# Math Camp

Justin Grimmer

Professor
Department of Political Science
Stanford University

September 6th, 2018

# Optimization

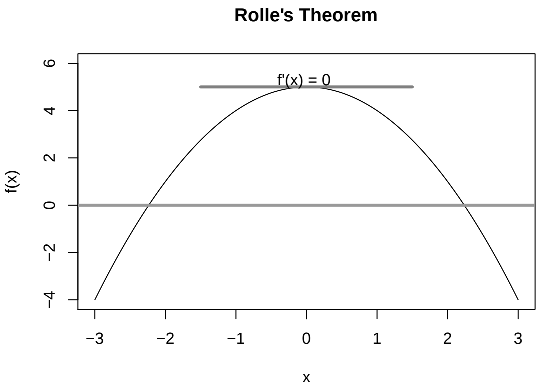
Social scientists are often concerned with finding extrema: maxima or minima

- Given data, most likely value of a parameter
- Game theory: given other player's strategy, action that maximizes utility
- Across substantive areas: what is the optimal action, strategy, prediction?

How to Optimize

- When functions are well behaved and known $\rightsquigarrow$ analytic solutions
    - Differentiate, set equal to zero, solve
    - Check end points and use second derivative test
- More difficult problems $\rightsquigarrow$ computational solutions

# Intuition: Optimization with Derivatives, Known well behaved functions




- Rolle's theorem guarantee's that, at some point, $f^{'}(x) = 0$
- Intuition from proof—what happens as we approach from the left?
- Intuition from proof—what happens as we approach from the right?
- critical intuition first, second derivatives

# Second Derivatives

**Definition**

*Suppose* $f : \Re \rightarrow \Re$ *is differentiable. Recall we write this as* $f^{'}$ *and suppose that* $f^{'} : \Re \rightarrow \Re$*. Then if the limit,*

$$\lim_{x \rightarrow x_0} R(x) = \frac{f^{'}(x) - f^{'}(x_0)}{x - x_0}$$

*exists, we call this the second derivative at* $x_0, f^{''}(x_0)$*.*

# Example of Second Derivatives

$$\begin{aligned} f(x) &= -x^2 + 20 \\ f^{'}(x) &= -2x \\ f^{''}(x) &= -2 \end{aligned}$$

# Approximating functions and second order conditions

Theorem

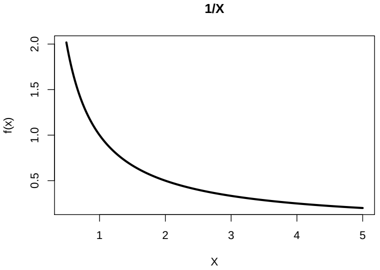
**Taylor's Theorem** *Suppose* $f : \Re \rightarrow \Re$, $f(x)$ *is infinitely differentiable function. Then, the taylor expansion of* $f(x)$ *around* $a$ *is given by*

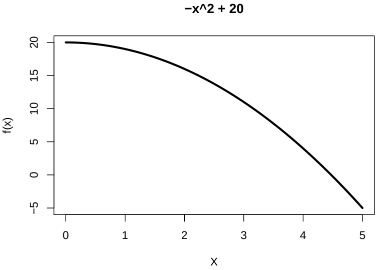
$$\begin{aligned} f(x) &= f(a) + \frac{f^{'}(a)}{1!}(x-a) + \frac{f^{''}(a)}{2!}(x-a)^2 + \frac{f^{'''}(a)}{3!}(x-a)^3 + \ldots \\ f(x) &= \sum_{n=0}^{\infty} \frac{f^n(a)}{n!}(x-a)^n \end{aligned}$$

R Code!

