Numerical Methods-Lecture V-Newton's Method

Trevor Gallen

Fall, 2015

Goals

Aim is to teach numerical methods, give you the tools you need to write down, solve, and estimate models

1. Interpolation
2. Numerical derivatives
3. Maximization/minimization
    - Deterministic, stochastic
    - Derivative-based, derivative-free
    - Local, global
4. Numerical integration/quadrature
5. Bellman equations

Motivation

- We have some function $f(x)$
- Want to find x^* such that $f(x^*) = 0$.
- Examples:
  - Equilibrium conditions: $X^S(p) - X^D(p) = 0$
  - Maximization conditions: $u'(c) - \lambda = 0$

Newton's Method: Idea

1. Take a linear approximation of function, find slope and intercept
2. Given equation of line, solve for zero
3. Take that new point, repeat.

Newton's Method Derivation

1. Take Taylor expansion: $f(x) \approx f(x_0) + \left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} (x - x_0)$
2. Set equal to zero: $0 = f(x_0) + \left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} (x - x_0)$
3. Solve for x: $x = x_0 - \frac{f(x_0)}{\left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0}}$
4. Call x x_0 repeat.

In other words, the interative procedure:

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

Example: $f(x) = sin(x),\ x_0 = 1$

1. $x_1 = sin(1) - \frac{sin(1)}{cos(1)} = 1 - \frac{0.8414}{0.5403} = -0.56$
2. $x_2 = -0.56 - \frac{sin(-0.56)}{cos(-0.56)} = -0.56 - \frac{0.53}{0.85} = 0.07$
3. $x_3 = 0.07 - \frac{sin(0.07)}{cos(0.07)} = 0.07 - \frac{0.07}{1} = -1e - 4$
4. And so on until $x_i \approx 0$.

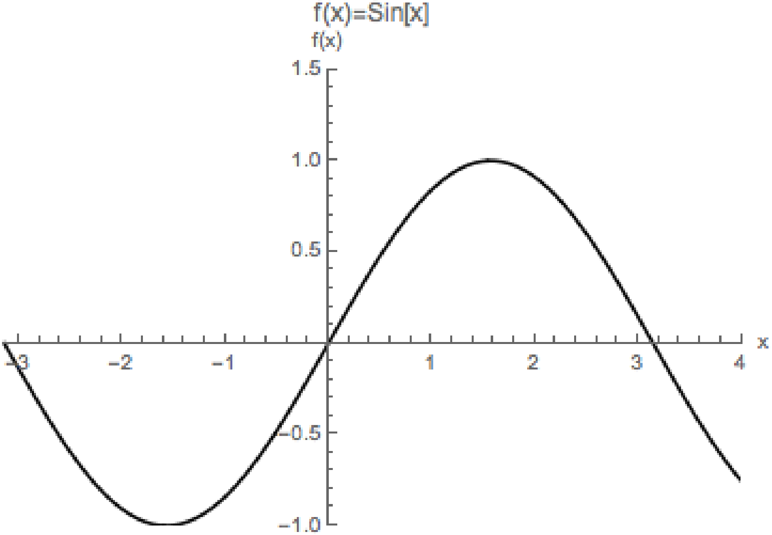
Note: I round.

Do different starting places matter?

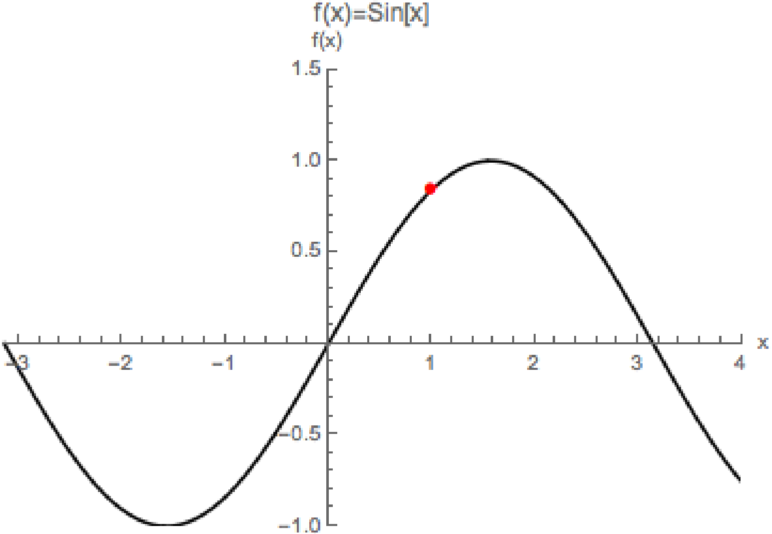
Iteration	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
0	1	0.5	2	3.5	1.4
1	-.56	-0.05	4.1	3.13	-4.40
2	0.07	3e-5	2.4	3.14	1.32
3	-1e-4	-1e-14	3.2	3.14	2.64
⋮	⋮	⋮	⋮	⋮	⋮
10	0	0	3.14...	3.14...	3.14...

