

# Artificial Intelligence - Linear regression

Martin Johansson dat13mjo and Robin Timan dat13rti

February 17, 2017

## 1 Find the file

The runnable file is found in the map dat13mjo/AI/project2. The source files exists in a zip file in the same map named project2.zip

## 2 Run the program

To run the executable just type "python3 batch.py" for the batch implementation of the gradient descent function, "python3 stochastic.py" for the stochastic version.

To run the Perceptron version write "python3 perceptron.py" and for the logistic regression version write "python3 logic.py"

## 3 Assignment

The assignment is to implement linear regression using gradient descent, known as the method of steepest decent. The use of gradient descent is to find a value that minimize the function, in this case the loss between the approximation of the function and the function values themselves. More or less one takes steps to get closer and closer to the minimum of the evaluated function. When the value does not change more than a predefined threshold for the change between two sets the evaluation stops and the approximation is given.

The perceptron algorithm is used for the classification of values, it makes prediction based on a linear regression line and classifies the values in the set on their location corresponding to this vector. In our case we want it to classify a book as either French or English depending on the frequency of "a" in the text. The perceptron algorithm sets the value as either 1 or 0

logistic regression is just a method to make the transaction between the 1 and 0 of the perceptron algorithm more smooth. The basic hard nature that it only provides 1 and 0 may cause problem for values close to the classification vector. So linear regression is used to soften this classification.

## 4 Implementation

The implementation of our code is more or less completely based on the approach found in the textbook on the sides 719-727 which is the sides presented in the problem text of the assignment. We choose to implement the stochastic version by comparing the difference between to adjacent values of the loss. To make the system a bit more robust to changing approach in the form of batch or stochastic we choose to implement the acceptable loss value as the square of the value for the batch function.

for

## 5 Result

The result of our linear regression shows that we obtain a good result for both batch and stochastic implementation of the regression function. The stochastic version is as expected faster than the batch version, and this is of course because we don't go through the entire set every time we evaluate the function.