Examples about Data Fitting

CoE 3SK3

February 13^{th} 2024

1 Line Fitting

Lecture notes, part 5, page 2-8

- a set of paired observations (x_1, y_1), (x_2, y_2), $\cdots$, (x_N, y_N);
- Model: $y = a_0 + a_1 x$
  - a_0 and a_1 are unknown parameters.
- Fitting error

$$\mathcal{L} = \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i)^2 \tag{1}$$

  is minimized when $\frac{\partial \mathcal{L}}{\partial a_0} = 0$ and $\frac{\partial \mathcal{L}}{\partial a_1} = 0$.

- $$\frac{\partial \mathcal{L}}{\partial a_0} = -2 \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i) = 0 \rightarrow \sum_{i=1}^{N} a_0 + \sum_{i=1}^{N} x_i a_1 = \sum_{i=1}^{N} y_i \tag{2}$$

- $$\frac{\partial \mathcal{L}}{\partial a_1} = -2 \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i) x_i = 0 \rightarrow \sum_{i=1}^{N} x_i a_0 + \sum_{i=1}^{N} x_i^2 a_1 = \sum_{i=1}^{N} x_i y_i \tag{3}$$

- Reformulate Eq.(2) and Eq.(3) into the form of Eq.(4).

$$\mathbf{X} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \mathbf{Y} \tag{4}$$

  where $\mathbf{X} \in \mathbb{R}^{2\times 2}$, $\mathbf{Y} \in \mathbb{R}^{2\times 1}$ and $\mathbf{Y} = [\sum_{i=1}^{N} y_i,\ \sum_{i=1}^{N} x_i y_i]^T$.

- The optimal a_0 and a_1 can be obtained by solving Eq.(4), i.e., $\mathbf{X}^{-1}\mathbf{Y}$

2 Quadratic Regression

Lecture notes, part 5, page 8-10

- a set of paired observations (x_1, y_1), (x_2, y_2), $\cdots$, (x_N, y_N);
- Model: $y = a_0 + a_1 x + a_2 x^2$
  - a_0, a_1 and a_2 are unknown parameters.
- Fitting error

$$\mathcal{L} = \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i - a_2 x_i^2)^2 \tag{5}$$

  is minimized when $\frac{\partial \mathcal{L}}{\partial a_0} = 0$, $\frac{\partial \mathcal{L}}{\partial a_1} = 0$ and $\frac{\partial \mathcal{L}}{\partial a_2} = 0$.

- $$\frac{\partial \mathcal{L}}{\partial a_0} = -2 \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i - a_2 x_i^2) = 0 \rightarrow \sum_{i=1}^{N} a_0 + \sum_{i=1}^{N} x_i a_1 + \sum_{i=1}^{N} x_i^2 a_2 = \sum_{i=1}^{N} y_i \tag{6}$$

- $$\frac{\partial \mathcal{L}}{\partial a_1} = -2 \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i - a_2 x_i^2) x_i = 0 \rightarrow \sum_{i=1}^{N} x_i a_0 + \sum_{i=1}^{N} x_i^2 a_1 + \sum_{i=1}^{N} x_i^3 a_2 = \sum_{i=1}^{N} x_i y_i \tag{7}$$

- $$\frac{\partial \mathcal{L}}{\partial a_2} = -2 \sum_{i=1}^{N} (y_i - a_0 - a_1 x_i - a_2 x_i^2) x_i^2 = 0 \rightarrow \sum_{i=1}^{N} x_i^2 a_0 + \sum_{i=1}^{N} x_i^3 a_1 + \sum_{i=1}^{N} x_i^4 a_2 = \sum_{i=1}^{N} x_i^2 y_i \tag{8}$$

- Reformuate Eq.(6), Eq.(7) and Eq.(8) into the form of Eq.(9).
- $$\mathbf{X} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \mathbf{Y} \tag{9}$$

  where $\mathbf{X} \in \mathbb{R}^{3\times 3}$, $\mathbf{Y} \in \mathbb{R}^{3\times 1}$ and $\mathbf{Y} = [\sum_{i=1}^{N} y_i,\ \sum_{i=1}^{N} x_i y_i,\ \sum_{i=1}^{N} x_i^2 y_i]^T$.

- The optimal a_0, a_1 and a_2 can be obtained by solving Eq.(9), i.e., $\mathbf{X}^{-1}\mathbf{Y}$