On Linear Regression

Harry Han

February 16, 2023

1 The Simplest: $y = ax + b$

Definitio 1.1 (R^2).

For a data set $\vec{x} = [x_0, x_1, x_2, \cdots]^T$ and $\vec{y} = [y_0, y_1, y_2, \cdots]^T$ with the same dimension, there exist a linear regression $y = ax + b$ such that the R^2 value is minimized.

The promised regression is:

$$a = \frac{m_2 n_1 + n_2}{1 - m_2 m_1}; b = \frac{m_2 n_1 + n_2}{1/m_1 - m_2} + n_1$$
$$m_1 = -\frac{\sum \vec{x}}{n}; n_1 = \frac{\sum \vec{y}}{n}; m_2 = -\frac{\sum \vec{x}}{|\vec{x}|^2}; n_2 = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}|^2} \tag{1}$$

Where $\sum \vec{x}$ denote the sum of all entries in the vector $\vec{x}$.