

Linear Regression - Exercise 2

1) Đọc hiểu code (file D5, D6, D7, D8) về cách train bài toán linear regression theo cách thông thường và vectorization cho m and N sample.

2) So sánh L1 loss (absolute error) và L2 loss (squared error). Nếu ưu nhược điểm của chúng.

3) Tìm hiểu về Huber loss và cách Huber loss khắc phục nhược điểm của 2 hàm loss trên. Cài đặt Huber loss cho các chương trình ở file D5, D6, D7, D8.

4) Xây dựng công thức linear regression cho N sample khi vector \mathbf{x} có dạng sau:

$$\mathbf{x} = \begin{bmatrix} sample1 \\ \dots \\ sampleN \end{bmatrix} \text{ (sắp xếp theo chiều ngang, khác với chiều dọc trong bài học).}$$

5) Cài đặt linear regression cho bài toán advertising theo 2 cách (cách thông thường và vectorization) dùng m sample (mini-batch gradient descent).

Data advertising:

<https://www.dropbox.com/s/ve57nlqwikgjzrh/advertising.csv?dl=0>