DS-GA 3001.03: Extended Syllabus Lecture 5

Optimization and Computational Linear Algebra for Data Science (Fall 2016)

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

October 11, 2016

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 the first part of this particular lecture (examples include the book suggested for the course [3]). For the second half I have written lecture notes that I reference here, Section 1.1-1.1.4 of these notes [1]

Please let me know if you find any typos!

- Given $A \in \mathbb{R}^{n\times m}$, the Frobenius norm of A, $\|A\|_F$ is defined as
$$\|A\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}^2.$$
- (shown in the homework): Given $A \in \mathbb{R}^{n\times m}$, $\|A\|_F^2 = \mathrm{Tr}(A^TA) = \mathrm{Tr}(AA^T)$.
- (shown in the homework): Given $A \in \mathbb{R}^{n\times m}$, $\|A\|_F^2 = \sum_{k=1}^{\min\{n,m\}} \sigma_k^2(A)$, where $\sigma_k(A)$ is the k-th singular value of A.
- (shown in the homework): 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.
- Given $A \in \mathbb{R}^{n\times m}$ the operator (or spectral) norm of A, $\|A\|_s$ is defined as
$$\|A\|_s = \max_{x\in\mathbb{R}^m\setminus\{0\}:\ \|x\|_2=1} \frac{\|Ax\|}{\|x\|}.$$
- (shown in the homework): Given $A \in \mathbb{R}^{n\times m}$, $\|A\|_s^2 = \max_{1\le k\le \min\{n,m\}} \sigma_k(A)^2$, where $\sigma_k(A)$ is the k-th singular value of A.
- (shown in the homework): If $A \in \mathbb{R}^{n\times n}$ symmetric, then $\|A\|_s^2 = \max_{1\le k\le n} |\lambda_k(A)|^2$, where $\lambda_k(A)$ is the k-th eigenvalue of A.

- (shown in the homework): For $A \in \mathbb{R}^{n \times n}$, we have

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

- Given $A \in \mathbb{R}^{n \times n}$ invertible, the condition number of A, $\kappa(A)$, is defined as

$$\kappa(A) = \|A\|_s \|A^{-1}\|_s.$$

- The condition number measures sensability to errors on the input, when solving linear systems.
- Given $A \in \mathbb{R}^{n \times n}$ invertible, we have

$$\kappa(A) = \frac{\max_k \sigma_k(A)}{\min_j \sigma_j(A)}$$

- Given $A \in \mathbb{R}^{n \times n}$ invertible and symmetric, we have

$$\kappa(A) = \frac{\max_k |\lambda_k(A)|}{\min_j |\lambda_k(A)|}$$

- Principal Component Analysis: I have written lecture notes for this part, see Section 1.1-1.1.4 in [1].

References

[1] Afonso S. Bandeira, *Ten Lectures and Forty-Two Open Problems in the Mathematics of Data Science*, available at `http://www.cims.nyu.edu/~bandeira/TenLecturesFortyTwoProblems.pdf`

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

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