UiT

THE ARCTIC UNIVERSITY OF NORWAY

2. Linear Regression

FYS-2021 Exercises

Department of Physics and Technology
Faculty of Science and Technology

Linear regression

Problem 1

(1a) Go to the website https://udlbook.github.io/udlfigures/, choose on the left "least squares loss" and play with this interactive example of linear regression. Find the best intercept and slope that minimize the loss.

Did you reach the minimum by changing the parameters doing small steps while looking at the loss value? did you change the slider direction when the loss was increasing? If yes, it looks like you intuitively performed gradient descent!

Problem 2

The data file `global-temperatures.csv` contains the average global annual temperatures spanning from the year 1880 to 2017.

(2a) Use the file from canvas and perform linear regression, with temperature as a function of years, and plot the results.

(2b) Briefly explain what the R^2 value tells you, and calculate the R^2 value for this model.

(2c) Assuming the regression model is on the form $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, then what is the interpretation of the estimator $\hat{\beta}_1$?

(2d) Plot the residuals and comment on the result with regards to the assumptions made in the regression model.

Problem 3

In this problem, we will use multiple linear regression to predict the fuel consumption of cars (measured in miles per gallon). The available predictors are: (i) cylinders, (ii) displacement, (iii) horsepower, (iv) weight, (v) acceleration and (vi) model year. The dataset is named `auto-mpg.csv`, and can be found in Canvas.

(3a) Implement your own function for estimating the linear-regression parameters, and use this to fit a regression model to the fuel-consumption data.

(3b) Explain why the magnitudes of the estimated coefficients are so different. What could be done to prevent this from happening?

(3c) Compute R^2 and use this to evaluate the quality of your model.

(3d) Experiment with removing different predictors from the model. Do you think all predictors are equally necessary? Explain.