# Concavity, Convexity, Inflections

Second derivatives provide further information about functions



# Concave Up/ Convex

Definition

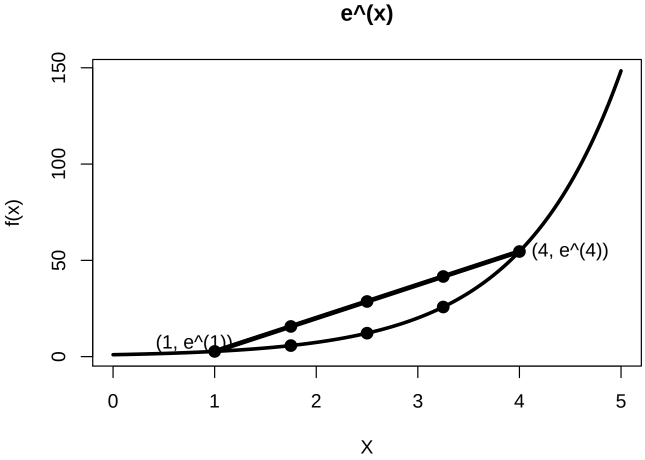
*Suppose* $f : [a, b] \rightarrow \Re$ *is a twice differentiable function. If, for all* $x \in [a, b]$ *and* $y \in [a, b]$ *and* $t \in (0, 1)$

$$f((1-t)x + ty) \quad < \quad (1-t)f(x) + tf(y)$$

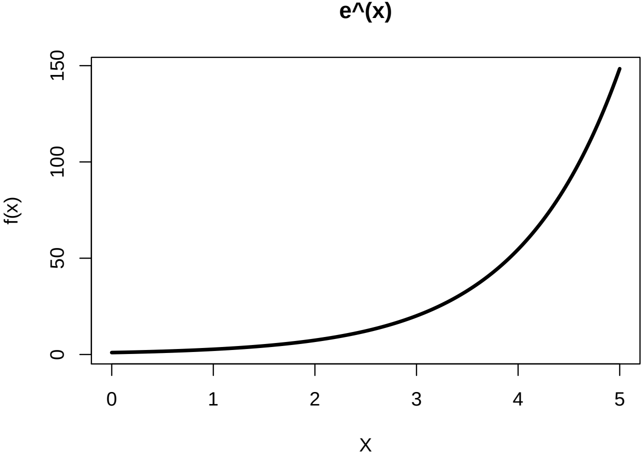
*We say that* $f$ *is strictly concave up or convex. Equivalently if* $f^{''}(x) > 0$ *for all* $x \in [a, b]$*, we say that* $f$ *is strictly concave up.*

# Concave Up, Graphical Test

$f(x) = e^x$, $[1, 4]$



# Concave Up, Second Derivative




$$\begin{aligned} f(x) &= e^x \\ f^{'}(x) &= e^x \\ f^{''}(x) &= e^x \end{aligned}$$

$e^x > 0$ for all $x \in [1, 4]$

# Concave Down

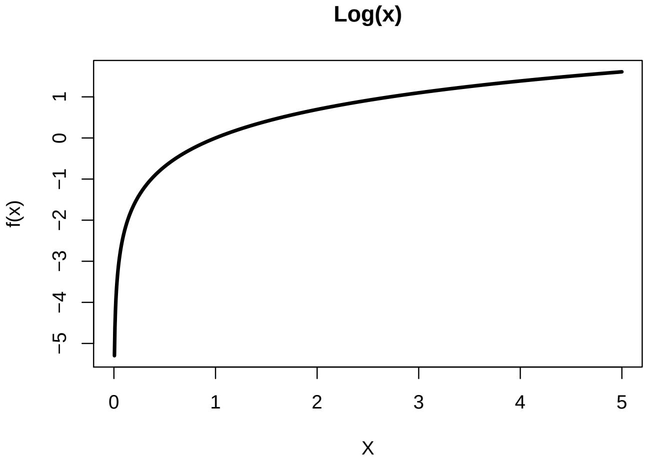
**Definition**

*Suppose* $f : [a, b] \rightarrow \Re$ *is a twice differentiable function. If, for all* $x \in [a, b]$ *and* $y \in [a, b]$ *and* $t \in (0, 1)$

$$f((1-t)x + ty) \quad > \quad (1-t)f(x) + tf(y)$$

*We say that* $f$ *is strictly concave down. Equivalently if* $f^{''}(x) < 0$ *for all* $x \in [a, b]$*, we say that* $f$ *is strictly concave down.*

# Concave Down




- Show Concave down with graph test for $x \in [1, 4]$
- Show concave down with second derivative test for $x \in [1, 4]$

