DS-GA 3001.03: Extended Syllabus Lecture 4

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

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

October 4, 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 this particular lecture, examples include the book suggested for the course [2]. Another place you can read about some of this is the Lecture Notes from last years DSGA1002 [1].

Please let me know if you find any typos!

- Given a matrix $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^n$ satistying
$$Av = \lambda v,$$
are called, respectively, an eigenvalue and an eigenvector. In this case, λ is the eigenvalue corresponding/associated to v and v is an eigenvector corresponding/associated to λ.

- If $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix, the maximizer of $\max_{v\in\mathbb{R}^n:\ \|v\|=1} \frac{v^TAv}{v^Tv}$ satisfying $Av = \lambda v$ for some $\lambda \in \mathbb{R}$ (was Homework).

- This implies: If $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix, then there exists $v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ such that $Av = \lambda v$. (was proved in the Homework).

- If $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $Av_1 = \lambda_2 v_1$ and $Av_2 = \lambda_2 v_2$ with $\lambda_1 \neq \lambda_2$ then $v_1^T v_2 = 0$. Proof: $\lambda_1 v_1^T v_2 = (Av_1)^T v_2 = v_1^T (A^T v_2) = v_1^T (Av_2) = \lambda_2 v_1^T v_2$.

- If λ is an eigenvalue of A then λ^2 is an eigenvalue of A^2. Proof: Take v such that $Av = \lambda v$, then $A^2 v = A(Av) = A(\lambda v) = \lambda(Av) = \lambda^2 v$.

- If $A \in \mathbb{R}^{n\times n}$ symmetric has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ then the corresponding n eigenvectors form an orthogonal basis.

- Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable in the sense $V^TAV = \Lambda$ where V satisfies $V^TV = I$ and Λ is a diagonal matrix. The columns of V are the eigenvectors of A and the diagonal entries of Λ the associated eigenvalues. The proof shown in the lecture is available in [2].

- Another way to write the point above: Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ has a spectral decomposition:
$$A = V\Lambda V^T,$$
where $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ satisfies $V^TV = I$ and $\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \lambda_n \end{bmatrix}$.

- Yet another way to write this: Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ can be written as
$$A = \sum_{k=1}^{n} \lambda_k v_k v_k^T.$$

- Notice that, in general, some of the eigenvalues of A may be repeated. In that case, the number of eigenvalues associated with a certain value λ can be defined as: 1) the number of times λ shows up in the diagonal of the spectral decomposition, or 2) the dimension of the subspace of vectors $v \in \mathbb{R}^n$ satisfying $Av = \lambda v$.

- If $A \in \mathbb{R}^{n\times n}$ is symmetric, the rank of$A$ is equal to the number of non-zero eigenvalues of A. The kernel of A is the subspace spanned by the eigenvectors of A associated to an eigenvalue 0.

- A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is said to be positive semidefinite ($A \succeq 0$) if all its eigenvalues are non-negative.

- Attention: Not every matrix has real eigenvalues, the example in class was:
$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$
.

- Non-symmetric (even non-square) matrices also have a decomposition, the **Singular Value Decomposition**.[1] Given a matrix $B \in \mathbb{R}^{n\times m}$ its Singular Value Decomposition, which always exists, is given by
$$B = U\Sigma V^T,$$
where $U \in \mathbb{R}^{n\times n}$ satisfies $U^TU = I$, $V \in \mathbb{R}^{n\times n}$ satisfies $V^TV = I$, and $\Sigma \in \mathbb{R}^{n\times m}$ is a diagonal matrix (in the sense that $\Sigma_{ij} = 0$ for $i \neq j$). The diagonal entries $\sigma_1, \ldots, \sigma_n$ of Σ are called singular values.

- The rank of B is the number of non-zero singular values of B.

- The Singular Value Decomposition is very useful for a variety of things such as: Data Compression, Low rank decompositions (for, for example, Recommendation Systems), Dimension Reduction of Data, etc. We'll some of these uses during the course.

[1] This is one of the most useful constructs in Linear Algebra

References

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

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