A little, and a lot.

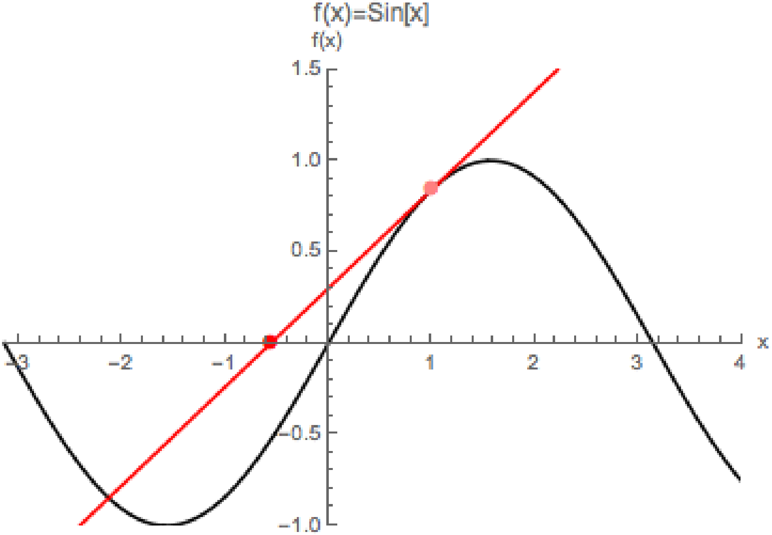
Newton's Method, Graphically-I



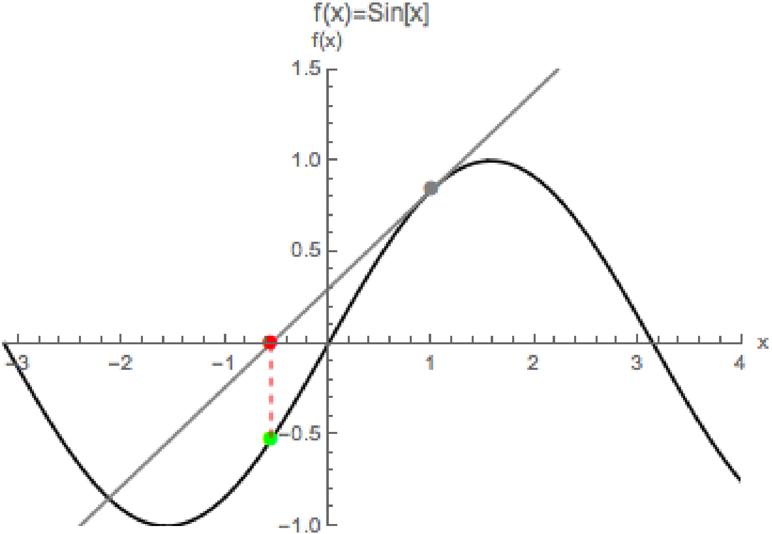
Newton's Method, Graphically-II



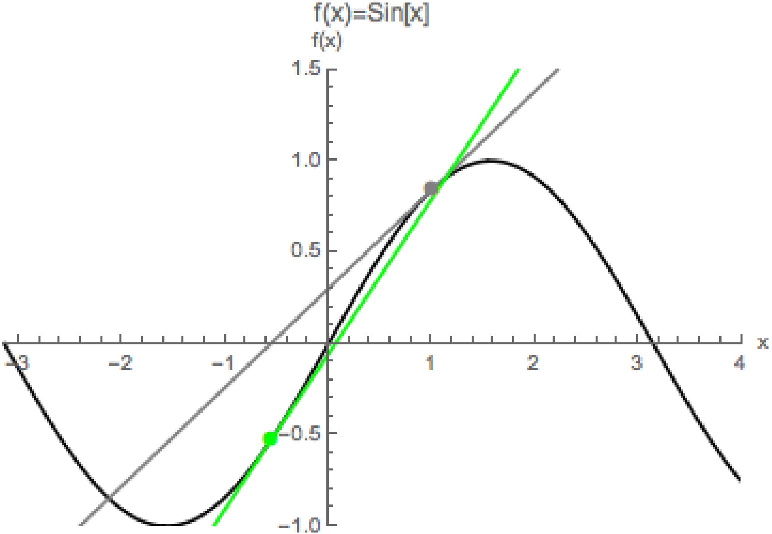
Newton's Method, Graphically-III



