AMATH 352: HOMEWORK 9

JOLIE TRAN

ptran32@uw.edu

1. INTRODUCTION AND OVERVIEW OF THE PROBLEM

This report investigates various fitting models to a dataset of fuel consumption (mpg) of various cars as a function of different attributes such as engine displacement, horsepower, and weight. The models are built by formulating an appropriate least squares problem through QR using the training data set to find the unknown parameters that define the models. We also assess the quality of the models with relative error, effective rank, and condition number to see if higher complexity results in better model performance.

2. THEORETICAL BACKGROUND AND DESCRIPTION OF ALGORITHM

2.1. **Condition Number:** A condition number measures how sensitive the output is to small perturbation in the input data. [1]. A well-conditioned problem is one where small perturbation to the input data lead to small perturbations of the output. An ill-conditioned problem is one where small perturbations of the input leads to large perturbations of the output. We can also think of the condition number as relative distance to singularity, which means that ill-conditioned problems are "close to singular." [1] If matrix A is singular, we define the condition number of A $\kappa(A) = \infty$. [2]

2.2. **Condition Number and Least Squares:** It is advised to solve least squares via QR because $\kappa(R) = \kappa(A)$ while forming the normal equations squares the condition number [3]. The condition number of A for least squares is

$$\kappa(A) = \kappa(R) = \sqrt{\kappa(A^T A)}$$

2.3. **Full Rank, Rank Deficiency:** The matrix is full rank, meaning that the rank is equal to the column of the matrix, then columns of the matrix is linearly independent. The matrix rank is less than the column of the matrix, then the columns are linearly dependent, and it is called rank deficiency. Rank deficiencncy in this context says there is insufficient information contained in our data to estimate the model that we desire and the solution of least squares problem is not unique.

3. COMPUTATIONAL RESULTS

3.1. Single-feature Models

First, we construct the function $y(\underline{x})$'s of that is only a function of one of the features. We want to predict the mpg y as a 2nd-degree order function of only one feature. For each of these models, we also consider higher degree order K-th for $K = 4,8,12$.

Figures 1, 2, and 3 each are the scatter plot of the testing data overlaid with 4-line plots of the model of 4 different degree orders. We can see that the higher the degree orders, the more oscillations it gets within the data range and the more off it gets outside the range.

Specifically, in Figure 1, while $K = 2$ and $K = 4$ have a similar decreasing behavior at almost the same rate, $K = 8$ and $K = 12$ have the oscillating behavior within the data range, and $K = 12$ oscillates at a higher rate. Both $K = 8$ and $K = 12$ get significantly off 'trend' when it gets out of the data range. Similarly, in Figure 2 and Figure 3, we also see similar patterns for K's of different values.

