Session 2: Linear Transformations & Matrices

Optimization and Computational Linear Algebra for Data Science

Marylou Gabrié (based on material by Léo Miolane)

Contents

Refs: Strang

1. Linear Transformations

Examples

You already know some linear transformations from high-school !

Symmetry

Rotation

1.1 Definition

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

$$\begin{array}{rccc} L: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & v & \mapsto & L(v), \end{array}$$

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

- $$\begin{array}{rccc} L: & \mathbb{R}^2 & \to & \mathbb{R}^3 \\ & (v_1, v_2) & \mapsto & (5v_1,\ 0,\ v_1 + v_2) \end{array}$$ is linear

An example of a non-linear map

The function $$\begin{array}{rccc} F: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & x^2 \end{array}$$ is **not** linear.

1.2 Properties: Basic properties

Proposition

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

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

Proof.

$\square$

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

$$\begin{array}{rccc} M \circ L : & \mathbb{R}^m & \to & \mathbb{R}^k \\ & v & \mapsto & M(L(v)) \end{array}$$

is also linear.

Proof.

$\square$

Questions?

Questions?

2. Matrices

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

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

2.1 Linear Maps & Matrices Definition

The key observation:

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

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

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

Conclusion: if you give me the vectors $L(e_1), \dots, L(e_m) \in \mathbb{R}^n$
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 $L(e_1), \dots, L(e_m)$:

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

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

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

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

Visualizing the formula

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

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

Example #1: identity matrix

The Identity map $$\begin{array}{rccc} \mathrm{Id}: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & x \end{array}$$ is linear.

Exercise: what is the canonical matrix of Id ?

Example #2: Homothety

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

$$\begin{array}{rccc} H_\lambda : & \mathbb{R}^n & \to & \mathbb{R}^n \\ & x & \mapsto & \lambda x \end{array}$$

is linear.

Exercise: what is the canonical matrix of H_λ ?

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 θ about the origin is linear.

Exercise: what is the canonical matrix of R_θ ?

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 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:

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

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:

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

Visualizing the formula

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

Proof

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 $R_a R_b$.

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: $A\,\mathrm{Id}_m =$
- Comutativity?

Can we divide two matrices ?

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

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

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 (cf HW2).

Example: orthogonal projection

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

4. Why do we care about this ?

Linear systems

Assume that we given a dataset:

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

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

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

4.1 Matrix Notation of Linear Systems

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.

4.2 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?