Practicals to Linear Regression

Aufgabe P 1. *Linear Regression, due 2. November, in class*

Before starting the exercise make sure you have installed Python and Tensorflow. In this practical you will use a dataset * that was taken from the StatLib library which is maintained at Carnegie Mellon University.The Boston house dataset has 506 samples and each sample contains 13 features. The goal is to learn a linear model to predict house prices.

In Tensorflow you first define the model, i.e. the computational graph. The calculations are done afterwards when you start a *Session*. Tensorflow has different ways to input the data into the graph. In our class you will mainly use the so called *placeholder* pipeline. This means you specify what kind of data and what the shape of your data is. The stubs contain information in this regard.

(a) Open the starter files and run them

(b) Split the data into training: first 400 rows, and test: rest. (5 Pts)

(c) Implement the Normal Equations in Python to have the optimal parameters for this model. (10 Pts)

(d) Implement with Tensorflow the linear regression by using in-built Tensorflow's GradientDescentOptimizer. (40 Pts)

(e) Show a scatterplot of the actual price versus predicted price (5 Pts)

(f) Implement with Tensorflow the linear regression by computing the gradient and update the parameters iteratively. (40 Pts)

---

*The Boston house-price data of Harrison, D. and Rubinfeld, D.L. 'Hedonic prices and the demand for clean air', J. Environ. Economics Management, vol.5, 81-102, 1978. Used in Belsley, Kuh Welsch, 'Regression diagnostics ...', Wiley, 1980.