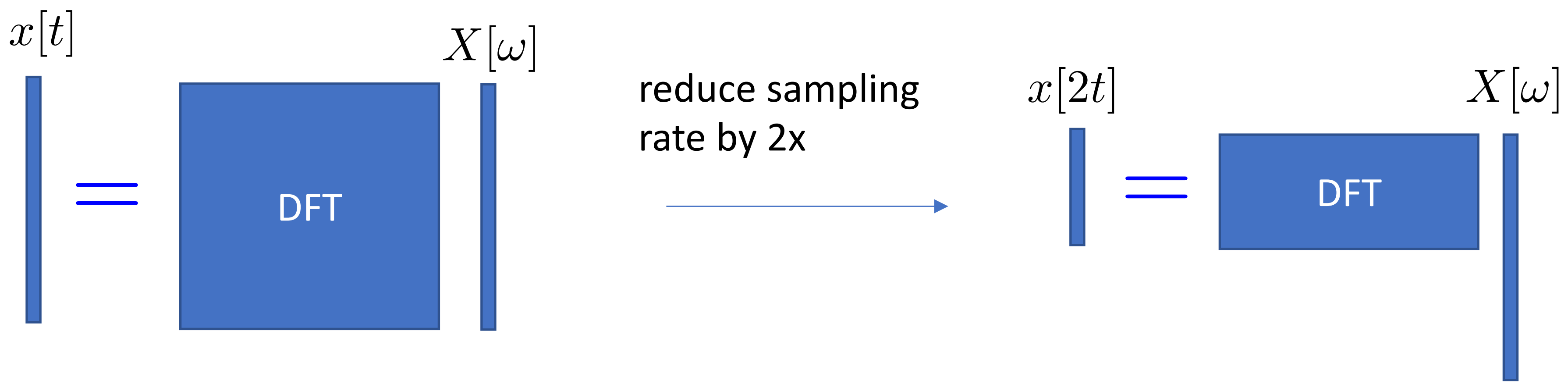
CS/ECE/ME 532
Period 18

- Today – Unit 5
  - Sparse solutions to least squares problems
- Today's activity
  - Understanding why L1 gives sparse solutions

Applications:
- breast cancer prediction
- Sub-Nyquist sampling



Proximal Gradient Descent for Least Squares with L1 regularization

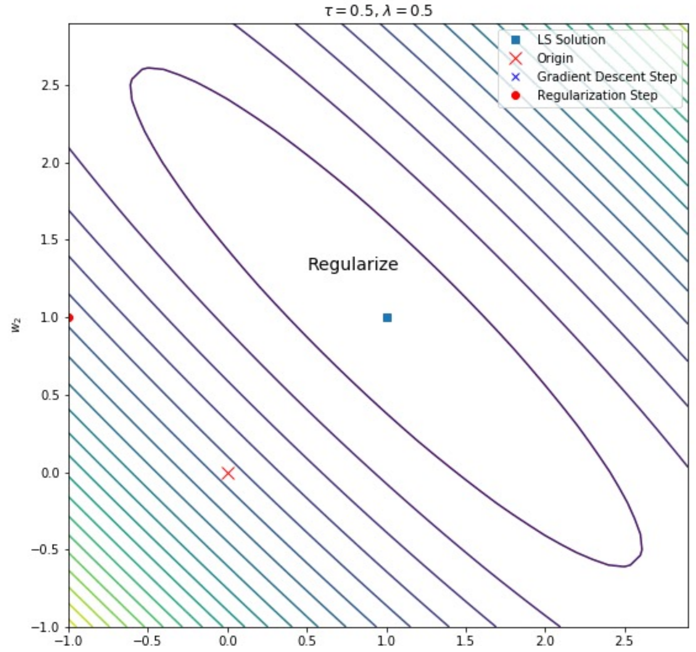
Training a model:

Loss function

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

L1 gives sparse solutions

No closed form solution.
Optimization.



Alternating Gradient Descent with regularization

$$\text{goal: } \min_{\boldsymbol{w}} \ell(\boldsymbol{w}) + \lambda r(\boldsymbol{w})$$

$$\text{for } k = 1 \ldots$$

$$\boldsymbol{z}^{(k)} = \boldsymbol{w}^{(k)} - \tau \nabla \ell(\boldsymbol{w}^{(k)})$$

Gradient Descent

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

Regularization Step

stay close to z, but regularize

$$\ell(\boldsymbol{w}) = ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2 \qquad r(\boldsymbol{w}) = ||\boldsymbol{w}||_1$$

$$\boldsymbol{w}^{(k+1)} = \arg\min_{\boldsymbol{w}} \sum (z_i^{(k)} - w_i)^2 + \lambda\tau\, |w_i|$$

$$w_i^{(k+1)} = (|z_i| - \lambda\tau/2)_+ \operatorname{sign}(z_i)$$

soft threshold