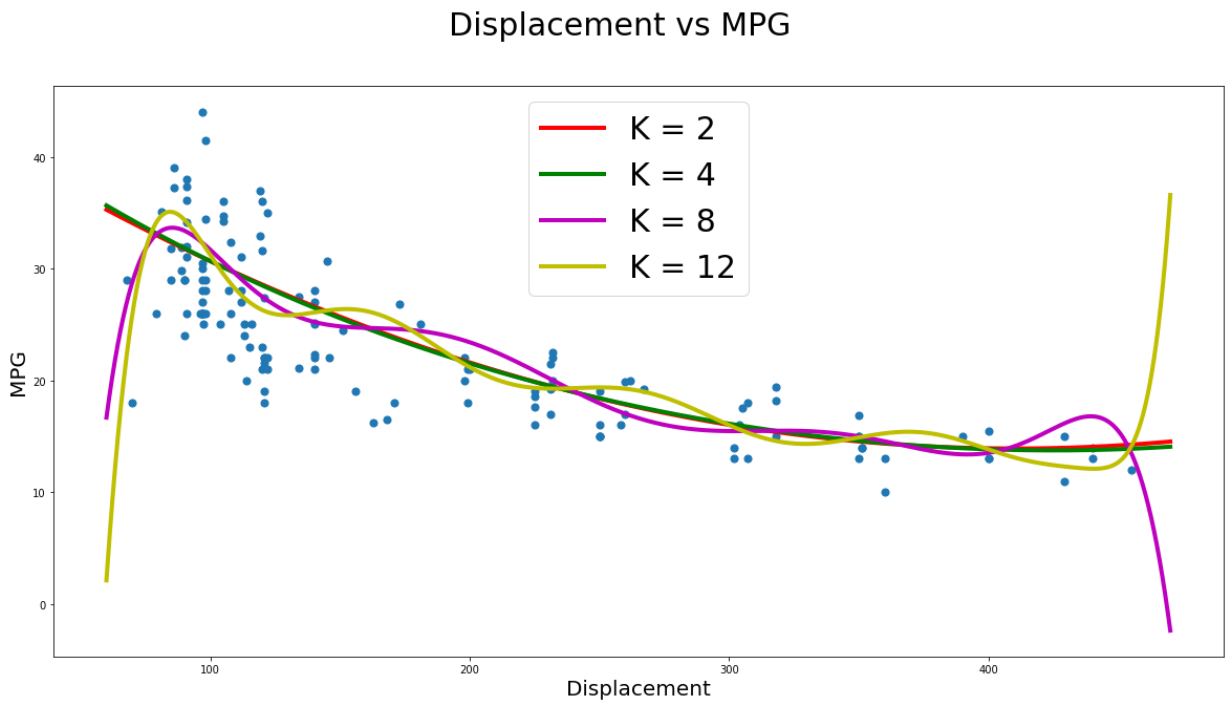

Figure 1: Displacement vs MPG Model with Different Degree Orders

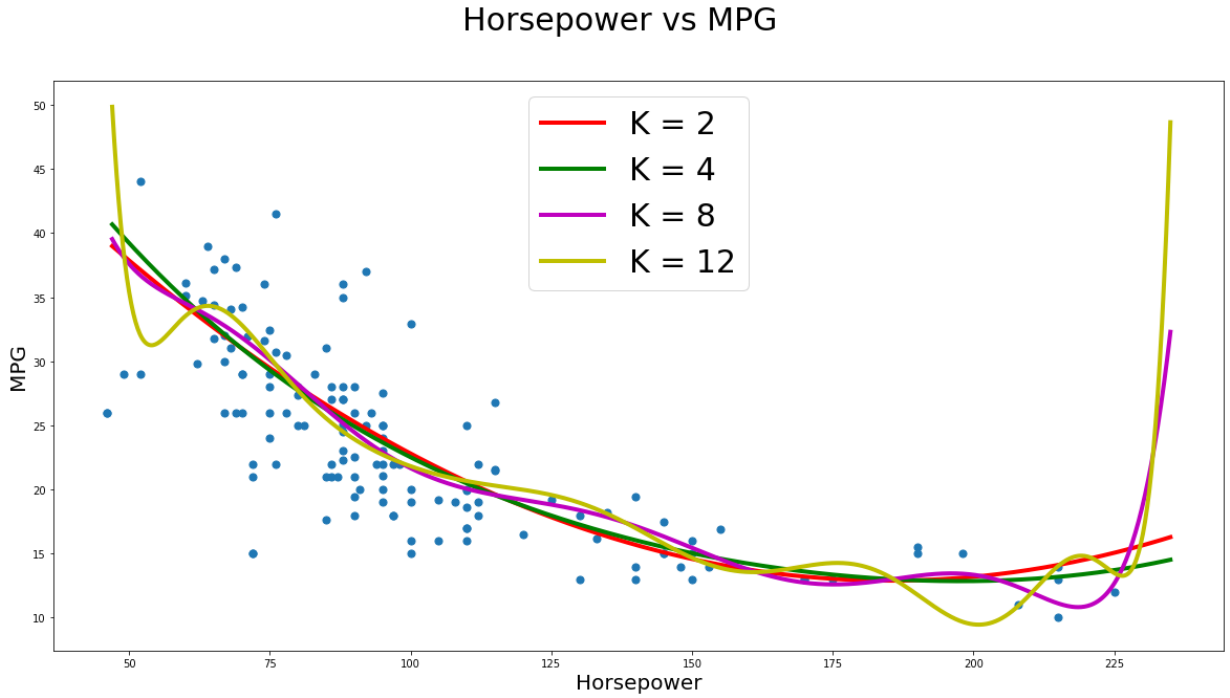

Figure 2: Horsepower vs MPG Model with Different Degree Orders

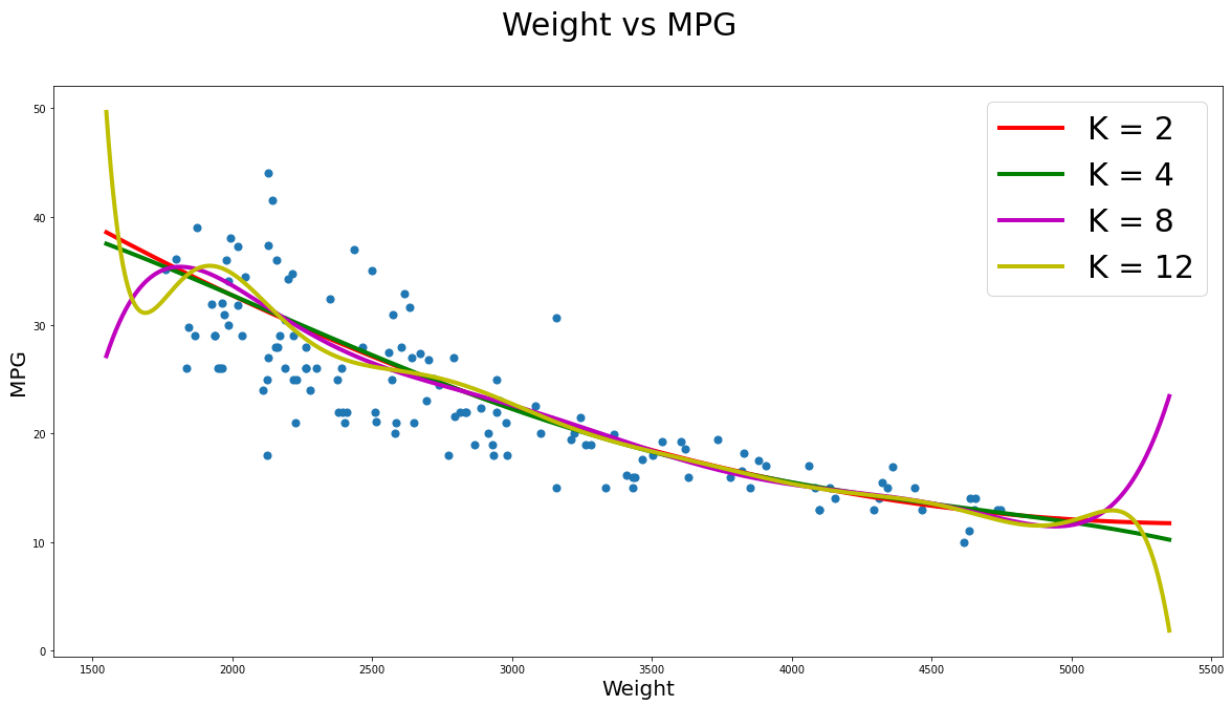

Figure 3: Weight vs MPG Model with Different Degree Orders

Looking at the figures themselves although give us an idea of the behaviors and patterns of different models and degree orders, we cannot really tell the accuracy and effectiveness of each model without looking at their relative errors. In Table 1, we see that the relative errors tend to be larger when we have larger K for all models except the Displacement's. Additionally, it also stands out to us that among 3 models, Weight model consistently gives us the best prediction of the fuel consumption (mpg), regardless of the degree orders.

	$K = 2$	$K = 4$	$K = 8$	$K = 12$
Displacement	0.03405	0.03390	0.03283	0.03287
Horsepower	0.03826	0.03917	0.03894	0.06231
Weight	0.02878	0.02879	0.02978	0.03047

Table 1: Relative Error of Single-feature Models with Different Degree Orders

While Weight model of degree order 2 has the lowest relative error out of all models, it surprisingly has the largest condition number of all models with the same degree order (Table 2). We can say that the error that we get depends on the structure