INFO251 – Applied Machine Learning

Lab 5

Suraj R. Nair

Announcements

- PS3 due next week
- Quiz 1 on Feb 27
- Lab schedule:
  - Today: Cross Validation, Normalization, Standardization + Gradient Descent Demo
  - Feb 21: Gradient Descent (~20 mins) + Review for Quiz 1 (~30 min)

Topics

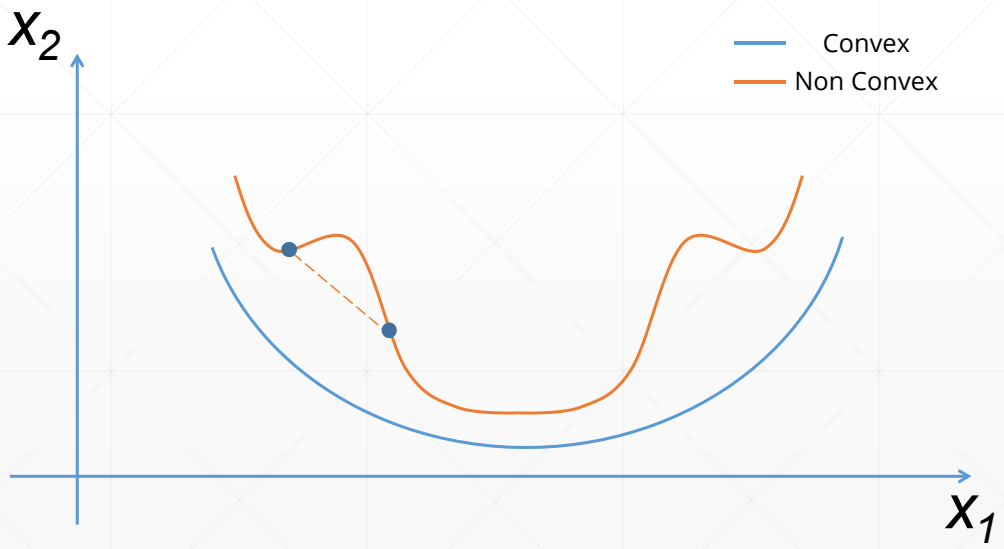
- Optimization
- Convexity
- Gradient descent
  - Random initialization, learning rate, iterations, stopping conditions

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**

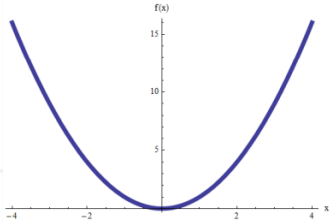
Convexity

- **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



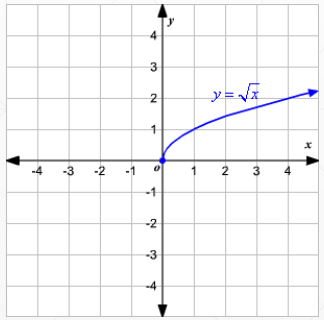
Convexity

$f(x) = x^2$

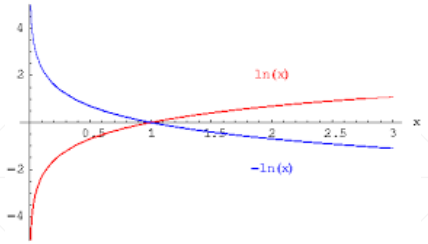


$f(x) = x^3$

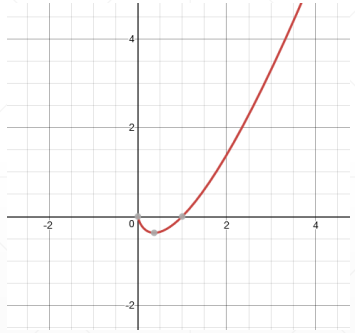
$f(x) = x^{1/2}$



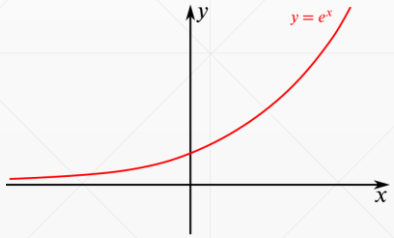
$f(x) = \ln(x)$



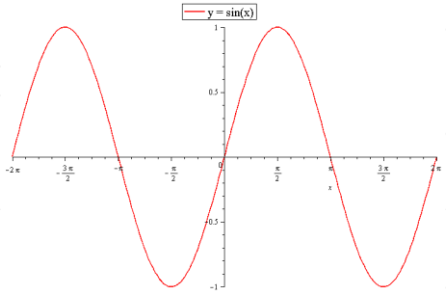
$f(x) = x\ln(x)$



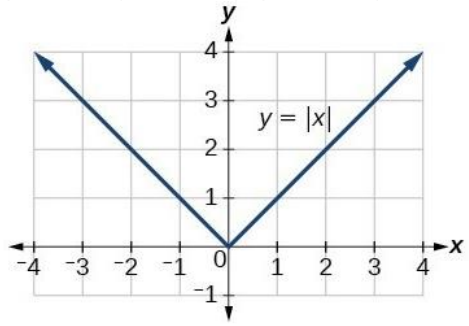
$f(x) = e^x$



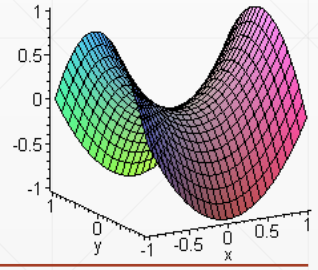
$f(x) = \sin(x)$



$f(x) = |x|$



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



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!

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