DS-GA 1014: Extended Syllabus Lecture 4

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

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

September 25, 2018

These are not meant to be Lecture Notes. They are simply extended syllabi with the most important definitions and results from the lecture. As such, they lack the intuition and motivation and so they are not a good place to learn the material the first time, just to briefly review it. These extended syllabi will also have references.

There are many amazing books about linear algebra and virtually all of them will contain the material for this particular lecture, examples include the book suggested for the course [2]. Another place you can read about some of these is the Lecture Notes for DSGA1002 [1].

Please let me know if you find any typos

- Given two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times p}$ we have
$$(AB)^T = B^T A^T.$$
- Given two matrices $A, B \in \mathbb{R}^{n\times n}$ both invertible (meaning that A^{-1} and B^{-1} exist) we have
$$(AB)^{-1} = B^{-1}A^{-1}.$$
- A matrix $L \in \mathbb{R}^{n\times n}$ is said to be symmetric if $L^T = L$. They arise very naturally in Data Science if, for example, entry i, j corresponds to some (symmetric) similarity measure (or, say, distance) between i and j.
- A matrix $L \in \mathbb{R}^{n\times n}$ is said to be anti-symmetric if $L^T = -L$.
- Systems of Linear Equations, matrix representation, Inverse matrix, Triangular Systems, Gaussian Elimination, LU factorization, See (for example) Chapter 2 of [2].
- Given a vector $x \in \mathbb{R}^n$ we define it'sthfd norm $\|x\|$ (distance to 0) as
$$\|x\|^2 = x^T x = \sum_{i=1}^n x_i^2.$$
- Given two vectors $x, y \in \mathbb{R}^n$ we define their (Euclidean) inner product as
$$\langle x, y\rangle = x^T y = \sum_{i=1}^n x_i y_i.$$

- Inner products can be defined in different ways. We say that $\langle x, y\rangle$ defines an inner product if the following holds

$$\begin{array}{rc} (i) & \forall_{x,y\in\mathbb{R}^n}\langle x, y\rangle = \langle y, x\rangle \\ (ii) & \forall_{x,y,z\in\mathbb{R}^n}\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle \\ (iii) & \forall_{x,y\in\mathbb{R}^n}\forall_{\alpha\in\mathbb{R}}\langle \alpha x, y\rangle = \alpha\langle x, y\rangle \\ (iv) & \forall_{x\in\mathbb{R}^n}\langle x, x,\rangle \geq 0 \quad \text{and} \quad \langle x, x,\rangle = 0 \text{ iff } x = 0 \end{array}$$

References

[1] Carlos Fernandez-Granda, *Lecture Notes of DSGA1002*, 2015 version available at `http://www.cims.nyu.edu/~cfgranda/pages/DSGA1002_fall15/notes.html`, 2015

[2] Gilbert Strang, *Introduction to Linear Algebra*, Fifth Edition, 2016