of the input matrix $X_{Vand(i,K)}$ and on the input vector Y_{test}. In Table 2, we also see a trend of increasing condition number as we increase the degree order K. We then consider the least squares problems that involve these matrices $X_{Vand(i,K)}$ where $K = 4,8,12$ are ill-conditioned.

	$K = 2$	$K = 4$	$K = 8$	$K = 12$
Displacement	$3.5 \times 10^5 \approx 10^5$	$2.4 \times 10^{11} \approx 10^{11}$	$1.9 \times 10^{23} \approx 10^{23}$	$2.0 \times 10^{35} \approx 10^{35}$
Horsepower	$1.3 \times 10^5 \approx 10^5$	$2.8 \times 10^{10} \approx 10^{10}$	$2.4 \times 10^{21} \approx 10^{21}$	$3.0 \times 10^{32} \approx 10^{32}$
Weight	$1.6 \times 10^8 \approx 10^8$	$3.1 \times 10^{16} \approx 10^{16}$	$1.8 \times 10^{33} \approx 10^{33}$	$1.8 \times 10^{50} \approx 10^{50}$

Table 2: Condition Number of Single-feature Models with Different Degree Orders

To better understand how the effective the model is, we also look at the rank of the matrix in the least squares problems. In Table 3, we see that the higher the degree order, the more likely the matrix will be rank-deficient. We thus could say that increasing the degree order does not necessarily increase the efficiency

because our effective rank tends to be smaller than the column size when $K \geq 4$ (except for the Displacement model and the Horsepower model).

