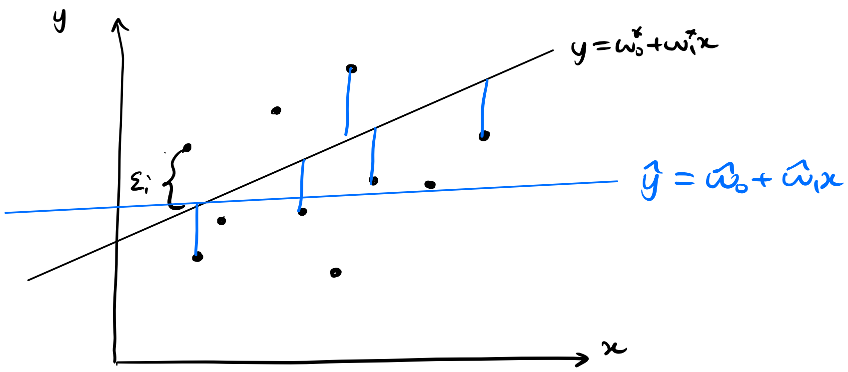
COMP9417 - Week 2 Tutorial notes

Linear Regression

training data: $(x_1, y_1), \ldots, (x_n, y_n)$ $\quad \varepsilon_i \sim (0, \sigma^2)$

$$y_i = w_0^* + w_1^* x_i + \varepsilon_i$$



$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - (\underline{w_0} + \underline{w_1} x_i))^2$$

$$L(w_0, \ldots, w_n) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$= \frac{1}{n}\|y - \hat{y}\|_2^2$$

Euclidean norm

$$a^T a = a_1^2 + \ldots + a_n^2 = \|a_n\|^2$$

$$(AB)^T = B^T A^T$$

$$= \frac{1}{n}\|y - Xw\|_2^2$$

$$= \frac{1}{n}(y - Xw)^T(y - Xw)$$

$$= \frac{1}{n}(y^T - w^T X^T)(y - Xw)$$

$$= \frac{1}{n}(y^T y - w^T X^T y - y^T X w + w^T X^T X w)$$

$$= \frac{1}{n}(y^T y - 2(X^T y)^T w + w^T X^T X w)$$

Hint 1: $\frac{\partial v^T u}{\partial u} = v$

Hint 2: $\frac{\partial u^T A u}{\partial u} = \underbrace{Au + A^T u}$

A is symmetric

$$\frac{\partial L}{\partial w} = \frac{1}{n}(-2X^T y + 2X^T X w)$$

$$\frac{\partial L}{\partial w} = 0$$

$$-2X^T y + 2X^T X w = 0$$

$$\cancel{2}X^T X w = \cancel{2}X^T y \Rightarrow \boxed{\hat{w} = (X^T X)^{-1} X^T y}$$