Optimization and Computational Linear Algebra for Data Science
Lecture 2: Linear transformations

Léo Miolane · leo.miolane@gmail.com

April 29, 2019

Warning: *This material is not meant to be lecture notes. It only gathers the main concepts and results from the lecture, without any additional explanation, motivation, examples, figures...*

1 Linear transformations

Definition 1.1 (*Linear transformation*)

A function $L : \mathbb{R}^m \to \mathbb{R}^n$ is linear if

(i) for all $v \in \mathbb{R}^m$ and all $\alpha \in \mathbb{R}$ we have $L(\alpha v) = \alpha L(v)$ and

(ii) for all $v, w \in \mathbb{R}^m$ we have $L(v + w) = L(v) + L(w)$.

Notice that $L : \mathbb{R}^m \to \mathbb{R}^n$ is linear if and only if $L(\alpha v + w) = \alpha L(v) + L(w)$ for all $v, w \in \mathbb{R}^m$ and all $\alpha \in \mathbb{R}$.

Proposition 1.1

The set $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ of all linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$ is a vector space.

Proposition 1.2

If $L : \mathbb{R}^m \to \mathbb{R}^n$ and $M : \mathbb{R}^n \to \mathbb{R}^k$ are two linear transformations, then the composite function $M \circ L : \mathbb{R}^m \to \mathbb{R}^k$ is also linear.

Theorem 1.1 (*Equality on a basis implies equality everywhere*)

Let L and M be two linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^m$. Let $(v_1, \dots, v_m)$ be a basis of $\mathbb{R}^m$ and supose that for all $i \in \{1, \dots, m\}$ we have

$$L(v_i) = M(v_i).$$

Then $L = M$, i.e. $L(v) = M(v)$ for all $v \in \mathbb{R}^m$.

2 Matrix representation

From Theorem 1.1 we that a linear transformation $L : \mathbb{R}^m \to \mathbb{R}^n$ is uniquely characterized by the image $L(v_1), \dots, L(v_m)$ of any basis $(v_1, \dots, v_m)$ of the input space.

We consider the canonical basis $(e_1, \dots, e_m)$ of $\mathbb{R}^m$ and encode L by a $n \times m$ matrix (that we will write also L) whose colums are $L(e_1), \dots, L(e_m)$:

$$L = \begin{pmatrix} | & | & & | \\ L(e_1) & L(e_2) & \cdots & L(e_m) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,m} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,m} \end{pmatrix}$$

where we write $L(e_j) = \begin{pmatrix} L_{1,j} \\ L_{2,j} \\ \vdots \\ L_{n,j} \end{pmatrix}$. The matrix L is called the (canonical) matrix of the linear transformation L.

Definition 2.1 (*Matrix product*)

Let $L \in \mathbb{R}^{n\times m}$ and $M \in \mathbb{R}^{k\times n}$. The product ML is the $k \times m$ matrix defined by

$$(ML)_{i,j} = \sum_{r=1}^{n} M_{i,r} L_{r,j} \quad \textit{for all} \quad 1 \leq i \leq k, \quad 1 \leq j \leq m.$$

Proposition 2.1 (*Matrix product means composition of linear transformations*)