ACS234
Maths and Data Modelling

Tutorial 5
Wednesday 1pm online

https://github.com/ineskris/ACS234/tree/master/Tutorial5

Done in Lecture (week 5/6)

- Polynomial Regression
- General Linear Models

Polynomial Regression

Polynomial Model $y = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m + e$

Estimation (least squares method) $Y = X\hat{a} + e$ $\hat{a} = (X'X)^{-1}X'Y$

General Polynomial Regression - 2 dimension $y = a_0 + a_1 X_1 + a_2 x_2 + a_3 X_1^2 + a_4 X_1 X_2 + a_5 X_2^2 + e$

Exercice 1

x	0	1	2	3
f(x)	2	7	14	23

Based on the data above, estimate the parameters a0,a1,a2 of the **polynomial regression model**. Calculate the MSE error.

Exercice 2

X1	0	1	2	3
X2	12	12.3	12.6	12.9
f(x)	2	-3.3	-3.2	2.3

Based on the data above, estimate the parameters a0,a1,a2, a3, a4, a5 of the **general polynomial regression** model. Calculate the MSE error.

Exercice 1 - solution

$$y = a_0 + a_1 x + a_2 x^2 \qquad Y = X\hat{a} \qquad \hat{a} = (X'X)^{-1}X'Y$$

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \qquad X'X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} = A$$

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} = \frac{1}{det(A)} \begin{pmatrix} 98 \times 14 - 36 \times 36 & 36 \times 14 - 6 \times 98 & 6 \times 36 - 14 \times 14 \\ 14 \times 36 - 6 \times 98 & 4 \times 98 - 14 \times 14 & 6 \times 14 - 4 \times 36 \\ 6 \times 36 - 14 \times 14 & 14 \times 6 - 4 \times 36 & 4 \times 14 - 6 \times 6 \end{pmatrix}$$

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} 76 & -84 & 20 \\ -84 & 196 & -60 \\ 20 & -60 & 20 \end{pmatrix}$$

$$det(A) = 4 \times 14 \times 98 + 6 \times 36 \times 14 + 14 \times 6 \times 36 - 14 \times 14 \times 14 - 36 \times 36 \times 4 - 98 \times 6 \times 6 = 80$$

$$X'Y = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ 14 \\ 23 \end{pmatrix} = \begin{pmatrix} 46 \\ 104 \\ 270 \end{pmatrix} \qquad \hat{a} = \begin{pmatrix} 0.95 & -1.05 & 0.25 \\ -1.05 & 2.45 & -0.75 \\ 0.25 & -0.75 & 0.25 \end{pmatrix} \begin{pmatrix} 46 \\ 104 \\ 270 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$$

General Linear Models

There are three components to any GLM:

- *Random Component* : noise model or error model. e
- *Systematic Component* - the linear predictor $\eta = X\beta$
- *Link Function*, *η or g(μ)* - specifies the link between random and systematic components. $E[Y] = g^{-1}(X\beta)$

General linear model (GLM) includes multiple linear regression.

Example - logistic regression

$$X\beta = ln(\frac{\mu}{1-\mu}) \qquad \mu = E[Y]$$