# Session 2: Linear Transformations & Matrices

Optimization and Computational Linear Algebra for Data Science

Marylou Gabrié (based on material by Léo Miolane)

#### **Contents**

- 1. Linear Transformations
  - 1.1 Definition
  - 1.2 Properties
- Matrices
  - 2.1 Linear maps & matrices definition
  - 2.2 Matrix-vector product
  - 2.3 Addition & scalar multiplication
  - 2.4 Matrix product
  - 2.5 Invertible matrices
- 3. Kernels and Images
- 4. Why do we care about all these things?

  Solving linear systems
  - 4.1 Matrix notation of linear systems
  - 4.2 Gaussian elimination

Refs: Strang

# 1. Linear Transformations

1. Linear Transformations 3/42

### **Examples**

You already know some linear transformations from high-school!

Symmetry Rotation

1. Linear Transformations 4/42

#### 1.1 Definition

Symmetries (about a line passing through the origin) and rotations (about the origin) are mappings

$$L: \mathbb{R}^2 \to \mathbb{R}^2$$

$$v \mapsto L(v),$$

that are "linear":

#### **Definition**

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

- 1. for all  $v, w \in \mathbb{R}^m$  we have L(v+w) = L(v) + L(w) and
- 2. for all  $v \in \mathbb{R}^m$  and all  $\alpha \in \mathbb{R}$  we have  $L(\alpha v) = \alpha L(v)$ .

#### An example

### An example of a non-linear map

The function

 $F: \mathbb{R} \to \mathbb{R}$ 

is **not** linear.

### 1.2 Properties: Basic properties

#### Proposition

If  $L: \mathbb{R}^m \to \mathbb{R}^n$  is linear, then

- L(0) = 0.
- $L\Big(\sum_{i=1}^k \alpha_i v_i\Big) = \sum_{i=1}^k \alpha_i L(v_i), \text{ for all } \alpha_i \in \mathbb{R}, v_i \in \mathbb{R}^m.$

Proof.

### 1.2 Properties: Composition

#### Proposition

If  $L:\mathbb{R}^m\to\mathbb{R}^n$  and  $M:\mathbb{R}^n\to\mathbb{R}^k$  are both linear, then the composite function

$$M \circ L : \mathbb{R}^m \to \mathbb{R}^k$$
  
 $v \mapsto M(L(v))$ 

is also linear.

Proof.

# **Questions?**

1. Linear Transformations 1.2 Properties of linear transformations

10/42

# **Questions?**

1. Linear Transformations 1.2 Properties of linear transformations

10/42

# 2. Matrices

2. Matrices 11/42

#### 2.1 Linear Maps & Matrices Definition

#### The key observation:

- Let  $L: \mathbb{R}^m \to \mathbb{R}^n$  be a linear transformation.
- Let  $(e_1,\ldots,e_m)$  be the canonical basis of  $\mathbb{R}^m$ .

Then, for all  $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ :

$$L(x) = L\left(\sum_{i=1}^{m} x_i e_i\right) = \sum_{i=1}^{m} x_i L(e_i).$$

#### 2.1 Linear Maps & Matrices Definition

#### The key observation:

- Let  $L: \mathbb{R}^m \to \mathbb{R}^n$  be a linear transformation.
- Let  $(e_1,\ldots,e_m)$  be the canonical basis of  $\mathbb{R}^m$ .

Then, for all  $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ :

$$L(x) = L(\sum_{i=1}^{m} x_i e_i) = \sum_{i=1}^{m} x_i L(e_i).$$

**Conclusion**: if you give me the vectors  $L(e_1), \ldots, L(e_m) \in \mathbb{R}^n$  then, I am able to compute L(x) for any  $x \in \mathbb{R}^m$ .

« One needs  $n \times m$  numbers to store the linear map L on a computer »

#### **Matrices**

#### **Definition**

A  $n \times m$  matrix is an array (of real numbers) with n rows and m columns. We denote by  $\mathbb{R}^{n \times m}$  the set of all  $n \times m$  matrices.

### Canonical matrix of a linear map

We can encode a linear map  $L: \mathbb{R}^m \to \mathbb{R}^n$  by a  $n \times m$  matrix.

#### **Definition**

The canonical matrix of L is the  $n \times m$  matrix (that we will write also L) whose columns are  $L(e_1), \ldots, L(e_m)$ :

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

where we write 
$$L(e_j) = egin{pmatrix} L_{1,j} \\ L_{2,j} \\ \vdots \\ L_{n,j} \end{pmatrix}$$
 .

### 1.3 Matrix-vector product

Consider a linear map  $L:\mathbb{R}^m\to\mathbb{R}^n$  and its associated matrix  $\widetilde{L}\in\mathbb{R}^{n\times m}.$ 

**Question:** Can we use the matrix  $\widetilde{L}$  to compute the image L(x) of a vector  $x \in \mathbb{R}^m$  ?

#### Proposition

For all  $x \in \mathbb{R}^m$  we have

$$L(x) = \widetilde{L}x$$

where the "matrix-vector" product  $\widetilde{L}x \in \mathbb{R}^n$  is defined by

$$(\widetilde{L}x)_i = \sum_{j=1}^m \widetilde{L}_{i,j} x_j$$
 for all  $i \in \{1, \dots, n\}$ .

# Visualizing the formula

$$(\widetilde{L}x)_i = \sum_{j=1}^m \widetilde{L}_{i,j} x_j = \widetilde{L}_{i,1} x_1 + \widetilde{L}_{i,2} x_2 + \dots + \widetilde{L}_{i,m} x_m$$

# Why do we have $L(x) = \widetilde{L}x$ ?

### Example #1: identity matrix

The Identity map  $\begin{array}{ccc} \operatorname{Id}: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & x \end{array} \quad \text{is linear.}$ 

**Exercise**: what is the canonical matrix of Id?

### **Example #2: Homothety**

Let  $\lambda \in \mathbb{R}$ . The homothety map of ratio  $\lambda$ :

$$H_{\lambda}: \mathbb{R}^n \to \mathbb{R}^n$$
$$x \mapsto \lambda x$$

is linear.

