A First Course on Kinetics and Reaction Engineering

Example 13.3

Problem Purpose

This example illustrates the analysis of CSTR data using linear least squares for a situation where there are two independent variables in the model being fit to the data.

Problem Statement

Aqueous reaction (1) was studied using a steady state CSTR operating at various space times and inlet concentrations of A and B. Table 1 presents the results from several experiments that were all conducted at the same temperature. Find the best power-law rate expression to describe the data presented in that table.
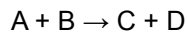
$$A + B \rightarrow C + D \tag{1}$$

Table 1

C_A^0 (M)	C_B^0 (M)	Space Time (s)	C_A (M)
1.00	1.00	94.7	0.10
1.00	0.75	20.1	0.29
1.00	0.50	2.9	0.58
1.00	0.25	1.3	0.83
0.75	1.00	0.9	0.50
0.75	0.75	1.1	0.52
0.75	0.50	3.8	0.40
0.75	0.25	1.1	0.63
0.50	1.00	2.4	0.25
0.50	0.75	1.0	0.37
0.50	0.50	2.3	0.29
0.50	0.25	2.4	0.36
0.25	1.00	21.1	0.05
0.25	0.75	1.2	0.18
0.25	0.50	4.4	0.13
0.25	0.25	20.8	0.09

Problem Analysis

This problem involves a chemical reaction taking place within a CSTR. Upon reading the question, we see that the rate expression is unknown; we are asked to generate one using the CSTR kinetics data that are provided. To answer, we will write the design equation for the reactor, substituting the rate expression being tested into it, fit the resulting equation to the experimental kinetics data, and decide whether or not the fit is satisfactory.

Problem Solution

Since only one chemical reaction is taking place and the reactor is isothermal, all that is needed to model the reactor is a single mole balance design equation, equation (2) written for i = any one of the reactants or products. The problem asks us to find the best power-law rate expression, so the rate expression is given by equation (3). The mole balance design equation for A, after substitution of the rate expression, is given by equation (4), where α and β represent the unknown reaction orders in A and B.

$$\dot{n}_i - \dot{n}_i^0 = V r_{i,j} \tag{2}$$

$$r_1 = k C_A^\alpha C_B^\beta \tag{3}$$

$$\dot{n}_A - \dot{n}_A^0 = -V k C_A^\alpha C_B^\beta \tag{4}$$

Equation (4) is not linear, but if the logarithm of its negative is taken, the resulting equation (5) is linear. This can be seen more easily if y, x_1, x_2, m_1, m_2 and b are defined as given in equations (6) through (11); upon substitution of these definitions into equation (5), equation (12) results, and equation (12) is clearly linear.

$$\ln\left(\dot{n}_A^0 - \dot{n}_A\right) = \alpha \ln\left(C_A\right) + \beta \ln\left(C_B\right) + \ln\left(Vk\right) \tag{5}$$

$$y = \ln\left(\dot{n}_A^0 - \dot{n}_A\right) \tag{6}$$

$$x_1 = \ln\left(C_A\right) \tag{7}$$

$$x_2 = \ln\left(C_B\right) \tag{8}$$

$$m_1 = \alpha \tag{9}$$

$$m_2 = \beta \tag{10}$$

$$b = \ln\left(Vk\right) \tag{11}$$

$$y = m_1 x_1 + m_2 x_2 + b \tag{12}$$

Since the model equation is linear, linear least squares can be used to fit it to the experimental data (see Supplemental Unit S3). Linear least squares fitting can be performed manually, using a calculator,

using a spreadsheet or using mathematics software. No matter which tool one chooses to use, it will be necessary to provide the following information and input data:

- the number of independent (x) variables
- whether or not the model includes an intercept (b)
- a set of experimental data points, each of which consists of a value for the dependent variable (y) and corresponding values for each of the independent variables (x_i)

In order to fit equation (12) to the experimental data, we need to calculate values of y, x_1 and x_2 corresponding to each of the data points given in the table in the problem statement. As seen in equation (6), in order to calculate values of y, we will need to calculate the inlet and outlet molar flow rates of A. We are given the inlet and outlet concentrations of A as well as the space time. None of these quantities is extensive, so we are free to choose a basis for our calculations. Here we will choose a reactor volume, V, of 1 L as our basis. Having done so, the inlet volumetric flow rate can be calculated from the space time for each experiment, as shown in equation (13). Since this system is liquid phase, the inlet and outlet volumetric flow rates may be assumed to be equal.

$$\dot{V}^0 = \frac{V}{\tau} = \dot{V} \tag{13}$$

Table 1 lists the inlet and outlet concentrations of A, and these can be used to calculate the inlet and outlet molar flow rates of A as in equations (14) and (15). Once these are known, the value of y can be calculated using equation (6). The value of C_A is given in the data table in the problem statement, so x_1 can also be calculated for each data point using equation (7).

