Università della Svizzera italiana | Institute of Computing CI

Numerical Computing **2023**

Student: Hun Rim **Discussed with:** FULL NAME

Bonus assignment

Due date: Wednesday, 22 November 2023, 11:59 PM

Numerical Computing 2023 — Submission Instructions

(Please, notice that following instructions are mandatory: submissions that don't comply with, won't be considered)

- Assignments must be submitted to iCorsi (i.e. in electronic format).
- Provide both executable package and sources (e.g. C/C++ files, MATLAB). If you are using libraries, please add them in the file. Sources must be organized in directories called:

  Project_number_lastname_firstname

  and the file must be called:

  project_number_lastname_firstname.zip
  project_number_lastname_firstname.pdf
- The TAs will grade your project by reviewing your project write-up, and looking at the implementation you attempted, and benchmarking your code's performance.
- You are allowed to discuss all questions with anyone you like; however: (i) your submission must list anyone you discussed problems with and (ii) you must write up your submission independently.

Exercise 1: Inconsistent systems of equations [10 points]

Consider the following inconsistent systems of equations:

(a) $A_1 x = b_1$, where

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} b_1 = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$$

(b) $A_2 x = b_2$, where

$$A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} b_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Find the least squares solution x^* and compute the Euclidean norm of the residual, SE and RMSE.

solution:
Least Square solution x^* can be obtained by solving the following equation:

$$A^T A x = A^T b \tag{1}$$

Then from the x^* obtained, we can get the residual vector as following and from it, we can calculate the Euclidean norm and proceed to SE (Sum of Square Residuals) and RMSE (Root Mean Squared Error):

$$r = Ax^* - b \tag{2}$$

$$EuclideanNorm = ||r||_2 \tag{3}$$

$$SE = ||r||_2^2 \tag{4}$$

$$RMSE = \sqrt{\frac{SE}{m}} \tag{5}$$

Where m is the number of rows in residual vector.

(a)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} x = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} x = \begin{bmatrix} 11 \\ 0 \end{bmatrix}$$

$$x^* = \begin{bmatrix} 3.6667 \\ 0 \end{bmatrix}$$

$$r^* = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3.6667 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1.3333 \\ 1.6667 \\ -0.3333 \end{bmatrix}$$

$$EuclideanNorm = ||r||_2 = \sqrt{(-1.3333)^2 + (1.6667)^2 + (-0.3333)^2} = \sqrt{4.6667} \approx 2.1602$$

$$SE = ||r||_2^2 = (-1.3333)^2 + (1.6667)^2 + (-0.3333)^2 \approx 4.6667$$

$$RMSE = \sqrt{\frac{||r||_2^2}{m}} \approx \sqrt{\frac{4.6667}{3}} \approx 1.2472$$

(b)

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 & 2 \\ 3 & 6 & 3 \\ 2 & 3 & 3 \end{bmatrix} x = \begin{bmatrix} 9 \\ 10 \\ 9 \end{bmatrix}$$

If we re-arrange the formula for x^* we get:

$$x^* = (A_2^T A_2)^{-1} A_2^T b_2 \tag{6}$$

and the resulting x^* will be

$$x^* \approx \begin{bmatrix} 2 \\ -0.3333 \\ 2 \end{bmatrix}$$

$$r = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -0.3333 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix} \approx \begin{bmatrix} -0.3333 \\ -0.3333 \\ 0.3333 \\ 0 \end{bmatrix}$$

$$EuclideanNorm = ||r||_2 \approx \sqrt{(-0.3333)^2 + (-0.3333)^2 + (0.3333)^2 + (0)^2} \approx \sqrt{0.3333} \approx 0.5774$$

$$SE = ||r||_2^2 = (-0.3333)^2 + (-0.3333)^2 + (0.3333)^2 + (0)^2 \approx 0.3333$$

$$RMSE = \sqrt{\frac{||r||_2^2}{m}} \approx \sqrt{\frac{0.3333}{4}} \approx 0.2887$$

Exercise 2: Polynomials models for least squares [20 points]

(a) Write *leastSquare.m* function which calculates least squares x^*, euclidean norm, SE and RMSE of a matrix A and vector b, and write a script *ex2a.m* which computes the result of exercise 1.

```
function [x, EuclideanNorm, SE, RMSE] = leastSquares(A, b)
  x = pinv(A) * b;
  r = A * x - b;
  EuclideanNorm = norm(r);
  SE = EuclideanNorm ^ 2;
  MSE = SE / length(b);
  RMSE = sqrt(MSE);
end
```

