

# Linear regression 2

Reminders: Gdrive Office hrs.  
Piazza  
Homework 2 released by Wed.

Q's about video? (gradients, normal eqns.)

↓  
Google drive, Canvas      WUWU credentials

- new video up → watch it before tomorrow
- check that python 3 runs

- your computer
- lab / linux

numpy  
scipy  
matplotlib }  
wget ...

# Lecture 1

$$\min_{\vec{\beta}} \|\underline{X} \vec{\beta} - \vec{y}\|^2 \xrightarrow[\text{video}]{\text{gradient} = 0} \underline{X}^T \underline{X} \vec{\beta} - \underline{X}^T \vec{y} = 0$$

normal eqns.

$$\underbrace{\vec{\hat{y}}}_{(n \times 1)} = \underbrace{\underline{X} \vec{\beta}}_{(n \times d) (d \times 1)}$$

$n = \#$  data pts

$d = \#$  dims.

$$\underbrace{\underline{X}^T \underline{X}}_{d \times d} \vec{\beta} = \underbrace{\underline{X}^T \vec{y}}_{d \times 1}$$

$$K \vec{\beta} = \vec{c}$$

linear system of eqns.  
for  $\vec{\beta}^*$

What happens for

- $n = d$
- $n > d$
- $n < d$
- $n > d$  but some data pts very similar rows/cols

$$\mathbf{X}^T \mathbf{X} \vec{\beta} = \mathbf{X}^T \vec{y}$$

$$\mathbf{K} \vec{\beta} = \vec{c}$$

$$\mathbf{K} = \mathbf{X}^T \mathbf{X}$$

$d \times d$        $d \times n$        $n \times d$

For tomorrow • 2 SVD videos  
optional  
• polynomial video

•  $n = d$  : just enough ... possible overfit

•  $n < d$  : not enough information, underdetermined

•  $n > d$  :  $\mathbf{K}$  singular  $\rightarrow$  no  $\mathbf{K}^{-1}$ , multiple solutions  
 $\mathbf{K}$  nonsingular unless  $\mathbf{X}$  has repeated rows ... could overfit

• similar data pts / features

-  $\vec{x}_1 \approx \vec{x}_2$  don't count as much

- bad numerical...  $\mathbf{X}$  ill-conditioned

$$\vec{x}_1 = 0.9 \vec{x}_2$$