Methods of Least Squares

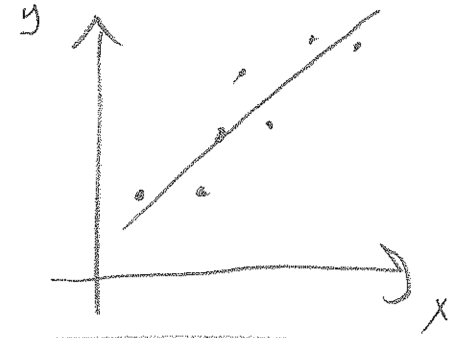
steps

① Data set

Sample:

x_i	y_i
x_1	y_1
x_2	y_2
⋮	⋮
x_n	y_n

② Plot



---

③

a = slope

b = y-intercept

n = count

x_i = Paricular sample value

y_i = Particular sample value

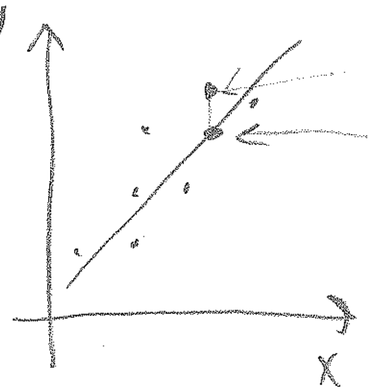
Equation = $y = ax + b$

---

④ predicted value = $y_i = a x_i + b$

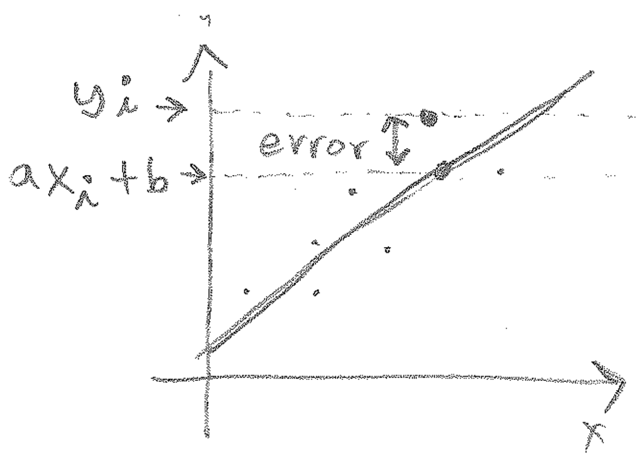
---

⑤



Particular sample point = (x_2, y_2) | (x_i, y_i)

Predicted value = $y_2 = a x_2 + b$ | $y_i = a x_i + b$



Error = $\underbrace{y_i}_{\text{particlar sample point}} - \underbrace{ax_i - b}_{\text{Predicted y-value}}$

(6) Square Error $(y_i - ax_i - b)^2$

---

(7) Mean Square Error $= E$

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - ax_i - b)^2$$

---

(8)

- we cannot calculate E because we don't know values of a & b (slope & intercept).
- we want to choose a & b in such a way to minimize E.
- Necessary conditions for this are when the 2 paritial derivaties:

$$\frac{dE}{da} = 0 \quad \text{and} \quad \frac{dE}{db} = 0$$

---

(9)

$$\frac{dE}{da} = \frac{d}{da}\frac{1}{n}\sum(y_i - ax_i - b)^2$$

$$= 2\frac{1}{n}\sum(y_i - ax_i - b)(-x_i)$$

Power Rule & chain Rule

(10) $$\frac{dE}{db} = \frac{d}{db}\frac{1}{n}\sum (y_i - ax_i - b)^2$$

(p.3)

$$= 2\frac{1}{n}\sum (y_i - ax_i - b)(-1)$$

Power Rule & Chain Rule

$$= -2\frac{1}{n}\sum (y_i - ax_i - b)$$

---

(11) set 2 equations equal to zero (slope for line) & solve for a & b

* min point is point where slope of line equals zero

[1] $$0 = 2\frac{1}{n}\sum (y_i - ax_i - b)(-x_i)$$

[2] $$0 = -2\frac{1}{n}\sum (y_i - ax_i - b)$$

---

(12) change this one first: [1]

Divide both sides by $2\frac{1}{n}$

$$0 = 2\frac{1}{n}\sum (y_i - ax_i - b)(-x_i)$$

$$0 = \sum (y_i - ax_i - b)(-x_i)$$

carry $-x_i$ in

carry $\sum$ in

$$0 = \sum (-x_i y_i + ax_i^2 + bx_i)$$

$$0 = -\sum x_i y_i + \sum ax_i^2 + \sum bx_i$$

Add term $\sum x_i y_i$ to each side

Bring constant in front of $\sum$

$$\sum x_i y_i = \sum ax_i^2 + \sum bx_i$$

$$\sum x_i y_i = a\sum x_i^2 + b\sum x_i$$

(13) Δ change formula [2]