The code above calculates least squares (*leastSquare.m*) using matlab library function and Euclidean norm is calculated by using the matlab norm function after calculating the residual vector. Sum of Square Residuals (SE) are calculated by directly squaring the Euclidean norm, and before the Root Mean Squared Error (RMSE) is calculated, the Mean Squared Error is calculated through dividing SE by length of vector b.

```
A_1 = [1, 0; 1, 0; 1, 0];
b_1 = [5; 2; 4];
A_2 = [1, 1, 0; 0, 1, 1; 1, 2, 1; 1, 0, 1];
b_2 = [2; 2; 3; 4];

[x_1, norm1, SE1, RMSE1] = leastSquares(A_1, b_1);
[x_2, norm2, SE2, RMSE2] = leastSquares(A_2, b_2);
```

This particular code (*ex2a.m*) defines the values of matrix A and vector b as given in the exercise 1, and it calculates and stores the least squares, Euclidean Norm, SE, RMSE using function in *leastSquare.m* as described above.

(b) Consider the linear model $y_i = \alpha_1 + \alpha_2 x_i$ and apply it to the crude oil and kerosene production data in the period 1980-2011. Write a script *linearModel.m* in which you use *leastSquares()* to compute the least squares solution x^* and the metrics of the residual. For each dataset, create a figure in which you plot the original data points and the linear model.

```
[year, production, ~] = readData(filePath);

duration = 2012 - 1980 + 1;

y = str2double(production);

x = [ones(duration, 1), year(:)];

[factors, ~, ~, ~] = leastSquares(x, y);

z = factors(1) + factors(2) * year(:); %based on y = mx + b;

% plot the line and scatter graph...
```

The code above simply follows the process of reading the data from the file, then converting 'x' and 'y' to viable format of matrix A and vector b, and recaclulating the linear model (line graph) according to the least squares returned and plotting them with the origin (scatter graph). During conversion, columns of the matrix A filled with $x_i^{columnNumber}$ (columnNumber starts from 0, and there are only 2 columns) and 'y' is placed as vector b.

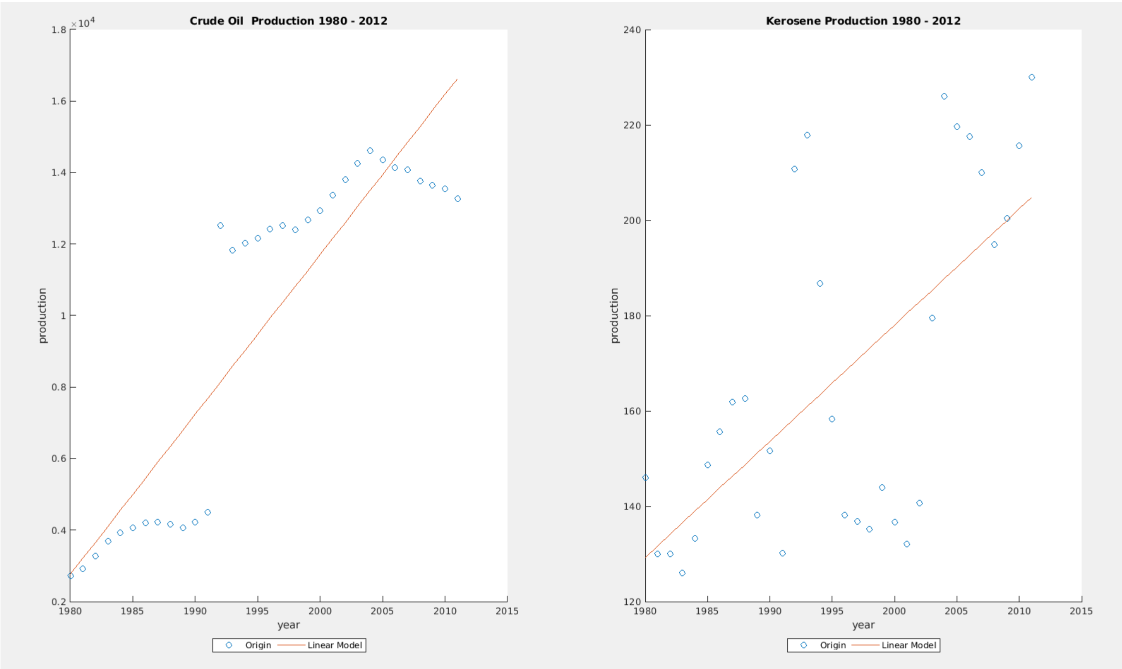

Figure 1: Linear Model and original data of crude oil and kerosene production from 1980 to 2011

(c) Consider the quadratic model $y_i = \alpha_1 + \alpha_2 x_i + \alpha_3 x_i^2$ and apply it to the crude oil and kerosene production data in the period 1980-2011. Write a script *quadraticModel.m* in which you use *leastSquares()* to compute the least squares solution x^* and the metrics of the residual. For each dataset, create a figure in which you plot the original data points and the quadratic model.

