112-1 Introduction to AI Linear Regression

學號：412770116　　　　姓名：許嘉隆

1. 假設有個資料點 $(x_1, y_1),\ (x_2, y_2),\ \dots,\ (x_n, y_n)$，將課堂中利用最小平方法(Least Squared Method)以求得簡單線性迴歸(Simple Linear Regression)公式 $f(x) = ax + b$ 中的係數(斜率)與(截距)，從已知的資料點來計算與的值。

$Let\ \hat{y} = f(x) = ax + b$

$Distance:$

$$\begin{aligned} d_n &= \hat{y}_n - y_n \\ &= f(x_n) - y_n \\ &= ax_n + b - y_n \end{aligned}$$

$Loss\ function:$

$$\begin{aligned} E &= \textstyle\sum_{i=1}^{n} (d_i)^2 \\ &= \textstyle\sum_{i=1}^{n} \left(ax_i + b - y_i\right)^2 \end{aligned}$$

$Least\ Squared\ Method:$

$$\frac{\partial E}{\partial a} = 0,\ \frac{\partial E}{\partial b} = 0$$

$$\begin{aligned} \frac{\partial E}{\partial a} &= \frac{\partial}{\partial a}\left(\textstyle\sum_{i=1}^{n} (ax_i + b - y_i)^2\right) &&= 0 \\ &= \textstyle\sum_{i=1}^{n} \frac{\partial}{\partial a}(ax_i + b - y_i)^2 &&= 0 \\ &= \textstyle\sum_{i=1}^{n} 2(ax_i + b - y_i)\cdot x_i &&= 0 \\ &= \textstyle\sum_{i=1}^{n} (ax_i + b - y_i)\cdot x_i &&= 0 \\ &= a\textstyle\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i &&= 0 \\ (1) \quad &= a\textstyle\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i &&= \textstyle\sum_{i=1}^{n} x_i y_i \end{aligned}$$

$$\begin{aligned} \frac{\partial E}{\partial b} &= \frac{\partial}{\partial b}\left(\textstyle\sum_{i=1}^{n} (ax_i + b - y_i)^2\right) &&= 0 \\ &= \textstyle\sum_{i=1}^{n} \frac{\partial}{\partial b}(ax_i + b - y_i)^2 &&= 0 \\ &= \textstyle\sum_{i=1}^{n} 2(ax_i + b - y_i)\cdot 1 &&= 0 \\ &= \textstyle\sum_{i=1}^{n} (ax_i + b - y_i) &&= 0 \\ &= a\textstyle\sum_{i=1}^{n} x_i + b\cdot n - \sum_{i=1}^{n} y_i &&= 0 \\ (2) \quad &= a\textstyle\sum_{i=1}^{n} x_i + b\cdot n &&= \textstyle\sum_{i=1}^{n} y_i \end{aligned}$$

$$Let\ Sx = \sum_{i=1}^{n} x_i\ ,\ Sxx = \sum_{i=1}^{n} x_i^2\ ,\ Sy = \sum_{i=1}^{n} y_i\ ,\ Sxy = \sum_{i=1}^{n} x_i y_i.\ Then\ \begin{cases} a\cdot Sxx + b\cdot Sx &= Sxy \quad (1) \\ a\cdot Sx + b\cdot n &= Sy \quad (2) \end{cases}$$

$$\begin{array}{rr} & a\cdot Sxx\cdot n + b\cdot Sx\cdot n = Sxy\cdot n \\ -) & a\cdot Sx\cdot Sx + b\cdot n\cdot Sx = Sy\cdot Sx \\ \hline & a(Sxx\cdot n - Sx\cdot Sx) = Sxy\cdot n - Sy\cdot Sx \end{array}$$

$$\begin{array}{rr} & a\cdot Sxx\cdot Sx + b\cdot Sx\cdot Sx = Sxy\cdot Sx \\ -) & a\cdot Sx\cdot Sxx + b\cdot n\cdot Sxx = Sy\cdot Sxx \\ \hline & b(Sx\cdot Sx - n\cdot Sxx) = Sxy\cdot Sx - Sy\cdot Sxx \end{array}$$

$$\begin{aligned} a &= \frac{Sxy\cdot n - Sy\cdot Sx}{Sxx\cdot n - Sx\cdot Sx} \\ &= \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} y_i \cdot \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i\ \cdot \sum_{i=1}^{n} x_i} \end{aligned}$$

$$\begin{aligned} b &= \frac{Sxy\cdot Sx - Sy\cdot Sxx}{(Sx\cdot Sx - n\cdot Sxx)} \\ &= \frac{\sum_{i=1}^{n} x_i y_i \cdot \sum_{i=1}^{n} x_i\ - \sum_{i=1}^{n} y_i \cdot \sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i\ \cdot \sum_{i=1}^{n} x_i\ - n\cdot \sum_{i=1}^{n} x_i^2\right)} \end{aligned}$$

