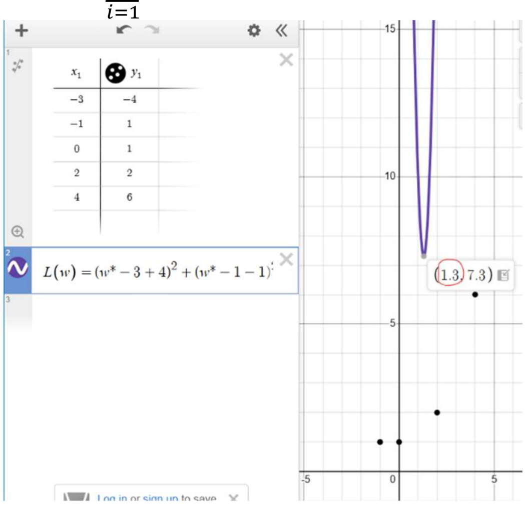
Introduction of Machine Learning
EX1-regression

Matthis Brocheton

Problem 1 - a

Euclidean standard square:

$$L(w) = \sum_{i=1}^{n} (w\, x_i - y_i)^2$$



Euclidean standard:

$$w * = \frac{\sum x_i y_i}{\sum x^2{}_i}$$

$\sum x_i y_i = \big((-3 * \ -4) + (-1 * 1) + (0 * 1) + (2 * 2) + (4 * 6)\big)$ = 39

$\sum x^2{}_i = ((-3)^2 + (-1)^2 + (0)^2 + 2^2 + 4^2)$ = 30

$$w * = \frac{39}{30} = \ 1.3$$

$w * = 1.3$

Problem 1 – b

$$f(x) = \ w0 \ + \ w1x$$

$$w \ = \ [w0, w1]T$$

Normal equation

$$w \ = \ (X^T X)^{-1} X^T y$$

Matrix

$$X = \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} y = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 0.2055 & -0.0137 \\ -0.0137 & 0.0342 \end{bmatrix}$$

$X^T y = \begin{bmatrix} 6 \\ 39 \end{bmatrix}$

$$(X^T X)^{-1} X^T y = \begin{bmatrix} 0.6987 \\ 1.2516 \end{bmatrix}$$

$$w = \begin{bmatrix} 0.6987 \\ 1.2516 \end{bmatrix}$$

Problem 1 – c

$model\ 1: f(x) = 1.3x$

$model\ 2: g(x) = 0.6987 + 1.2534x$

$f(1) = 1.3$

$g(1) = 1.9521$

Problem 1 – d

Loss functions

$$J(w) = \frac{1}{2m} \sum_{i=1}^{m} (w\, x_i - y_i)^2$$

m : This is the name of the data samples
x_i : is the entrance (valve of x to the i-eme point)
y_i : is the actual output associated with x_i
wx_i: is the model prediction

Update rule

$$w \coloneqq w - \alpha \frac{dJ}{dw}$$

$\alpha = 0.1$
$w = 0$

Matrix

$$y = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix}$$

$m = 5$

Derivative of $J(w)$

$$\frac{\mathrm{dJ}}{\mathrm{dw}} = \frac{1}{\mathrm{m}} \sum_{i=1}^{m} x_i(wx_i - y_i)$$

Calculation

$$\frac{\mathrm{dJ}}{\mathrm{dw}} = \frac{1}{5} \sum_{i=1}^{5} x_i(wx_i - y_i)$$

With w_0 we will calculate the gradient

$$\frac{dw}{dJ} = \frac{5}{1}[(-3)(0-(-4)) + (-1)(0-1) + (0)(0-1) + (2)(0-2) + (4)(0-6)]$$

$$\frac{dw}{dJ} = \frac{5}{1}(-39) = -7.8$$

gradient descent update with -7.8

$$w_1 = -\alpha \frac{dJ}{dw}$$

$$w_1 = 0 - (0.1)(-7.8) = 0.78$$

With w_1 we will calculate the gradient

$$\frac{\mathrm{dJ}}{\mathrm{dw}} = \frac{1}{5}[(-3)(0.78(-3) - (-4)) + (-1)(0.78(-1) - 1) + (0)(0.78(0) - 1) + (2)(0.78(2) - 2) + (4)(0.78(4 - 6)]$$

$$\frac{\mathrm{dJ}}{\mathrm{dw}} = \frac{5}{1}(-15.6) = -3.12$$

gradient descent update with -3.12

$$w_2 = -\alpha \frac{dJ}{dw}$$

$$w_2 = 0.78 - (0.1)(-3.12)$$

$$w_2 = 0.78 + 0.312 = 1.092$$

This is these 2 gradient descent iteration:

$w_1 = 0.78$

$w_2 = 1.092$

Problem 2 – a

$$y_1 = w_0 x_1 + w_1 x_2$$

$$y_2 = w'_0 x_1 + w'_1 x_2$$

Normal equation

$$w = (X^T X)^{-1} X^T Y$$

Matrix

$$X = \begin{bmatrix} 3 & 5 \\ 2 & 2 \\ 1 & 2 \\ 3 & 3 \end{bmatrix} Y = \begin{bmatrix} 7 & 10 \\ 4 & 5 \\ 2 & 9 \\ 6 & 6 \end{bmatrix}$$

Calculation

$$X^T X = \begin{bmatrix} 23 & 30 \\ 30 & 42 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 0.6364 & -0.4545 \\ -0.4545 & 0.3485 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 49 & 67 \\ 65 & 96 \end{bmatrix}$$

$$w = (X^T X)^{-1} X^T Y = \begin{bmatrix} 0.6364 & -0.4545 \\ -0.4545 & 0.3485 \end{bmatrix} \begin{bmatrix} 49 & 67 \\ 65 & 96 \end{bmatrix}$$

$$w = \begin{bmatrix} 1.6359 & -0.9090 \\ 0.3550 & 3.0025 \end{bmatrix}$$

Problem 2 – b

$$w = \begin{bmatrix} 1.6359 & -0.9090 \\ 0.3550 & 3.0025 \end{bmatrix}$$

$x_1 = 7$
$x_2 = 4$

$y_1 = w_0 x_1 + w_1 x_2$
$y_2 = w'_0 x_1 + w'_1 x_2$

$y1 = (1.6359 \times 7) + (-0.9090 \times 4) = 11.4513 - 3.6360 = 7.8153$
$y2 = (0.3550 \times 7) + (3.0025 \times 4) = 2.4850 + 12.0100 = 14.4950$

$y1 \approx 7.8153$
$y2 \approx 14.4950$