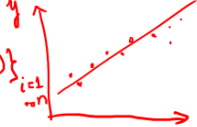


HUỖNH THIÊN TÙNG (19522492)

Câu 1: Tại sao ko dùng 2 hàm loss như trong tình huống a) và b) mà lại dùng công thức c) ?

Linear Regression

Tập dữ liệu $\{(x_i, y_i)\}_{i=1}^n$



1. Thiết kế $\hat{y} = f_{\theta}(x) = \theta_1 x + \theta_0$ (x : scalar)
đơn

~~$f_{\theta}(x) = \theta^T \begin{bmatrix} 1 \\ x \end{bmatrix}$ (x : vector)~~

2. Thiết hàm loss :

$\mathcal{L}(\theta) = (\theta_1 x + \theta_0 - y)^2$ (1 mẫu)
đơn

$\mathcal{L}(\theta) = \frac{1}{2n} \sum_{i=1}^n (\theta_1 x_i + \theta_0 - y_i)^2$

3. Giải $\mathcal{L}(\theta)$ nhỏ nhất,

$\frac{\partial \mathcal{L}}{\partial \theta_0}, \frac{\partial \mathcal{L}}{\partial \theta_1} \Rightarrow \text{gradient descent.}$

$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}$

$\tilde{y} \approx y$

a) $\mathcal{L}(\theta) = \tilde{y} - y = \theta_1 x + \theta_0 - y$

b) $\mathcal{L}(\theta) = |\tilde{y} - y| = |\theta_1 x + \theta_0 - y|$

c) $\mathcal{L}(\theta) = (\tilde{y} - y)^2 = (\theta_1 x + \theta_0 - y)^2$

Trả lời:

- Với công thức a), chúng ta không đánh giá được loss là giá trị âm. Bản thân hàm mất mát loss hoặc còn gọi là cost function thì giá trị phải là dương thì mới đánh giá đúng và hợp lý được.
 - Với công thức b), đã khắc phục được điểm yếu của công thức a) nhưng khi triển khai tính đạo hàm rất là phức tạp. Đạo hàm của một biểu thức trong dấu giá trị tuyệt đối thì rất phức tạp vì phải xét trường hợp biểu thức bên trong mang giá trị âm hay dương. Điều này gây khó khăn trong quá trình tính toán.
- ⇒ Vì thế, công thức c) đã giải quyết được các vấn đề trên. Tính đạo hàm dễ và thể hiện rõ được vai trò hàm loss