f(a) \Rightarrow$ solution in $[m, b]$.
    - $f(m)$ has the same sigh as $f(b) \Rightarrow$ solution in $[a, m]$.
- Linear convergence with rate 1/2.

Pros and cons

- Pros
  - Easy to implement.
  - Guarantee to converge with guaranteed convergent rate.
  - No derivative required.
  - Cost per iteration (function value evaluation) is very cheap.
- Cons
  - Slow convergence.
  - Do not work for double roots, like solving $(x-1)^2 = 0$

Regula falsi (false position) LVF pp.57-59

- Straight line approximation + intermediate value theorem

- Given two points $(a, f(a)), (b, f(b))$, $a \neq b$, the line equation

$$L(x) = y = f(b) + \frac{f(a) - f(b)}{a - b}(x - b),$$

  and its root, $L(s) = 0$, is $s = b - \frac{a-b}{f(a)-f(b)} f(b)$.

- Use intermediate value theorem to determine $x^* \in [a, s]$ or $x^* \in [s, b]$

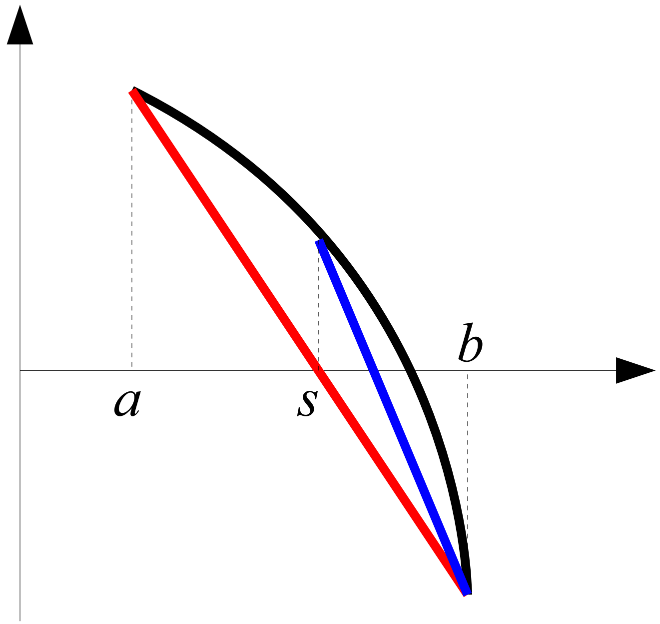
Convergence of regula falsi

Consider a special case: $(b, f(b))$ is fixed.

- Note $[s, b]$ may not go to zero. (compare to bisection method.)
- Change measurement

$$\frac{|s - x^*|}{|a - x^*|} = \frac{|(b - s) - (b - x^*)|}{|(b - a) - (b - x^*)|}$$

- $b - s = \dfrac{-f(b)}{f(a) - f(b)}(b - a).$
- Let $m = \dfrac{-f(b)}{f(a) - f(b)} < 1.$



$$\frac{|s - x^*|}{|a - x^*|} = \frac{|m(b - a) - (b - x^*)|}{|(b - a) - (b - x^*)|} < 1$$

- Linear convergence

Newton's method LVF pp.66-71

- Approximate $f(x)$ by the tangent line $f(x_k) + (x - x_k)f'(x_k)$.
- Find the minimum of the square error

$$\min_x |f(x) - 0|^2 \Longleftrightarrow d(f(x))^2/dx = 0$$

- The minimizer is $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$
- Convergent conditions
  - $f(x), f'(x), f''(x)$ are continuous near x^*, and $f'(x) \neq 0$.
  - x_0 is sufficiently close to x^*. $\left[\frac{\max|f''|}{2\min|f'|}|x_0 - x^*| < 1\right]$.

