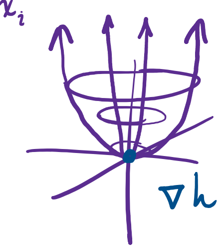
A1 Written

Tuesday, January 30, 2024 12:34 PM

$$\sum_{i=1}^{N}\left(y_i - (\theta_0 + \theta_1 x_i)\right)^2$$

$a = \theta_0 + \theta_1 x_i$



$\nabla h(\theta_0, \theta_1) = \vec{0}$

$$\sum (y_i - a)^2$$

$$\sum y_i^2 - 2y_i a + a^2$$

$$\sum y_i^2 - 2\sum y_i a + \sum a^2$$

$$N\overline{y^2} - 2\left(\sum y_i(\theta_0 + \theta_1 x_i)\right) + \sum (\theta_0 + \theta_1 x_i)^2$$

$$N y^2 - 2\left(\sum \theta_0 y_i + \theta_1 x_i y_i\right) + \sum \theta_0^2 + 2\theta_0\theta_1 x_i + \theta_1^2 x_i^2$$

$$N\left(\overline{y^2} - 2\left(\theta_0\bar{y} + \theta_1\overline{xy}\right) + \theta_0^2 + 2\theta_0\theta_1\bar{x} + \theta_1^2\overline{x^2}\right) = h(\theta_0, \theta_1)$$

$$N\left[\overline{y^2} - 2\theta_0\bar{y} - 2\theta_1\overline{xy} + \theta_0^2 + 2\theta_0\theta_1\bar{x} + \theta_1^2\overline{x^2}\right] =$$

$$\frac{\partial h}{\partial \theta_0} = -2N\bar{y} + 2N\theta_0 + 2N\theta_1\bar{x}$$

$$\frac{\partial h}{\partial \theta_1} = -2N\overline{xy} + 2N\theta_0\bar{x} + 2N\overline{x^2}\theta_1$$

Critical Points

$0 =$

$0 = -\bar{y} + \theta_0 + \theta_1\bar{x}$

$\bar{y} = \theta_0 + \theta_1\bar{x}$

$\theta_0 = \bar{y} - \theta_1\bar{x}$

$\theta_1 = \dfrac{\overline{xy} - \theta_0\bar{x}}{\overline{x^2}}$

sub

$0 =$

$0 = -\overline{xy} + \theta_0\bar{x} + \overline{x^2}\theta_1$

$\dfrac{\overline{xy} - \theta_0\bar{x}}{\overline{x^2}} = \theta_1$

$$\theta_1 = \frac{\overline{xy} - \theta_0 \overline{x}}{\overline{x^2}}$$

sub

$$\theta_1 = \frac{\overline{xy} - \overline{x}\left[\overline{y} - \theta_1 \overline{x}\right]}{\overline{x^2}}$$

$$\theta_1 = \frac{\overline{xy} - \overline{x}\,\overline{y} + \theta_1 \overline{x}^2}{\overline{x^2}}$$

$$\left(\overline{x^2} - \overline{x}^2\right)\theta_1 = \overline{xy} - \overline{x}\,\overline{y}$$

$$\boxed{\theta_1 = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2}}$$

sub

$$\theta_0 = \overline{y} - \theta_1 \overline{x}$$

$$\theta_0 = \overline{y} - \left[\frac{\overline{x}\left(\overline{xy} - \overline{x}\,\overline{y}\right)}{\overline{x^2} - \overline{x}^2}\right]$$

$$= \frac{\overline{y}\left(\overline{x^2} - \overline{x}^2\right) - \left(\overline{x}\,\overline{xy} - \overline{x}^2\overline{y}\right)}{\overline{x^2} - \overline{x}^2}$$

$$= \frac{\overline{x^2}\,\overline{y} - \cancel{\overline{x}^2\overline{y}} - \overline{x}\,\overline{xy} + \cancel{\overline{x}^2\overline{y}}}{\overline{x^2} - \overline{x}^2}$$

$$\boxed{\theta_0 = \frac{\overline{x^2}\,\overline{y} - \overline{x}\,\overline{xy}}{\overline{x^2} - \overline{x}^2}}$$

B)

Critical Point at (θ_0, θ_1) when

$$\theta_0 = \frac{\overline{x^2}\,\overline{y} - \overline{x}\,\overline{xy}}{\overline{x^2} - \overline{x}^2}, \quad \theta_1 = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2}$$

This means the point exists when $\overline{x^2} \neq \overline{x}^2$ (cannot divide by 0.

