CSC 865 AI

HW5

KHALID MEHTAB KHAN

SFSU ID: 923673423

1. Regularization

» $f(x) = \sin(\pi x)$ $x \in [-1, 1]$, no noise

a) The sum of squared errors cost function

$$C(w) = \sum_{i=1}^{n} (y_i - (w_0 + w_1 x_i))^2$$

for only 2 points.

$$C = (y_1 - (w_0 + w_1 x_1))^2 + (y_2 - (w_0 + w_1 x_2))^2$$

i) $$\frac{dC(w)}{dw_0} = \frac{d}{dw_0}(y_1 - (w_0 + w_1 x_1))^2 + (y_2 - (w_0 + w_1 x_2))^2$$

$$= 2(y_1 - (w_0 + w_1 x_1))(-1) + 2(y_2 - (w_0 + w_1 x_2))(-1)$$

$$= -2y_1 + 2w_0 + 2w_1 x_1 - 2y_2 + 2w_0 + 2w_1 x_2$$

$$= 4w_0 + 2w_1(x_1 + x_2) - 2(y_1 + y_2)$$

ii) $$\frac{\partial C(w)}{\partial w_1} = \frac{\partial}{\partial w_1}((y_1 - (w_0 + w_1 x_1))^2 + (y_2 - (w_0 + w_1 x_2))^2$$

$$= 2(y_1 - (w_0 + w_1 x_1))(-x_1) + 2(y_2 - (w_0 + w_1 x_2))(-x_2)$$

$$= -2x_1 y_1 + 2w_0 x_1 + 2w_1 x_1^2 - 2x_2 y_2 + 2w_0 x_2 + 2w_1 x_2^2$$

$$= 2w_0(x_1 + x_2) + 2w_1(x_1^2 + x_2^2) - 2(x_1 y_1 + x_2 y_2)$$

b) adding l2 regularization term

$$\tilde{C}(\omega) = \sum_{i=1}^{n} (y_i - (\omega_0 + \omega_1 x_i))^2 + \lambda \sum_{j=0}^{p} \omega_j^2$$

for two points:

$$\tilde{C}(\omega) = (y_1 - (\omega_0 + \omega_1 x_1))^2 + (y_2 - (\omega_0 + \omega_1 x_2))^2 + \lambda(\omega_0^2 + \omega_1^2)$$

i) $$\frac{\partial \tilde{C}(\omega)}{\partial \omega_0} = \frac{\partial (y_1 - (\omega_0 + \omega_1 x_1))^2 + (y_2 - (\omega_0 + \omega_1 x_2))^2 + \lambda(\omega_0^2 + \omega_1^2)}{\partial \omega_0}$$

$$= 2(y_1 - (\omega_0 + \omega_1 x_1))(-1) + 2(y_2 - (\omega_0 + \omega_1 x_2))(-1) + 2\lambda\omega_0 + 0$$

$$= -2y_1 + 2\omega_0 + 2\omega_1 x_1 - 2y_2 + 2\omega_0 + 2\omega_1 x_2 + 2\lambda\omega_0$$

$$2\omega_0(2 + \lambda) + 2\omega_1(x_1 + x_2) - 2(y_1 + y_2)$$

ii) $$\frac{\partial \tilde{C}(\omega)}{\partial \omega_1} = \frac{\partial (y_1 - (\omega_0 + \omega_1 x_1))^2 + (y_2 - (\omega_0 + \omega_1 x_2))^2 + \lambda(\omega_0^2 + \omega_1^2)}{\partial \omega}$$

$$= 2(y_1 - (\omega_0 + \omega_1 x_1))(-x_1) + 2(y_2 - (\omega_0 + \omega_1 x_2))(-x_2) + 2\lambda\omega_1 + 0$$

$$= -2x_1 y_1 + 2x_1\omega_0 + 2\omega_1 x_1^2 - 2x_2 y_2 + 2x_2\omega_0 + 2\omega_1 x_2^2 + 2\lambda\omega_1$$

$$= 2\omega_0(x_1 + x_2) + 2\omega_1(x_1^2 + x_2^2 + \lambda) - 2(x_1 y_1 + x_2 y_2)$$