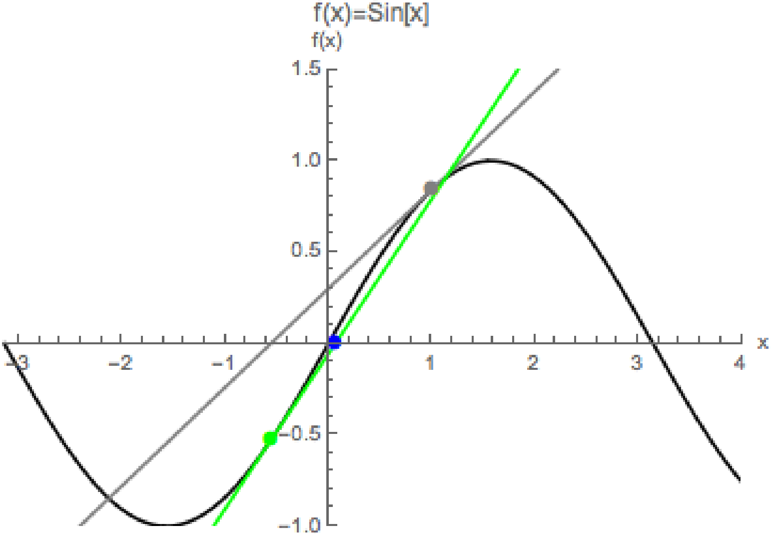
Newton's Method, Graphically-IV



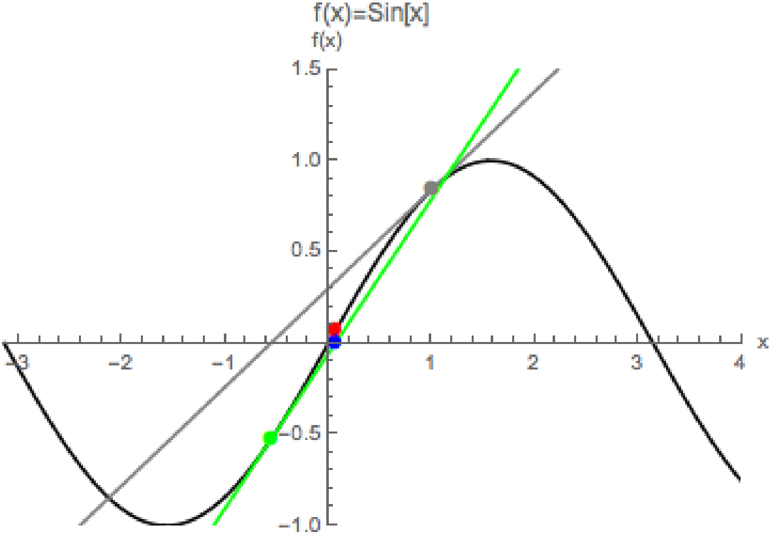
Newton's Method, Graphically-V



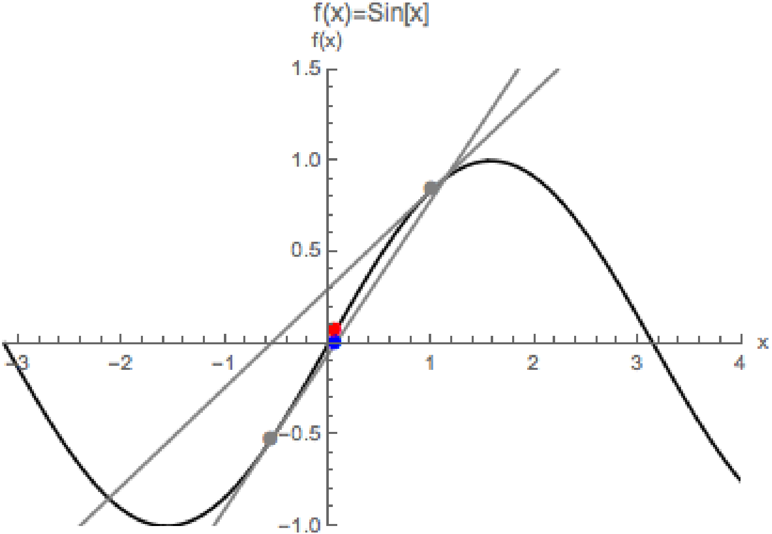
Newton's Method, Graphically-VI



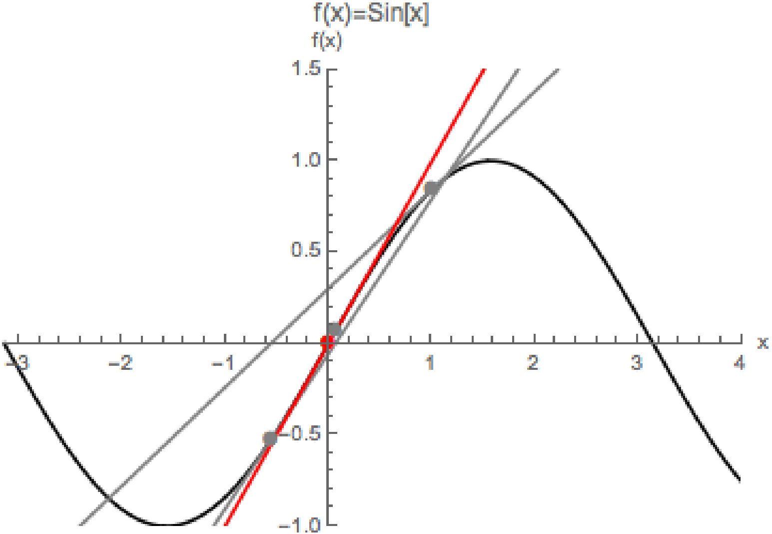
