Naman Jain Vimal Kumar
1001670153

Problem 1 - Exercise 1.1

Consider the sum-of-squares error function given by (1.2) in which the function y(x, w) is given by the polynomial (1.1). Show that the coefficients w = {w_i} that minimize this error function are given by the solution to the following set of linear equations

$$\sum_{j=0}^{M} A_{ij} w_j = T_i$$

where,

$$A_{ij} = \sum_{n=1}^{N} {x_n}^{i+j} \text{ and } T_i = \sum_{n=1}^{N} {x_n}^{i} * t_n$$

Here a suffix i or j denotes the index of a component, whereas x^i denotes x raised to the power of i.

Solution

We know that,

$$y(x,w) = w_0 + w_1 x + w_2 x^2 + \cdots + w_m x^m = \sum_{j=0}^{M} w_j x^j \quad \ldots (1.1)$$

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, w) - t_n\}^2 \quad \ldots (1.2)$$

Using 1.1, finding the derivative with respect to w_j:

$$y(x,w) = \sum_{j=0}^{M} w_j x^j$$

$$\frac{dy(x,w)}{dw_j} = x^j \qquad \textit{As all other will equate to } 0. \qquad \ldots (1)$$

Using 1.2, expanding the equation:

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

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, w)^2 - (2 * y(x_n, w) * t_n) + (t_n)^2\}$$

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, w)^2\} - \frac{1}{\cancel{2}} * \cancel{2} * \sum_{n=1}^{N} \{y(x_n, w) * t_n\} + \frac{1}{2} \sum_{n=1}^{N} \{t_n\}^2$$

Taking derivative with respect to w_i:

$$\frac{dE(w)}{dw_i} = \frac{1}{2} \sum_{n=1}^{N} \left\{2 * y(x_n, w) * \frac{dy(x_n, w)}{dw_i}\right\} - \sum_{n=1}^{N} \left\{\frac{dy(x_n, w)}{dw_i} * t_n\right\} + 0$$

$$\frac{dE(w)}{dw_i} = \frac{1}{\cancel{2}} * \cancel{2} * \sum_{n=1}^{N} \left\{y(x_n, w) * \frac{dy(x_n, w)}{dw_i}\right\} - \sum_{n=1}^{N} \left\{\frac{dy(x_n, w)}{dw_i} * t_n\right\} \ldots (2)$$

Substituting (1) and (1.1) in the above equation

$$\frac{dE(w)}{dw_i} = \sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^j * {x_n}^i\right\} - \sum_{n=1}^{N}\{{x_n}^i * t_n\}$$

To minimize the error, we need to equate the derivative to 0. Therefore,

$$\sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^j * {x_n}^i\right\} - \sum_{n=1}^{N}\{{x_n}^i * t_n\} = 0$$

$$\sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^{i+j}\right\} - \sum_{n=1}^{N}\{{x_n}^i * t_n\} = 0$$

$$\sum_{j=0}^{M}\left\{w_j * \sum_{n=1}^{N} {x_n}^{i+j}\right\} - \sum_{n=1}^{N}\{{x_n}^i * t_n\} = 0$$

Substituting T_i and A_i in the above equation:

$$\sum_{j=0}^{M}\{w_j * A_{ij}\} - T_i = 0$$

$$\boldsymbol{\sum_{j=0}^{M}\{A_{ij} * w_j\} = T_i}$$

Problem 2

Show that when M=1, the results of Problem1 is identical the results of linear regression.

Solution

We know that linear regression follows the below equation:

$$y = mx + c$$

Here, ‘m’ is the slope and ‘c’ is the x-intercept

Equating M = 1 in Equation 1.1:

$$y(x, w) = \sum_{j=0}^{1} w_j x^j = w_0 x^0 + w_1 x^1$$

$$y(x, w) = w_0 + w_1 x \qquad \text{... (3)}$$

And the derivation with respect to w_i remains the same. That is:

$$\frac{dy(x, w)}{dw_j} = x^j \qquad \text{... (4)}$$

Let us continue from step (2) of Problem 1:

$$\frac{dE(w)}{dw_i} = \sum_{n=1}^{N}\left\{y(x_n, w) * \frac{dy(x_n, w)}{dw_i}\right\} - \sum_{n=1}^{N}\left\{\frac{dy(x_n, w)}{dw_i} * t_n\right\} \quad ... (2)$$

Substituting (3) and (4) in the above equation (2):

$$\frac{dE(w)}{dw_i} = \sum_{n=1}^{N}\{w_0 + w_1x_n * x_n{}^i\} - \sum_{n=1}^{N}\{x_n{}^i * t_n\}$$