$$\dot{n}_A = C_A \dot{V} \tag{14}$$

$$\dot{n}_A^0 = C_A^0 \dot{V}^0 \tag{15}$$

In order to calculate x_2, we will need to first calculate C_B. The concentration of B is related to the outlet molar flow rate of B and the outlet volumetric flow rate as given in equation (16). A mole table, or the definition of the extent of reaction, can be used to derive the relationship between the outlet molar flow rates of A and B as given in equation (17). The inlet molar flow rate of B can be found using equation (18), and once that is done, C_B can be calculated using equation (16), and the result can be used in equation (8) to calculate x_2.

$$C_B = \frac{\dot{n}_B}{\dot{V}} \tag{16}$$

$$\dot{n}_B = \dot{n}_B^0 - \dot{n}_A^0 + \dot{n}_A \tag{17}$$

$$\dot{n}_B^0 = C_B^0 \dot{V} \tag{18}$$

Thus, we can generate a set of (x_1,x_2,y) data as just described. The model has two independent variables, x_1 and x_2, and it *does* include an intercept. With this information and input data we have everything needed to fit the model to the data. Upon doing so, using whichever linear least squares fitting tool one chooses to employ, the resulting output shows that the correlation coefficient, r^2, is equal to 0.9768, the best value of the slope, m_1, is equal to 1.39 ± 0.13, the best value of the slope, m_2, is equal to 0.47 ± 0.1 and the best value of the y-intercept, b, is equal to -0.30 ± 0.27 (95% confidence limits based upon the data given in the problem statement). The fitting tool may also produce a parity plot like that shown in Figure 1 and residuals plots like those shown in Figures 2 and 3. If these plots are not created, they can be generated easily.
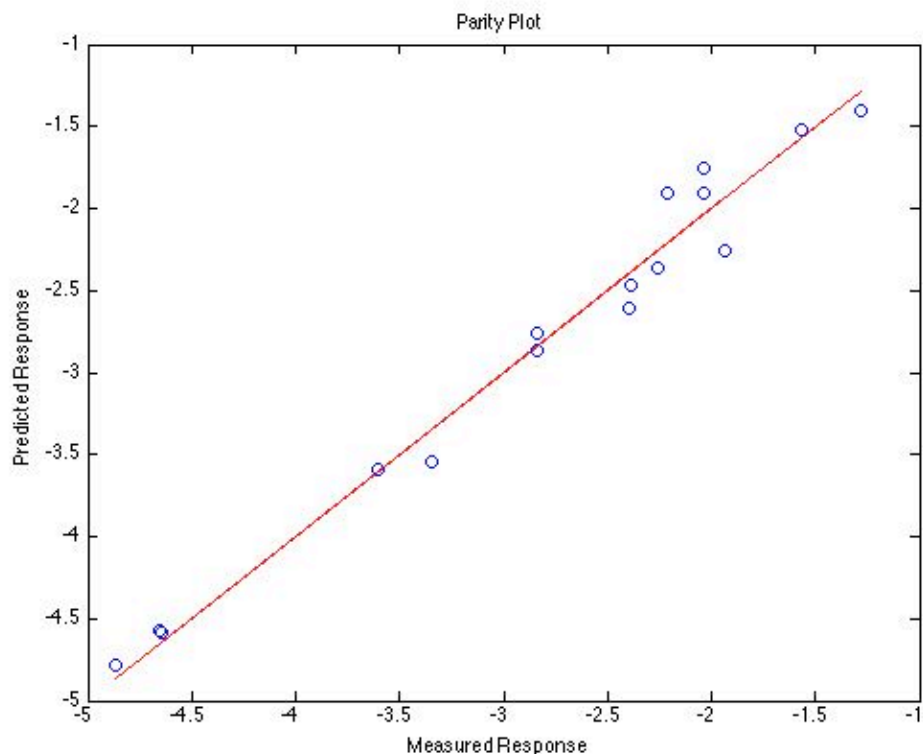