Newton's Method, Graphically-VII



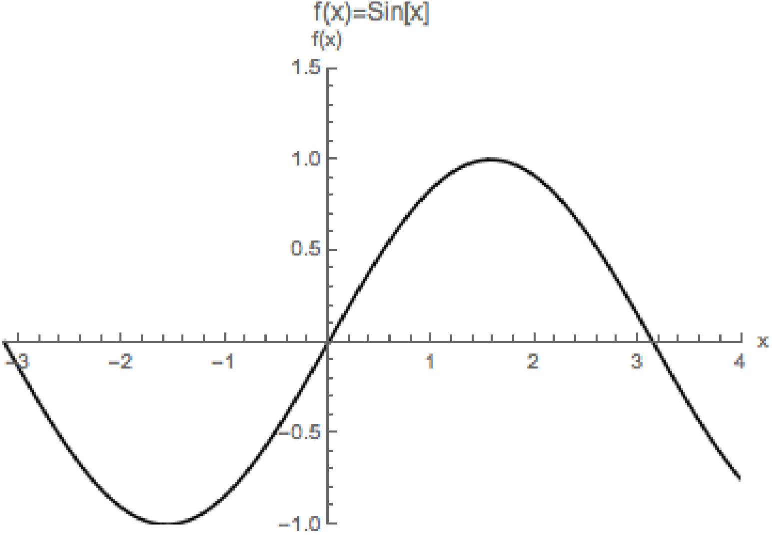
Newton's Method, Graphically-VIII



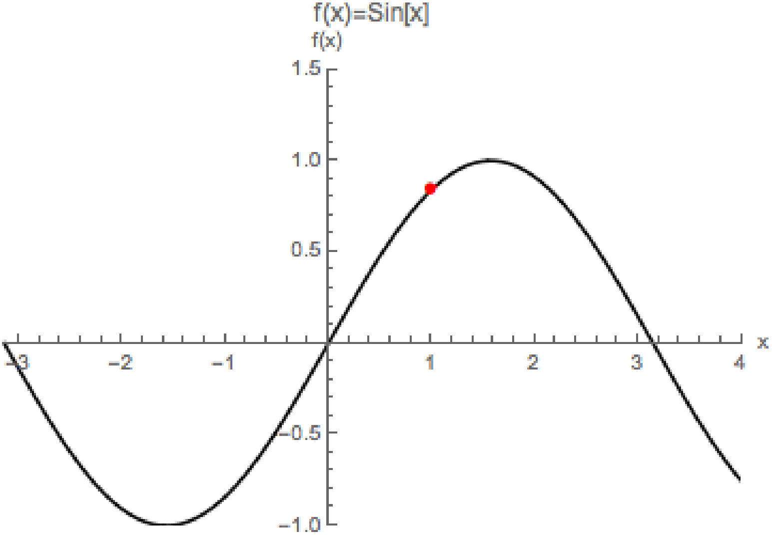
Newton's Method, Graphically-IX



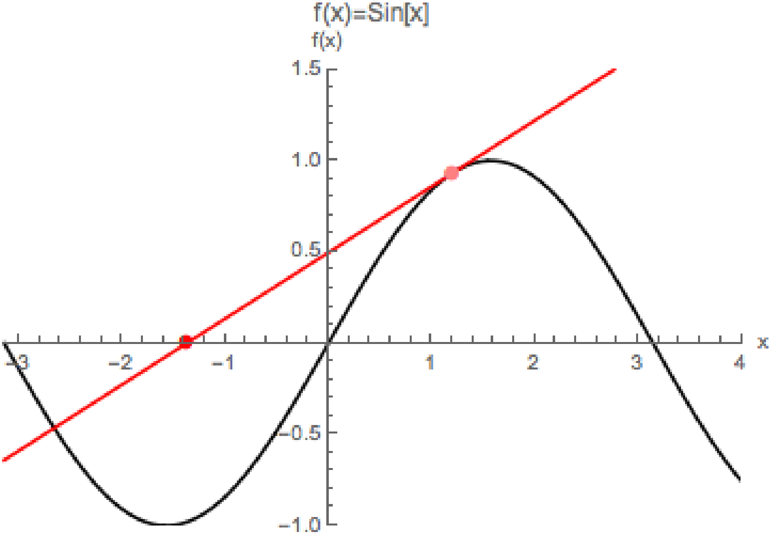
Newton's Method, Graphically, New Starting Point