# Optimization

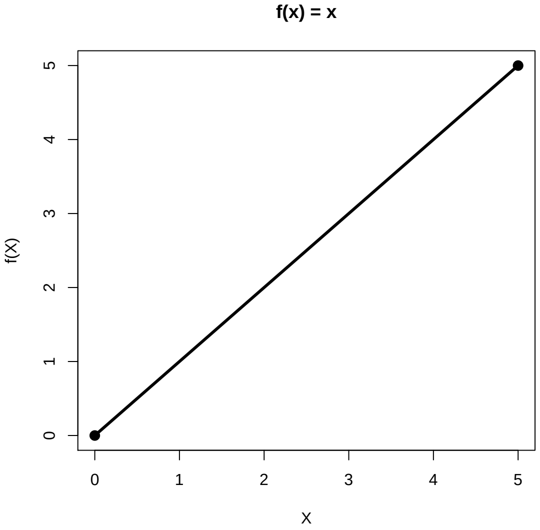
**Theorem**

*Extreme Value Theorem Suppose $f : [a, b] \rightarrow \Re$ and that $f$ is continuous. Then $f$ obtains its extreme value on $[a, b]$.*

**Corollary**

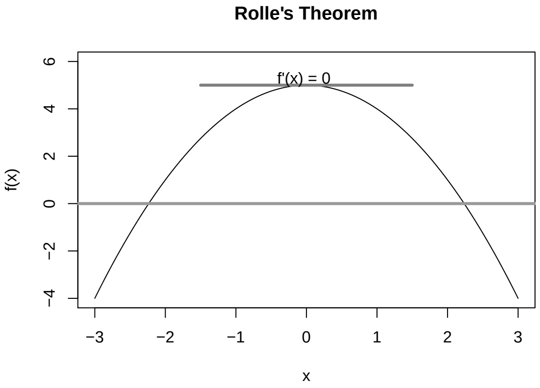
*Suppose $f : [a, b] \rightarrow \Re$, that $f$ is continuous and differentiable, and that $f(a)$ nor $f(b)$ is the extreme value. Then $f$ obtains its maximum on $(a, b)$ and if $f(x_0)$ is the extreme value of $f$ $x_0 \in (a, b)$ then, $f^{'}(x_0) = 0$.*

