CSCI 145 Problem 6: Leave One Out Without Fitting Again

October 27, 2026

Submission Instructions

Please upload your work to Gradescope by **11:59pm on Wednesday, September 30**.

- Turn in a single PDF of *handwritten* solutions. Typed (including LaTeX) solutions will not be accepted.
- A complete, good-faith attempt earns submission credit. The grader may select any recent assigned problem for a brief conversation, so keep your work and be prepared to explain each solution and answer a related question. Scheduling details will be announced in class.

Problem 6: Leave One Out Without Fitting Again

To evaluate a prediction rule, we would like to test it on observations whose labels it did not use for training. In *leave-one-out evaluation*, we remove each observation in turn, fit least squares to the remaining observations, and predict the omitted label. This appears to require fitting n separate models.

Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ be the design matrix and $\mathbf{y} \in \mathbb{R}^n$ the labels. Assume the columns of $\mathbf{X}$ remain linearly independent after deleting any row. The full-data fit is

$$\mathbf{y}^\star = \mathbf{H}\mathbf{y}, \qquad \mathbf{H} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{X}^\top.$$

1. How can we compute the leave-one-out residual on observation i, its label minus its prediction from the model trained without it, using only y_i, $y_i^\star$, and H_{ii}? Prove your formula, showing why no additional least-squares fits are needed.

Hint: For any label vector $\mathbf{z} \in \mathbb{R}^n$, the fitted predictions on the same design matrix are $\mathbf{H}\mathbf{z}$. Under the stated independence assumption, $0 \le H_{ii} < 1$ for every i.

Takeaway: *For least squares, we can compute all the leave-one-out errors from one fit. This gives the same evaluation as fitting a separate model for each omitted point, with much less repeated work.*