Math 342W/642/742W

Recitation – Day #8 (2.27.25)

I. OLS Basics

(i) Define $\boldsymbol{b}$ in terms of the solution to the *normal equations* to the *least squares* problem?

(ii) What does finding $\boldsymbol{b}$ accomplish in the context of the overall machine learning problem of *regression*?

(iii) Define the "*hat*" matrix, denoted by H.

(iv) What are the two properties H? What do we call such matrices?

(v) What does H do to the given vector of responses/labels $\boldsymbol{y}$?

II. Preliminaries for Geometric Interpretation of OLS

We have $X \in \mathbb{R}^{n \times (p+1)}$.

(i) Define colsp$[X]$.

(ii) What is the rank of X? How so? What does this mean for the rank of H?

(iii) Unpack mathematically what $\hat{\boldsymbol{y}} \in$ colsp$[X]$ means.

III. Geometric Interpretation of OLS

(i) Give an illustration of OLS with $\boldsymbol{y}, \hat{\boldsymbol{y}}, \boldsymbol{e}, \mathbb{R}^n$ and colsp$[X]$.

(ii) In what space does the *residual vector*, $\boldsymbol{e}$ reside in? What is its dimension?

(iii) What is the matrix that projects $\boldsymbol{y}$ onto the *residual space*? What is its rank?

(iv) In what way can the "*full space*", $\mathbb{R}^n$, be decomposed in?

IV. QR-Decomposition

(i) What is the goal behind the QR-decomposition/factorization of a matrix? What is accomplished?

(ii) What is the process that creates the QR-decomposition of a matrix called?

(iii) How will we now define the orthogonal projection matrix H?