# Extrema on End Points



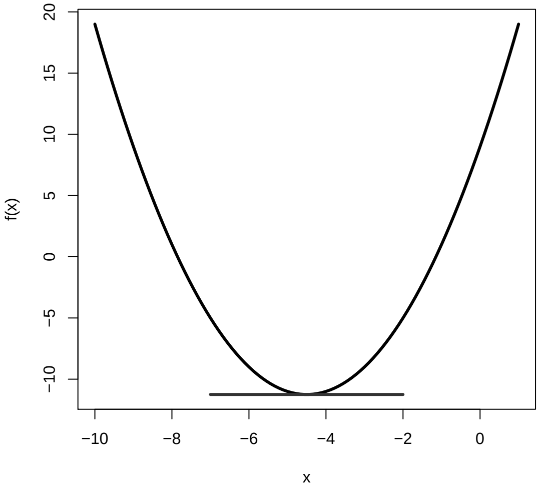
# Maximum in Middle, Concave Down

$f(x) = -x^2 + 5.$



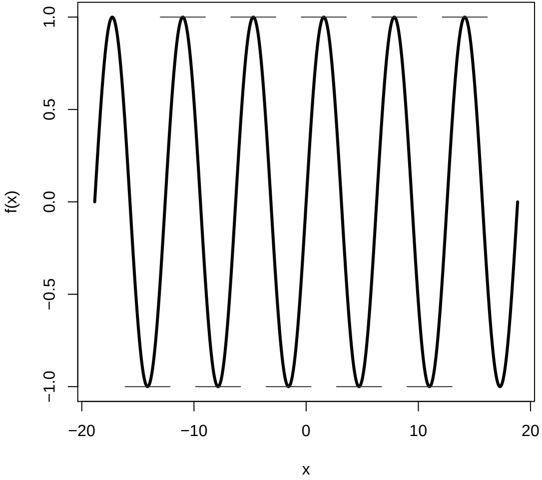
# Minimum in Interior, Concave Up

$f(x) = x^2 + 9x + 9$



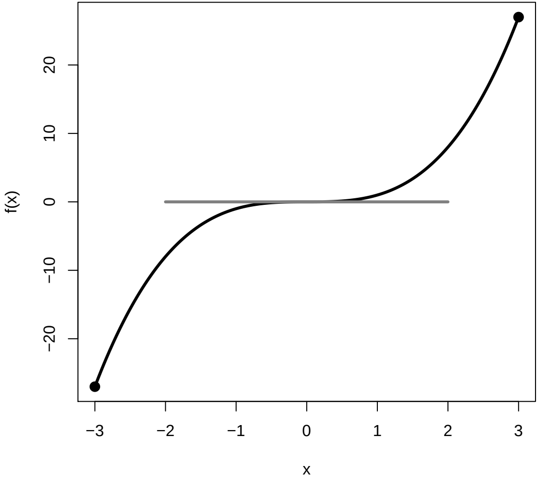
# Local Optima

$f(x) = \sin(x)$



# Inflection points

$f(x) = x^3$



# Framework for Optimization

Recipe for optimization

- Find $f^{'}(x)$.
- Set $f^{'}(x) = 0$ and solve for $x$. Call all $x_0$ such that $f^{'}(x_0) = 0$ critical values.
- Find $f^{''}(x)$. Evaluate at each $x_0$.
    - If $f^{''}(x) > 0$, Concave up, local minimum
    - If $f^{''}(x) < 0$, Concave down, local maximum
    - If $f^{''}(x) = 0$, No knowledge—local minimum, maximum, or inflection point
- Check End Points!

# Example 1: $f(x) = -x^2$, $x \in [-3, 3]$

1) Critical Value:

$$
\begin{aligned}
f^{'}(x) &= -2x \\
0 &= -2x^{*} \\
x^{*} &= 0
\end{aligned}
$$

2) Second Derivative:

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

$f^{''}(x) < 0$, local maximum

# Example 1: $f(x) = -x^2$, $x \in [-3, 3]$

3) Check end points

$$\begin{aligned} f(0) &= -0^2 = 0 \\ f(-3) &= -(-3)^2 = -9 \\ f(3) &= -(3)^2 = -9 \end{aligned}$$

# Example 2: $f(x) = x^3$, $x \in [-3, 3]$

1) Critical Values:

$$\begin{aligned} f^{'}(x) &= 3x^2 \\ 0 &= 3(x^*)^2 \\ x^* &= 0 \end{aligned}$$

2) Second Derivative:

$$\begin{aligned} f^{''}(x) &= 6x \\ f^{''}(0) &= 0 \end{aligned}$$

No information

# Example 2: $f(x) = x^3$, $x \in [-3, 3]$

3) Check End Points:

$$\begin{aligned} f(0) &= 0^3 = 0 \\ f(-3) &= -3^3 = -27 \\ f(3) &= 3^3 = 27 \end{aligned}$$

Neither maximum nor minimum, saddle point

# Example 3: Spatial Model

A large literature in Congress supposes legislators and policies can be situated in policy space

Suppose legislator $i$ and policies $x, i \in \Re$.

Define legislator $i$'s utility as, $U : \Re \rightarrow \Re$,

$$\begin{aligned} U_i(x) &= -(x-\mu)^2 \\ U_i(x) &= -x^2 + 2x\mu - \mu^2 \end{aligned}$$

What is $i$'s optimal policy over the range $x \in [\mu - 2, \mu + 2]$?

$$\begin{aligned} U_i^{'}(x) &= -2(x-\mu) \\ 0 &= -2x^{*} + 2\mu \\ x^{*} &= \mu \end{aligned}$$

Second Derivative Test

$$U_i^{''}(x) = -2 < 0 \rightarrow \text{Concave Down}$$

We call $\mu$ legislator $i$'s ideal point

# Example 3: Spatial Model

$$\begin{aligned} U_i(\mu) &= -(\mu - \mu)^2 = 0 \\ U_i(\mu - 2) &= -(\mu - 2 - \mu)^2 = -4 \\ U_i(\mu + 2) &= -(\mu + 2 - \mu)^2 = -4 \end{aligned}$$

Maximize utility at $\mu$

# Example 4: Maximum Likelihood Estimation

In statistics classes we'll learn about parameters from data.
Here is an example likelihood function: We want to find the Maximum likelihood estimate

$$
\begin{aligned}
f(\mu) &= \prod_{i=1}^{N} \exp\left(\frac{-(Y_i - \mu)^2}{2}\right) \\
&= \exp\left(-\frac{(Y_1 - \mu)^2}{2}\right) \times \ldots \times \exp\left(-\frac{(Y_N - \mu)^2}{2}\right) \\
&= \exp\left(-\frac{\sum_{i=1}^{N}(Y_i - \mu)^2}{2}\right)
\end{aligned}
$$

