ML Problem Solving

Least squares

1. Consider the following dataset, where each row is one input and the measured output is the last column.

D = [[6, 5, 1], [5, 6, 0], [4, 3, 1], [3, 4, 0], [2, 3, 0], [3, 2, 1]]

Draw the data points from X in a coordinate system.

2. Now let's try to manually choose $\beta \in \mathbb{R}^3$ to minimize the Residual Sum of Squares (RSS). Choose at least 3 different values for β and calculate the associated RSS.

β_a = [_____ , _____ , _____]

β_b = [_____ , _____ , _____]

β_c = [_____ , _____ , _____]

3. Draw the decision boundaries for each model defined by your β. Do it for all three values you chose.

4. Calculate $\beta = (X^T X)^{-1} X^T y$

5. Considering that $RSS(\beta) = (y - X\beta)^T (y - X\beta)$, derive a closed-form β that minimizes RSS. (optional)