10-601 Introduction to Machine Learning

Machine Learning Department
School of Computer Science
Carnegie Mellon University

Linear Regression / Optimization for ML

Matt Gormley
Lecture 7
Sep. 19, 2018

Q&A

Reminders

- **Homework 2: Decision Trees**
  - **Out: Wed, Sep 05**
  - **Due: Wed, Sep 19 at 11:59pm**
- **Homework 3: KNN, Perceptron, Lin.Reg.**
  - **Out: Wed, Sep 19**
  - **Due: Wed, Sep 26 at 11:59pm**
- Matt's office hours on Thu are cancelled this week

LINEAR REGRESSION AS FUNCTION APPROXIMATION

Regression Problems

Chalkboard

- Definition of Regression
- Linear functions
- Residuals
- Notation trick: fold in the intercept

Linear Regression as Function Approximation

Chalkboard

- Objective function: Mean squared error
- Hypothesis space: Linear Functions

OPTIMIZATION IN CLOSED FORM

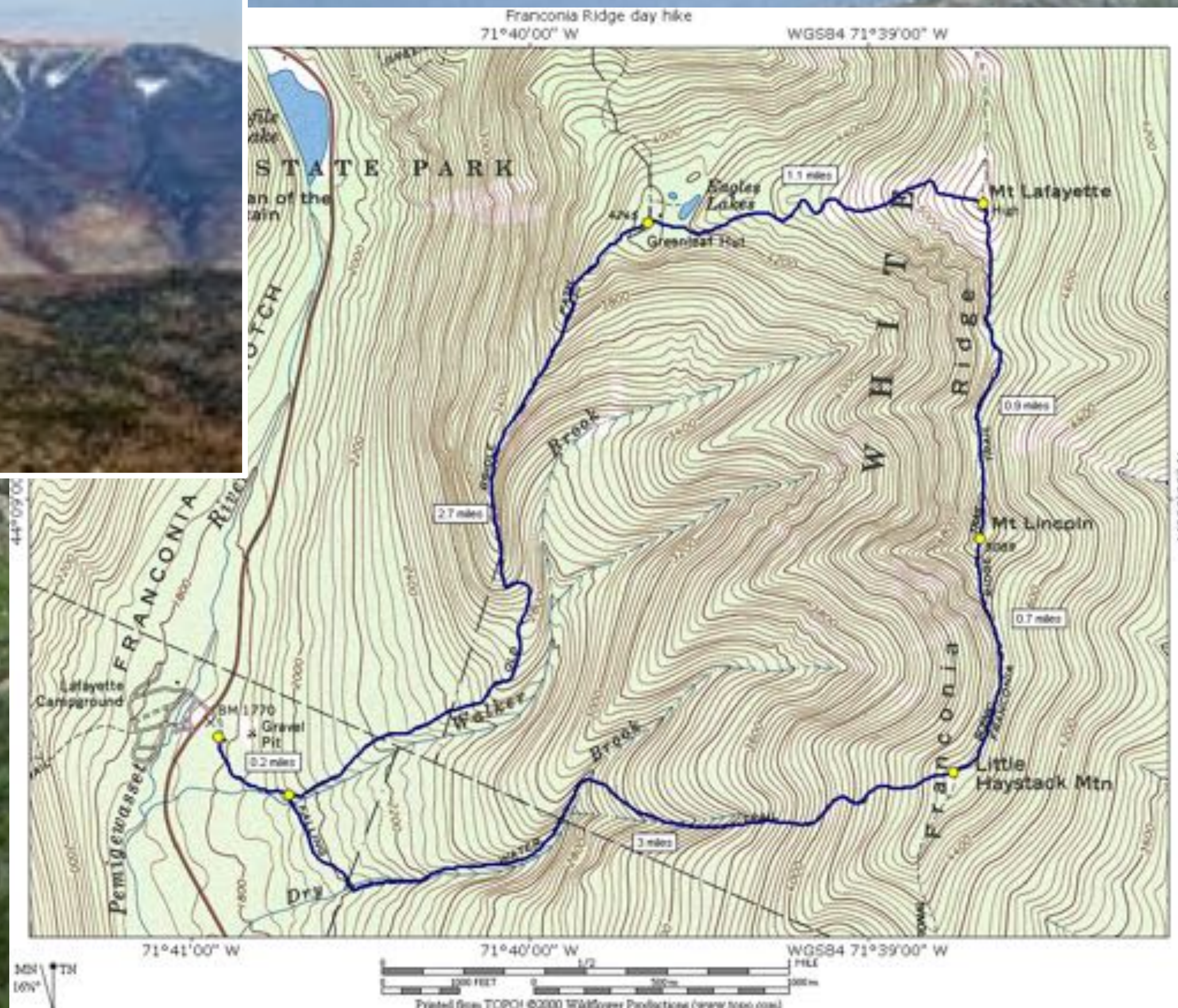
Optimization for ML

Not quite the same setting as other fields…

- Function we are optimizing might not be the true goal
  (e.g. likelihood vs generalization error)
- Precision might not matter
  (e.g. data is noisy, so optimal up to 1e-16 might not help)
- Stopping early can help generalization error
  (i.e. “early stopping” is a technique for regularization – discussed more next time)

Topographical Maps

Topographical Maps



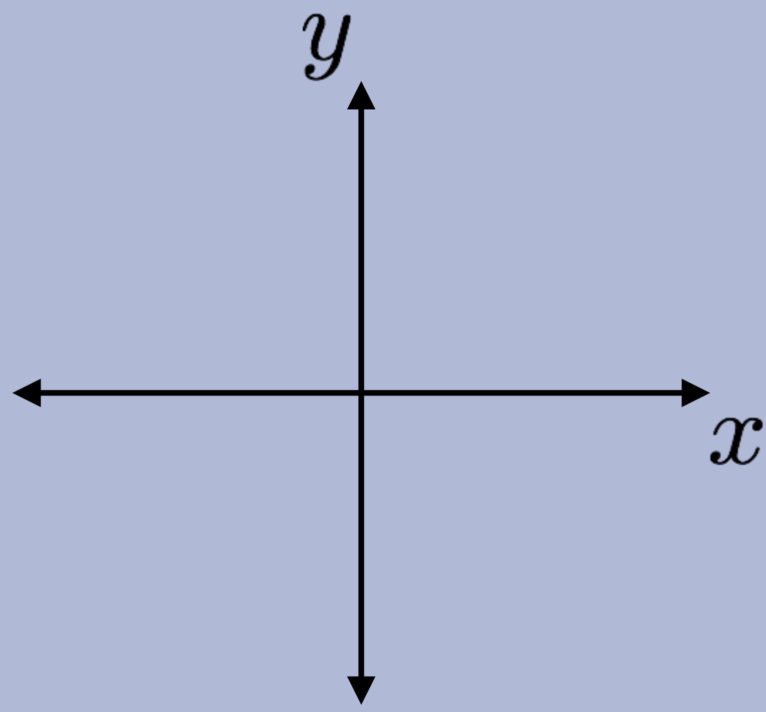
Calculus

In-Class Exercise

Plot three functions:

1. $f(x) = x^3 - x$
2. $f'(x) = \frac{\partial y}{\partial x}$
3. $f''(x) = \frac{\partial^2 y}{\partial x^2}$

Answer Here:



Optimization for ML

Chalkboard

- Unconstrained optimization
- Convex, concave, nonconvex
- Derivatives
- Zero derivatives
- Gradient and Hessian

