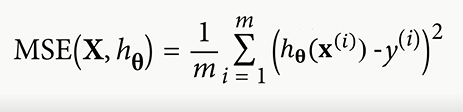
1. Hypothesis functions là gì?

Hypothesis functions là model (đường màu đỏ hoặc xanh) là chính là công cụ để giúp những chương trình Machine Learning dự đoán và tìm các trọng số tối ưu thông qua Cost Function sẽ giúp các dự đoán này chính xác hơn

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

Cost functions là hàm dùng để đo sai số của dự đoán.

Mean Squared Error (MSE)

Công thức tính:

Trong đó: h là hypothesis function  
 y là lable

h(x) là prediction

i là sample

m là tổng số sample

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

Có 2 cách: cả 2 cách này đều dùng phương pháp nào đó để tính theta  
+ Normal Equation: tìm theta bằng cách tính đạo hàm theo **⊖ (theta)**

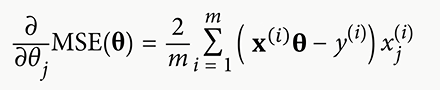


\*Khác nhau: dùng đạo hàm

\*Ưu điểm: dễ làm bằng cách tính đạo hàm, dòng lệnh ít, đẹp

\*Nhược điểm: đôi khỉ nghịch đảo không được, số lượng feature nhiều thì thời gian chạy lâu hoặc rất lâu, không dùng được cho dữ liệu lớn.

+ Gradient descent: tìm theta bằng cách lặp tầng lặp

J’ = 

\*Khác nhau: bằng cách khởi tạo theta rồi dùng tích chất của đạo hàm để chỉ hướng cho theta để tìm theta tốt nhất, dùng gradient để hướng cho parameters’ update.

\*Ưu điểm: chạy nhanh hơn Normal Equation, có cost function thì không cần biết nó ở dạng nào thì vẫn chạy oke, chỉ dùng phép toán cơ bản như cộng, trừ, nhân, chia không lấy nghịch đảo của ma trận nên chạy nhanh.  
\*Nhược điểm: Khi m quá lớn thì Gradient descent không thực hiện được.