M20 - Introduction to Computer Programming with MATLAB
Instructor: Prof. Enrique López Droguett, Ph.D.
Teacher Assistants: M. Fidansoy, G. San Martín, M. Pishahang, V. Vela.
Fall 2023 – UCLA
Student: *Alex Lie*
UCLA ID: *905901892*

HOMEWORK 8

Task 1: Nonlinear Regression

Introduction

I am using MATLAB to create a general nonlinear regression to model the reaction velocity in enzyme kinetics (the Michaelis-Menten equation). I will find values of the maximum velocity and the Michaelis constant that will make the equation fit the data well. I will then plot the enzyme kinetics data from the Excel spreadsheet and my regression model in the same graph. After, I will calculate the coefficient of determination and evaluate how well my nonlinear regression models the data from the enzyme kinetics.

Model and Theory

1) $v = \frac{V_{max} \cdot S}{K_M + S}$

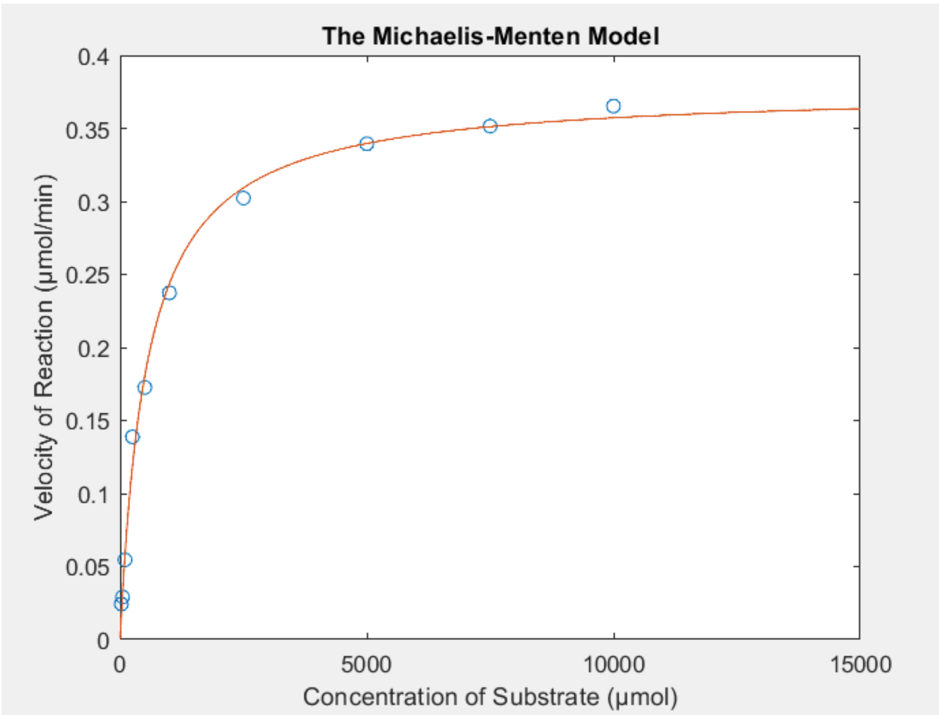
Methodology

I first cleared the workspace and command window. I then imported the data from the Excel spreadsheet using the readmatrix() function and declaring variables. I then declared an options variable that sets the options for when I find the minimum of a multivariable function which includes the tolerance of the independent variable, the tolerance of the value of the function, and the maximum number of iterations. I then used the fminsearch() function to find the values of the Michaelis constant and the maximum velocity that would minimize the residual sum of squares of the reaction velocity. The parameters of the fminsearch() function are the MichaelisMentis() function I made, my guesses for the coefficients (the Michaelis constant and the maximum velocity), and the options variable I declared. I created my MichaelisMentis() function at the bottom of my code, which takes in the reaction velocity data, concentration of substrate data, the Michaelis constant, and the maximum velocity as inputs. My function calculates the theoretical reaction velocity based off equation 1 using the data inputs, and finds the residual sum of squares by finding the difference of the theoretical reaction velocity by the actual reaction velocity, squaring it, and adding up all of the squared differences for each value of the reaction velocity. After running the fminsearch() function, I stored the output (the coefficients that would minimize the residual sum of squares) in two separate variables. Afterwards, I created a range of the concentration of substrate values, as well as the corresponding reaction velocity values I calculated using equation 1 and the coefficients I found from the fminsearch() function. I then plotted the original data points, as well as the nonlinear regression model on the same graph. I also added axis labels and a title. I then calculated the coefficient of determination. I first calculated the residual sum of squares using my MichaelisMentis() function with the original data and new coefficients as inputs. I then calculated the total sum of squares by finding the differences of each reaction velocity value and the average reaction velocity, squaring the results, and summing all of the squared differences. I then set the coefficient of determination to 1 minus the residual sum of squares divided by the total sum of squares, and outputted the results. Lastly, I added comments to make my code easier to understand for users.

M20 - Introduction to Computer Programming with MATLAB
Instructor: Prof. Enrique López Droguett, Ph.D.
Teacher Assistants: M. Fidansoy, G. San Martín, M. Pishahang, V. Vela.
Fall 2023 – UCLA
Student: *Alex Lie*
UCLA ID: *905901892*

Calculations and results

2)



3)

```
Vmax_LS =

    0.3767


KM_LS =

  544.9243


R_squared =

    0.9958
```

The graph shows my nonlinear regression plotted with the data points. The coefficient of determination for my nonlinear regression model is shown above, as well as the optimized maximum velocity and Michaelis constant.