	$K = 2$	$K = 4$	$K = 8$	$K = 12$
Displacement	Rank: 3, Column: 3	Rank: 5, Column: 5	Rank: 5, Column: 9	Rank: 4, Column: 13
Horsepower	Rank: 3, Column: 3	Rank: 5, Column: 5	Rank: 5, Column: 9	Rank: 5, Column: 13
Weight	Rank: 3, Column: 3	Rank: 4, Column: 5	Rank: 3, Column: 9	Rank: 3, Column: 13

Table 3: Matrix Rank and Column Size from Single-feature Models with Different Degree Orders

3.2. Multi-feature Models

Next, we consider models that depend on multiple features at a time. Looking at Table 4, we see that the model of Displacement & Weight has the smallest relative error, and the Displacement & Horsepower model has the largest. However, the Table 5 shows us the opposite result for condition number – the model of Displacement & Horsepower (10^5) is the least sensitive to the input matrix and all other models with condition number (10^8) are the more ill-conditioned models. Again, the relative error depends on the structure of input matrix X_i and on input vector Y_{test}. Addtionally, looking at Table 6, we see that the matrices from our multi-feature models are all full-rank. This implies that these matrices in the least squares problem give us unique solutions.

Displacement & Horsepower	Displacement & Weight	Horsepower & Weight	Displacement, Horsepower & Weight
0.03069	0.02909	0.03068	0.03011

Table 4: Relative Error of Multi-feature Models

Displacement & Horsepower	Displacement & Weight	Horsepower & Weight	Displacement, Horsepower & Weight
$7.1 \times 10^5 \approx 10^5$	$2.4 \times 10^8 \approx 10^8$	$1.6 \times 10^8 \approx 10^8$	$2.0 \times 10^8 \approx 10^8$

Table 5: Condition Number of Multi-feature Models

Displacement & Horsepower	Displacement & Weight	Horsepower & Weight	Displacement, Horsepower & Weight
Rank: 6, Column: 6	Rank: 6, Column: 6	Rank: 6, Column: 6	Rank: 7, Column: 7

Table 6: Matrix Rank and Column Size from Multi-feature Models

3.3. Models Comparison & Best Performance Model

As we compare the single-feature models and multiple-feature models with the relative errors, having more features in our model definitely increase the accuracy of our model. However, we know that the relative error depends on the structure of the input matrix and the input vector.

So, in choosing the best performance model, instead of just looking at the relative error, we want to choose the model that have both a small condition number and a large rank from a small-size matrix. We want to choose the model that have the large rank and small-sized matrix because we want to reduce the cost of computing. We ackowledge that the smallest condition number that we have for all of our models is 10^5 so we consider it an indicator for the well-conditioned problem, and our best matrix rank is 3. This narrows us down to the Displacement model and the Horsepower model. Then we want to filter the option with the smaller relative error, and thus we get the Displacement as our best performance model. Therefore, we could say that adding more features does not necessarily improve the performance of the model.

4. SUMMARY AND CONCLUSIONS

With the training dataset of fuel consumption (mpg) of various cars as a function of different attributes such as engine displacement, horsepower, and weight, we construct and evaluate models and, in the end, pick out the best performance based on the relative error, condition number, matrix rank, and matix size. We find that (1) when constructing a least squares problem, increasing the degree order does not necessarily increase the accuracy, and it potentially result in very ill-conditioned problem, (2) when computing relative error, our error depends on the structure of our input data, (3) the best performance model does not necessarily have the smallest error, but it must be constructed from a well-conditioned problem and have the large rank from a small-sized matrix, and (4) adding more features does not necessarily improve the performance of the model.

REFERENCES

[1] Bindel, D., 2012: Least Squares Reminder, Intro to Scientific Computing (CS 3220), Department of Computer Science - Cornell University, Accessed 09 December 2022, https://www.cs.cornell.edu/~bindel/class/cs3220-s12/notes/lec11.pdf

[2] Carrier, E., Meng, Y., Olson, L., 2017, Condition Numbers, CS 357: Numerical Methods I, Accessed 10 December 2022, https://courses.engr.illinois.edu/cs357/fa2021/notes/ref-10-condition.html

[3] Moler, C., 2017: What is the Condition Number of a Matrix? MATLAB, Accessed 09 December 2022, https://blogs.mathworks.com/cleve/2017/07/17/what-is-the-condition-number-of-a-matrix