INFO251 – Applied Machine Learning

Lab 5
Emily Aiken

Announcements

- PS2 grades released

- PS3 solution will be posted tomorrow

- PS4 posted, due Monday March 14. Start early!

- Quiz 1 on Tuesday March 1, 9:40am-10:20am
  - Via bcourses, you can take it in class or from home
  - Course staff will be available in person and on zoom for clarifications
  - Multiple choice and simple short answer
  - Closed book
  - **Contact us ASAP if you need an alternate quiz time or have a DSP accommodation**

Topics

- Optimization

- Convexity

- Gradient descent
  - Random initialization, learning rate, iterations, stopping conditions

Optimization

- **Optimization**: Finding the **global minimum** of a function

- Methods for optimization
  - Naïve grid search
  - Gradient descent
  - Linear programming, quadratic programming
  - Newton's method
  - …many, many more

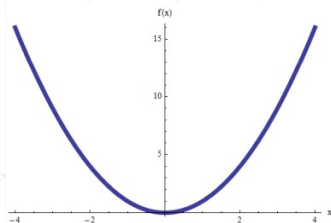- More on optimization**: EECS 127** and **EECS 227**

Convexity

- **Convex function:** Second derivative is always nonnegative

- **Graphical interpretation**: Line segment between any two points on the graph of the function does not lie below the graph between the two points
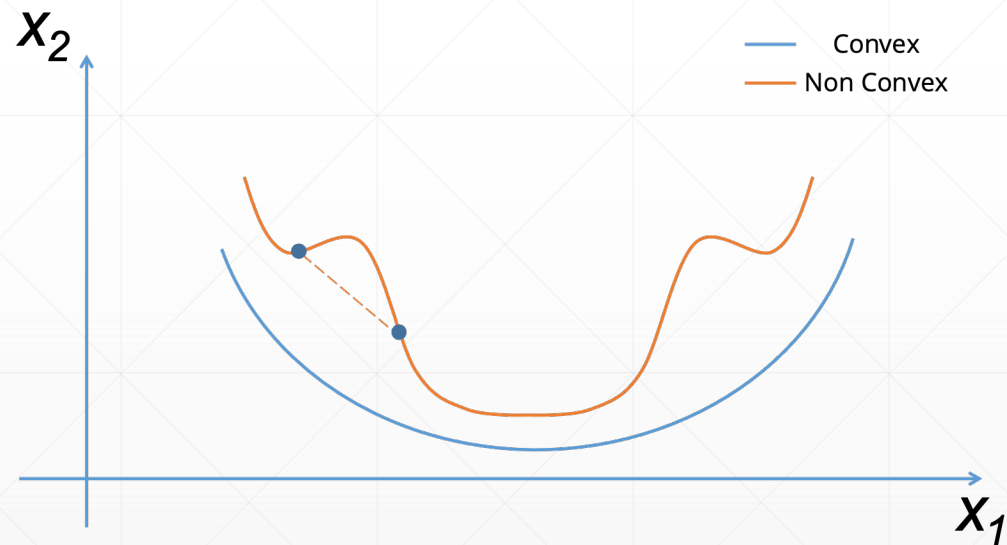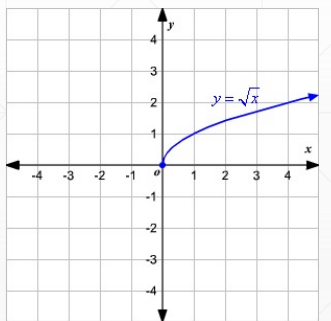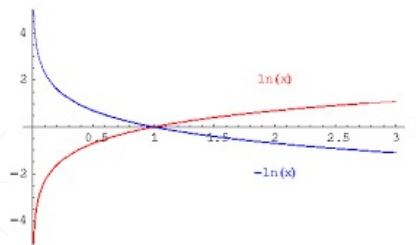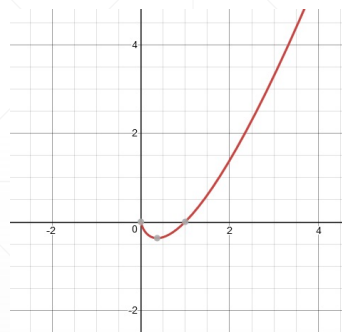
Convexity

$f(x) = x^2$

$f(x) = \ln(x)$

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

$f(x) = x^3$
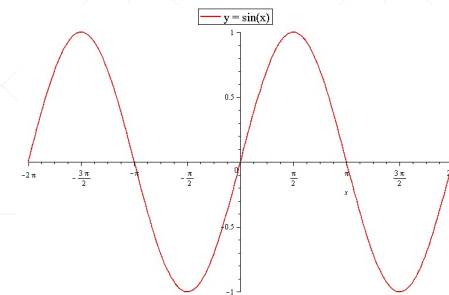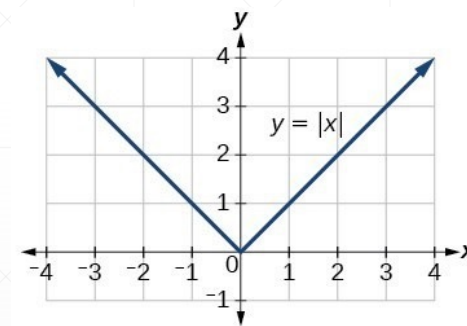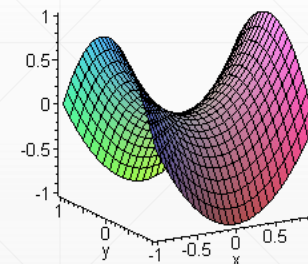
$f(x) = x\ln(x)$

$f(x) = |x|$

$f(x) = x^{1/2}$
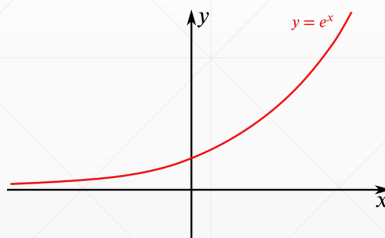
$f(x) = e^x$

$f(x) = ax^2 - by^2$

Gradient Descent

1. Begin at a random point

2. Calculate the function value at the point and the gradient (partial derivatives)

3. Pick a new point, by moving in the direction of the gradient. The size of the step is governed by the **learning rate**.

4. Repeat!

$$\mathbf{b} = \mathbf{a} - \gamma \, \nabla \, \mathbf{f}(\mathbf{a})$$