DS-GA 1014: Extended Syllabus Lecture 3

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

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

September 18, 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

- Kernel and Nullspace: Given a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ its Kernel (or Nullspace) $\mathrm{Ker}(L)$ is given by all vectors $v \in \mathbb{R}^m$ such that $L(v) = 0$. It is a subspace.

- Image and Column Space: Given a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ its Image (or Column space) $\mathrm{Im}(L)$ is given by all vectors $u \in \mathbb{R}^n$ for which there exist $v \in \mathbb{R}^m$ such that $L(v) = u$. It is a subspace. It is also the subpspace spanned by the columns of the matrix representation of L.

- (Part of) Fundamental Theorem of Linear Algebra: Given a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ we have $\dim(\mathrm{Ker}(L)) + \dim(\mathrm{Im}(L)) = m$.

- Given a matrix $L \in \mathbb{R}^{n\times m}$ (which means it is a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$) the rank of L is
$$\mathrm{rank}(L) = \dim(\mathrm{Im}(L)).$$

- A matrix $L \in \mathbb{R}^{n\times n}$ has an inverse L^{-1} if and only if $\mathrm{rank}(L) = n$.

- Rank is a very important concept in recommendation systems. As we will see later, data matrices naturally arising in recommendation systems tend to be (approximately) low rank, and this can be leveraged to make meaningful recommendations.

- Given a matrix $L \in \mathbb{R}^{n\times m}$ we define its transpose $L^T \in \mathbb{R}^{m\times n}$ as
$$\left(L^T\right)_{ij} = L_{ji}.$$

- Given a matrix $L \in \mathbb{R}^{n \times m}$ we have

$$\operatorname{rank}(L) = \operatorname{rank}(L^T).$$

- Given two matrices $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times p}$ we have

$$(AB)^T = B^T A^T.$$

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