Discussions and Conclusions

As seen in figure 2, as the concentration of the substrate increases, the velocity of the reaction increases. The coefficient of determination is 0.9958 (3), which means there is a strong relationship between the concentration of the substrate and the velocity of the reaction. This coefficient of determination tells us the nonlinear model I developed is a really good fit for the data.

Task 2: Multivariate Nonlinear Regression

Introduction

I am using MATLAB to develop multivariate nonlinear regression functions and find the coefficients in the equations that minimize the residual sum of squares. The first case has two independent variables (W=6), and the second case has three independent variables. I will plot the nonlinear regression model from the first case, as well as the original data points, on a 3D plot. I will then evaluate how well the equation relates the independent variables and the dependent variables.

Model and Theory

M20 - Introduction to Computer Programming with MATLAB
Instructor: Prof. Enrique López Droguett, Ph.D.
Teacher Assistants: M. Fidansoy, G. San Martín, M. Pishahang, V. Vela.
Fall 2023 – UCLA
Student: *Alex Lie*
UCLA ID: *905901892*

1) $Y(D, W, H) = D^{a} \cdot W^{b} \cdot H^{c} + d$

Methodology

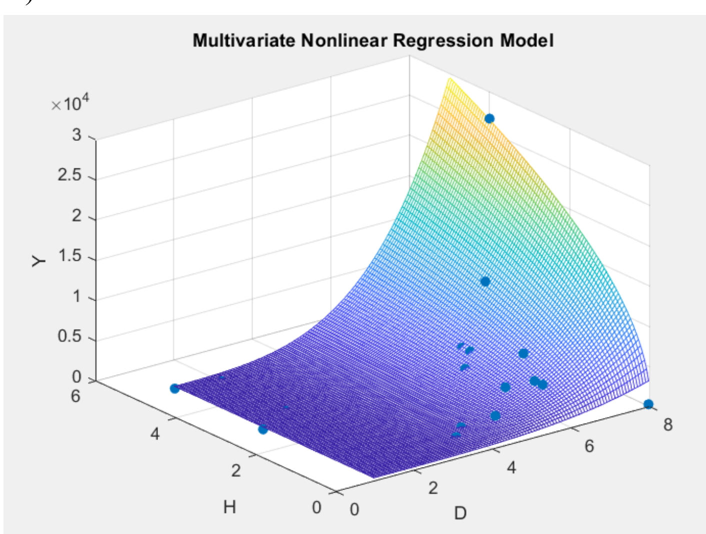
I first cleared the workspace and command window. I then imported the data from the Excel spreadsheet using the readmatrix() function and declaring variables. For the first case, I created another variable W_c that is set to the value of 6. I then declared an options variable that sets the options for when I find the minimum of a multivariable function which includes the tolerance of the independent variable, the tolerance of the value of the function, and the maximum number of iterations. I then used the fminsearch() function to find the values of the coefficients a, b, c, and d that would minimize the residual sum of squares of the value of the function. The parameters of the fminsearch() function are the Yfunction() function I made, my guesses for the coefficients (a, b, c, and d), and the options variable I declared. I created my Yfunction() function at the bottom of my code, which takes in the Y, D, W, and H data, and a vector of coefficients as inputs. My function calculates the value of Y based off equation 1 using the data inputs, and then finds the residual sum of squares by finding the difference of the calculated values of Y by the actual values of Y, squaring it, and adding up all of the squared differences for each value of Y. After running the fminsearch() function with W equaling 6 for the first case (two independent variables), I outputted all four coefficients (a, b, c, and d) that would minimize the residual sum of squares. After, I created a range of values for D and H, as well as calculated values of Y for each combination of D and H using linspace(), meshgrid(), the coefficients from fminsearch(), and equation 1. I then plotted the nonlinear regression model from the first case and the original data points in a 3D graph using scatter3() and mesh(). I also added axis labels and a title. I then ran the fminsearch() function again for the second case (with three independent variables), but instead of W equaling 6, it is now set as the W data from the spreadsheet. The rest of the parameters are the same except my guesses are slightly different. I then outputted my new coefficients (a, b, c, and d) for the second case. Lastly, I added comments to make my code easier to understand for users.

Calculations and results

2)

```
a_1 =

    4.2138

b_1 =

    0.4126

c_1 =

    0.4838

d_1 =

  329.5281
```

3)



4)

```
a_2 =

    4.4622

b_2 =

    0.2139

c_2 =

    0.5571

d_2 =

  384.9126
```

The coefficients on the left are for the nonlinear regression model with two independent variables. The 3D graph in the middle shows the nonlinear regression model with two independent variables, as well as the original data points. The coefficients on the right are for the nonlinear regression model with three independent variables.

Discussions and Conclusions

The coefficients from the first case (with two independent variables) (2), and the second case (with three independent variables) (4) do not seem to vary drastically. The nonlinear regression model seems to model the data well, since the data points are near the mesh surface (3). Based on results, I think the equation is a good assumption for the variable relationship between inputs and the output. There may even be another equation that can model the relationship between the inputs and the outputs even better.