DS-GA 3001.03: Extended Syllabus Lecture 1

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

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

September 13, 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 these is the Lecture Notes from last years DSGA1002 [1].

Please let me know if you find any typos

- Span: Given vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ the span of $v_1, \ldots, v_m$ is the set of the vectors that can be written as linear combinations of these:

$$\begin{aligned} \operatorname{Span}(v_1, \ldots, v_m) &= \langle v_1, \ldots, v_m \rangle \\ &= \left\{ v \in \mathbb{R}^n : v = \alpha_1 v_1 + \cdots + \alpha_m v_m \text{ for some } \alpha_1, \ldots, \alpha_m \in \mathbb{R} \right\}. \end{aligned}$$

- Linear dependency: vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ are linear dependent if there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ not all zero such that $v \in \mathbb{R}^n : v = \alpha_1 v_1 + \cdots + \alpha_m v_m$. Otherwise, they are linearly independent.

- Basis: A basis for a vector space is a set of vectors that span the whole vector space (say $\mathbb{R}^n$) and that is linearly independent.

- Dimension: The dimension of a vector space V is the number of vectors in a basis of that vector space (the dimension of $\mathbb{R}^n$ is n), we call this $\dim(V)$.

- Axiomatic definition of vector space: see, for example, here [1]

- Subspace: Given a vector space V, a subspace S is a subset $S \subset V$ such that: (i) for all $v \in S$ and $\alpha \in \mathbb{R}$ we have $\alpha v \in S$, and (ii) for all $v_1, v_2 \in S$ we have $v_1 + v_2 \in S$.

- Linear Transformation: Given two vectors spaces (let us focus in $\mathbb{R}^m$ and $\mathbb{R}^n$ for now) a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ is a function that satisfies: (i) for all $v \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$ we have $L(\alpha v) = \alpha L(v)$ and (ii) for all $v_1, v_2 \in \mathbb{R}^m$ we have $L(v_1 + v_2) = L(v_1) + L(v_2)$.

- Matrix Representation: Let $e_1, \ldots, e_m$ denote the canonical basis of $\mathbb{R}^m$ and $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Given a vector $x \in \mathbb{R}^m$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = x_1 e_1 + \cdots + x_m e_m, \text{ we have } L(x) = x_1 L(e_1) + \cdots + x_n L(e_n).$$

  This means that every linear transformation is characterized by $L(e_1), \ldots, L(e_m)$ (or the image through L of any basis). For this reason we neatly represent L by a $n \times m$ matrix

$$L = \begin{bmatrix} | & | & & | \\ L(e_1) & L(e_2) & \cdots & L(e_m) \\ | & | & & | \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \cdots & L_{nm} \end{bmatrix}.$$

  Note that we write $L(e_k) = \begin{bmatrix} L_{1k} \\ \vdots \\ L_{nk} \end{bmatrix}$.

- Matrix Vector Product: In this representation, given an $n \times m$ matrix L and $x \in \mathbb{R}^m$, matrix vector product is simply $Lx = L(x)$.

- Matrix Product: Given two linear transformations $L : \mathbb{R}^m \to \mathbb{R}^n$ and $U\mathbb{R}^p \to \mathbb{R}^m$, the product LU of the matrix L and U is simply the matrix representation of the linear transformation $L \circ U : \mathbb{R}^p \to \mathbb{R}^n$, which is a $n \times p$ matrix.

- Identity Matrix: The $n \times n$ identity matrix is the matrix representation of the linear transformation $I : \mathbb{R}^n \to \mathbb{R}^n$ such that $L(x) = x$ for all $x \in \mathbb{R}^n$. It is a diagonal matrix given by

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

- Kernel and Nullspace: Given a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ its Kernel (or Nullspace) $\text{Ker}(L)$ 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) $\text{Im}(L)$ 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 $\dim(\text{Ker}(L)) + \dim(\text{Im}(L)) = m$.

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