DS-GA 1014: Homework Problem Set 6

Optimization and Computational Linear Algebra for Data Science (Fall 2018)

Afonso S. Bandeira
bandeira@cims.nyu.edu
http://www.cims.nyu.edu/~bandeira

Brett D. Bernstein
brettb@
~brettb

Due on Friday October 19, 2018

This homework problem set is due on October 19 on NYU Classes.

If you have questions about the homework feel free to contact Brett Bernstein (brett.bernstein@nyu.edu**) or myself, or stop by our office hours.**

Unless otherwise stated all answer must be mathematically justified, even questions that ask for a computation.

Try not to look up the answers, you'll learn much more if you try to think about the problems without looking up the solutions. If you need hints, feel free to ask on Piazza.

You can work in groups but each student must write his/her own solution based on his/her own understanding of the problem. Please list, on your submission, the students you work with for the homework (this will not affect your grade).

Late submissions will be graded with a penalty of 10% per day late. Weekend days do not count, from Friday to Monday counts only 1 day.

If you need to impose extra conditions on a problem to make it easier (or consider specific cases of the question, like taking n to be 2, e.g.), state explicitly that you have done so. Solutions where extra conditions were assume, or where only special cases where treated, will also be graded (probably scored as a partial answer).

Problems with a $(*)$ are extra credit, they will not (directly) contribute to your score of this homework. However, for every 4 (four) extra credit questions successfully answered you get a homework "bye": your lowest homework score (or one you did not hand in) gets replaced by a perfect score.

Problem 6.1 *Let $Q \in \mathbb{R}^{n\times n}$. Prove that Q is orthogonal if and only if $\|Qx\| = \|x\|$ for all $x \in \mathbb{R}^n$. [Hint: Try $x = e_i$ and $x = e_i + e_j$.]*

Problem 6.2 *For any subspace U of $\mathbb{R}^n$ prove that $I - 2P_U$ is orthogonal, where P_U is the (orthogonal) projection onto U.*

Problem 6.3 *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $\lambda > 0$.*

1. *Prove there is a matrix $B \in \mathbb{R}^{(m+n)\times n}$ and vector $y \in \mathbb{R}^{m+n}$ such that*

$$\|Ax - b\|^2 + \lambda\|x\|^2 = \|Bx - y\|^2$$

   for all $x \in \mathbb{R}^n$.

2. *Write the minimizer of*

$$\operatorname{argmin}_x \|Ax - b\|^2 + \lambda\|x\|^2$$

   as a solution to a linear system (in terms of A and b). Justify your answer; no need to solve the linear system.

Problem 6.4 *In each of the following questions, it is intended that you solve the problem using the numpy/scipy libraries in Python, but you can use any programming language. Please include just the answer in your writeup, and include the code files in your submission (i.e., do not include the code in your writeup). A good idea is to test your code on some small simple instances you create by hand to make sure it is correct. Note that calculations on the computer do not use exact arithmetic, so you should allow for some margin of error in your code.*

The file `HW6matrix1.txt` contains a 200×2 matrix A. Each line of each file represents a row of the corresponding matrix, and the values on each line are space-delimited. The ith row of A contains the ith data point $(x_i, y_i) = (A_{i,1}, A_{i,2})$. The data is such that $x_i \in [0, 3]$ for $i = 1, \dots, 200$. For any $\alpha, \beta \in \mathbb{R}^3$ we define the function $f_{\alpha,\beta} : [0, 3] \to \mathbb{R}$ by

$$f_{\alpha,\beta}(x) = \begin{cases} \alpha_1 + \beta_1 x & \text{if } x \in [0,1] \\ \alpha_2 + \beta_2 x & \text{if } x \in (1,2] \\ \alpha_3 + \beta_3 x & \text{if } x \in (2,3]. \end{cases}$$

1. *Compute the α and β that minimize*

$$\sum_{i=1}^{200} (f_{\alpha,\beta}(x_i) - y_i)^2.$$

   Give the coordinates of α and β to two decimal places.

2. *Compute the α and β that minimize*

$$\sum_{i=1}^{200} (f_{\alpha,\beta}(x_i) - y_i)^2$$

with the added constraint that $f_{\alpha,\beta}$ must be continuous on $[0,3]$. In other words, we must have

$$\alpha_1 + \beta_1 = \alpha_2 + \beta_2 \quad \textit{and} \quad \alpha_2 + 2\beta_2 = \alpha_3 + 2\beta_3.$$

Give the coordinates of α and β to two decimal places.

In both parts, you may find it helpful to superimpose your plotted function on top of a scatter plot of the data.

$(*)$ **Problem 6.5 (For Extra Credit)** *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Show that*

$$R(x) = \frac{x^T A x}{x^T x},$$

has a maximum $x^\natural$ over $\mathbb{R}^n \setminus \{0\}$, and that this maximum satisfies:

$$Ax^\natural = \lambda x^\natural,$$

where $\lambda = R(x^\natural)$.