DS-GA 3001.03: Homework Problem Set 9

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 November 29, 2016

This homework problem set is due on November 29, 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.

Problem 9.1 *Let* $A \in \mathbb{R}^{n \to m}$*, show that the function* $g : \mathbb{R}^m \to \mathbb{R}^n$ *given by*

$$g(x) = Ax.$$

Show that g *is Lipschitz with constant* $L = \sigma_{\max}(A)$*.*

Problem 9.2 *Given* $f\mathbb{R}^n \to \mathbb{R}$ *convex and differentiable and* $x \in \mathbb{R}^n$*, show that for all* $\alpha > 0$

$$f(x - \alpha \nabla f(x)) \geq f(x) - \alpha \|\nabla f(x)\|^2.$$

(you are welcome to use the fact that the linear approximation of a convex function is below it)

Problem 9.3 *Given* $f\mathbb{R}^n \to \mathbb{R}$ *convex and differentiable and* $x \in \mathbb{R}$*, show that there exists* $\alpha > 0$ *such that*

$$f(x - \alpha \nabla f(x)) \leq f(x) - \frac{1}{2}\alpha \|\nabla f(x)\|^2$$

Problem 9.4 *Given* $Q \in \mathbb{R}^{n \times n}$ *symmetric and positive definite,* $a \in \mathbb{R}^n$*, and* $c \in \mathbb{R}$ *what is the minimizer of*

$$f(x) = c + a^T x + \frac{1}{2} x^T Q x$$

in $\mathbb{R}^n$*?*

$(*)$ **Problem 9.5 (For Extra Credit)** *Suppose you want to represent* n *data points in* $\mathbb{R}^d$ *and all you are given is estimates for their Euclidean distances* $\delta_{ij} \approx \|x_i - x_j\|_2^2$*. Multiimensional scaling attempts to find an* d *dimensions that agrees, as much as possible, with these estimates. Organizing* $X = [x_1, \ldots, x_n]$ *and consider the matrix* Δ *whose entries are* δ_{ij}*.*

1. *Show that, if* $\delta_{ij} = \|x_i - x_j\|_2^2$ *then there is a choice of* x_i *(note that the solution is not unique, as a translation of the points will preserve the pairwise distances, e.g.) for which*

   $$X^T X = -\frac{1}{2} H \Delta H,$$

   where $H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$*.*

2. *If the goal is to find points in* $\mathbb{R}^d$*, how would you do it (keep part 1 of the question in mind)?*

(The procedure you have just derived is known as Multidimensional Scaling)

This motivates a way to embed a graph in d dimensions. Given two nodes we take δ_{ij} *to be the square of some natural distance on a graph such as, for example, the geodesic distance (the distance of the shortest path between the nodes) and then use the ideas above to find an embedding in* $\mathbb{R}^d$ *for which Euclidean distances most resemble geodesic distances on the graph. This is the motivation behind a dimension reduction technique called ISOMAP (J. B. Tenenbaum, V. de Silva, and J. C. Langford, Science 2000).*