DS-GA 3001.03: Homework Problem Set 5

Optimization and Computational Linear Algebra for Data Science

(Fall 2016)

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

Due on October 18, 2016

This homework problem set is due on October 18, before class at the homework dropbox in the CDS reception.

If you have any questions about the homework post them on the piazza page for the course, contact Vlad Kobzar at vkobzar@cims.nyu.edu **or myself at** bandeira@cims.nyu.edu**, or stop by our office hours.**

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 email me or Vlad.

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).

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.

Consult the Extended Syllabus 5 for the definitions needed in this homework, it is available at `http://www.cims.nyu.edu/~bandeira/2106_DS_ExtendedSyllabus5.pdf`

Problem 1.1 *Given $A \in \mathbb{R}^{n\times m}$ show:*

- *(a)* $\|A\|_F^2 = \mathrm{Tr}(A^T A) = \mathrm{Tr}(AA^T)$.
- *(b)* $\|A\|_F^2 = \sum_{k=1}^{\min\{n,m\}} \sigma_k^2(A)$, *where* $\sigma_k(A)$ *is the* k*-th singular value of* A.
- *(c) If* $A \in \mathbb{R}^{n\times n}$ *symmetric, then* $\|A\|_F^2 = \sum_{k=1}^{n} \lambda_k^2(A)$, *where* $\lambda_k(A)$ *is the* k*-th eigenvalue of* A.

Problem 1.2 *Given $A \in \mathbb{R}^{n\times m}$ show:*

(a) $\|A\|_s = \max_{1\le k\le \min\{n,m\}} \sigma_k(A)$, *where* $\sigma_k(A)$ *is the* k*-th singular value of* A.

(b) If $A \in \mathbb{R}^{n\times n}$ *symmetric, then* $\|A\|_s = \max_{1\le k\le n} |\lambda_k(A)|$, *where* $\lambda_k(A)$ *is the* k*-th eigenvalue of* A.

(c) For $A \in \mathbb{R}^{n\times n}$, *we have*

$$\|A\|_s \le \|A\|_F \le \sqrt{n}\|A\|_s.$$

Problem 1.3 *In class we saw linear regression on one variable, but in practice we may believe that a certain parameter y is a linear function of k different variables $x^1, \dots, x^k$, not just 1.*

Say you believe that the average rating of an ice cream is a linear function of: how much sugar the ice cream has, how much fat the ice cream has, and the price of the ice cream. You decide to investigate n different ice creams and collect their average rating, sugar content, fat content, and price. (a) How would you find the linear coefficients? (write the least squares problems and describe well what each matrix is, and which values it has). (b) Let's say you collect only data about 2 ice creams and see a perfect fit on your linear model, what can you deduct from this?

Problem 1.4 *Consider $x_1, \dots, x_n \in \mathbb{R}^p$. The variance of these dates points is*

$$\Gamma = \min_{\mu\in\mathbb{R}^p} \sum_{k=1}^{n} \|x_k - \mu\|^2 .$$

- *(a) How much variance of the points is it possible to capture on a d dimensional projection? (relate to eigenvalues of a certain matrix, justify your answer)*
- *(b) How much is the smallest possible sum of squared distances from these points to a d-dimensional approximation?*
- *(c) When is it that all of the variance is captured on a d dimensional projection?*
- *(d) When is it that there exists a d-dimensional approxiamtion that is exact?*

$(*)$ **Problem 1.5 (For Extra Credit)** *Show the resut we used in class: If* $M \in \mathbb{R}^{n \times n}$ *is a symmetric matrix and* $d \leq n$ *then*

$$\max_{\substack{U \in \mathbb{R}^{n \times d} \\ U^T U = \mathrm{I}_{d \times d}}} \mathrm{Tr}\left(U^T M U\right) = \sum_{k=1}^{d} \lambda_k^{(+)}(M),$$

where $\lambda_k^{(+)}$ *is the largest* k*-th eigenvalue of* M*.*