CS/ECE/ME 532
Period 17

- Assignment 7 due Monday

- Today:
  - proximal gradient descent

Proximal Gradient Descent for Least Squares

loss function

regularizer:
L2 gives solutions with small 2 norm

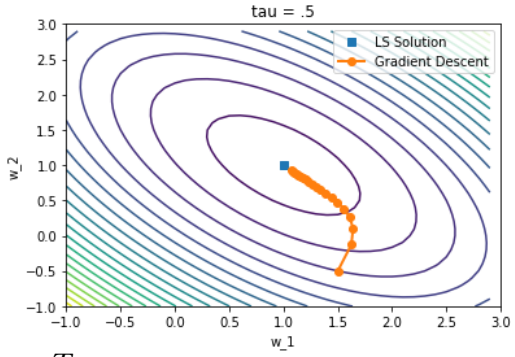
$$\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||^2 + \lambda||\boldsymbol{w}||_2^2$$

Different regularizer or loss function?

L1 gives sparse solutions

$$\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||^2 + \lambda||\boldsymbol{w}||_1$$

No closed form solution.
Optimization.



Gradient Descent for LS

goal: $\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$

for $k = 1 \ldots$

$$\boldsymbol{w}^{(k+1)} = \boldsymbol{w}^{(k)} - \tau(\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{w} - \boldsymbol{X}^T\boldsymbol{y})$$

step size

Main idea: use the gradient to head downhill

Alternating Gradient Descent for LS with regularization

goal: $\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2 + \lambda r(\boldsymbol{w})$

set $\boldsymbol{w}_0$

for $k = 1 \ldots$

Gradient Descent

$$\boldsymbol{z}^{(k)} = \boldsymbol{w}^{(k)} - \tau\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{w}^{(k)} - \boldsymbol{y})$$

Regularization Step

$$\boldsymbol{w}^{(k+1)} = \arg\min_{\boldsymbol{w}} ||\boldsymbol{z}^{(k)} - \boldsymbol{w}||_2^2 + \lambda\tau\, r(\boldsymbol{w})$$

stay close to z, but regularize (L2 – shrink)