Figure 1. Parity plot comparing the experimental value of y to the value predicted by the fitted model.
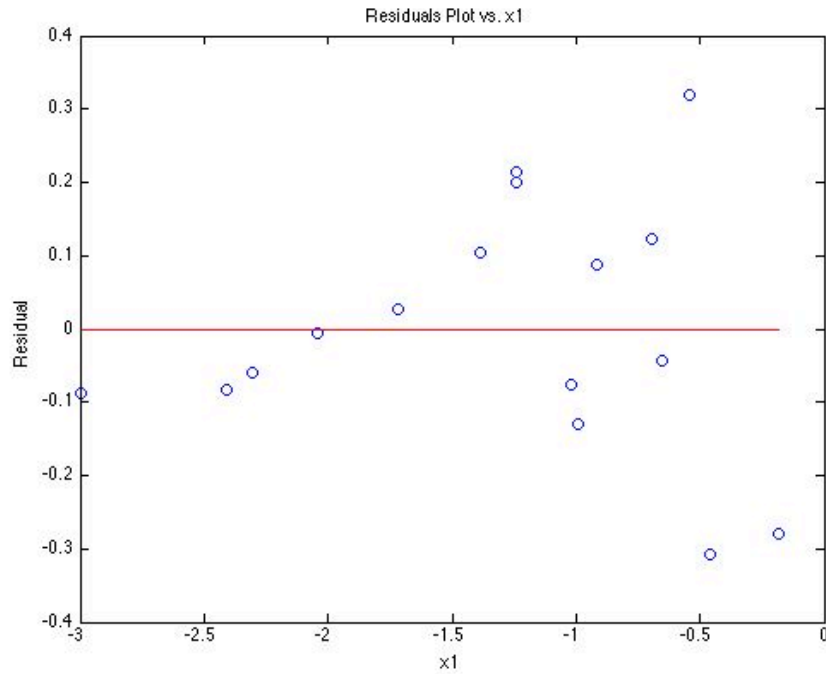
Figure 2. Residuals plot showing the difference between the experimental and model-predicted values of y as a function of x_1.
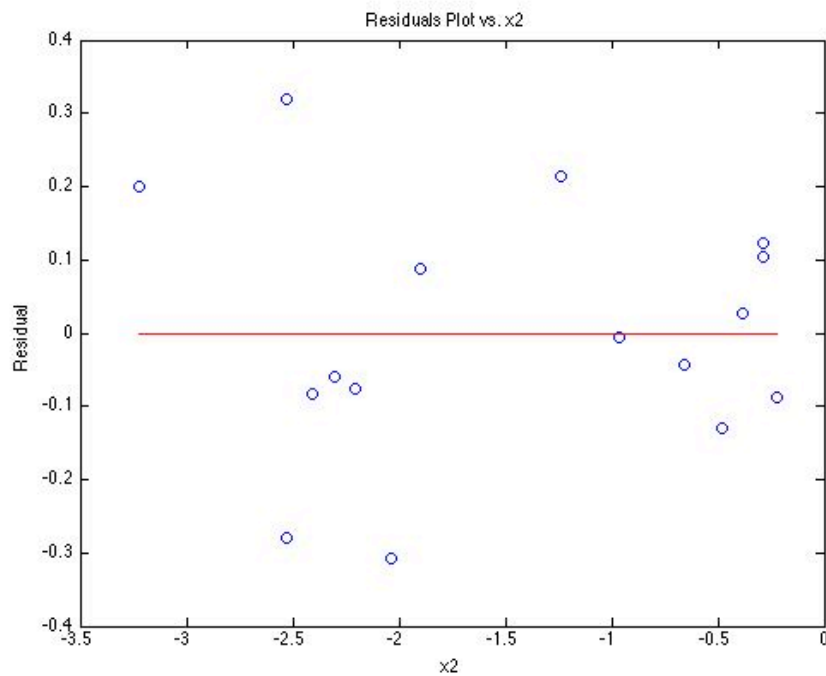

Figure 3. Residuals plot showing the difference between the experimental and model-predicted values of y as a function of x_2.

Before the best values for m_1, m_2 and b can be accepted, one must decide whether the final model is sufficiently accurate. In this case, the accuracy of the model can be assessed using the correlation coefficient, the parity plot and the residuals plots. The closer the correlation coefficient is to a value of 1.0, the better the fit of the model to the data. In this case, the correlation coefficient of 0.9768 indicates a reasonably good fit. Additionally, if the fit is accurate, then the deviation of the data points in the parity plot from the diagonal line should be small. This can be seen to be the case for this problem by examination of Figure 1. Finally, if the fit is accurate, the residuals should scatter randomly about zero in the residuals plots; there should not be any apparent trends in the scatter of the data about the line in the figures. Examination of Figures 2 and 3 shows this to be true for the present problem. Thus, the model does appear to offer a sufficiently accurate representation of the experimental data, and the values of the slopes and intercept can accepted.

