Session 2: Linear transformations and matrices

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

Questions?

Linear maps & matrices

Two sides of the same coin

Linear map

$$L : \mathbb{R}^m \to \mathbb{R}^n$$

Matrix

$$L \in \mathbb{R}^{n \times m}$$

Two sides of the same coin

Linear map

$$L : \mathbb{R}^m \to \mathbb{R}^n$$

Matrix

$$L \in \mathbb{R}^{n \times m}$$

Rotations in $\mathbb{R}^2$

Let $\theta \in \mathbb{R}$. The rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ of angle θ about the origin is linear.

Exercise: what is the canonical matrix of R_θ ?

Operations on matrices

Addition and scalar multiplication

- Sum of two matrices of the **same** dimensions:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,m}+b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1}+b_{n,1} & \cdots & a_{n,m}+b_{n,m} \end{pmatrix}$$

- Multiplication by a scalar λ:

$$\lambda \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,m} \\ \vdots & \ddots & \vdots \\ \lambda a_{n,1} & \cdots & \lambda a_{n,m} \end{pmatrix}$$

A new vector space!

Proposition

- $\mathbb{R}^{n\times m}$ is a vector space.
- $\dim(\mathbb{R}^{n\times m}) =$

Proof.

$\square$

Product of two matrices

Warning:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} \times \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix} \neq \begin{pmatrix} a_{1,1} \times b_{1,1} & \cdots & a_{1,m} \times b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} \times b_{n,1} & \cdots & a_{n,m} \times b_{n,m} \end{pmatrix}$$

Matrix product

Let $L \in \mathbb{R}^{n \times m}$ and $M \in \mathbb{R}^{m \times k}$.

Definition (Matrix product)

The matrix product LM is the $n \times k$ matrix of the linear map $L \circ M$.

Matrix product

Theorem

Let $L \in \mathbb{R}^{n\times m}$ and $M \in \mathbb{R}^{m\times k}$.
The entries matrix product LM are given by

$$(LM)_{i,j} = \sum_{\ell=1}^{m} L_{i,\ell} M_{\ell,j}, \qquad \text{for } 1 \le i \le n \text{ and } 1 \le j \le k.$$

Proof

Rotations in $\mathbb{R}^2$

The R_a and R_b denote respectively the matrices of the rotations of angles a and b about the origin, in $\mathbb{R}^2$.

Exercise: Compute the product $R_a R_b$.

Matrix product properties

Can we divide two matrices ?

Invertible matrices

Definition (Matrix inverse)

A **square** matrix $M \in \mathbb{R}^{n \times n}$ is called *invertible* if there exists a matrix $M^{-1} \in \mathbb{R}^{n \times n}$ such that

$$MM^{-1} = M^{-1}M = \mathrm{Id}_n.$$

Such matrix M^{-1} is unique and is called the *inverse* of M.

Exercise: Let $A, B \in \mathbb{R}^{n \times n}$. Show that if $AB = \mathrm{Id}_n$ then $BA = \mathrm{Id}_n$.

Kernel and image

Definitions

Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation.

Definition (Kernel)

The kernel $\mathrm{Ker}(L)$ (or nullspace) of L is defined as the set of all vectors $v \in \mathbb{R}^m$ such that $L(v) = 0$, i.e.

$$\mathrm{Ker}(L) \stackrel{\text{def}}{=} \{v \in \mathbb{R}^m \mid L(v) = 0\}.$$

Definition (Image)

The image $\mathrm{Im}(L)$ (or column space) of L is defined as the set of all vectors $u \in \mathbb{R}^n$ such that there exists $v \in \mathbb{R}^m$ such that $L(v) = u$.

Picture

Remarks

Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation.

Proposition

- $\mathrm{Ker}(L)$ is a subspace of $\mathbb{R}^m$.
- $\mathrm{Im}(L)$ is a subspace of $\mathbb{R}^n$.

Remark: $\mathrm{Im}(L)$ is also the Span of the columns of the matrix representation of L.

Example: orthogonal projection

Consider $L : \mathbb{R}^2 \to \mathbb{R}^2$ to be the orthogonal projection onto the x-axis.

Why do we care about this ?

Linear systems

Assume that we given a dataset:

$$a_i = (a_{i,1}, \dots, a_{i,m}) \in \mathbb{R}^m, \quad y_i \in \mathbb{R} \qquad \text{for} \quad i = 1, \dots, n.$$

We would like to find $x \in \mathbb{R}^m$ such that

$$x_1 a_{i,1} + \cdots + x_m a_{i,m} = y_i \qquad \text{for all} \quad i \in \{1, \dots, n\}.$$

Matrix notation

Let's write

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} \in \mathbb{R}^{n\times m} \qquad \text{and} \qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n.$$

Let's find all solutions !

Conclusion: 3 possible cases

1. $y \notin \mathrm{Im}(A)$: there is no solution to $Ax = y$.

2. $y \in \mathrm{Im}(A)$, then there exists $x_0 \in \mathbb{R}^m$ such that $Ax_0 = y$. The set of solutions in then

$$S = \left\{ x_0 + v \,\middle|\, v \in \mathrm{Ker}(A) \right\}.$$

   - If $\mathrm{Ker}(A) = \{0\}$, then $S = \{x_0\}$: x_0 is the unique solution.
   - If $\mathrm{Ker}(A) \neq \{0\}$, then $\mathrm{Ker}(A)$ contains infinitely many vectors: there are infinitely many solutions.

Gaussian elimination

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{n\times m} \qquad \text{and} \qquad y = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \in \mathbb{R}^n.$$

Gaussian elimination

Gaussian elimination

Gaussian elimination

Questions?

Questions?

Questions?