Convergence of Newton's method LVF pp.70-71

- Taylor expansion: for some η between x^* and x_k

$$f(x^*) = f(x_k) + (x^* - x_k)f'(x_k) + \frac{(x^* - x_k)^2}{2} f''(\eta) = 0$$

$$x^* = x_k - f(x_k)/f'(x_k) - (x^* - x_k)^2 \frac{f''(\eta)}{2f'(x_k)}$$

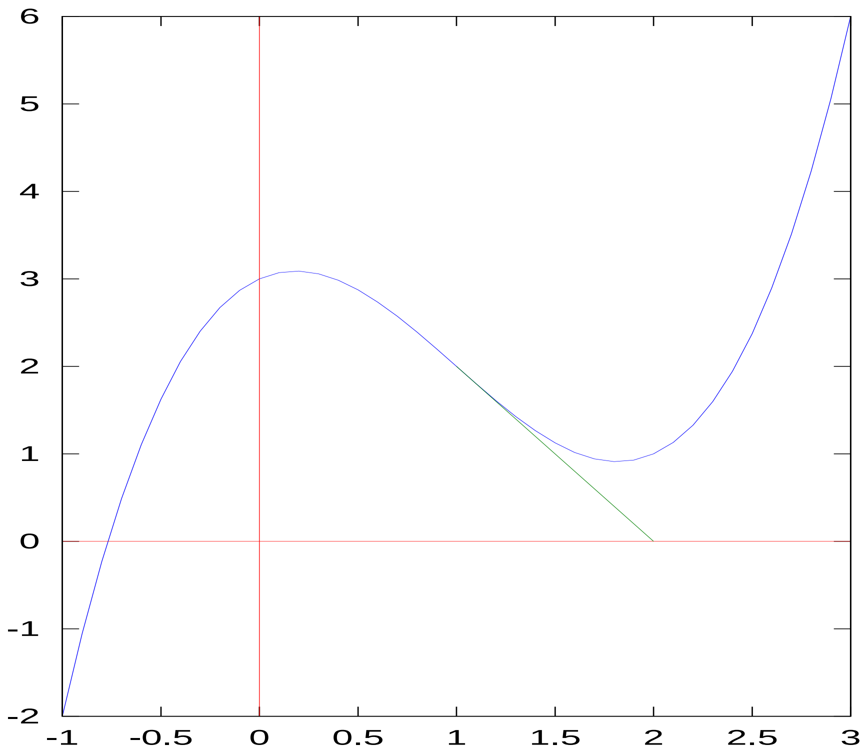
- Substitute Newton's step $x_k - f(x_k)/f'(x_k) = x_{k+1}$.

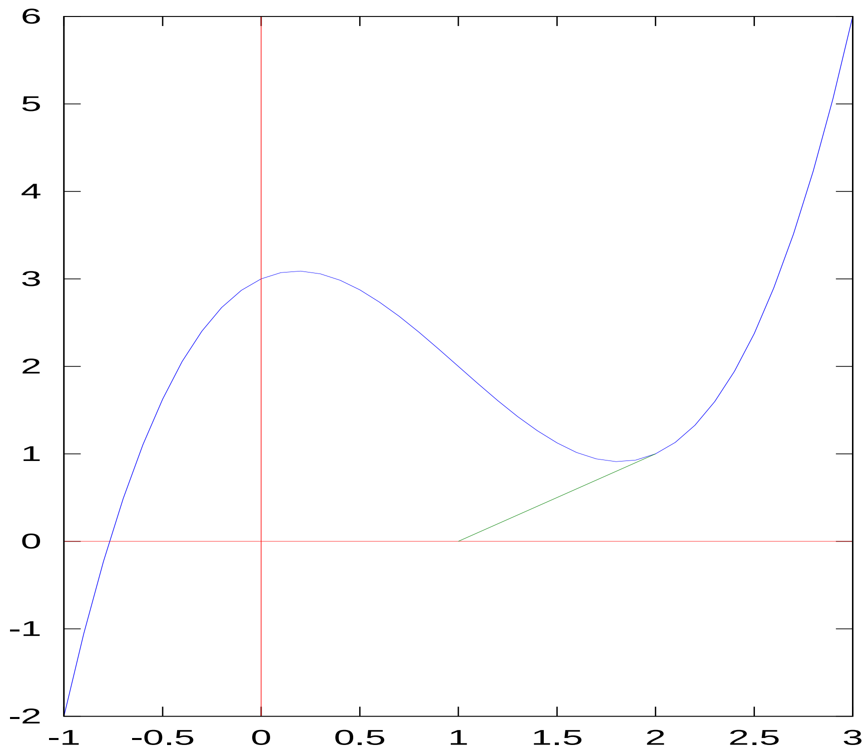
$$x^* - x_{k+1} = -(x^* - x_k)^2 \frac{f''(\eta)}{2f'(x_k)}$$

