

MAD Assignment 2

Ask Jensen

1. december 2021

Indhold

1	Problem 1	1
1.1	(a)	1
2	Problem 2	2
3	Problem 3	2
4	Problem 4	2
5	Problem 5	2

1 Problem 1

1.1 (a)

$$\begin{aligned}
 \mathcal{L} &= \frac{1}{N} \sum_{n=1}^N A(\mathbf{w}^T \mathbf{x}_n - t_n)^2 \\
 &= A((\mathbf{X}\mathbf{w} - \mathbf{t})^T (\mathbf{X}\mathbf{w} - \mathbf{t}))^T \\
 &= A((\mathbf{X}\mathbf{w})^T - \mathbf{t}^T)(\mathbf{X}\mathbf{w} - \mathbf{t})^T \\
 &= A\left(\frac{1}{N}(\mathbf{X}\mathbf{w})^T \mathbf{X}\mathbf{w} - \frac{1}{N} \mathbf{X}\mathbf{w} \mathbf{t}^T - \frac{1}{N} (\mathbf{X}\mathbf{w})^T \mathbf{t} + \frac{1}{N} \mathbf{t}^T \mathbf{t}\right) \\
 &= A\left(\frac{1}{N} \mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w} - \frac{2}{N} \mathbf{w}^T \mathbf{X}^T \mathbf{t} + \frac{1}{N} \mathbf{t}^T \mathbf{t}\right)
 \end{aligned}$$

using "case 4" from the table 1.14

$$\mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w} = 2\mathbf{X}^T \mathbf{X} \mathbf{w}$$

and "case 1" gives me

$$2\mathbf{w}^T \mathbf{X}^T \mathbf{t} = 2\mathbf{X}^T \mathbf{t}$$

thus

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 2\mathbf{X}^T \mathbf{X} \mathbf{w} - 2\mathbf{X}^T \mathbf{t} = 0$$

+2 and -2 cancel out leaving me with

$$\mathbf{X}^T \mathbf{X} \mathbf{w} - \mathbf{X}^T \mathbf{t} = 0$$

$$\mathbf{X}^T \mathbf{X} \mathbf{w} = \mathbf{X}^T \mathbf{t}$$

multiplying both sides with $(\mathbf{X}^T \mathbf{X})^{-1}$ which can be rewritten as \mathbf{I} which is the denotation of the Identity matrix

$$\mathbf{I} \mathbf{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{t}$$

Multiplying the vector \mathbf{w} with the Identity matrix, will simply return the vector \mathbf{w} thus the result is

$$\hat{\mathbf{w}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{t}$$

The total training loss is the average loss multiplied by the number of instances in the dataset. which is exactly the same as the value we obtain from the average loss function

2 Problem 2

3 Problem 3

4 Problem 4

5 Problem 5