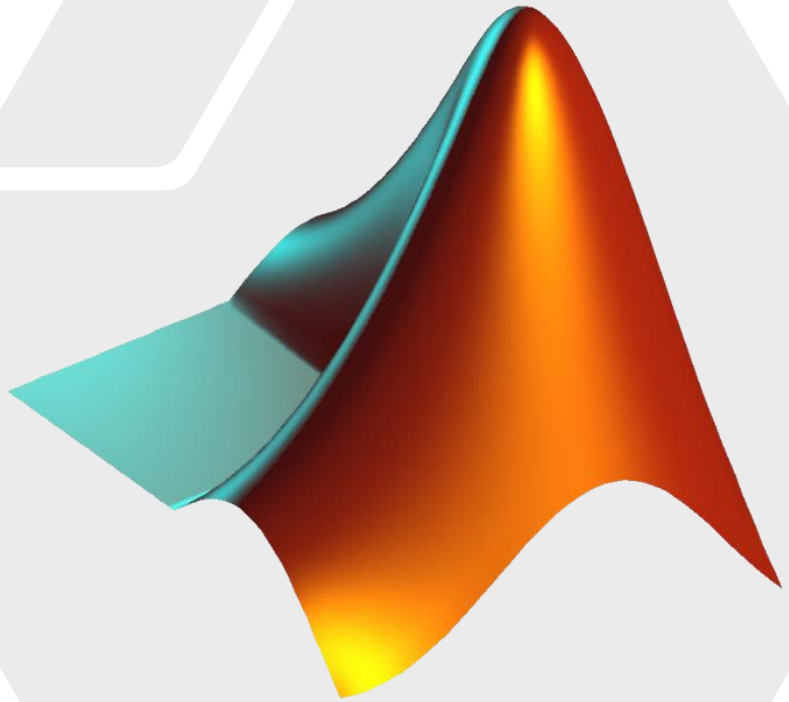
Data Analysis 1

Week 6 (16:00, 24/11/2021)



Vivatsathorn Thitasirivit

Student. Faculty of Engineering, Chulalongkorn University.

Week 6 Outlines

Data Analysis in MATLAB

Assignments

Week 6 Data Sample

MATLAB Programming: File *(Week 5)*

Reading a matrix/an array from a file

In MATLAB, you can open a file to read, write, or append text/number to it. The format of an array in a text form can vary. One of the most used format is Comma-delimited value (CSV) format. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

can be written in text format as:

```
1,2,3
4,5,6
7,8,9
```

There are other formats as well, e.g., tab-delimited, character-delimited, etc.

Reading a matrix from delimited file

```
M = readmatrix('filename.csv');
```

Writing a matrix to a CSV-formatted file

```
writematrix(M, 'filename.csv');
```

MATLAB also supports reading/writing a matrix from/to Excel spreadsheet file (.xls, .xlsx, .xlsm)

A CSV format will be autogenerated. If you wish to use other formats, consider the following command:

```
writematrix(M, 'filename.csv', 'Delimiter', 'tab');
```

MATLAB Programming: Correlation Coefficient

Calculating corr. coef.

In MATLAB, use

```
R = corrcoef(A, B) or corrcoef(A)
```

when A is the first random variable and B is the second random variable.
By term *random variable* means observed or measured variables.
For example, generate a data:

```
>> C = [1,2,3,4;2,3,2,4;3,4,1,4;4,5,0,3]
```

REMINDER
As you have studied correlation coefficient, we will get over what it is and start with how to do it in MATLAB.

We see that,

```
C =

     1     2     3     4
     2     3     2     4
     3     4     1     4
     4     5     0     3
```

If we use command *corrcoef(C)*, a weird matrix output is generated as the following.

```
COR =

    1.0000    1.0000   -1.0000   -0.7746
    1.0000    1.0000   -1.0000   -0.7746
   -1.0000   -1.0000    1.0000    0.7746
   -0.7746   -0.7746    0.7746    1.0000
```

MATLAB Programming: Correlation Coefficient

Understanding the matrix

```
COR =

    1.0000    1.0000   -1.0000   -0.7746
    1.0000    1.0000   -1.0000   -0.7746
   -1.0000   -1.0000    1.0000    0.7746
   -0.7746   -0.7746    0.7746    1.0000
```

```
C =

     1     2     3     4
     2     3     2     4
     3     4     1     4
     4     5     0     3
```

The output matrix is a matrix of “correlation of each pair”.

1.0000 in the diagonal means the correlation of column pair (1,1), (2,2), (3,3), and (4,4)

for other column coordinates such as (1,4) *as well as* (4,1) with value of –0.7746 means the correlation of data column 1 and 4

แล้วถ้าอยากรู้แค่ (1,2), (1,3), (1,4) ต้องดูแค่ส่วนไหน?

DUE: 6 DECEMBER 2021

Data Analysis Assignment 1

Given a CSV file:

dat_wk6.csv

Column 1 is an X-axis, i.e., constraining data and Col. 2–9 are Y-axes data (sampling data).
Therefore, least required matching data for plotting must be: (1,2), (1,3), … , (1,9)

Do the following:

1. ( ) Generate scatter plots of all data
2. ( ) Generate lines connecting the scatters
3. (★ ★) Generate linear fitting for each data
4. (★) Calculate correlation coefficients
5. (★ ★ ★) Generate a 5.1. variance-covariance and 5.2. correlation matrix
6. (*Bonus*) Generate plots and calculate correlation coefficients for other pairs, e.g. (2,3), (2,4), … , (8,9), etc. (just wonder for yourselves)

ทำนายว่าข้อมูลแต่ละคู่หลัก (Column pair) อาจเป็นข้อมูลที่วัดมาเป็นอะไรได้บ้าง

เช่น กราฟลักษณะนี้ อาจเป็นความสูงเทียบกับเวลา

Submit in only one *.m* file with all required information output able to be generated in it.