## CS 4850: Introduction to Machine Learning

© Fall 2018

## Homework 2: Gradient Descent

INSTRUCTOR: DANIEL L. PIMENTEL-ALARCÓN

DUE 09/24/2018

In this homework you will code gradient descent and use it to find the minimizer of a function, i.e., to solve an optimization problem of the form:

$$\underset{\mathbf{x} \in \mathbb{R}^{\mathrm{D}}}{\operatorname{arg\,min}} \ f(\mathbf{x}).$$

Recall that given a function  $f(\mathbf{x})$  to be minimized, gradient descent can be summarized as follows:

- (i) Select an initial point  $\mathbf{x}_0$ .
- (ii) For every t > 0, repeat:

$$\mathbf{x}_{\mathrm{t}} = \mathbf{x}_{\mathrm{t-1}} - \eta \nabla f(\mathbf{x}_{\mathrm{t-1}})$$

until convergence, where  $\eta > 0$  is a step size (parameter).

Implement gradient descent in your favorite coding language, and use it to identify the solution to

$$\underset{\mathbf{x} \in \mathbb{R}^5}{\arg\min} \ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2,$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \\ 1 & 4 & 16 & 64 & 256 \\ 1 & 5 & 25 & 125 & 625 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 5 \\ 31 \\ 121 \\ 341 \\ 781 \end{bmatrix}.$$

Deliver your code, solution, and choice of  $\eta$ .