# DS-GA 1014: Extended Syllabus Lecture 5

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

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

October 2 and 3, 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 a matrix $L \in \mathbb{R}^{n\times m}$, the matrix $L^T \in \mathbb{R}^{m\times n}$ is the matrix such that, for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ we have:
$$\langle x, Ly\rangle = \langle L^T x, y\rangle$$
- Interesting example: If we take the vector space of smooth functions from $[0,1]$ to $\mathbb{R}$ that satisfy $f(0) = 1$ and $f(1) = 1$ and define the inner product
$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt,$$
then integration by parts shows that the derivative (as a linear transformation) is anti-symmetric. You can see example in, for example, Section 2.7. of [2].
- Cauchy-Schwartz inequality: $\forall_{x,y\in\mathbb{R}^n}$ we have
$$|x^T y| \leq \|x\|\|y\|.$$
To prove this, notice that $\left\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\right\|^2 \geq 0$ and $\left\|\frac{x}{\|x\|} + \frac{y}{\|y\|}\right\|^2 \geq 0$ and expand the norm square of the sum...
- Triangular inequality: $\|x + y\| \leq \|x\| + \|y\|$
- The Projection of $y$ onto span of $x$ is given by
$$P_{\langle x\rangle}(y) = \frac{x^T y}{\|x\|^2}x.$$

- If $\theta$ is the angle between $x$ and $y$ we have

$$\cos(\theta) = \frac{x^T y}{\|x\|\|y\|}$$

Note that, since $-1 \leq \cos(\theta) \leq 1$, this is an alternative proof of the Cauchy-Schwartz inequality.

- If $x^T y = 0$ then the angle between them is $\frac{\pi}{2}$ and we say that the vectors are orthogonal. The observation that in that case $\|x - y\|^2 = \|x\|^2 + \|y\|^2$ is Pythagoras theorem.

- More generally, given $y \in \mathbb{R}^n$ and a subspace $U \subset \mathbb{R}^n$ the projection of $y$ to $U$, $P_U(y)$, is given by the vector $u \in U$ that minimizes the distance to $y$, $\|u - y\|$.

- An orthogonal basis of a vector space is a set of vectors that span the vector space and are pairwise orthogonal. If, moreover, each vector has norm 1 it is called an orthonormal basis.

- Given an orthonormal basis $v_1, \ldots, v_m \in \mathbb{R}^n$ of a subspace $U$, the projection $P_U(y)$ of a vector $y \in \mathbb{R}^n$ to $U$ is easily writable as a linear combination of the basis as $P_U(y) = \langle y, v_1 \rangle v_1 + \cdots + \langle y, v_n \rangle v_n$, where $\langle y, v_k \rangle = y_k^v$ denotes the Euclidean inner product.

- If $v_1, \ldots, v_n \in \mathbb{R}^n$ are an orthonormal basis of $\mathbb{R}^n$ then $u \in \mathbb{R}^n$ is easily writable as a linear combination of the basis as $u = \langle u, v_1 \rangle v_1 + \cdots + \langle u, v_n \rangle v_n$, where $\langle u, v_k \rangle = u_k^v$ denotes the Euclidean inner product.

# 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