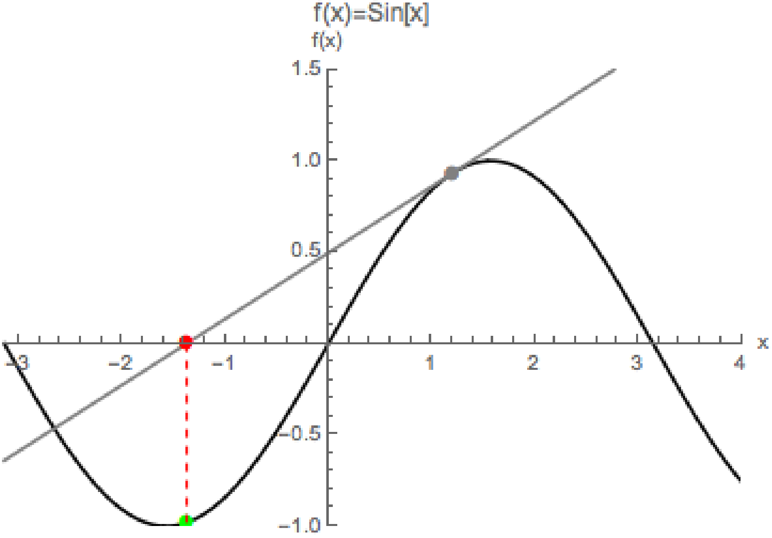
Newton's Method, Graphically-XI



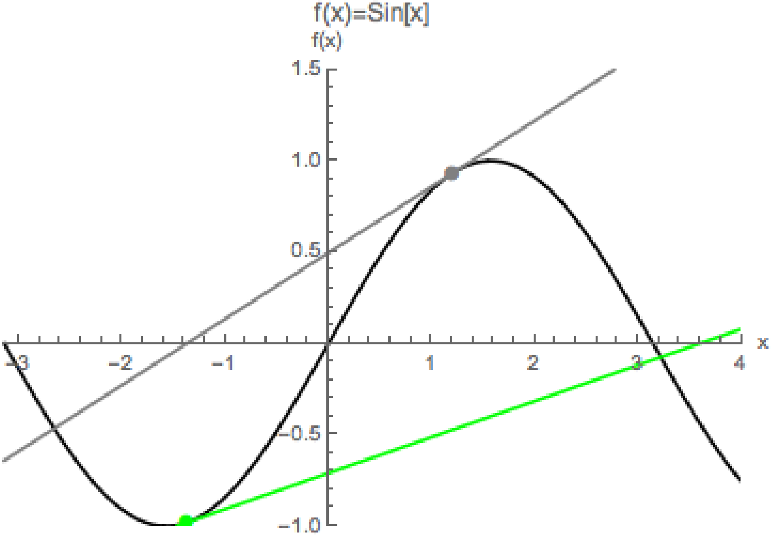
Newton's Method, Graphically-XII



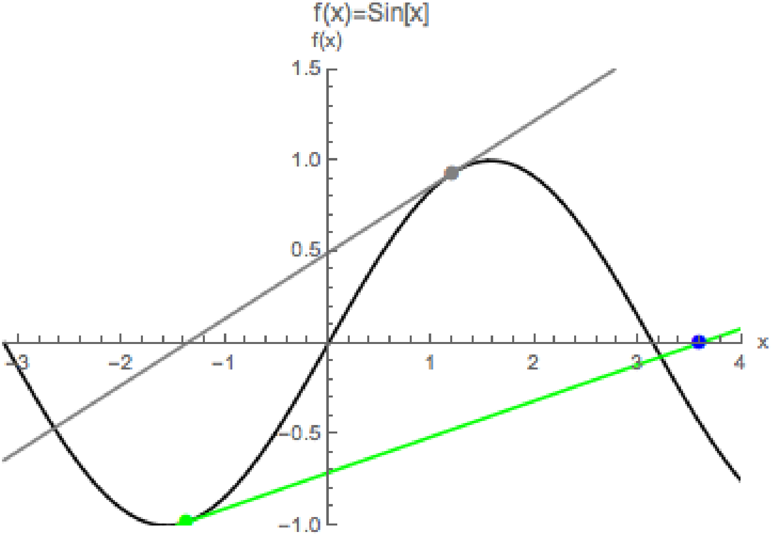
Newton's Method, Graphically-XIII



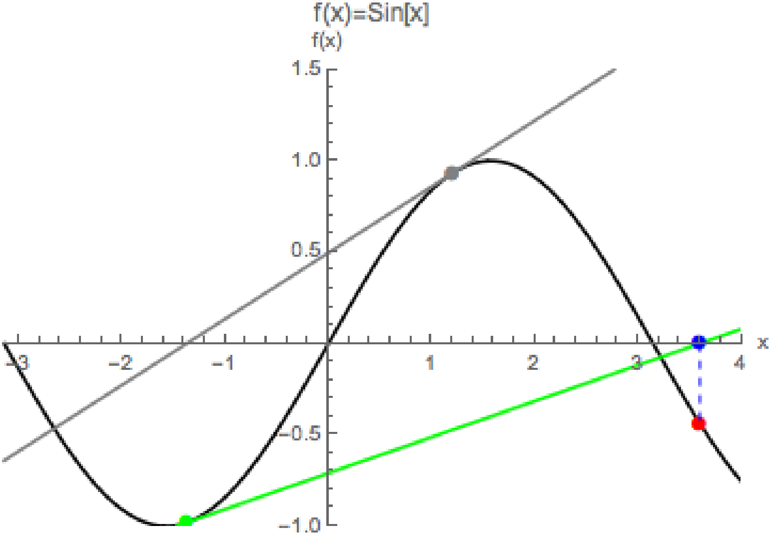
Newton's Method, Graphically-XIV



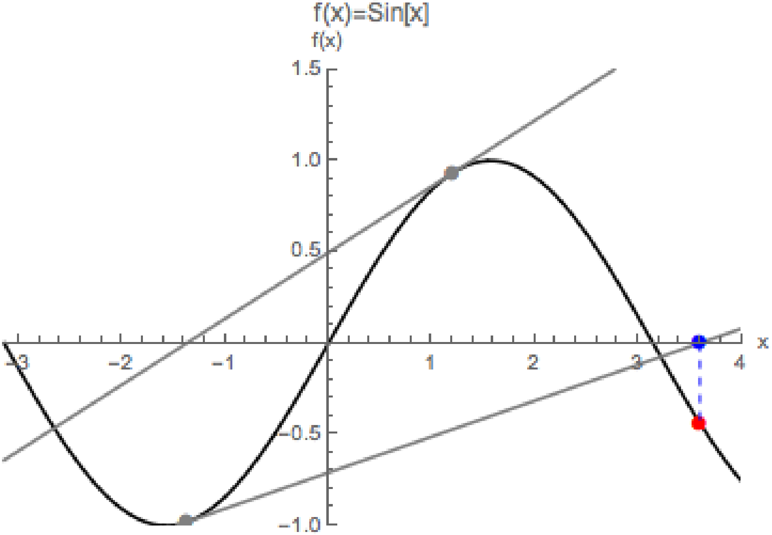
Newton's Method, Graphically-XV



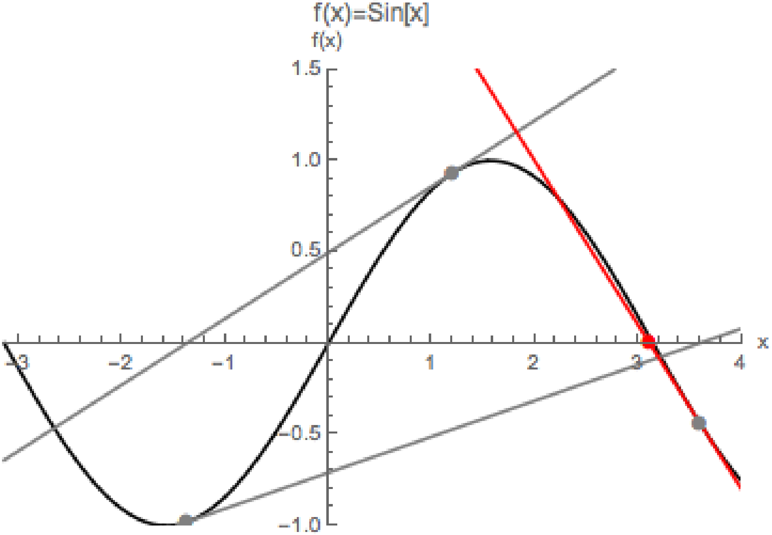
Newton's Method, Graphically-XVI



Newton's Method, Graphically-XVII



Newton's Method, Graphically-XVIII



Benefits, Costs

- Newton's method is quite rapid (locally quadratically convergent)
- Requires smoothness/derivative
- Can get trapped at local minima
- Can be sensitive to starting points

One dimension is easy...how about two?

Before, scalar equation:

$$f(x) = 0$$

Now:

$$\begin{bmatrix} f^{[1]}(x_1, x_2) \\ f^{[2]}(x_1, x_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Newton's Method: Vector of Equations

$$\begin{bmatrix} f^{[1]}(x_1, x_2) \\ f^{[2]}(x_1, x_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The Jacobian (matrix of first derivatives):

$$Df = \begin{bmatrix} \frac{\partial f^{[1]}}{\partial x_1} & \frac{\partial f^{[1]}}{\partial x_2} \\ \frac{\partial f^{[2]}}{\partial x_1} & \frac{\partial f^{[2]}}{\partial x_2} \end{bmatrix}$$