Theorem

*Suppose* $f : \Re \rightarrow (0, \infty)$*. If* $x_0$ *maximizes* $f$*, then* $x_0$ *maximizes* $\log(f(x))$*.*

# Example 4: Maximum LIkelihood Estimation

$$
\begin{aligned}
\log f(\mu) &= \log\left(\exp(-\frac{\sum_{i=1}^{N}(Y_i-\mu)^2}{2})\right) \\
&= -\frac{\sum_{i=1}^{N}(Y_i-\mu)^2}{2}) \\
&= -\frac{1}{2}\left(\sum_{i=1}^{N} Y_i^2 - 2\mu\sum_{i=1}^{N} Y_i + N\times\mu^2\right) \\
\frac{\partial \log f(\mu)}{\partial\mu} &= -\frac{1}{2}\left(-2\sum_{i=1}^{N} Y_i + 2N\mu\right)
\end{aligned}
$$

# Example 4: Maximum Likelihood Estimation

$$
\begin{aligned}
0 &= -\frac{1}{2}\left(-\sum_{i=1}^{N} Y_i + 2N\mu^*\right) \\
2\sum_{i=1}^{N} Y_i &= 2N\mu^* \\
\frac{\sum_{i=1}^{N} Y_i}{N} &= \mu^* \\
\bar{Y} &= \mu^*
\end{aligned}
$$

Second Derivative Test

$$
\begin{aligned}
f^{'}(\mu) &= -\frac{1}{2}\left(-2\sum_{i=1}^{N} Y_i + 2N\mu\right) \\
f^{''}(\mu) &= -N
\end{aligned}
$$

# Example 5: IR Bargaining (from Jim Fearon, Part 1)

Countries fight wars, usually to get stuff.

- Suppose two countries $1, 2$ are fighting for something they value at $v$.
- Each country decides to invest $a_1 \in [0,1]$ and $a_2 \in [0,1]$.
- The probability of country 1 winning the war is

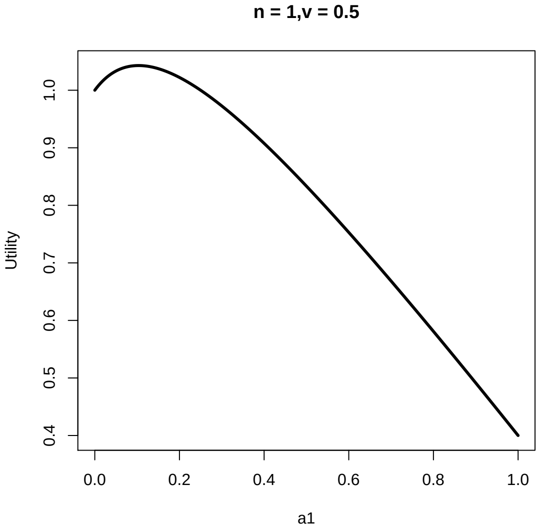
$$p(a_1, a_2) = \frac{a_1^n}{a_1^n + a_2^n}$$

- Country 1's utility is given by

$$\begin{aligned} U_1(a_1) &= \underbrace{1 - a_1}_{\text{cost}} + \underbrace{p(a_1, a_2)v}_{\text{Expected Benefit}} \\ &= 1 - a_1 + \frac{a_1^n}{a_1^n + a_2^n} v \end{aligned}$$

- Suppose country 2 selected value $x$. What should country 1 invest to maximize utility?

