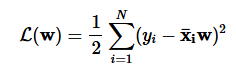
**Bài: Linear Regression**

- một kỹ thuật học có giám sát dùng để dự đoán giá trị thực từ các biến đầu vào (ví dụ: dự đoán giá nhà từ diện tích, số phòng ngủ, khoảng cách đến trung tâm)

- Phương trình tuyến tính cơ bản:

f(x) = w1x1+w2x2+w3x3+w0  
- Sai số trên 1 mẫu: 1/2 (y - xw)^2

- Sai số trên toàn bộ dữ liệu: 

- Hàm số L(w)L(w) được gọi là hàm mất mát (loss function) của bài toán Linear Regression. Chúng ta luôn mong muốn rằng sự mất mát (sai số) là nhỏ nhất, điều đó đồng nghĩa với việc tìm vector hệ số ww sao cho giá trị của hàm mất mát này càng nhỏ càng tốt.  Giá trị của ww làm cho hàm mất mát đạt giá trị nhỏ nhất được gọi là điểm tối ưu (optimal point), ký hiệu: 

- Cách phổ biến nhất để tìm nghiệm cho một bài toán tối ưu là giải phương trình đạo hàm (gradient) bằng 0. Đạo hàm theo ww của hàm mất mát là:  
  
- Với khái niệm giả nghịch đảo, điểm tối ưu của bài toán Linear Regression có dạng: 