Problem Set 05 · Non-Linear Regression

Instructions

1. Read each problem carefully before starting your work. You are responsible for following all instructions within each problem. Remember that all code submissions must follow the course programming standards.
2. Below are the expected deliverables for each problem. Use problem-specific templates when provided. Name your files to match the format in the table below. Note that you must include in your deliverables any data files required to run your code, your published code, and your original m-file.

Item	Type	Deliverable to include in Submission
Problem 1: Uranium Recovery from Seawater	Paired	☐ PS05_uranium_*login1*_*login2*.m ☐ PS05_uranium_ *login1*_*login2*_report.pdf ☐ Data file loaded into your m-file
Problem 2: Fuel Spray Ignition Delay	Individual	☐ PS05_ignition_delay_*yourlogin*.m ☐ PS05_ignition_delay _*yourlogin*_report.pdf ☐ Data files loaded into your m-file

3. Save all files to your Purdue career account in a folder specific to PS05.
4. When you are ready to submit your assignment,
   - Compress all the deliverables into one zip file and name it **PS05_yourlogin.zip**. Be sure that you
     i. Only compress files using **.zip** format. No other compression format will be accepted.
     ii. Only include deliverables. Do **not** include the problem document, blank templates, etc.
   - Submit the zip file to the Blackboard drop box for PS05 before the due date.
5. After grades are released for this assignment, access your feedback via the assignment rubric in the My Grades section of Blackboard.

Notes Before You Begin

Formatting Reminder

Always format your text, plots, and numerical outputs in a professional manner.

- Numerical values must have a reasonable number of decimal places. Include units when necessary.
- Displayed text should be descriptive and professional. Use complete sentences.
- Plots require a title, x- and y-axis labels, and gridlines. Multiple data sets on a single plot require a legend.

Problem 1: Uranium Recovery from Seawater

Paired

Learning Objectives

Your work on this problem may be assessed using any of the following learning objectives:

- Create and execute a script
- Assign and manage variables
- Manipulate arrays (vectors or matrices)
- Manage text output
- Import numeric data stored in .csv and .txt files
- Create and evaluate x-y plots suitable for technical presentation
- Perform linear regression
- Identify function types from graphs of bivariate data, specifically linear, power, exponential, and logarithmic
- Confirm function identification using a combination of linear and log transformations of the independent and dependent data variables
- Create plots with linear and/or log axis scales (MATLAB)
- Linearize and plot data appropriately
- Linearize power, exponential, and logarithmic functions
- Determine the linear and general forms of the equations for linear, power, exponential, and logarithmic functions
- Use the function to make predictions only when appropriate

Problem Setup

Uranium is the key material necessary for nuclear power. Earth's oceans are the largest source of uranium on the planet, but the concentration is very low (approximately 3 parts per billion). While the oceans are believed to have a vast uranium supply, it is expensive and time-consuming to extract using current methods. The primary method to extract uranium from seawater is to use an adsorbent material to attract and collect the uranium ions from the water. The adsorbent is made of a plastic fiber that contains a compound called amidoxime, which attracts uranium.

As a nuclear engineer, you are part of a team that is testing an amidoxime-based fabric for uranium adsorption. Your team submerged several samples of the fiber in a testing tank filled with a uranium-spiked solution. They removed each fiber sample at a different time and measured the sample's uranium uptake, measured in micrograms of uranium per gram of fiber ($\mu g/g$). They compiled the data into a spreadsheet named **Data_uranium_adsorption.csv**. You must now model the experimental data.

Problem Steps

1. Open **PS05_uranium_template.m** and complete the header. Save it using the name indicated in the Deliverables list. Use programming standards to place code in the appropriate sections within the template.

    a. Create a figure with 2x2 subplots (i.e., 2 rows and 2 columns) that contains plots of the data on linear, semilogx, semilogy, loglog scales. Format the subplots for technical presentation, paying particular attention to the labels on the x- and y-axes.

    Hint: You can control the font size of all subplot elements by using the FontSize property on the current axis (`gca`) within the set command. For example, the command to use 8-point font on one subplot is

    ```
    set(gca,'FontSize',8).
    ```

    If you chose to use the `set` command above to manage the font size, then you will need that line of code for each subplot.

    Hint: You can display Greek letters in MATLAB graph formatting.

    b. Use the subplots to diagnose the type of function that best represents the form of the relationship. In the `ANALYSIS` section, answer the following:

    Q1: What type of function best represents the relationship between the data? Justify your answer by making reference to the plots and the axes' scaling.

    c. In the `LINEARIZATION` section,

    i. Linearize the data using the technique that is most appropriate for the function type you diagnosed above.

    ii. Find the linearized form of the equation using least squares regression.

    iii. Print the linearized form of the equation to the Command Window. Use clear, appropriate variable names in place of x and y in the equation.

    iv. Plot the linearized data and the trend line on a single plot in a second figure window. Be sure to label the plot for technical presentation, paying particular attention to the labels on the x- and y-axes

    d. In the `UPTAKE MODEL` section,

    i. Determine the general form of the best-fit equation for the relationship.

    ii. Clearly display the equation in the MATLAB Command Window. Use clear, appropriate variable names in place of x and y in the equation.

    iii. Plot, on a single plot and in a third figure window, the original data with the best-fit curve.