**Exercise**: what is the canonical matrix of  $H_{\lambda}$ ?

# Example #3: rotations in $\mathbb{R}^2$

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

**Exercise**: what is the canonical matrix of  $R_{\theta}$ ?

### 2.3 Addition & 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  $\lambda$ :

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

### A new vector space!

#### Proposition

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

#### Proof.





#### Product of two matrices

#### Warning:

2. Matrices 2.4 Matrix product

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

$$\vdots$$
 $l_{n,m}$  )

$$\left\langle \left( egin{matrix} dots \ b_n, \end{matrix} 
ight.$$

$$b_n$$

$$\binom{a_1}{a_1}$$

$$\times b$$

23/42

### 2.4 Matrix product

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

#### **Definition - Proposition**

- The matrix product LM is the  $n \times k$  matrix of the linear map  $L \circ M$ .
- Its coefficients are given by the formula:

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

### Visualizing the formula

$$(LM)_{i,j} = \sum_{i=1}^{m} L_{i,\ell} M_{\ell,j} = L_{i,1} M_{1,j} + \dots + L_{i,m} M_{m,j}$$

# **Proof**

2. Matrices 2.4 Matrix product

26/42

# **Example:** 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  $R_aR_b$ .

### **Matrix product properties**

Let 
$$A, B \in \mathbb{R}^{n \times m}$$
 and  $C, D \in \mathbb{R}^{m \times k}$ ,

$$(A+B)C =$$

$$A(C+D) =$$

- Multiplication by the identity:  $A \operatorname{Id}_m =$
- Comutativity?

#### Can we divide two matrices?

For instance, if we have AB = AC, do we have B = C?

29/42

#### 2.5 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  $BA = \mathrm{Id}_n$ .

# 3. Kernel and image

3. Kernel and image 31/42

#### **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.

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

#### **Definition (Image)**

The image  $\operatorname{Im}(L)$  (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.

3. Kernel and image 32/42

# Picture

3. Kernel and image

33/42

#### Remarks

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

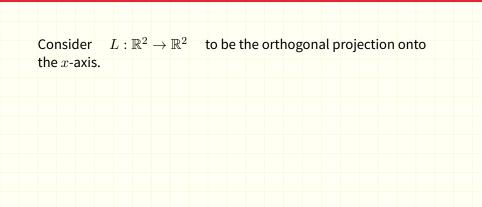
#### Proposition

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

**Remark:**  $\operatorname{Im}(L)$  is also the Span of the columns of the matrix representation of L (cf HW2).

3. Kernel and image 34/42

## **Example: orthogonal projection**



# 4. Why do we care about this?

#### **Linear systems**

Assume that we given a dataset:

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

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

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

#### 4.1 Matrix Notation of Linear Systems

$$/a_{1,1}$$
 ···  $a_1$ 

$$\cdots a_{1,m}$$

$$\in \mathbb{R}$$

$$n$$
.

$$A =$$

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

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



### Let's find all solutions!

39/42

4. Why do we care about this? 4.1 Matrix Notation of Linear Systems

#### **Conclusion: 3 possible cases**

- 1.  $y \notin \text{Im}(A)$ : there is no solution to Ax = y.
- 2.  $y \in \text{Im}(A)$ , then there exists  $x_0 \in \mathbb{R}^m$  such that  $Ax_0 = y$ . The set of solutions in then

$$S = \{x_0 + v \mid v \in \operatorname{Ker}(A)\}.$$

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

#### 4.2 Gaussian elimination

$$A=egin{pmatrix}1&-1&0&1\2&0&1&-1\-1&5&2&0\end{pmatrix}\in\mathbb{R}^{n imes m} \quad ext{ and }\quad y=egin{pmatrix}1\1\4\end{pmatrix}\in\mathbb{R}^n.$$

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{pmatrix}$$

$$=\begin{pmatrix}1\\1\\\end{pmatrix}$$

$$\in \mathbb{R}^n$$

### **Gaussian elimination**

42/42

4. Why do we care about this? 4.2 Gaussian elimination

### **Gaussian elimination**

42/42

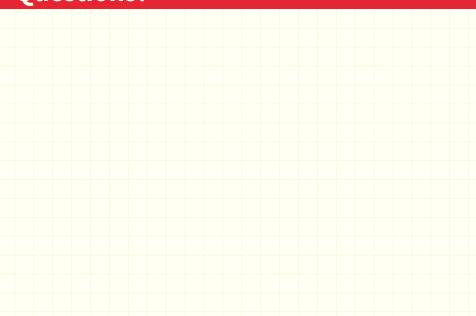
4. Why do we care about this? 4.2 Gaussian elimination

### **Gaussian elimination**

42/42

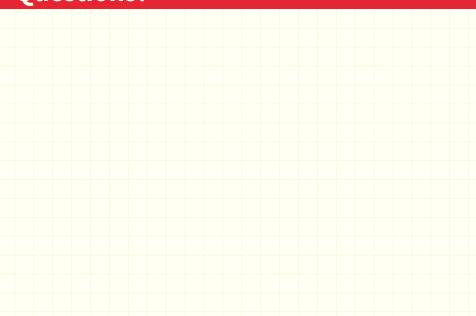
4. Why do we care about this? 4.2 Gaussian elimination

# **Questions?**



43/42

# **Questions?**



43/42

# **Questions?**

