Optimisation for Linear Regression

The loss $$L = \frac{1}{N}\sum_n (t_n - w_0 - w_1 x_n)^2$$

Steps

1. Multiply out
2. Differentiate
3. Set to zero and solve

1

$$L = \frac{1}{N}\sum_n \left(t_n^2 - 2t_n w_0 - 2t_n w_1 x_n + w_0^2 + 2w_0 w_1 x_n + w_1^2 x_n^2\right)$$

$$= \frac{1}{N}\left(\sum_n t_n^2\right) - \frac{2w_0}{N}\left(\sum_n t_n\right) - \frac{2w_1}{N}\left(\sum_n t_n x_n\right)$$
$$+ w_0^2 + \frac{2w_0 w_1}{N}\left(\sum_n x_n\right) + \frac{w_1^2}{N}\sum_n x_n^2$$

to make things neater define:

$$\overline{f} = \frac{1}{N}\sum_n f_n$$

So $\overline{t^2} = \frac{1}{N}\sum_n t_n^2$ $\quad \overline{x} = \frac{1}{N}\sum_n x_n$

$$\overline{x^2} = \frac{1}{N}\sum_n x_n^2 \quad (\neq \overline{x}^2)$$

$$\overline{xt} = \frac{1}{N}\sum_n x_n t_n$$

So $L = \overline{t^2} - 2w_0\overline{t} - 2w_1\overline{tx} + w_0^2$

$$+ 2w_0 w_1 \overline{x} + w_1^2 \overline{x^2}$$

2 $\dfrac{\partial L}{\partial w_0} = -2\overline{t} + 2w_0 + 2w_1\overline{x}$

$$\frac{\partial L}{\partial w_1} = -2\overline{tx} + 2w_0\overline{x} + 2w_1\overline{x^2}$$

Set to 0 and solve

$$-\bar{t} + w_0 + w_1\bar{x} = 0$$

$$-\overline{tx} + w_0\bar{x} + w_1\overline{x^2} = 0$$

$$\boxed{w_0 = \bar{t} - w_1\bar{x}}$$

$$-\overline{tx} + (\bar{t} - w_1\bar{x})\bar{x} + w_1\overline{x^2} = 0$$

$$-\overline{tx} + \bar{t}\bar{x} - w_1\bar{x}\bar{x} + w_1\overline{x^2} = 0$$

$$\boxed{\frac{\bar{t}\bar{x} - \overline{tx}}{\bar{x}\bar{x} - \overline{x^2}} = w_1}$$

Using the chain rule is <u>much</u> quicker.

$$L = \frac{1}{N} \sum_n \left(t_n - w_0 - w_1 x_n \right)^2$$

$$\partial/\partial w_0 = \frac{-2}{N} \sum_n t_n - w_0 - w_1 x_n$$

$$\partial/\partial w_1 = \frac{-2}{N} \sum_n \left(t_n - w_0 - w_1 x_n \right) x_n$$

<u>etc</u>