ECE 4960

Spring 2017

Lecture 13

Nonlinear Equations and Optimization:
The Newton-Raphson Method

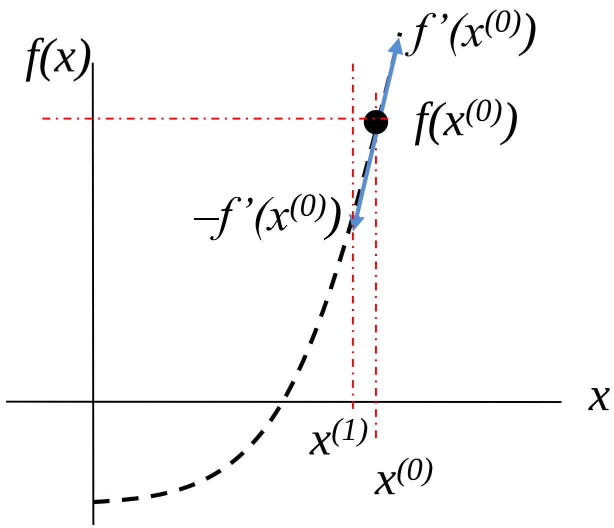
Edwin C. Kan

School of Electrical and Computer Engineering

Cornell University

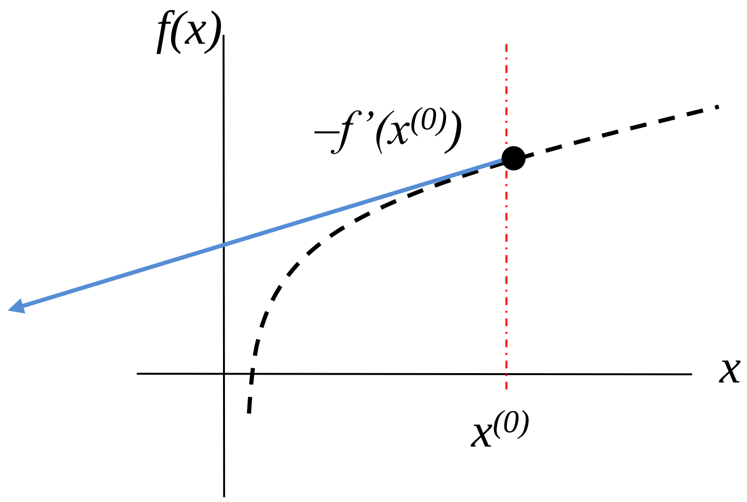
Finding the Root of 1D Nonlinear Equation

- For a problem of $y = f(x) = 0$, we can find the "root" by using the **slope** to direct the search of $f(x) = 0$.
- In addition to using the sign at the present solution to know which section to continue search in bisection, we evaluate the slope J to make predictions about what the next guessed solution is.
- If $f(x)$ and $f'(x)$ are both positive, we know that we should decrease x by the amount of $f(x)/f'(x)$ for the next solution. We should achieve the solution in this step if $f(x)$ is close to be linear.



Slope-Based Iterative Methods

- Good news: if we are close to the solution (i.e., Δx is small), all continuous equations will be close to linear, as can be seen from the Taylor series: $f(x) = f(x_0) + f'(x_0)\Delta x + O(\Delta x^2)$.
- Bad news: for a highly nonlinear problem, we can "**overshoot**" the correction to bring it to a region with very large error. e^x is not "convergent" for $x \rightarrow \infty$; $\log(x)$ not convergent for $x \rightarrow 0$.



Convergence in Iterative Methods

- We make an initial guess of $x^{(0)}$, and evaluate $\Delta x^{(0)}$ by $f(x^{(0)})$ for $x^{(1)} = x^{(0)} + \Delta x^{(0)}$, and the process goes on until the **absolute residue** approaches 0:

$$\lim_{k\to\infty} f\left(x^{(k)}\right) \to 0$$

- The **relative residue** will also approach 0 at the same time:

$$\lim_{k\to\infty} \Delta x^{(k)} \to 0$$

- For multi-variate cases:

