# CS/ECE/ME 532 Matrix Methods in Machine Learning



Welcome!

## Announcements



- Mid Course Feedback Survey
- This week:
  - Last class: finish Unit 4
  - Today: start Unit 5
- Next week: spring break
- After spring break:
  - Tuesday: Assessment 3 review
  - Thursday: Assessment 3
- Assignment 7 is due the week after Assessment 3, but:
  - Q1 on the Eigendecomposition (attempt before the assessment)

# Activity 16

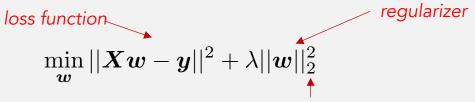
### Hints for 1: video lecture 2.8

$$\boldsymbol{w}^* = \boldsymbol{V} \Sigma^{-1} \boldsymbol{U}^T \boldsymbol{y}$$

 $\mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w} = c$  define an ellipse with  $\mathbf{X} = \mathbf{U} \Sigma \mathbf{V}^T$  then columns of  $\mathbf{V}$  define axis



### Before:



(This encourages solutions with small 2-norm)

closed form solution 
$$oldsymbol{w}^* = (oldsymbol{X}^Toldsymbol{X} + \lambda oldsymbol{I})^{-1}oldsymbol{X}^Toldsymbol{y}$$

What if we want some other regularizer?

(maybe only a few non-zero entries)

$$\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||^2 + \lambda r(\boldsymbol{w})$$

Or different a loss function all together?

$$\min_{m{w}} f(m{w})$$

In general: no closed-form solution - Interative methods

### **Gradient Descent**

Main idea: use the gradient to head downhill

Gradient descent for least-squares:

$$egin{aligned} & \operatorname{goal:} \ \min_{oldsymbol{w}} ||oldsymbol{X} oldsymbol{w} - oldsymbol{y}||_2^2 \ & \operatorname{for} \ k = 1 \dots \ & oldsymbol{w}^{(k+1)} = oldsymbol{w}^{(k)} - au(oldsymbol{X}^T oldsymbol{X} oldsymbol{w}^{(k)} - oldsymbol{X}^T oldsymbol{y}) \end{aligned}$$

