

Multiple Linear Regression

1. Multiple features

Trong thực tế, bài toán đòi hỏi mô hình học và đưa ra kết quả dựa trên nhiều biến (features) hơn.

- Trước đây: mô hình đơn biến

$$\hat{y} = wx + b$$

- Giờ đây: nhiều đặc trưng

$$\hat{y} = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b$$

Trong đó:

- $x(i)$: vector đặc trưng của mẫu huấn luyện thứ i
- x_j : đặc trưng thứ j
- $x_j^{(i)}$: đặc trưng thứ j của mẫu huấn luyện thứ i
- w : vector trọng số (weights)
- b : bias (hệ số chệch)

2. Vectorization

Vectorization (vector hóa) là kỹ thuật viết code sử dụng các **phép toán vector/matrix** thay vì các vòng lặp thủ công. Giúp:

- Code **ngắn gọn hơn** (ít dòng hơn)
- Chạy **nhANH hơn** (tận dụng tối đa phần cứng hiện đại như CPU/GPU)

Ví dụ

- Giả sử:
 - Vector trọng số: $w = [w_1, w_2, w_3]$
 - Vector đặc trưng: $x = [x_1, x_2, x_3]$
 - Bias: b

- **Không dùng vector hóa** (code thủ công):

```
f = w[0]*x[0] + w[1]*x[1] + w[2]*x[2] + b
```

- **Dùng vòng lặp:**

```
f = 0
for j in range(n):
    f += w[j] * x[j]
f += b
```

- **Dùng vector hóa với NumPy** (rút gọn 1 dòng):

```
f = np.dot(w, x) + b
```

Trong Numpy, **Song song hóa** là chìa khóa tốc độ — thay vì nối tiếp từng phép toán, máy tính tính hàng loạt trong cùng 1 thời điểm.

- Với **số lượng nhỏ (16 features)** → Tốc độ cải thiện chưa rõ.
- Với **hàng nghìn features và dữ liệu lớn**:
 - Không vector hóa có thể **mất hàng giờ**
 - Vector hóa có thể **chạy trong vài phút**

Đây là lý do **vector hóa là điều kiện bắt buộc** để các thuật toán học máy hiện đại có thể mở rộng ra **tập dữ liệu lớn**.

3. Normal Equation

- **Normal Equation** là một cách **giải hệ phương trình** tuyến tính để tìm **w** và **b** **mà không cần lặp (no iterations)**.
- Dùng công thức:

$$w = (X^T X)^{-1} X^T y$$

(với **x** là ma trận đặc trưng đã thêm cột 1 cho **b** nếu cần)

- **Ưu điểm:** Không cần lặp, có lời giải đóng.
- **Nhược điểm:**

- Không áp dụng được cho mô hình phức tạp như **logistic regression**, **neural networks**, ...
- **Chậm và tốn bộ nhớ** với số lượng đặc trưng lớn (n lớn)
- Không linh hoạt như gradient descent