Optimization: Closed form solutions

Chalkboard

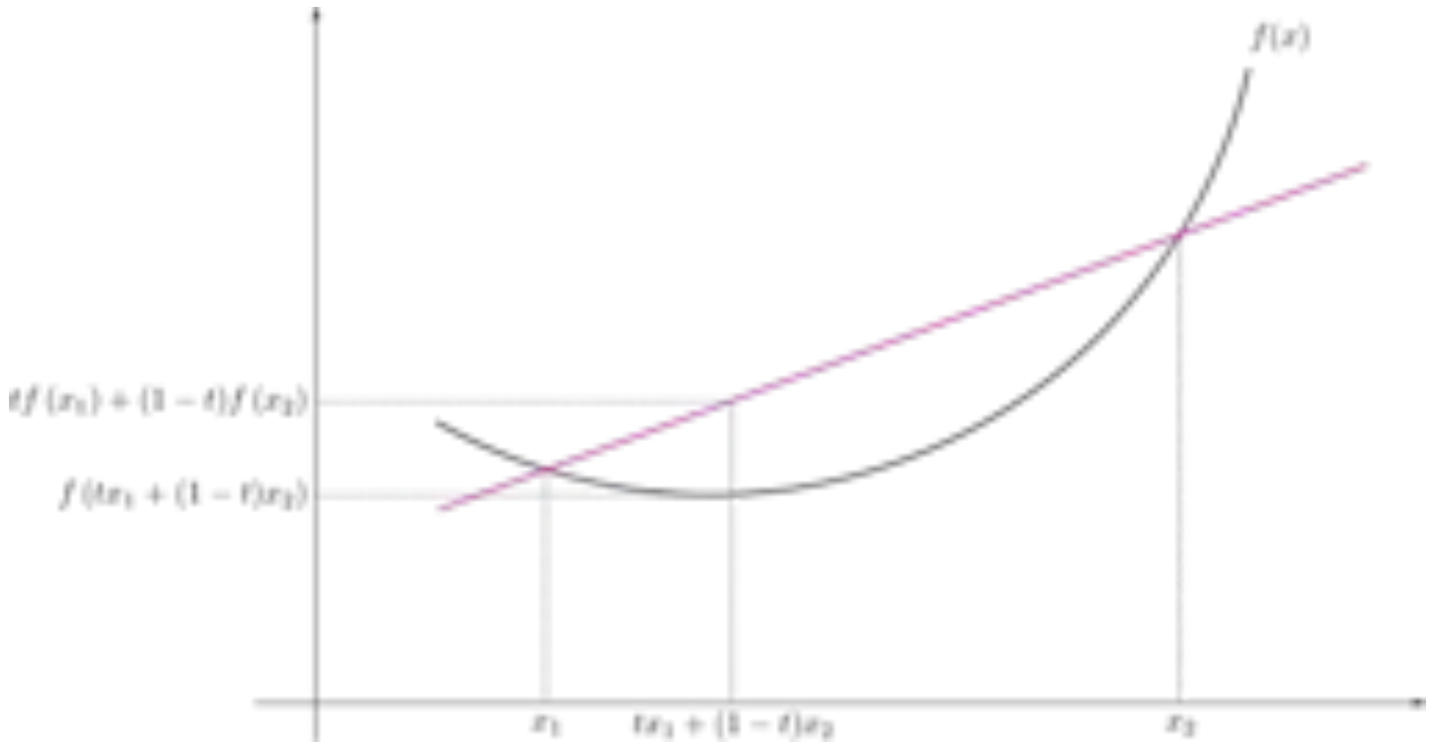
- Example: 1-D function
- Example: higher dimensions

Convexity

Function $f : \mathbb{R}^M \rightarrow \mathbb{R}$ is **convex**
if $\forall\ \mathbf{x}_1 \in \mathbb{R}^M, \mathbf{x}_2 \in \mathbb{R}^M, 0 \leq t \leq 1$:

$$f(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) \leq t f(\mathbf{x}_1) + (1-t) f(\mathbf{x}_2)$$

There is only one local optimum if the function is *convex*



Slide adapted from William Cohen

CLOSED FORM SOLUTION FOR LINEAR REGRESSION

Optimization for Linear Regression

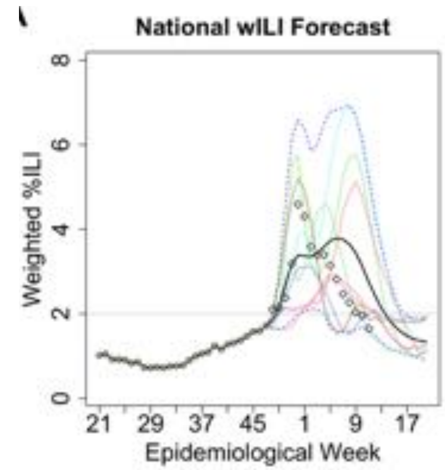
Chalkboard

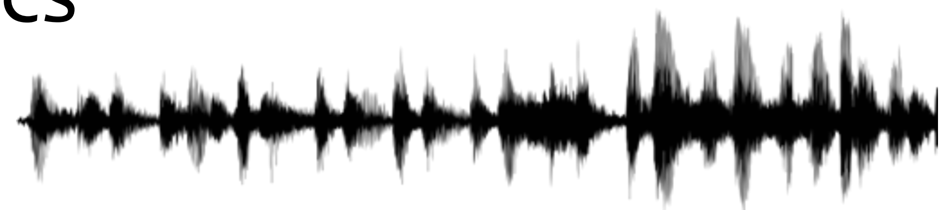
- Closed-form (Normal Equations)
- Computational complexity of Closed-form Solution
- Stability of Closed-form Solution

Regression

Example Applications:

- Stock price prediction
- Forecasting epidemics
- Speech synthesis
- Generation of images (e.g. *Deep Dream*)
- Predicting the number of tourists on Machu Picchu on a given day





Function Approximation

Chalkboard

- The Big Picture