

Câu 1: Linear Regression

- Tại sao không sử dụng hàm hiệu?
 - Vì hàm hiệu không có cực trị (đạo hàm của hàm hiệu là một hằng số) nên không thể tìm giá trị nhỏ nhất của hàm số.
 - Ví dụ: $L = t - t_h \Rightarrow L'(t) = 1$
- Tại sao không sử dụng hàm trị tuyệt đối?
 - Vì hàm trị tuyệt đối không có đạo hàm tại điểm mà tử số của đạo hàm bằng 0 nên khi sử dụng thuật toán Gradient Descent có thể xảy ra lỗi
 - Ví dụ: $y = |x| \Rightarrow y'(x) = \frac{x}{|x|}$, ta thấy tại điểm $x=0$, $y'(x)$ không xác định

Câu 2: Tại sao không sử dụng hàm bình phương mà sử dụng Cross Entropy trong Logistic Regression?

Nếu sử dụng hàm bình phương $L(\theta) = (y - \tilde{y})^2$, lúc này $y \in \{0, 1\}$ và $\tilde{y} \in (-\infty, +\infty)$ thì $L(\theta)$ sẽ có giá trị lớn. Do đó hàm $L(\theta)$ sẽ phụ thuộc vào \tilde{y} hơn là phụ thuộc vào y tức là khi sử dụng Gradient Descent, thuật toán sẽ tập trung làm sao cho \tilde{y} nhỏ nhất hơn là làm cho \tilde{y} xấp xỉ bằng y .

Hàm Cross Entropy có thể giải quyết được vấn đề trên do đó người ta sử dụng Cross Entropy thay vì sử dụng hàm bình phương cho bài toán Logistic Regression.