c) Prove the critical point is a minimum

Hessian Conditions for minimum:

$$\frac{\partial^2 h}{\partial \theta_0^2} > 0 \quad \text{and} \quad \frac{\partial^2 h}{\partial \theta_0^2} \cdot \frac{\partial^2 h}{\partial \theta_1^2} - \left(\frac{\partial^2 h}{\partial \theta_0 \partial \theta_1}\right)^2 > 0$$

at $$(\theta_0, \theta_1) = \left(\frac{\overline{x^2}\,\overline{y} - \overline{x}\,\overline{xy}}{\overline{x^2} - \overline{x}^2}, \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} \right)$$

$$\frac{\partial h}{\partial \theta_0} = -\overline{y} + \theta_0 + \theta_1 \overline{x}$$

$$\frac{\partial h}{\partial \theta_1} = -\overline{xy} + \theta_0 \overline{x} + \overline{x^2}\,\theta_1$$

$$\frac{\partial^2 h}{\partial \theta_0^2} = 1$$

$$\frac{\partial^2 h}{\partial \theta_1^2} = \overline{x^2}$$

$$\frac{\partial^2 h}{\partial\theta_0^2} = 1 \qquad \frac{\partial^2 h}{\partial\theta_1^2} = \overline{x^2}$$

$$\frac{\partial^2 h}{\partial\theta_0 \partial\theta_1} = \overline{x}$$

$$\frac{\partial^2 h}{\partial\theta_0^2} > 0$$

$$1 > 0$$

$$\frac{\partial^2 h}{\partial\theta_0^2} \cdot \frac{\partial^2 h}{\partial\theta_1^2} - \left(\frac{\partial^2 h}{\partial\theta_0 \partial\theta_1}\right)^2 > 0$$

$$1 \cdot \overline{x^2} - \overline{x}^2 > 0$$

$$\overline{x^2} - \overline{x}^2 > 0$$

Critical points exist when $\overline{x^2} \neq \overline{x}^2$.

When it exists, it has to be a minimum because $\frac{\partial^2 h}{\partial\theta_0^2} = 1$, and $1 > 0$. $\overline{x^2}$ must be greater than $\overline{x}^2$

Look up Cauchy-Schwartz

to prove $\overline{x^2} > \overline{x}^2$.

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{N} x_i^2$$

$$\overline{x}^2 = \left(\frac{1}{n}\sum_{i=1}^{N} x_i\right)^2$$

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2 > \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)^2$$

$$\frac{1}{N}\sum_{i=1}^{N} x_i^2 > \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)^2$$

$$\sum_{i=1}^{N} x_i^2 > \frac{1}{N^2}\left(\sum_{i=1}^{N} x_i\right)^2$$

$$N\left(1^2 + 2^2 \ldots + x_N^2\right) > \left(1 + 2 + \ldots x_N\right)^2$$

$$N\left(1^2+2^2+\dots+x_N^2\right) > \left(1+2+\dots x_N\right)^2$$

Cauchy Schwartz Inequality:

Let

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\left(|\vec{a}| \cdot |\vec{b}|\right)^2 \geq |\vec{a} \cdot \vec{b}|^2$$

$$(a_1^2 + a_2^2 + \dots + a_n^2)(b_1^2 + b_2^2 + \dots + b_n^2) \geq (a_1b_1 + a_2b_2 + \dots + a_nb_n)^2$$

Applying this, we can set $\vec{a}$ to the set of all data points, and $\vec{b}$ to a vector of 1's.

$$\vec{a} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ set of data}$$

$$\vec{b} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

This gives us

$$(x_1^2 + x_2^2 + \dots + x_n^2)(1+1+\dots 1) \geq (x_1 \cdot 1 + x_2 \cdot 1 + \dots + x_n \cdot 1)^2$$

$$N(x_1^2 + x_2^2 + \dots + x_n^2) \geq (x_1 + x_2 + \cdots x_n)^2$$

Because this is an extension of Cauchy-Schwartz, we know that $\overline{x^2} \geq \overline{x}^2$.

Because $\overline{x^2} \geq \overline{x}^2$ will always be true, and $\overline{x^2} \neq \overline{x}^2$ must be true if a critical point exists, this means that when a critical point exists, it must be a minimum.

d)

Critical point does not exist when $\overline{x^2} = \overline{x}^2$, so solve

$$\sum_{i=1}^{N} x_i^2 = \left(\sum_{i=1}^{N} x_i\right)^2$$

$$N(x_1^2 + x_2^2 + \cdots x_N^2) = N^2(x_1 + x_2 + \cdots + x_N)^2$$

$$(x_1^2 + x_2^2 + \cdots + x_N^2) = N(x_1 + x_2 + \cdots + x_N)^2$$

$$x_1^2 + x_2^2 + \cdots + x_N^2 = N\overline{x}^2$$

This can only be true if

$$x_1 = x_2 = \cdots = x_N = \overline{x}.$$

All data points must be identical, which is highly unlikely.