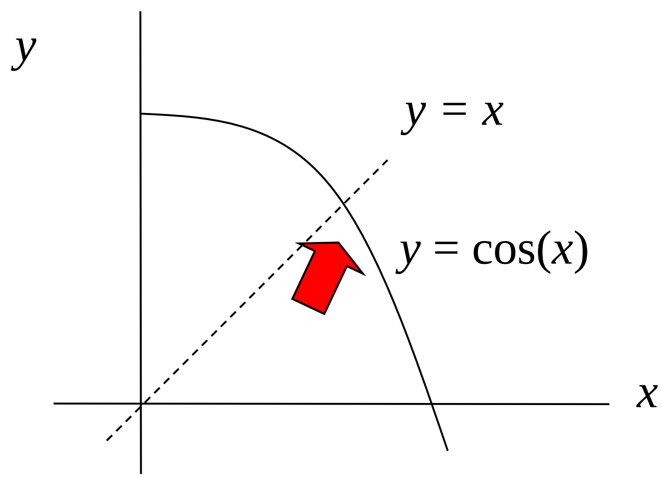
$$\lim_{k\to\infty} \left\| \vec{f}\left(\vec{x}^{(k)}\right) \right\|_2 \to 0$$

$$\lim_{k\to\infty} \frac{\left\|\Delta\vec{x}^{(k)}\right\|_2}{\left\|\vec{x}^{(k)}\right\|_2} \to 0$$

Self normalized!

Linear Convergence of Bisection

- For example, $f(x) = x - \cos(x)$, where we know there is a solution in $(0, \pi/2)$, but we also know that this transcendental equation does not have closed-form solution and have to usean iterative scheme.
- We can use the bisection method in 1D to search the region of $(0, \pi/2)$, which will be sufficiently efficient, and k digits of precision will need $k \cdot \log_2 10$ steps of search. This is called "**linear convergence**" (it is actually exponential, but the improvement is proportional to $\exp(\Delta x)$ instead of $\exp(\Delta x^2)$).



The Newton Raphson Method

1. Set up the variables and evaluation of $f(x)$ for solving x that satisfis the nonlinear equation $f(x) = 0$

2. Make an **initial guess** $x^{(k)}$, where $k = 0$ initially Notice that this is a difficult choice and has dominant influence on the convergence behavior.

3. Evaluate $f(x^{(k)})$ and its slope $f'(x^{(k)})$ or the Jacobian matrix J for the multi-variate case. Calculate the update vector:
$\Delta x^{(k)} = -f(x^{(k)})/f'(x^{(k)})$ or $\Delta\vec{x}^{(k)} \cong -\left[J\left(\vec{x}^{(k)}\right)\right]^{-1} \cdot \vec{f}\left(\vec{x}^{(k)}\right)$

4. Evaluate the norm of $\|\Delta x^{(k)}\|_2$ and $\|f(x^{(k)})\|_2$. Stop if $\|\Delta x^{(k)}\|_2$ or $\|f(x^{(k)})\|_2 <$ tolerance (often set between 10^{-7} to 10^{-9}).

5. Update $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$, k++, and return to Step 3 to iterate.

1D Newton-Raphson Method Example

- $f(x) = x - \cos x$; Initial guess: $x^{(0)} = 0$.
- $f'(x) = 1 + \sin x$;
- $\Delta x^{(k)} = -\left[f'\left(x^{(k)}\right)\right]^{-1} f\left(x^{(k)}\right) = -\frac{x^{(k)} - \cos x^{(k)}}{1 + \sin x^{(k)}}$
- $x^{(0)} = 0; \Delta x^{(0)} = 1.$
- No line search use. If we do use, we will search $\Delta x^{(0)} = 2, 1, 0.5, 0.25, 0.125$ to see which one gives the smallest $f(x^{(1)})$.
- $x^{(1)} = 1; \Delta x^{(1)} = -0.24.$
- $x^{(2)} = 0.76; \Delta x^{(2)} = \ldots.$

Isaac Newton
1643 - 1726

?

Joseph Raphson
1648 - 1715

Quadratic Convergence of Newton Method

Step size or residual	Linear convergence	Quadratic convergence
$\frac{\left\|\Delta x^{(k)}\right\|}{x^{(k)}}$ or $\left\|f\left(x^{(k)}\right)\right\|$	0.1	0.1
$\frac{\left\|\Delta x^{(k+1)}\right\|}{x^{(k+1)}}$ or $\left\|f\left(x^{(k+1)}\right)\right\|$	0.01	0.01
$\frac{\left\|\Delta x^{(k+2)}\right\|}{x^{(k=2)}}$ or $\left\|f\left(x^{(k+2)}\right)\right\|$	10^{-3}	10^{-4}
$\frac{\left\|\Delta x^{(k+3)}\right\|}{x^{(k+3)}}$ or $\left\|f\left(x^{(k+3)}\right)\right\|$	10^{-4}	10^{-8}
$\frac{\left\|\Delta x^{(k+4)}\right\|}{x^{(k+4)}}$ or $\left\|f\left(x^{(k+4)}\right)\right\|$	10^{-5}	10^{-16}

Hacker Practice

Use the Newton method to solve the following nonlinear equation:

$$f(x) = e^{100x} - 1 = 0$$

$$\Delta x^{(k)} = -\left[f'\left(x^{(k)}\right)\right]^{-1} f\left(x^{(k)}\right)$$

Report $x^{(k)}$, $\Delta x^{(k)}$, $f(x^{(k)})$.

