Projects

00103335: Deep Learning and Reinforcement Learning
Fall 2026

Objective. Use one simulation, adapted from Figure 2 of Hastie et al. (2022), to observe the interpolation threshold and double descent in least squares.

Problem 1. Double-descent experiment

Generate data from

$$y = X\beta + \varepsilon, \qquad X_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \qquad \varepsilon \sim \mathcal{N}(0, I_n), \qquad n = 200, \quad \beta = \sqrt{5}\, e_1,$$

and vary the number of features over

$$p \in \{20, 40, 60, 80, 100, 120, 140, 160, 180, 199, 200, 220, 240, 260, \\ 300, 400, 600\}.$$

For each value of p, fit the minimum-norm least-squares estimator $\widehat{\beta} = X^{+}y$. Generate an independent test matrix $X_{\text{test}} \in \mathbb{R}^{2000\times p}$ with i.i.d. standard Gaussian entries, and estimate the prediction risk by

$$R = \frac{1}{2000} \left\| X_{\text{test}}\widehat{\beta} - X_{\text{test}}\beta \right\|_2^2.$$

(a) Repeat the experiment 20 times independently for every p, recording R in each repetition.

(b) Make two separate plots of R against p: one showing the median and one showing the sample mean. Use each plot to identify the value of p at which the peak occurs.

(c) Repeat the same experiment, including both plots, for $\beta = \sqrt{2}\, e_1$ and $\beta = e_1$. Use the same values of p and the same number of repetitions.

Reference. T. Hastie, A. Montanari, S. Rosset, and R. J. Tibshirani, “Surprises in high-dimensional ridgeless least squares interpolation,” *The Annals of Statistics*, 50(2):949–986, 2022. doi:10.1214/21-AOS2133.