COS 516: Final Project Report
Proving the Correctness of the Normal Equations with Dafny

Samuel Sanft

December 11, 2024

Honor Statement

This assignment represents my own work in accordance with university policy.

1 Introduction

The Normal Equations refer to the following equation and its associated solution:

$$\begin{aligned} A^T A x &= A^T b \\ x &= (A^T A)^{-1} A^T b \end{aligned}$$

2 Overview

3 Project Tasks

4 Results

5 Discussion

6 Conclusions

7 Appendix

7.1 Proof of Optimality of Normal Equations

In this section, a proof is provided that the solution to the Normal Equations is an optimal solution to the OLS problem. The proof roughly matches the structure of the proof that is written in Dafny. Since the squared euclidean norm is a convex function, this is normally done by showing that the solution to the Normal Equations satisfies the first-order optimality conditions for the OLS problem. However, in order to verify the proof in Dafny without implementing symbolic or automatic differentiation, it was necessary to find a proof that did not require differentiation. That proof is given here.

In order to show that the Normal Equations provide an optimal solution to the OLS problem, it is sufficient to show (let $x^* := (A^TA)^{-1}A^Tb$):

$$\forall x, ||Ax^* - b||^2 \leq ||Ax - b||^2$$

7.1.1 Lemma 1

$$||Ax^*||^2 = -\langle Ax^*, b\rangle$$

$$\begin{aligned} ||Ax^*||^2 &= \langle Ax^*, Ax^*\rangle \\ &= \langle x^*, A^TAx^*\rangle \\ &= \langle x^*, A^TA(A^TA)^{-1}A^Tb\rangle \\ &= \langle x^*, A^Tb\rangle \\ &= \langle Ax^*, b\rangle \end{aligned}$$

7.1.2 Lemma 2

Using Lemma 1:

$$||Ax^* - b||^2 = -\langle Ax^*, b\rangle + ||b||^2$$

$$\begin{aligned} ||Ax^* - b||^2 &= ||Ax^*||^2 - 2\langle Ax^*, b\rangle + ||b||^2 \\ &= \langle Ax^*, b\rangle - 2\langle Ax^*, b\rangle + ||b||^2 \\ &= -\langle Ax^*, b\rangle + ||b||^2 \end{aligned}$$

7.1.3 Lemma 3

Using Lemma 1:

$$||Ax - b||^2 = ||Ax - Ax^*||^2 - \langle Ax^*, b\rangle + ||b||^2$$

$$\begin{aligned} ||Ax - b||^2 &= ||Ax||^2 - 2\langle Ax, b\rangle + ||b||^2 \\ &= ||Ax||^2 - 2\langle A(A^TA)^{-1}A^TAx, b\rangle + ||b||^2 \\ &= ||Ax||^2 - 2\langle Ax, A(A^TA)^{-1}A^Tb\rangle + ||b||^2 \\ &= ||Ax||^2 - 2\langle Ax, Ax^*\rangle + ||Ax^*||^2 - ||Ax^*||^2 + ||b||^2 \\ &= ||Ax - Ax^*||^2 - ||Ax^*||^2 + ||b||^2 \\ &= ||Ax - Ax^*||^2 - \langle Ax^*, b\rangle + ||b||^2 \end{aligned}$$

7.1.4 Proof

Now the original inequality can be demonstrated, using Lemmas 2 and 3 and the fact that the euclidean norm is always nonnegative:

$$
\begin{aligned}
||Ax^* - b||^2 &\leq ||Ax - b||^2 \\
-\langle A^*, b\rangle + ||b||^2 &\leq ||Ax - Ax^*||^2 - \langle A^*, b\rangle + ||b||^2 \\
0 &\leq ||Ax - Ax^*||^2
\end{aligned}
$$