- Quadratic convergence with $\lambda = \left| \frac{f''(x^*)}{2f'(x^*)} \right|$.

Oscillations in Newton's method LVF pp.71

- Solve $f(x) = x^3 - 3x^2 + x + 3 = 0$ with $x_0 = 1$.





Newton's method for repeated roots LVF pp.72

- If x^* is a repeated root, Newton's method converges linearly.
- Newton's method can be regarded as a fixed-point iteration.

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

  - Convergence of fixed-point iteration: LVF pp.22-23.

- Taylor expansion of $g(x)$ about x_n near x^*

$$x_{n+1} = g(x_n) = g(x^*) + g'(x^*)(x_n - x^*) + \frac{g''(\xi)}{2}(x_n - x^*)^2.$$

  - Quadratic convergence if $g'(x^*) = 0$.

case 1 If $f(x^*)$ is a simple root, $(f'(x^*) \neq 0)$

- $g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{(f'(x))^2} = 1 - 1 + \frac{f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$.
- $g'(x^*) = 0$

case 2 If $f(x^*)$ is a repeated root, $(f'(x^*) = 0)$

- Assume $f(x) = (x - x^*)^2 h(x)$ where $h(x^*) \neq 0$.
- $f'(x) = 2(x - x^*)h(x) + (x - x^*)^2 h'(x)$.
- $g(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x - x^*)h(x)}{2h(x) + (x - x^*)h'(x)}$.
- Let $a(x) = 2h(x) + (x - x^*)h'(x)$. (we will use that to simply the proof).

– $g'(x) = 1 - \frac{\big(h(x)+(x-x^*)h'(x)\big)a(x)-(x-x^*)h(x)a'(x)}{(a(x))^2}$

– $a(x^*) = 2h(x^*) + \cancel{(x^* - x^*)h'(x^*)} = 2h(x^*) \neq 0$

$$\begin{aligned} g'(x^*) &= 1 - \frac{(h(x^*) + \cancel{(x^* - x^*)h'(x^*)})a(x^*) - \cancel{(x^* - x^*)h(x^*)a'(x^*)}}{a(x^*)^2} \\ &= 1 - \frac{h(x^*)}{a(x^*)} = \frac{h(x^*)}{2h(x^*)} = 1 - 1/2 \neq 0. \end{aligned}$$

$\Rightarrow$ When x^* is a repeated root, convergence is linear.

- How to modify it to restore the quadratic convergence?

  – For $f(x) = (x - x^*)^2 h(x)$, let $g(x) = x - {\color{red}2}\frac{f(x)}{f'(x)} \Rightarrow g'(x^*) = 0$.

