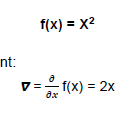
**Workshop 2: Linear Regression using Gradient Descent.**

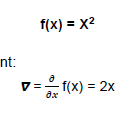
In this practical, we address the problem of Linear Regression using Gradient Descent. Model answers are provided in the weekly lectures.

# **Task 1**

**We will minimize a toy example function:**



which has the following gradient:



## **Run a gradient descent algorithm for four iterations, starting with an arbitrary point: x= -2 and a step size of 0.1 [5 marks]**

**Example iteration 1:**

**X (next step) = X – η\* ∇(f(X))**

Starting point is: X= -2

X (next step) = (-2) – 0.1\*(2X)

X (next step) = (-2) – 0.1 (-4)

X (next step) = -2 + (0.1)4

X (next step) = -2 + 0.4

X (next step) = -1.6

**Iteration 2:**

**X (next step) = X – η\* (∇)f(X)**

X = X – (η)(2X)

X = -1.6 – (0.1)(2(-1.6))

X = -1.6 + 0.32

X = -1.28

**Iteration 3:**

**X (next step) = X – η\* (∇)f(X)**

X = X – (η)(2X)

X = -1.28 – (0.1)(2(-1.28)

X = -1.28 + 0.256

X = -1.024

**Iteration 4:**

**X (next step) = X – η\* (∇)f(X)**

X = X – (η)(2X)

X = -1.024 – (0.1)(2(-1.024))

X = -1.024 + 0.2048

X = -0.8192

# **Task 2: Develop, train, and test a Linear Regression Model with Boston housing dataset using Gradient Descent. Goal is to predict the houses price in Boston.**

## **2.1. Please run the jupyter notebook provided on Canvas to plot the predictions after 100, 1000, 10000, and 40000 iterations, and explain the results. [2 marks]**

## **2.2. After 40000 iterations, compare the results with WS1-task 1 and explain the difference in your own words [2 marks]**

**For example, after 10 iterations the difference looks something like is:**

| **Week 1 workshop** | **Current workshop** |
| --- | --- |
| Chart  Description automatically generated | Chart  Description automatically generated |

A**fter 4000 iterations the difference looks something like is:**

## **2.3 After 40000 iterations, print network weights and identify the inputs with the highest weight. Compare with the results of WS1-task 2.2 [1 mark]**

Use the table below to compare the weights.

| **Input** |  | **Weights**  **Wk 1** | **Theta\_Best (10000 its.) Wk 2** |
| --- | --- | --- | --- |
| CRIM | *Per capita crime rate by town* | 88 |  |
| ZN | *Proportion of residential land zoned for lots over 25000 sq.ft* |  |  |
| INDUS | *Proportion of non-retail business acres per town* |  |  |
| CHAS | *Charles River dummy variable* |  |  |
| NOX | *Nitric oxides concentration (parts per 10 million)* |  |  |
| RM | *Average number of rooms per dwelling* |  |  |
| AGE | *Proportion of owner-occupied units built prior to 1940* |  |  |
| DIS | *Weighted distances to five Boston employment centres* |  |  |
| RAD | *Index of accessibility to radial highways* |  |  |
| TAX | *Full-value property-tax rate per $10000* |  |  |
| PTRATIO | *Pupil-teacher ratio by town* |  |  |
| B | *1000(Bk-0.63)^2 where Bk is the proportion of blacks by town* |  |  |
| LSTAT | *% lower status of the population* |  |  |