Make $x^{(0)} = 1$, and then recompute using $x^{(0)} = 10$. When do you observe quadratic convergence?

Multi-Variate Newton's Method

$$\vec{f}\left(\vec{x}+\Delta\vec{x}\right)-\vec{f}\left(\vec{x}\right)=\left[J\left(\vec{x}\right)\right]\cdot\Delta\vec{x}+O\left(\left\|\Delta x\right\|^{2}\right)$$

J: Jacobian matrix

$$\left[J\left(\vec{x}\right)\right]\cdot\Delta\vec{x}\cong-\vec{f}\left(\vec{x}\right)$$

$$\Delta\vec{x}\cong-\left[J\left(\vec{x}\right)\right]^{-1}\cdot\vec{f}\left(\vec{x}\right)$$

Newton iteration

Multi-Variate Newton's Method Example

$$f_1: \quad 3x_1^2 + x_2 - 4 = 0$$
$$f_2: \quad x_1^2 - 3x_2 + 2 = 0$$

Nonlinear to x_1; linear to x_2.
Two solutions at (1, 1) and (–1, 1).

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 6x_1 & 1 \\ 2x_1 & -3 \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{(ad - bc)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\vec{x}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad \Delta\vec{x}^{(0)} = -\begin{bmatrix} 6 & 1 \\ 2 & -3 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} -3 & -1 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \qquad \Delta\vec{x}^{(1)} = -\begin{bmatrix} 6 & 1 \\ 2 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Converged!

Influence of Initial Guess

$$f_1: \quad 3x_1^2 + x_2 - 4 = 0$$
$$f_2: \quad x_1^2 - 3x_2 + 2 = 0$$

Nonlinear to x_1; linear to x_2.
Two solutions at (1, 1) and (–1, 1).

$$x^{(0)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}; \qquad \Delta x^{(0)} = -\begin{bmatrix} 12 & 1 \\ 4 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 6 \end{bmatrix} = \frac{1}{40} \begin{bmatrix} -3 & -1 \\ -4 & 12 \end{bmatrix} \begin{bmatrix} 8 \\ 6 \end{bmatrix} = \begin{bmatrix} -0.75 \\ 1 \end{bmatrix}$$

$$x^{(1)} = \begin{bmatrix} 1.25 \\ 1 \end{bmatrix}; \qquad \Delta x^{(1)} = \ldots$$

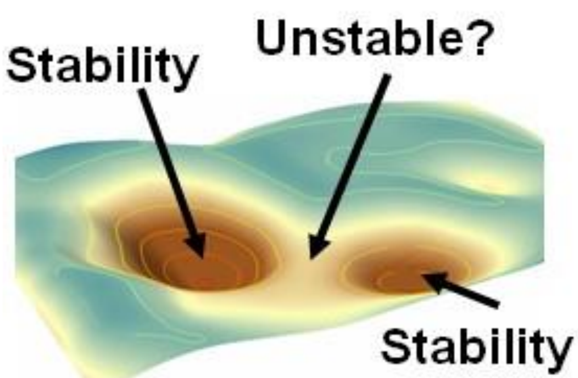
x_2 correct in the first step and x_1 needs more iterations

Line Search in Newton's Method

- Newton's method will converge only in the proximity of the solution: **basin of attraction**.
  1. $||\Delta x||$ is sufficiently small.
  2. $f(x + \Delta x)$ is close to zero!
  3. Δx is NOT zero (or else we cannot make any improvement).
  4. J^{-1} will not stretch the f vector by much, i.e., J^{-1} is not ill-conditioned.

- **Line search:** $\left\| \vec{f}\left(\vec{x} + t\Delta\vec{x}\right) \right\|$ is minimum for all scarlar t

- Example: Bisection search for t
- Δx can be zero at the **deflection point**!



Basin of Attraction

Hacker Practice

Use the Newton method with line search to solve the same nonlinear equation by making $x^{(0)} = 10$.

$$f(x) = e^{100x} - 1 = 0$$

$$\Delta x^{(k)} = -\left[f'\left(x^{(k)}\right)\right]^{-1} f\left(x^{(k)}\right)$$

Report $x^{(k)}$, $\Delta x^{(k)}$, $f(x^{(k)})$.

What is the change in the beginning and end of the convergence behavior?