ACS234
Maths and Data Modelling

Tutorial 6
Wednesday 1pm online

https://github.com/ineskris/ACS234/tree/master/Tutorial6

Done in Lecture (week 6/7)

- General Linear Regression
- Non Linear Regression

General Linear Regression

$$Y = X\hat{a} + e \qquad \hat{a} = (X'X)^{-1}X'Y$$

Linear means the response Y is a linear function with the unknown parameter a.

Linear Regression with Nonlinear Terms, one example : $y = a_0 + a_1 x + a_2 e^{-x} = Xa$ $\quad X = \begin{pmatrix} 1 & x_0 & e^{-x_0} \\ 1 & x_1 & e^{-x_1} \\ 1 & x_2 & e^{-x_2} \\ 1 & x_3 & e^{-x_3} \end{pmatrix}$

Exercice 1

Based on each dataset, estimate the parameters of the **general linear regression model**. Calculate the MSE error.

a) $f_1(x) = a_0 x + a_1 sin(x)$

x	0	3	10	20
y	0	3.28	8.91	21.8

b) $f_2(x) = a_0 + a_1 \log(x)$

x	1	3	10	20
y	10	8.9	7.7	7

c) $f(x) = a_1 sin(2\pi x) + a_2 cos(2\pi x)$

x	0	1
y	-3	-3

For questions b and c, don't you see a simpler way to estimate the parameters?

Exercice 2

Use a **polynomial regression model of degree 2** to estimate the parameters from ex1a). What is the MSE error ?

Solutions can be found in

https://github.com/ineskris/ACS234/blob/master/Tutorial6/Tutorial6.ipynb

a) $\hat{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

b) $\hat{a} = \begin{pmatrix} 10 \\ -1 \end{pmatrix}$

c) $\hat{a} = \begin{pmatrix} anything \\ -3 \end{pmatrix}$

Non Linear Regression

the response Y is a non linear function with the unknown parameter a

$$\hat{a} = (X'X)^{-1}X'Y$$

Non Linear Regression with Nonlinear Terms, one example $y = a_0 + e^{-a_1 x}$

We want to minimise the difference between the response y and the estimated y induced by the model.

Exercice 3

Estimate the parameter η of the Weibull Distribution given below using the data points and Matlab or Python.
Write the function to minimise.

$$f(x) = k \times x^{k-1} \exp^{-(x^k)}$$

x	0.1	0.6	1.4	4
y	1.17	0.35	0.15	0.03

Tips :
In Matlab, you can use :

- lsqnonlin
- fminsearch

Equivalent in Python :

```
from scipy.optimize import curve_fit
```

Exercice 3 - solution

x	0.1	0.6	1.4	4
y	1.17	0.35	0.15	0.03

Function to minimise.

$$\hat{k} = argmin_k \sum_{i=1}^{4} y_i - f(x_i) = \sum_{i=1}^{4} y_i - k \times x_i^{k-1} \exp^{-(x_i^k)}$$

$$\hat{k} \approx 0.59$$