1. Hypothesis functions là gì?

Hypothesis functions là funtion mô tả tốt nhất mục tiêu của học máy có giám sát. The hypothesis là một thuật toán sẽ phụ thuộc vào dữ liệu và cũng phụ thuộc vào các hạn chế và thiên vị mà ta đã áp dụng đối với dữ liệu.

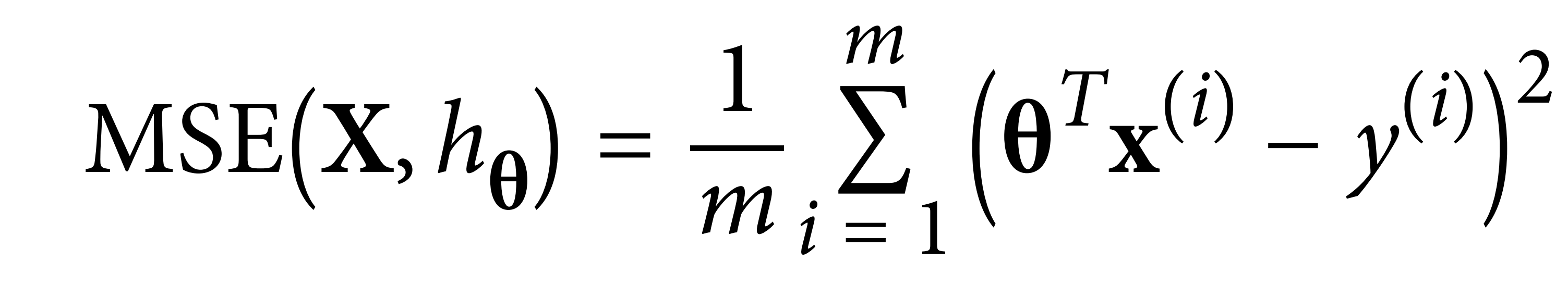
Công thức của Hypothesis Function là:

hθ(x)=θ1 x+θ0

Ký hiệu θ ở đây ta đọc là "theta". Vây θ1 đọc là "theta một" và θ0 đọc là "theta không", các phần tử này tượng trưng cho các trọng số của chúng ta

1. Cost functions là gì? Mô tả MSE cost functions.

Cost function là một cơ chế được sử dụng trong học máy có giám sát, nó trả về sai số giữa kết quả dự đoán so với kết quả thực tế.



*) -*

Trong đó:

* m : số lượng mẫu dữ liệu thu thập được
* x : giá trị đầu vào (các dữ kiện mà ta dùng để đưa dự đoán)
* y : giá trị thực thu được trong dữ liệu (vd: giá thật được bán của miếng đất)
* x(i) : giá trị đầu vào thứ i trong các mẫu dữ liệu
* y(i) : giá trị thực thứ i trong các mẫu dữ liệu
* (x,y) : một bản ghi đầy đủ của dữ liệu mà ta thu thập được

*) là giá trị dự đoán thứ i và là giá trị thực tế thứ i*

Với mỗi cặp dữ liệu x, y ta sử dụng x để tạo ra dự đoán và tìm chênh lệch với y. Sau đó bình phương lên để mất dấu rồi cộng tất cả các sai số của từng dữ liệu và lấy trung bình cộng.

1. So sánh normal equation và gradient decent (điểm giống nhau, khác nhau, ưu điểm, khuyết điểm).

**Giống nhau**: normal equation và gradient decent đều tìm tham số 𝛉 để cho cost function có giá trị nhỏ nhất.

**Khác nhau:**

|  |  |
| --- | --- |
| **Gradient Descent** | **Normal Equation** |
| Cần phải chọn alpha. | Không cần chọn alpha. |
| Cần lặp nhiều bước tối ưu. | Chỉ thực hiện 1 bước. |
| Độ phức tạp O(kn2). | Độ phức tạp O(n3), cần tìm ma trận khả nghịch. |
| Hoạt động tốt với training set lớn. | Xử lí chậm với training set lớn (tìm ma trận khả nghịch làm chậm tốc độ khá nhiều). |

# **TÀI LIỆU THAM KHẢO**

<https://www.geeksforgeeks.org/ml-understanding-hypothesis/>

<https://rootonchair.blogspot.com/2019/02/nhap-mon-machine-learning-bai-3-linear.html>

<https://www.howkteam.vn/course/machine-learning-co-ban-voi-numpy/normal-equation-cho-linear-regression-3999>