Divide both sides by $-2\frac{1}{n}$

$$0 = -2\frac{1}{n}\sum(y_i - ax_i - b)$$

$$0 = \sum(y_i - ax_i - b)$$

Bring $\sum$ in

$$0 = \sum y_i - \sum ax_i - \sum b$$

Bring constant in front of $\sum$

$$0 = \sum y_i - a\sum x_i - \sum b$$

$\sum$ of b from 1 to n $= nb$

Add $a\sum x_i$ to both sides

$$0 = \sum y_i - a\sum x_i - nb$$

$$\sum y_i - a\sum x_i = nb$$

Divde both sides by n

$$b = \frac{\sum y_i - a\sum x_i}{n}$$

(14) substitute b into equation [1]

[1] $$\sum x_i y_i = a\sum x_i^2 + b\sum x_i$$

$$\sum x_i y_i = a\sum x_i^2 + \left[\frac{\sum y_i - a\sum x_i}{n}\right]\sum x_i$$

multiply all by n

$$\sum x_i y_i n = a\sum x_i^2 n + \left[\sum y_i - a\sum x_i\right]\sum x_i$$

carry $\sum x_i$ in

$$\sum x_i y_i n = a\sum x_i^2 n + \sum x_i \sum y_i - a\left(\sum x_i\right)^2$$

P.5

$$\sum x_i y_i n - \sum x_i \sum y_i = a\sum x_i^2 n - a\left(\sum x_i\right)^2$$

Subtract $\sum y_i \sum x_i$ both sides

$$\sum x_i y_i n - \sum x_i \sum y_i = a\left(\sum x_i^2 n - \left(\sum x_i\right)^2\right)$$

Factor out a

Divide both sides by term

$$\frac{\sum x_i y_i n - \sum x_i \sum y_i}{\sum x_i^2 n - \left(\sum x_i\right)^2} = a$$

mutiply Left term by $\dfrac{\frac{1}{n}}{\frac{1}{n}}$

$$\frac{\sum x_i y_i - \dfrac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \dfrac{\left(\sum x_i\right)^2}{n}} = a$$

---

(15) Transform numerator & denominator separately

$$a = \frac{\sum x_i y_i - \dfrac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \dfrac{\left(\sum x_i\right)^2}{n}} \qquad \begin{matrix}[3] \\ \\ [4]\end{matrix}$$

(16) Goal is to end up with Formula like this:

$$q = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \quad [5]}{\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad [6]}$$

---

(17) Need:

$$[3] = [5]$$

$$\sum_{i=1}^{n} x_i y_i - \frac{\sum x_i \sum y_i}{n} = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

AND

$$[4] = [6]$$

$$\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum X_i\right)^2}{n} = \sum_{i=1}^{n}(X_i - \bar{X})^2$$

(17.1)

$$\sum (X_i - \bar{X})^2 \qquad [6]$$

$$\sum \left(X_i - \frac{\sum X_i}{n}\right)^2$$

$$\sum \left(X_i^2 - 2X_i \frac{\sum X_i}{n} + \left(\frac{\sum X_i}{n}\right)^2\right)$$

$$\sum X_i^2 - 2\sum X_i \frac{\sum X_i}{n} + \sum \left(\frac{\sum X_i}{n}\right)^2$$

$$\sum X_i^2 - \frac{2}{n}\left(\sum X_i\right)^2 + \sum \left(\frac{\sum X_i}{n}\right)^2$$

$$\sum X_i^2 - \frac{2}{n}\left(\sum X_i\right)^2 + \sum_{i=1}^{n} \frac{\left(\sum X_i\right)^2}{n^2}$$

just number * n

$$\sum X_i^2 - \frac{2}{n}\left(\sum X_i\right)^2 + \frac{n\left(\sum X_i\right)^2}{n^2}$$

$$\sum X_i^2 - \frac{2}{n}\left(\sum X_i\right)^2 + \frac{\left(\sum X_i\right)^2}{n}$$

$$\sum X_i^2 - 2\frac{\left(\sum X_i\right)^2}{n} + \frac{\left(\sum X_i\right)^2}{n}$$

$$\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n} \qquad [4]$$

Done! $[6] = [4]$

17.2

P.7

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) \quad [5]$$

$$\sum \left(X_i - \frac{\sum X_i}{n}\right)\left(Y_i - \frac{\sum Y_i}{n}\right)$$

$$\sum \left(X_i Y_i - X_i \frac{\sum Y_i}{n} - Y_i \frac{\sum X_i}{n} + \frac{\sum X_i}{n} \frac{\sum Y_i}{n}\right)$$

$$\sum X_i Y_i - \sum X_i \frac{\sum Y_i}{n} - \sum Y_i \frac{\sum X_i}{n} + \sum_{i=1}^{n} \frac{\sum X_i}{n} \frac{\sum Y_i}{n}$$

Just number times n

$$\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} - \frac{\sum Y_i \sum X_i}{n} + n \frac{\sum X_i}{n} \frac{\sum Y_i}{n}$$

(1 n cancel)

$$\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \underbrace{- \frac{\sum Y_i \sum X_i}{n} + \frac{\sum X_i \sum Y_i}{n}}_{0}$$

$$\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \quad \boxed{3}$$

Done $[5] = [3]$

18 Thus:

$$a = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Formula for slope

PROVED!!

19 For b, use: $\boxed{b = \bar{Y} - a\bar{X}}$