PHYS 177
winter 2019

lecture 18: loss functions and regression

learning goals
- interpret loss functions
- solve a simple regression problem

statistical inference tells us how to infer parameters from data

common approach is through minimization of a loss function

$\min_{\theta} L(\theta|x)$ parameters θ, data x

example: linear regression

say we are measuring force on a spring vs. distance
relationship should be linear,

$$\underset{\text{force}}{Y} = -k\underset{\text{stretch}}{X} + b$$

if measurements were perfect we would only need a single point, but in reality there is noise

find the spring constant by making multiple measurements and finding the best fit line

$$\underbrace{\min_{k} \sum_i |Y_i + kX_i|^2}_{\text{"ordinary least squares" - OLS}}$$

in this special case we can solve exactly

$$\underbrace{\hat{k} = (X^T X)^{-1} X^T Y}_{\text{OLS estimate of } k}$$

(general form valid if measuring multiple variables)

regularization

what if our data is very noisy, so that we don't have too much confidence in small # of measurements?

we can regularize to control the estimate   how could we do this?

example: add parameter estimates to the loss function

$$\min_{k} \sum_i |Y_i + kX_i|^2 + \lambda k^2 \rightarrow \text{ solution is } \hat{k} = (X^T X + \lambda I)^{-1} X^T Y$$

Tikhonov or L_2 regularization

note scaling, as # measurements $\rightarrow \infty$ effect of regularization vanishes

* notebook example