So we can take the multivariate taylor expansion, in matrix form:

$$\begin{bmatrix} f^{[1]}(x_1, x_2) \\ f^{[2]}(x_1, x_2) \end{bmatrix} \approx \begin{bmatrix} f^{[1]}(\overline{x_1}, \overline{x_2}) \\ f^{[2]}(\overline{x_1}, \overline{x_2}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f^{[1]}}{\partial x_1} & \frac{\partial f^{[1]}}{\partial x_2} \\ \frac{\partial f^{[2]}}{\partial x_1} & \frac{\partial f^{[2]}}{\partial x_2} \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \overline{x_1} \\ \overline{x_2} \end{bmatrix} \right)$$

Newton's Method: Vector of Equations

Set it equal to zero:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} f^{[1]}(\overline{x_1}, \overline{x_2}) \\ f^{[2]}(\overline{x_1}, \overline{x_2}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f^{[1]}}{\partial x_1} & \frac{\partial f^{[1]}}{\partial x_2} \\ \frac{\partial f^{[2]}}{\partial x_1} & \frac{\partial f^{[2]}}{\partial x_2} \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \overline{x_1} \\ \overline{x_2} \end{bmatrix} \right)$$

Solving for $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ assuming and inverse matrix:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \overline{x_1} \\ \overline{x_2} \end{bmatrix} - \begin{bmatrix} \frac{\partial f^{[1]}}{\partial x_1} & \frac{\partial f^{[1]}}{\partial x_2} \\ \frac{\partial f^{[2]}}{\partial x_1} & \frac{\partial f^{[2]}}{\partial x_2} \end{bmatrix}^{-1} \begin{bmatrix} f^{[1]}(\overline{x_1}, \overline{x_2}) \\ f^{[2]}(\overline{x_1}, \overline{x_2}) \end{bmatrix}$$

Or, with obvious notation for the vector X, $\bar{X}$, F, and the jacobian of F, J:

$$X_{n+1} = \bar{X}_n - J^{-1} F(\bar{X}_n)$$

Example-I

$$\begin{bmatrix} f^{[1]}(x_1, x_2) \\ f^{[2]}(x_1, x_2) \end{bmatrix} = \begin{bmatrix} \exp(x_1) - 2\exp(x_2) \\ x_1 + x_2 - 3 \end{bmatrix}$$

So we can write:

$$J(X) = \begin{bmatrix} \exp(x_1) & -2\exp(x_2) \\ 1 & 1 \end{bmatrix}$$

$$J(X)^{-1} = \frac{1}{\exp(x_1) + 2\exp(x_2)} \begin{bmatrix} 1 & 2\exp(x_2) \\ -1 & \exp(x_1) \end{bmatrix}$$

Example-II

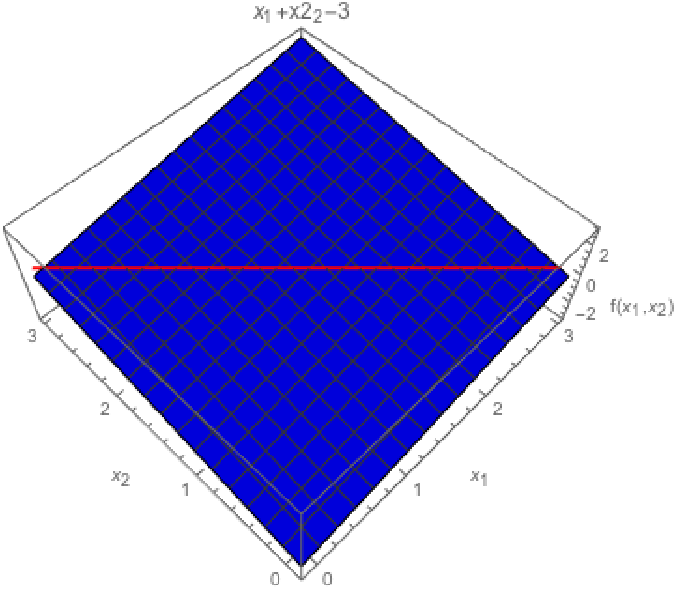
Start at $(0,0)$:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0.33 & 0.66 \\ -0.33 & 0.66 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 2.33 \\ 0.66 \end{bmatrix}$$

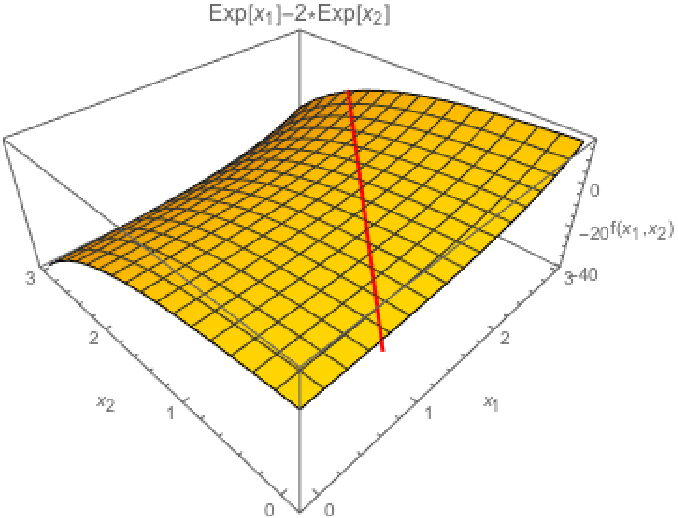
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.33 \\ 0.66 \end{bmatrix} - \begin{bmatrix} 0.07 & 0.27 \\ -0.07 & 0.72 \end{bmatrix}^{-1} \begin{bmatrix} 6.41 \\ 0 \end{bmatrix} = \begin{bmatrix} 2.88 \\ 1.11 \end{bmatrix}$$

