Homework 2 for Math 173A - Fall 2019

A Foray into Least Squares

1. For each of the functions below, find the gradient. Remember to show your work

   (a) $f : \mathbb{R}^n \to \mathbb{R}$, with $f(x) = \|x\|_2^2$.

   (b) $f : \mathbb{R}^n \to \mathbb{R}$, with $f(x) = \|Ax\|_2^2$, where A is a matrix of size $m \times n$.
   Hint: The chain rule is your friend. Write $f(x) = g(h(x))$, with $g(z) = \|z\|_2^2$ and $h(w) = Aw$.

   (c) $f : \mathbb{R}^n \to \mathbb{R}$, with $f(x) = \|Ax - b\|_2^2$ where A is a matrix of size $m \times n$, and $b \in \mathbb{R}^m$.

   (d) $f : \mathbb{R}^n \to \mathbb{R}$, with $f(x) = \|Ax - b\|_2^2 + \gamma\|x\|_2^2$ where A is a matrix of size $m \times n$, $b \in \mathbb{R}^m$, and γ is a positive constant.

2. For each of the functions f in Question 1 above, write a gradient descent algorithm to solve the optimization problem

$$\min_{x \in \mathbb{R}^n} f(x).$$

   Note: Here, I am *not asking you to write a computer program*, just to write down the gradient descent iterations (and initialization) associated with the functions from Question 1.

3. Show that

   (a) the function $g : \mathbb{R}^n \to \mathbb{R}$, given by $g(x) = \|x\|_2^2$, is convex.

   (b) if $f : \mathbb{R}^m \to \mathbb{R}$ is convex, then $g : \mathbb{R}^n \to \mathbb{R}$ given by $g(x) = f(Ax - b)$ is also convex. A here is an $m \times n$ matrix, and $b \in \mathbb{R}^m$ is a vector.
   You may use any of the results we covered in class (but the definition of convexity may be an easy way to do this, and gives the most general result).

   (c) Conclude that the function given by $f(x) = \|Ax - b\|_2^2$ is convex.

4. Use the results of Question 3, and the fact that the gradient of a convex function is zero at the global minimizer, to write down a *closed form* expression for the solution x^* to the optimization problem

$$\min_{x \in \mathbb{R}^n} f(x).$$

   for each function f from Question 1.

5. **Computer Problem:** For this problem, recall that the equation of an ellipse in $\mathbb{R}^2$ is

$$a_1 x^2 + a_2 y^2 = 1. \tag{1}$$

Now, suppose that somebody gives you the coordinates of N points in $\mathbb{R}^2$ that you suspect may lie on an ellipse (or close to it) and you would like to find the coefficients a and b in Equation (1), so your ellipse best fits the data.

If the data points are $(u_i, v_i) \in \mathbb{R}^2, i = 1, ..., N$, this basically means we would like to solve the optimization problem:

$$\min_a f(a),$$

where

$$f(a) := f(a_1, a_2) = \sum_{i=1}^{N} (a_1 u_i^2 + a_2 v_i^2 - 1)^2$$

(a) Show that $f(a) = \|Aa - b\|_2^2$ for an appropriate choice of A and b. What are they, in terms of the data points (u_i, v_i)?

(b) Write a gradient descent algorithm for solving the optimization problem.

(c) Write a computer code (function) for implementing your algorithm. Use a step-size of $\mu = \frac{1}{\|A^T A\|}$.

(d) Using the data provided on the course Homework page, run your gradient descent algorithm to approximate the optimal $a = (a_1, a_2)$.

(e) Using the results of Question 4 (i.e., the closed form expression for the optimal solution), estimate the optimal a, and compare the results.

(f) Provide a scatter plot of the data $(u_i, v_i), i = 1, ..., N$ and plot an ellipse with the a_1, a_2 you computed in parts (d) and (e).