If advance demand is unexpectedly high, actual demand is equal to the advance demand (months 11, 13, 15, 18, 19 and 22). We also know that actual demand cannot be less than the advance demand Suppose we regress the demand values on the prior month’s demand values (AR1: autoregressive model with one lag) and add the constraint the demand predictions cannot be less than advance demand. 𝑌 is the vector of demand values from the second month to the last month (month 25). 𝑋 is the vector of demand values from the first month to the twenty-fourth month. 𝐿 is the vector of advance demand values from the second month to the last month (month 25). Objective is to minimize the sum of squared residuals (𝑌 −𝑋Β)T \* (𝑌 −𝑋Β), subject to: 𝑋Β ≥ 𝐿. The 𝛽REG=(𝑋T𝑋)-1(𝑋T𝑌+0.5𝑋T𝜆), where 𝜆=2(𝑋(𝑋T𝑋)-1 𝑋T)-1𝐿−2𝑌.

To solve this issue, ridge regularization has been utilized. Regression that regularizes, constricts, or decreases coefficient estimates in the direction of zero is known as regularization. To prevent the risk of overfitting, this strategy inhibits learning a more sophisticated or flexible model. Second-degree Regularization (L2) is modified by adding the shrinkage quantity to provide ridge regression. This function is minimized to estimate the coefficients. The tuning parameter λ determines how much we want to penalize our model's flexibility. If we want to minimize the aforementioned function, then these coefficients must be tiny because an increase in a model's coefficients represents an increase in its flexibility. The Ridge regression method does this by limiting the upward trend of coefficients.

After applying the ridge regression, the matrix yielded three values, or roughly 3.41e-13, -1.78e-15, and 1, respectively. The first column has a number of X ones, and the first answer is 3.41e-13. Because of the positive association, if the value of those 1s rises, Y will rise along with it. Although they have a weak distribution for predicting Y, those 1s can aid in the calculation.

The second response, 1.78e-15, is also a modest value and is found in the column X, which contains the demand values for months 1 through 24. Because of the negative correlation, the expected Y value will fall as the X value rises. The low number indicates that X's predictive distribution Y is poor.

The last answer is 1, which is relatively large. This column, L, contains the advanced demand values from the second to the last month as a vector. So, L has a one-to-one correlation when predicting Y, and it is the most important factor to predict Y. The data of advanced demand therefore has a greater impact than the data of demand. In anticipating Y, advance demand is more significant.

Creating a predictive model that would provide a prediction for each point is the aim of the task. I do this by using the Lagrange multiplier's lambda information, which was covered in class. The lambda value was -1.64e+02, -6.4e+01, -7.8e+01, -1.0e+02, -6.8e+01, -8.4e+01, -6.8e+01, -4.4e+01, -6.4e+01, 2.5e-12, -6.8e+01, 2.5e-12, -6.6e+01 ,-6.0e+01, -7.0e+01, -9.0e+01, 4.5e-13, -2.0e-12, -8.8e+01, -1.1e+02, -1.2e-12, -1.1e+02, -1.2e+02, -9.6e+01. Most lambda values for each month are negative, therefore, the constraint is not binding, and the optimal solution for the unconstrained model is also the optimal value for the constrained model. Lambda equal to the negative value will cause overfitting, so try to solve this problem, we will change negative lambda values to zero, and generate a new result for the beta prediction to avoid overfitting and high variance.

So, I changed all negative lambda values to zero and keep the positive lambda to make a more accurate prediction. In the end, the new beta value times X will end up with the new month’s demand for the most time (with a few times miscalculated/errors, but it is acceptable). For example, the prediction of the second entry is 111.13577478, which is close to the truth label 112. This means that the predictive model is relatively accurate.