To minimize the error, we need to equate the derivative to 0. Therefore,

$$\sum_{n=1}^{N}\{(w_0 + w_1x_n) * x_n{}^i\} - \sum_{n=1}^{N}\{x_n{}^i * t_n\} = 0$$

$$\sum_{n=1}^{N}\{(w_0 + w_1x_n) * x_n{}^i\} - \sum_{n=1}^{N}\{x_n{}^i * t_n\} = 0$$

$$\sum_{n=1}^{N}\{w_0x_n{}^i + w_1x_n{}^{i+1} - t_nx_n{}^i\} = 0$$

$$\sum_{n=1}^{N}\{x_n{}^i * (w_0 \ + w_1x_n \ - t_n)\} = 0$$

Dividing both the sides by x_n^i:

$$\sum_{n=1}^{N}\{w_0 \ + w_1x_n \ - t_n\} = 0$$

$$\sum_{n=1}^{N} w_0 + \sum_{n=1}^{N} w_1x_n - \sum_{n=1}^{N} t_n = 0$$

$$\sum_{n=1}^{N} t_n = \sum_{n=1}^{N} w_0 + \sum_{n=1}^{N} w_1x_n$$

Therefore, for every i in n, the equation can be written as:

$$t_i = w_1x_i + w_0$$

Here, slope 'm' can be taken as 'w_1' and constant c as 'w_0'. Thus, representing the linear regression equation.

Problem 3 – Exercise 1.2

Write down the set of coupled linear equations, analogous to (1.122), satisfied by the coefficients w_i which minimize the regularized sum-of-squares error function given by (1.4)

Solution

We know that,

$$y(x, w) = w_0 + w_1x + w_2x^2 + \cdots + w_mx^m = \sum_{j=0}^{M} w_jx^j \quad ... (1.1)$$

$$E(w) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2 + \frac{\lambda}{2}\|w\| \qquad \text{... (1.4)}$$

Using 1.1, finding the derivative with respect to w_j:

$$y(x, w) \quad = \sum_{j=0}^{M} w_j x^j$$

$$\frac{dy(x, w)}{dw_j} = x^j \qquad \textit{As all other will equate to } 0. \qquad \text{... (5)}$$

Using 1.4, expanding the equation:

$$E(w) \quad = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2 + \frac{\lambda}{2}\|w\|$$

$$E(w) \quad = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2 + \frac{\lambda}{2}w^2$$

$$E(w) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w)^2 - (2 * y(x_n, w) * t_n) + (t_n)^2\} + \frac{\lambda}{2}w^2$$

$$E(w) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, w)^2\} - \frac{1}{\cancel{2}} * \cancel{2} * \sum_{n=1}^{N}\{y(x_n, w) * t_n\} + \frac{1}{2}\sum_{n=1}^{N}\{t_n\}^2 + \frac{\lambda}{2}w^2$$

Taking derivative with respect to w_i:

$$\frac{dE(w)}{dw_i} = \frac{1}{2}\sum_{n=1}^{N}\left\{2 * y(x_n, w) * \frac{dy(x_n, w)}{dw_i}\right\} - \sum_{n=1}^{N}\left\{\frac{dy(x_n, w)}{dw_i} * t_n\right\} + 0 + \frac{\lambda}{\cancel{2}} * \cancel{2}w$$

$$\frac{dE(w)}{dw_i} = \frac{1}{\cancel{2}} * \cancel{2} * \sum_{n=1}^{N}\left\{y(x_n, w) * \frac{dy(x_n, w)}{dw_i}\right\} - \sum_{n=1}^{N}\left\{\frac{dy(x_n, w)}{dw_i} * t_n\right\} + \lambda w$$

Substituting (1) and (1.1) in the above equation

$$\frac{dE(w)}{dw_i} = \sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^j * {x_n}^i\right\} - \sum_{n=1}^{N}\left\{{x_n}^i * t_n\right\} + \lambda w$$

To minimize the error, we need to equate the derivative to 0. Therefore,

$$\sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^j * {x_n}^i\right\} - \sum_{n=1}^{N}\left\{{x_n}^i * t_n\right\} + \lambda w = 0$$

$$\sum_{n=1}^{N}\left\{\sum_{j=0}^{M} w_j {x_n}^{i+j}\right\} - \sum_{n=1}^{N}\left\{{x_n}^i * t_n\right\} + \lambda w = 0$$

$$\sum_{j=0}^{M}\left\{w_j * \sum_{n=1}^{N} {x_n}^{i+j}\right\} - \sum_{n=1}^{N}\left\{{x_n}^i * t_n\right\} + \lambda w = 0$$

Naman Jain Vimal Kumar
1001670153

Substituting T_i and A_i in the above equation:

$$\sum_{j=0}^{M} \{w_j * A_{ij}\} - T_i + \lambda w = 0$$

$$\boldsymbol{\sum_{j=0}^{M} \{A_{ij} * w_j\} + \lambda w = T_i}$$

References:

- Probability and Statistics for Computer Scientists, Second Edition, Michael Baron.