# Example 5: IR Bargaining (from Jim Fearon, Part 1)



# Example 5: IR War (from Jim Fearon, Part 1)

$$\begin{aligned}\frac{\partial U_1(a_1)}{\partial a_1} &= -1 + \frac{na_1^{n-1}(a_1^n + x^n) - (na_1^{n-1}a_1^n)}{(a_1^n + x^n)^2}v \\ &= -1 + \frac{na_1^{n-1}x^n}{(a_1^n + x^n)^2}v\end{aligned}$$

Set $n = 1$ (for simplicity)

$$\begin{aligned}0 &= -1 + \frac{x}{(a_1 + x)^2}v \\ a_1^* &= \sqrt{v}\sqrt{x} - x\end{aligned} \tag{0.1}$$

Second derivative!

$$U_1^{''}(a_1) = \frac{-2vx}{(a_1 + x)^3}$$

# Example 5: IR Bargaining (from Jim Fearon, Part 1)

One more—check endpoints

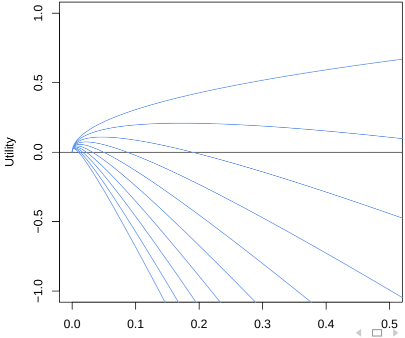
$$
\begin{aligned}
a_1^* &= 0, \text{ if } \sqrt{v}\sqrt{x} - x < 0 \\
a_1^* &= 0, \text{ if } \sqrt{v} < \sqrt{x} \\
a_1^* &= \sqrt{v}\sqrt{x} - x \text{ otherwise}
\end{aligned}
$$

# Optimization Challenge Problem

- Suppose a candidate is attempting to mobilize voters. Suppose that for each investment of $x \in [0, \infty)$ the candidate receives return of $x^{1/2}$, but incurs cost of $ax$. So, candidate utility is,

$$U_i = x^{1/2} - ax$$

What is the optimal investment $x^{*}$?



# Computational Optimization Approaches

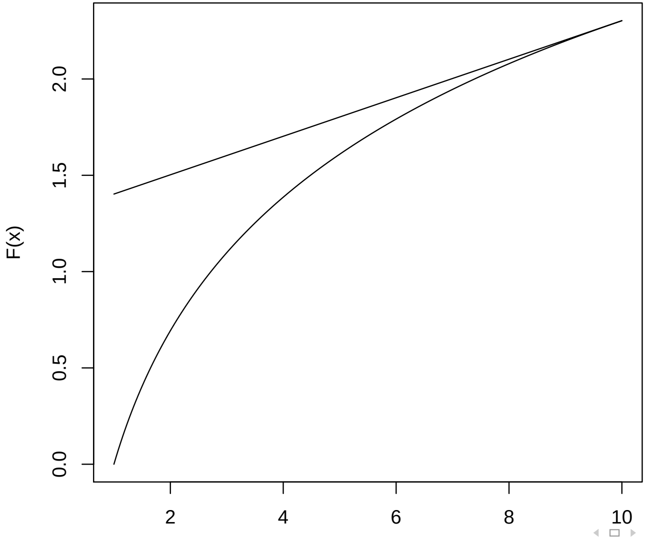
Analytic (Closed form) $\rightsquigarrow$ Often difficult, impractical, or unavailable
Computational $\rightsquigarrow$ iterative algorithm that converges to a solution (hopefully the right one!)

- Methods for optimization:
  - Newton's method and related methods
  - Gradient descent (ascent)
  - Expectation Maximization
  - Genetic Optimization
  - Branch and Bound ...

# Newton-Raphson Method

Iterative procedure to find a root
Often solving for $x$ when $f(x) = 0$ is hard$\rightsquigarrow$ complicated function
Solving for $x$ when $f(x)$ is linear$\rightsquigarrow$ easy
Approximate with tangent line, iteratively update

# Tangent Line



# Tangent Line

Formula for Tangent line at $x_0$:

$$g(x) = f^{'}(x_0)(x - x_0) + f(x_0)$$

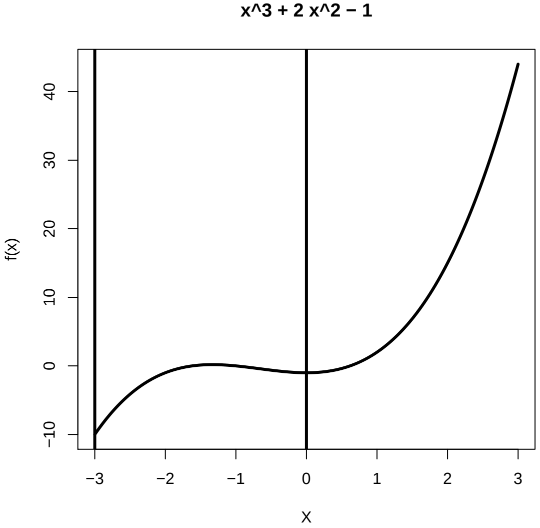
# Newton-Raphson Method

Suppose we have some initial guess $x_0$. We're going to approximate $f^{'}(x)$ with the tangent line to generate a new guess