And so on until the solution, $x_1 = 1.84$ and $x_2 = 1.15$

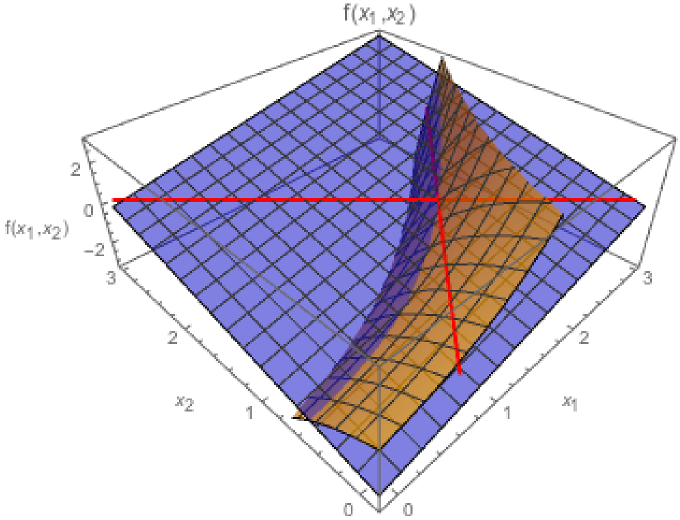
Example-II: First equation and zeros



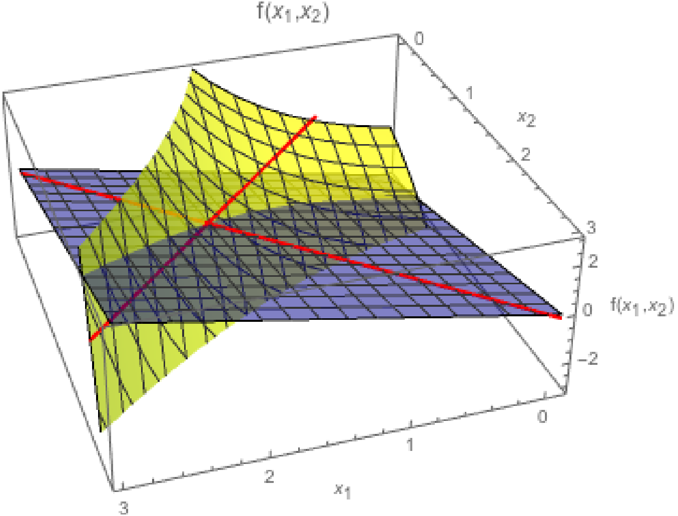
Example-II: Second equation and zeros



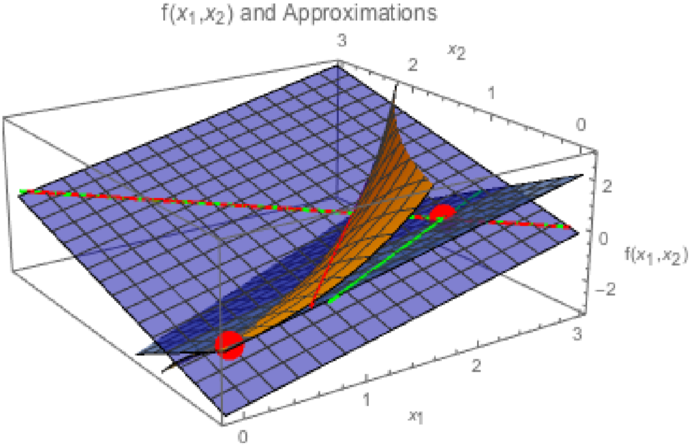
Example-II: Both Equations and zeros



Example-II: Both Equations and zeros (rotated)



Equations, approximations, and zeros



Start at (0,0), take tangent planes, find where planes intersect with zero (green lines), find where lines intersect (new point). In this case, red and green line are same because linear expansion of linear plane is plane.

Minimization

Conceptually, minimization is the same: just find the zeros of the derivative:

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

and:

$$X_{n+1} = X_n - H^{-1} f'(x_n)$$

Extensions

- 80% of the minimization and solving methods you meet will be derivative-based twists of Newton's Method
- Halley's Method, Householder Methods: Higher-order derivatives
- Quasi-Newton Methods: If you can't take J, approximate it
- Secant Method: Newton's Method with finite differences
- Gauss-Newton: Minimize sum of squared functions
- Lecture 2 talks about alternatives