The general flow of the logic is very similar, but the matrix A is extended via the same logic so it would produce bigger least square vector (there are 3 columns now). Then instead of $y_i = mx_i + b$, quadratic modelling equation $y_i = a + bx_i + cx_i^2$ is used.

```
same as before ...
x = [ones(duration, 1), year(:), year(:).^2];

[factors, ~, ~, ~] = leastSquares(x, y);

z = factors(1) + factors(2) * year(:) + factors(3) * year(:).^2;
same as before ...
```

As a result of that change, resulting graph (quadratic) became more fitting than the previous linear graph shown in figure 1.

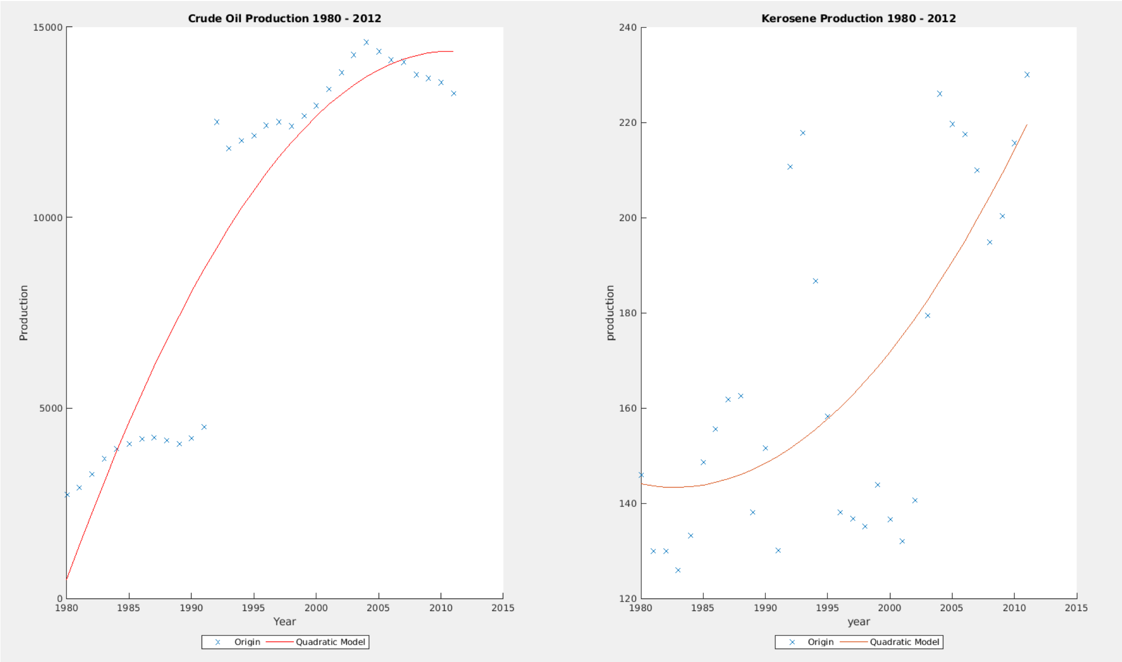

Figure 2: Quadratic Model and original data of crude oil and kerosene production from 1980 to 2011

(d) Consider the cubic model $y_i = \alpha_1 + \alpha_2 x_i + \alpha_3 x_i^2 + \alpha_4 x_i^3$ and apply it to the crude oil and kerosene production data in the period 1980-2011. Write a script cubicModel.m in which you use leastSquares() to com- pute the least squares solution x and the metrics of the residual. For each dataset, create a figure in which you plot the original data points and the cubic model.

Exercise 3: Analysis of periodic data [20 points]

Exercise 4: Data linearization and Levenberg-Marquardt method for the exponential model [20 points]

Exercise 5: Tikhonov regularization [15 points]