2.已知個資料點如上題，改採用非線性迴歸的拋物線，其公式 $f(x) = ax^2 + bx + c$。試嘗試利用相同的最小平方法來求得 a, b, c 的係數值。

$Let\ \hat{y} = f(x) = ax^2 + bx + c$

$Distance:$

$$\begin{aligned} d_n &= \hat{y}_n - y_n \\ &= f(x_n) - y_n \\ &= ax_n^2 + bx_n + c - y_n \end{aligned}$$

$Loss\ function:$

$$E = \textstyle\sum_{i=1}^{n} \left(ax_i^2 + bx_i + c - y_i\right)^2$$

$Least\ Squared\ Method:$

$$\frac{\partial E}{\partial a} = 0,\ \frac{\partial E}{\partial b} = 0,\ \frac{\partial E}{\partial c} = 0$$

$$
\begin{aligned}
\frac{\partial E}{\partial a} &= \frac{\partial}{\partial a}\left(\textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)^2\right) &&= 0\\
&= \textstyle\sum_{i=1}^{n}\frac{\partial}{\partial a}\left(ax_i^2+bx_i+c-y_i\right)^2 &&= 0\\
&= \textstyle\sum_{i=1}^{n}2\left(ax_i^2+bx_i+c-y_i\right)\cdot x_i^2 &&= 0\\
&= \textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)\cdot x_i^2 &&= 0\\
&= a\textstyle\sum_{i=1}^{n}x_i^4+b\sum_{i=1}^{n}x_i^3+c\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}x_i^2y_i &&= 0\\
(1) \quad &= a\textstyle\sum_{i=1}^{n}x_i^4+b\sum_{i=1}^{n}x_i^3+c\sum_{i=1}^{n}x_i^2 &&= \textstyle\sum_{i=1}^{n}x_i^2y_i
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E}{\partial b} &= \frac{\partial}{\partial b}\left(\textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)^2\right) &&= 0\\
&= \textstyle\sum_{i=1}^{n}\frac{\partial}{\partial b}\left(ax_i^2+bx_i+c-y_i\right)^2 &&= 0\\
&= \textstyle\sum_{i=1}^{n}2\left(ax_i^2+bx_i+c-y_i\right)\cdot x_i &&= 0\\
&= \textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)\cdot x_i &&= 0\\
&= a\textstyle\sum_{i=1}^{n}x_i^3+b\sum_{i=1}^{n}x_i^2+c\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}x_iy_i &&= 0\\
(2) \quad &= a\textstyle\sum_{i=1}^{n}x_i^3+b\sum_{i=1}^{n}x_i^2+c\sum_{i=1}^{n}x_i &&= \textstyle\sum_{i=1}^{n}x_iy_i
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E}{\partial c} &= \frac{\partial E}{\partial c}\left(\textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)^2\right) &&= 0\\
&= \textstyle\sum_{i=1}^{n}\frac{\partial E}{\partial c}\left(ax_i^2+bx_i+c-y_i\right)^2 &&= 0\\
&= \textstyle\sum_{i=1}^{n}2\left(ax_i^2+bx_i+c-y_i\right)\cdot 1 &&= 0\\
&= \textstyle\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right) &&= 0\\
&= a\textstyle\sum_{i=1}^{n}x_i^2+b\sum_{i=1}^{n}x_i+c\cdot n-\sum_{i=1}^{n}x_iy_i &&= 0\\
(3) \quad &= a\textstyle\sum_{i=1}^{n}x_i^2+b\sum_{i=1}^{n}x_i+c\cdot n &&= \textstyle\sum_{i=1}^{n}y_i
\end{aligned}
$$

$$
Let\ Sx=\sum_{i=1}^{n}x_i\ ,\ Sx^2=\sum_{i=1}^{n}x_i^2\ ,\ Sx^3\ =\sum_{i=1}^{n}x_i^3,\ Sx^4=\sum_{i=1}^{n}x_i^4,\ Sy=\sum_{i=1}^{n}y_i\ ,\ Sxy=\sum_{i=1}^{n}x_iy_i\ ,\ \ Sx^2y=\sum_{i=1}^{n}x_i^2y_i.\ \ Then\ \begin{cases} a\cdot Sx^4+b\cdot Sx^3+c\cdot Sx^2 &= \quad Sx^2y \quad (1)\\ a\cdot Sx^3+b\cdot Sx^2+c\cdot Sx &= \quad Sxy \quad (2)\\ a\cdot Sx^2+b\cdot Sx+c\cdot n &= \quad Sy \quad (3)\end{cases}
$$