We are interested in the values of the parameters, α, β and k, not the slopes and intercepts. In this problem, however the slopes, m_1 and m_2, are equal to α and β, so their values are known: α = 1.39 ± 0.13 and β = 0.47 ± 0.1. The value of k can be computed from the intercept, b, by rearranging equation (11), as shown in equation (19). A differential error analysis shows that if a model parameter, p, is related to the slopes, m_i, and intercept, b, of a linearized form of the model, as in equation (20), then the uncertainty in that parameter, λ_p, is related to the slopes, intercept and their uncertainties, λ_{m_i} and λ_b, according to equation (21). Applying that relationship to the present problem shows that the uncertainty in k should be calculated using equation (22). Applying equations (19) and (22) shows that k = 0.738 ± 0.200 $L^{0.86}$ $mol^{-0.86}$ s^{-1}.

$$k = \frac{\exp(b)}{V} \qquad (19)$$

$$p = f(\underline{m}, b) \qquad (20)$$

$$\lambda_p = \sqrt{\sum_i \left(\frac{\partial f}{\partial m_i}\right)^2 \lambda_{m_i}^2 + \left(\frac{\partial f}{\partial b}\right)^2 \lambda_b^2} \qquad (21)$$

$$\lambda_k = \frac{\exp(b)}{V} \lambda_b \qquad (22)$$

Calculation Details Using MATLAB

Three MATLAB script files are provided with Supplemental Unit S3. The file names indicate the number of independent variables and whether or not the model has an intercept. The script named FitLinmSR is used when the model has one independent variable (x) and *does not* include the intercept (b). FitLinmbSR is used when the model has one independent variable and *does* include the intercept, and FitLinSR is used when the model has two or more independent variables. (With MATLAB, when the model has two or more independent variables, it must have an intercept; Supplemental Unit S3 describes

how to convert a model without an intercept into a model that has an intercept.) In this problem the model has two independent variables, x_1 and x_2, so the script file named FitLinSR will be used. To do so, the script file must be located in the current MATLAB working directory or in the MATLAB search path.

Before executing FitLinSR, the experimental values of x_1 and x_2 must be stored in a matrix named x, and the experimental values of y must be stored in a vector named y_hat. More specifically, the first column in matrix x must be a vector containing the values of x_1, the second column in matrix x must be a vector containing the values of x_2 and the last column in matrix x must contain the value 1.0 in every row, as described in Supplemental Unit S3. Upon execution of the script, it will return the correlation coefficient as r_squared, the slopes as a vector named m, the 95% confidence limits on the slopes as a vector named m_u, the intercept as b and the 95% confidence limits on the intercept as b_u. It will also generate a parity plot and two residuals plots (Figures 1 through 3). Once these values are available, the best values for α, β and k and their uncertainties can be computed using equations (9), (10), (19) and (22). The commands for performing all these tasks can be entered at the MATLAB command prompt, but here they have been recorded in the MATLAB file named Example_13_3.m which accompanies this solution. Listing 1 shows the code from Example_13_3.m and Listing 2 shows the output that it produces, except for the plots.

```
% MATLAB file used in the solution of Example 13.3 of "A First Course on
% Kinetics and Reaction Engineering."

% Enter data provided in the problem statement in consistent units

V = 1; % L (basis)
CA0=[1.00 1.00 1.00 1.00 0.75 0.75 0.75 0.75 0.50 0.50 0.50 0.50 0.25 0.25 0.25
     0.25]'; % mol per L
CB0=[1.00 0.75 0.50 0.25 1.00 0.75 0.50 0.25 1.00 0.75 0.50 0.25 1.00 0.75 0.50
     0.25]'; % mol per L
tau=[94. 20.1 2.9 1.3 0.9 1.1 3.8 1.1 2.4 1.0 2.3 2.4 21.1 1.2 4.4 20.8]'; % a
CA=[0.10 0.29 0.58 0.83 0.50 0.52 0.40 0.63 0.25 0.37 0.29 0.36 0.05 0.18 0.13
    0.09]'; % mol per L

% Calculate corresponding values of x1, x2 and y_hat
VFR = V./tau;
nA = VFR.*CA;
nA0 = VFR.*CA0;
nB0 = VFR.*CB0;
nB = nB0 - nA0 + nA;
CB = nB./VFR;
x1 = log(CA);
x2 = log(CB);
y_hat = log(nA0-nA);
% Create the matrix x
x = [x1 x2 ones(length(x1),1)];

% Use the MATLAB script file "FitLinSR.m" from "A First Course on
% Kinetics and Reaction Engineering" to fit a straight line to the data.
FitLinSR

% Calculate k, alpha and beta and their uncertainties
alpha = m(1)
alpha_u = m_u(1)
beta = m(2)
beta_u = m_u(2)
k = exp(b)/V
k_u = b*k*b_u
```

Listing 1. Contents of the file Example_13_3.m used to solve this problem using MATLAB.

```
>> Example_13_3

r_squared =

    0.9768

m =

    1.3893
    0.4717

m_u =

    0.1317
    0.1033

b =

   -0.3041

b_u =

    0.2709

alpha =

    1.3893

alpha_u =

    0.1317

beta =

    0.4717

beta_u =

    0.1033

k =

    0.7378

k_u =

   -0.0608
```

Listing 2. Output generated upon execution of Example_13_3.m.