2. In the `ANALYSIS` section of your code, answer the following questions:

    Q2: The lead engineer wants to know the predicted uranium uptake after 10 hours, 100 hours, and 250 hours. Perform the calculations in the `PREDICTIONS` section of the template. In the `Q2` section, write a short paragraph explaining your results along with any limitations of the model that the lead engineer needs to know.

3. Publish your script as a PDF and name it as required in the Deliverables List.

Reference: Byers, Maggie; Schneider, Erich; Landsberger, Sheldon; Eder, Sebastian (2017), "Data for: Neutron Activation Analysis for the Characterization of Seawater Uranium Adsorbents ", Mendeley Data, v1 http://dx.doi.org/10.17632/yzxndgfvtf.1

Problem 2: Fuel Spray Ignition Delay

Individual

Learning Objectives

Your work on this problem may be assessed using any of the following learning objectives:

- Create and execute a script
- Assign and manage variables
- Manipulate arrays (vectors or matrices)
- Manage text output
- Import numeric data stored in .csv and .txt files
- Create and evaluate x-y plots suitable for technical presentation
- Perform linear regression
- Identify function types from graphs of bivariate data, specifically linear, power, exponential, and logarithmic
- Confirm function identification using a combination of linear and log transformations of the independent and dependent data variables
- Create plots with linear and/or log axis scales (MATLAB)
- Linearize and plot data appropriately
- Linearize power, exponential, and logarithmic functions
- Determine the linear and general forms of the equations for linear, power, exponential, and logarithmic functions
- Use the function to make predictions only when appropriate

Problem Setup

Turbine engines are used in aviation, ground transportation, and heavy machinery. These engines inject fuel into a compressor where the fuel spontaneously ignites due to the high temperature of the compressed air. There is an ignition delay between when the fuel is injected and when the fuel starts to combust. Minimizing that delay can reduce pollution and make the engine more efficient. One way to affect the delay is by changing the air temperature in the compressor.

You are an engineer working for an engine maufacturer that is studying ignition delay. Your colleagues have run experiments on two different fuel types, Diesel #2 and Jet-A. They have provided you with two data files, one for the diesel results (**Data_diesel_ignition_delay.csv**) and one for the jet fuel results (**Data_jetA_ignition_delay.csv**). The data files each have two data columns. One column is related to temperature and is represented as 1000 divided by air temperature in degrees Kelvin (1000/T (K^{-1})). The other column is the ignition delay in milliseconds (msec). Your assignment is to model the relationship between the measured temperature data and the ignition delays.

Problem Steps

1. Open **PS05_ignition_delay_template.m** and complete the header. Save it using the name indicated in the Deliverables list.

2. Write a script that will

a. Create a figure or figures of subplots that contain plots of the data on linear, semilogx, semilogy, loglog scales. Format the subplots for technical presentation, paying particular attention to the labels on the x- and y-axes.

   You need to decide how to display the plots. For example, you could show both data sets in one figure of 2x2 subplots as in Problem 1 or you could create two figures of subplots, with one for each data set.

   Hint: you can include superscripts in plot labels.

b. Use the subplots to diagnose the type of function that best represents the form of relationship. In the `ANALYSIS` section, answer the following:

   Q1: What type of function best represents the relationship between the data? Justify your answer by making reference to the plots and the axes' scaling.

c. Linearize the data for both fuels using the technique that is most appropriate for the function type you diagnosed in Q1.

d. Find the linearized form of the equations using least squares regression.

e. Print the linearized form of the equations to the Command Window. Use clear, appropriate variable names in place of x and y in the equation.

f. Plot the linearized data and the trend lines on a single plot in a new figure window. Be sure to label the plot for technical presentation, paying particular attention to the labels on the x- and y-axes.

g. Determine the general form of the best-fit equations for each relationship.

h. Clearly display the equations in the MATLAB Command Window. Use clear, appropriate variable names in place of x and y in the equation.

   Hint: you can display very small values in scientific notation using the `%d` formatting character. You still need to control the number of decimal places shown.

i. For each set of fuel data: plot, on a single plot and in a new figure window, the original data with its best-fit curve.

j. Follow good programming standards and place code in appropriately named sections of the script. You may need to add or rename sections in the template.

3. Publish your script as a PDF and name it as required in the Deliverables List.

Reference

https://www.sciencedirect.com/science/article/pii/0010218082900220?via%3Dihub