$$
\begin{array}{c}
\begin{array}{ccc} a & b & c \end{array}\\
\left[\begin{array}{ccc|c} Sx^4 & Sx^3 & Sx^2 & Sx^2y\\ Sx^3 & Sx^2 & Sx & Sxy\\ Sx^2 & Sx & n & Sy \end{array}\right]
\end{array}
=
\left[\begin{array}{ccc|c}
Sx^4\cdot Sx^3\cdot Sx^2 & Sx^3\cdot Sx^3\cdot Sx^2 & Sx^2\cdot Sx^3\cdot Sx^2 & Sx^2y\cdot Sx^3\cdot Sx^2\\
Sx^3\cdot Sx^4\cdot Sx^2 & Sx^2\cdot Sx^4\cdot Sx^2 & Sx\cdot Sx^4\cdot Sx^2 & Sxy\cdot Sx^4\cdot Sx^2\\
Sx^2\cdot Sx^3\cdot Sx^4 & Sx\cdot Sx^3\cdot Sx^4 & n\cdot Sx^3\cdot Sx^4 & Sy\cdot Sx^3\cdot Sx^4
\end{array}\right]
$$

$$
=\left[\begin{array}{ccc|c}
Sx^4\cdot Sx^3\cdot Sx^2 & Sx^3\cdot Sx^3\cdot Sx^2 & Sx^2\cdot Sx^3\cdot Sx^2 & Sx^2y\cdot Sx^3\cdot Sx^2\\
0 & Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2 & Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2 & Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2\\
0 & Sx\cdot Sx^3\cdot Sx^4-Sx^3\cdot Sx^3\cdot Sx^2 & n\cdot Sx^3\cdot Sx^4-Sx^2\cdot Sx^3\cdot Sx^2 & Sy\cdot Sx^3\cdot Sx^4-Sx^2y\cdot Sx^3\cdot Sx^2
\end{array}\right]
$$

$$
=\left[\begin{array}{ccc|c}
Sx^4\cdot Sx^3\cdot Sx^2 & Sx^3\cdot Sx^3\cdot Sx^2 & Sx^2\cdot Sx^3\cdot Sx^2 & Sx^2y\cdot Sx^3\cdot Sx^2\\
0 & Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4 & Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4 & Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\\
0 & Sx\cdot Sx^3\cdot Sx^4-Sx^3\cdot Sx^3\cdot Sx^2+Sx^2\cdot Sx^4\cdot Sx^2 & n\cdot Sx^3\cdot Sx^4-Sx^2\cdot Sx^3\cdot Sx^2+Sx^2\cdot Sx^4\cdot Sx^2 & Sy\cdot Sx^3\cdot Sx^4-Sx^2y\cdot Sx^3\cdot Sx^2+Sx^2\cdot Sx^4\cdot Sx^2
\end{array}\right]
$$

$$
=\left[\begin{array}{ccc|c}
Sx^4\cdot Sx^3\cdot Sx^2 & Sx^3\cdot Sx^3\cdot Sx^2 & Sx^2\cdot Sx^3\cdot Sx^2 & Sx^2y\cdot Sx^3\cdot Sx^2\\
0 & Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4 & Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4 & Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\\
0 & 0 & n\cdot Sx^3\cdot Sx^4+Sx^2\cdot Sx^4\cdot Sx^2-\left(Sx\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right) & Sy\cdot Sx^3\cdot Sx^4+Sx^2\cdot Sx^4\cdot Sx^2-\left(Sxy\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)
\end{array}\right]
\begin{array}{c} \\ (4)\\ (5)\end{array}
$$

by (5)

$$
c=\frac{Sy\cdot Sx^3\cdot Sx^4+Sx^2\cdot Sx^4\cdot Sx^2-\left(Sxy\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)}{n\cdot Sx^3\cdot Sx^4+Sx^2\cdot Sx^4\cdot Sx^2-\left(Sx\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)}
$$

$$
=\frac{Sy\cdot Sx^3\cdot Sx^4-\left(Sxy\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)}{n\cdot Sx^3\cdot Sx^4-\left(Sx\cdot Sx^4\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)}
$$

$$
=\frac{Sy\cdot Sx^3\cdot Sx^4-Sxy\cdot Sx^4\cdot Sx^2-Sx\cdot Sx^3\cdot Sx^4}{n\cdot Sx^3\cdot Sx^4-Sx\cdot Sx^4\cdot Sx^2-Sx\cdot Sx^3\cdot Sx^4}
$$

$$
=\frac{Sy\cdot Sx^3\cdot Sx^4-Sxy\cdot Sx^4\cdot Sx^2}{n\cdot Sx^3\cdot Sx^4-Sx\cdot Sx^4\cdot Sx^2}
$$

by (4)

$$
b\left(Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)+c\left(Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)=Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4
$$

$$
b\left(Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)=Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4-c\left(Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)
$$

$$
b=\frac{Sxy\cdot Sx^4\cdot Sx^2-Sx^2y\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4-c\left(Sx\cdot Sx^4\cdot Sx^2-Sx^2\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4\right)}{Sx^2\cdot Sx^4\cdot Sx^2-Sx^3\cdot Sx^3\cdot Sx^2+Sx\cdot Sx^3\cdot Sx^4}
$$

by (3)

$$
a\cdot Sx^2+b\cdot Sx+c\cdot n\ =Sy
$$

$$
a\cdot Sx^2=Sy-c\cdot n-b\cdot Sx
$$

$$
a=\frac{Sy-c\cdot n-b\cdot Sx}{Sx^2}
$$