AI5000 - Foundation of Machine Learning
Assignment-1 Question-1

Kranthi Kumar P (AI20RESCH14002) and Zeeshan MD (SM20MTECH12011)

October 2020

Question-1[Theory]: Linear Regression

Solution:

This is the case of multivariate regression.
$(y_0, y_1, \ldots, y_n)$ are dependent variables and $(x_0, x_1, \ldots, x_n)$ are independent variables.
All the dependent variables depends on some set of independent variables. So that every independent noise is added to each dimension of the independent variable set.
The model can be written as

$$y'(x_n, w) = w_o + \sum_{i=1}^{N} w_i(x_{ni} + \epsilon_{ni}) \tag{0.0.1}$$

$$= w_o + \sum_{i=1}^{N} w_i x_{ni} + \sum_{i=1}^{N} w_i \epsilon_{ni} \tag{0.0.2}$$

$$= y(x_n, w) + \sum_{i=1}^{N} w_i \epsilon_{ni} \tag{0.0.3}$$

The Error function is

$$E'_D(w) = \frac{1}{2} \sum_{n=1}^{N} \{y'(x_n, w) - t_n\}^2 \tag{0.0.4}$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left\{ y(x_n, w) + \sum_{i=1}^{N} w_i \epsilon_{ni} - t_n \right\}^2 \tag{0.0.5}$$

expanding the eq (0.0.5)

$$= \frac{1}{2} \sum_{n=1}^{N} \left\{ (y(x_n, w) - t_n)^2 + 2\,(y(x_n, w) - t_n) \left(\sum_{i=1}^{N} w_i \epsilon_{ni} \right) + \left(\sum_{i=1}^{N} w_i \epsilon_{ni} \right)^2 \right\} \tag{0.0.6}$$

Taking the expectation of eq (0.0.6)

$$\mathbb{E}[E'_D] = \frac{1}{2}\sum_{n=1}^{N}\left\{(y(x_n, w) - t_n)^2 + 2\,(y(x_n, w) - t_n)\left(\sum_{i=1}^{N} w_i \mathbb{E}\left[\epsilon_{ni}\right]\right) + \mathbb{E}\left[\left(\sum_{i=1}^{N} w_i \epsilon_{ni}\right)^2\right]\right\} \tag{0.0.7}$$

given

$$\mathbb{E}[\epsilon_{ni}] = 0 \tag{0.0.8}$$

Now expanding the $\mathbb{E}\left[\left(\sum_{i=1}^{N} w_i \epsilon_{ni}\right)^2\right]$ we get

$$\mathbb{E}\left[\left(\sum_{i=1}^{N} w_i \epsilon_{ni}\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \epsilon_{ni}\epsilon_{nj}\right] \tag{0.0.9}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \mathbb{E}\left[\epsilon_{ni}\epsilon_{nj}\right] \tag{0.0.10}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \delta_{ij}\sigma^2 \tag{0.0.11}$$

$$= \sum_{i=1}^{N} w_i^2 \tag{0.0.12}$$

substituting eq (0.0.12) and (0.0.8) in eq (0.0.7)

$$\mathbb{E}\left[E'_D(w)\right] = \frac{1}{2}\sum_{n=1}^{N}\left[y(x_n, w) - t_n\right]^2 + \sum_{i=1}^{N} w_i^2 \tag{0.0.13}$$

$$= E_D(w) + \frac{N}{2}\sum_{i=1}^{N} w_i^2 \tag{0.0.14}$$

$\therefore$ we get a L_2 regularization term without the bias parameter w_o.