Lecture 3 Outline:

Textbook Reading: ESL Chapter 3, Sections 3.2 and 3.4.1 and/or ISL Chapter 3 and Chapter 6.2.1.

1. Review: Least squares when \( p >> n \).

2. Ridge regression (Tikhonov regularization).
   - \( \ell_2 \)-norm constrained least squares.
   - Lagrangian form.
   - Closed form solution.
   - In practice: Intercepts, Scaling covariates.
   - Add a little bias to (hopefully) greatly decrease the variance.
     - Calculate Bias and Variance. Show Variance always less than that of least squares estimator.
   - Better at prediction than least squares - MSE Existence Theorem.
   - Degrees of Freedom.
   - Real data example.

3. Review SVD / PCA.
   - Properties of the SVD.
   - Find orthogonal linear projections that maximize the sample variance.
   - Real data example.

   - Ridge solution as a function of the SVD.
   - Interpretations.
   - Ridge regression for correlated predictors.

5. Least Squares on Derived Inputs.
   - Principal Components Regression. Least squares on PCs.
   - Partial Least Squares Regression.
     - PCA on cross-covariance between \( \mathbf{X} \) and \( \mathbf{Y} \) (orthogonalized with respect to \( \mathbf{X} \)).
     - Least squares constrained to the Krylov subspace.