$$\begin{aligned} g(x) &= f^{''}(x_0)(x - x_0) + f^{'}(x_0) \\ 0 &= f^{''}(x_0)(x_1 - x_0) + f^{'}(x_0) \\ x_1 &= x_0 - \frac{f^{'}(x_0)}{f^{''}(x_0)} \end{aligned}$$

# Example Function

$f(x) = x^3 + 2x^2 - 1$ find $x$ that maximizes $f(x)$ with $x \in [-3, 0]$

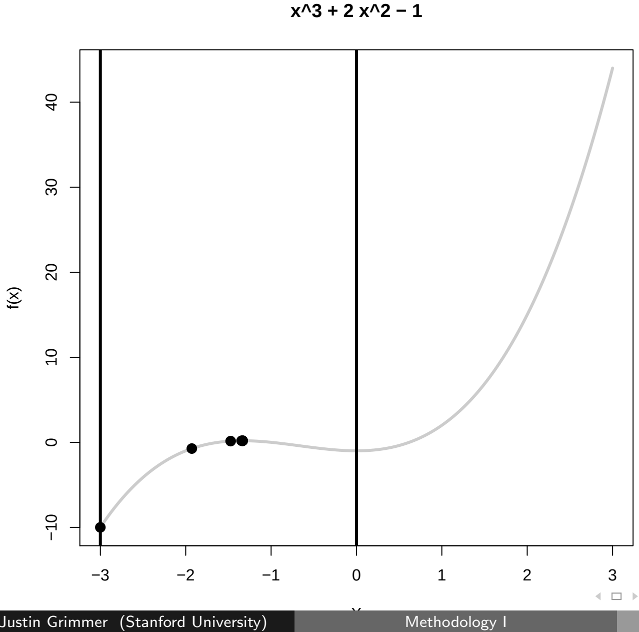


$$
\begin{aligned}
f^{'}(x) &= 3x^2 + 4x \\
f^{''}(x) &= 6x + 4
\end{aligned}
$$

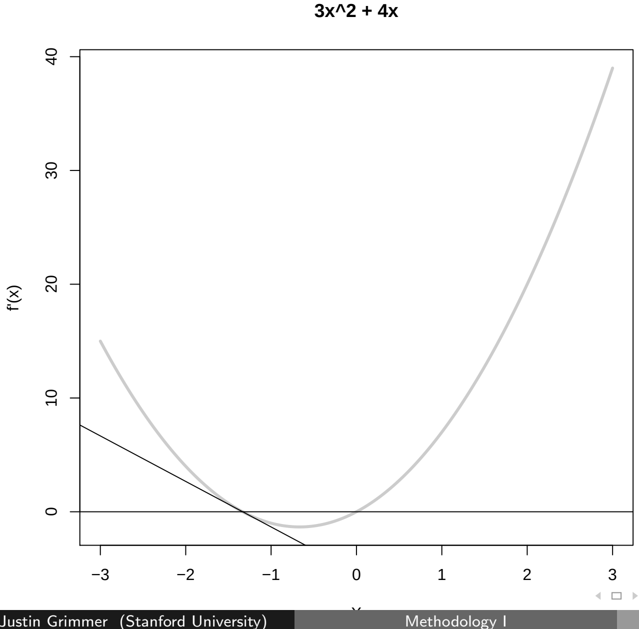
Suppose we have guess $x_t$ then the next step is:

$$
x_{t+1} = x_t - \frac{3x_t^2 + 4x_t}{6x_t + 4}
$$

$x^* = -1.3333$



# What is Happening with the Roots



To the R Code!

# Today/Tomorrow

- A Framework for optimization
  - Analytic: pencil and paper math
  - Computational: iterative algorithm that aids in solution
- Integration: antidifferentation/area finding