  – The algorithm becomes $x_{k+1} = x_k - {\color{red}2}\frac{f(x_k)}{f'(x_k)}$

Secant method LVF pp.60-65

- Newton's method requires derivative at each step.
- $f'(x_k)$ can be approximated by $\frac{f(x_{k-1})-f(x_k)}{x_{k-1}-x_k}$, which make

$$x_{k+1} = x_k - \frac{x_{k-1} - x_k}{f(x_{k-1}) - f(x_k)} f(x_k).$$

- Convergent conditions
  - $f(x), f'(x), f''(x)$ are continuous near x^*, and $f'(x) \neq 0$.
  - Initial guesses x_0, x_1 are *sufficiently close* to x^*.
    $\max(M|x_0 - x^*|, M|x_1 - x^*|) < 1$, where $M = \max|f''|/2\min|f'|$

Convergence of the secant method

- Let $e_k = x_k - x^*$

$$
\begin{aligned}
e_{k+1} &= x_{k+1} - x^* \\
&= x_k - \frac{x_{k-1} - x_k}{f(x_{k-1}) - f(x_k)} f(x_k) - x^* \\
&= \frac{(x_{k-1} - x^*) f(x_k) - (x_k - x^*) f(x_{k-1})}{f(x_{k-1}) - f(x_k)} \\
&= \frac{e_{k-1} f(x_k) - e_k f(x_{k-1})}{f(x_{k-1}) - f(x_k)}
\end{aligned}
$$

- Using Taylor expansion

$$
\begin{aligned}
f(x_k) &= \cancelto{0}{f(x^*)} + e_k f'(x^*) + e_k^2 f''(x^*)/2 + O(e_k^3) \\
f(x_{k-1}) &= \cancelto{0}{f(x^*)} + e_{k-1} f'(x^*) + e_{k-1}^2 f''(x^*)/2 + O(e_{k-1}^3)
\end{aligned}
$$

$$
\begin{aligned}
f(x_{k-1}) - f(x_k) &= (e_{k-1} - e_k) f'(x^*) + (e_{k-1}^2 - e_k^2) f''(x^*)/2 + O(e_{k-1}^3) \\
&\approx (e_{k-1} - e_k) f'(x^*)
\end{aligned}
$$

(We assume e_k is small enough so that $|e_k|^3 \ll |e_k|^2 \ll |e_k|$.)

$$
\begin{aligned}
e_k f(x_{k-1}) - e_{k-1} f(x_k) &= (e_{k-1} e_k - e_k e_{k-1}) f'(x^*) + \\
&\quad (e_k e_{k-1}^2 - e_k^2 e_{k-1}) f''(x^*)/2 + O(e_{k-1}^3) \\
&\approx e_k e_{k-1} (e_{k-1} - e_k) f''(x^*)/2
\end{aligned}
$$

- Summarizing above equations

$$
\begin{aligned}
e_{k+1} &= \frac{e_{k-1} f(x_k) - e_k f(x_{k-1})}{f(x_{k-1}) - f(x_k)} \\
&= \frac{e_k e_{k-1} (e_{k-1} - e_k) f''(x^*)/2}{(e_{k-1} - e_k) f'(x^*)} \\
&= \frac{e_{k-1} e_k f''(x^*)}{2 f'(x^*)}
\end{aligned}
$$

- We want to prove $|e_{k+1}| = C|e_k|^\alpha$
- $\left|\dfrac{e_{k-1}e_k f''(x^*)}{2f'(x^*)}\right| = C|e_k|^\alpha$
- Recursively, $|e_k| = C|e_{k-1}|^\alpha$.

  $$\left|\frac{Ce_{k-1}^{1+\alpha} f''(x^*)}{2f'(x^*)}\right| = C^{1+\alpha}|e_{k-1}|^{\alpha^2} \Rightarrow \left|\frac{f''(x^*)}{2f'(x^*)}\right| = C^\alpha |e_{k-1}|^{\alpha^2-\alpha-1}$$

- $|e_{k-1}|^{\alpha^2-\alpha-1}$ equals to a constant, $\alpha^2 - \alpha - 1 = 0$.
  $\alpha = (1+\sqrt{5})/2 = 1.618$
- $C = \left|\frac{f''(x^*)}{2f'(x^*)}\right|^{1/\alpha} \approx \left|\frac{f''(x^*)}{2f'(x^*)}\right|^{0.618}$
- Superlinear convergence with $\lambda = \left|\frac{f''(x^*)}{2f'(x^*)}\right|^{0.618}$

Muller's method LVF pp.73-77

- Approximate $f(x)$ by a parabola.
- A parabola passes $(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3))$ is

$$P(x) = f(x_3) + c_2(x - x_3) + d_1(x - x_3)(x - x_2),$$

$$c_1 = \frac{f(x_1) - f(x_3)}{x_1 - x_3}, c_2 = \frac{f(x_2) - f(x_3)}{x_2 - x_3}, d_1 = \frac{c_1 - c_2}{x_1 - x_2}.$$

- We want to find a solution closer to x_3. Let $y = x - x_3$ and rewrite $P(x)$ as a function of y.

$$\begin{aligned} P(x) &= f(x_3) + c_2(x - x_3) + d_1(x - x_3)(x - x_2) \\ &= f(x_3) + c_2(x - x_3) + d_1(x - x_3)(x - x_3 + x_3 - x_2) \\ &= f(x_3) + c_2 y + d_1 y(y + x_3 - x_2) \\ &= f(x_3) + (c_2 + d_1(x_3 - x_2))y + d_1 y^2 \end{aligned}$$

- Let $s = c_2 + d_1(x_3 - x_2)$. The solution is

$$y = \frac{-s \pm \sqrt{s^2 - 4d_1 f(x_3)}}{2d_1}, \; x = x_3 - \frac{s \pm \sqrt{s^2 - 4d_1 f(x_3)}}{2d_1}$$

- Let x_4 be the solution closer to x_3, $x_4 = x_3 - \frac{s - \text{sign}(s)\sqrt{s^2 - 4d_1 f(x_3)}}{2d_1}$, which equals to (in a more stable way)

$$x_4 = x_3 - \frac{2f(x_3)}{s + \text{sign}(s)\sqrt{s^2 - 4f(x_3)d_1}}.$$

- x_4 is the a better approximation to x^* than x_3.

- Use $(x_2, f(x_2)), (x_3, f(x_3)), (x_4, f(x_4))$ as next three parameters, and continue the process until converging.

Properties of Muller's method

- No derivative needed
- Can find complex roots
- Fails if $f(x_1) = f(x_2) = f(x_3)$, when x is a repeated root.
- Superlinear convergence, $p \approx 1.84$, with

  $$\lambda = |f'''(x^*)|^\beta / |2f'(x^*)|^\beta,$$

  where $\beta = (p-1)/2$. The proof is similar to the secant method's.

Accelerating convergence

Aitken's Δ^2 method

- Accelerate the convergence of a linearly convergent sequence.
- Suppose $\{p_k\}_{k=0}^{\infty} \to p$ linearly, and $(p_{k+1} - p)/(p_k - p) > 0$ for $k > N$, where N is some constant. Then the sequence

$$q_k = p_k - \frac{(p_{k+1} - p_k)^2}{p_{k+2} - 2p_{k+1} + p_k}$$

converges to p, with better convergence order than p_k,

$$\lim_{k \to \infty} \frac{q_k - p}{p_k - p} = 0.$$

LVF pp.197, also check last year's notes.

Sketch of the proof

- Since $\lim_{k\to}(p_{k+1}-p)/(p_k-p) = \lambda > 0$, for large k

$$\frac{p_{k+1}-p}{p_k-p} \approx \frac{p_{k+2}-p}{p_{k+1}-p}.$$

- Expanding the terms yields

$$p \approx p_k - \frac{(p_{k+1}-p_k)^2}{p_{k+2}-2p_{k+1}+p_k} = q_k.$$

- Comparing $q_k - p$ and $p_k - p$ for large k gives

$$\lim_{k\to\infty} \frac{q_k-p}{p_k-p} = 0.$$