Session 1: Vector spaces

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

Logistics

The teaching team

- **Lecturer**: Léo Miolane – *lm4271nyu.edu*
  `leomiolane.github.io/linalg-for-ds.html`

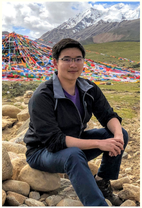
The teaching team

- **Lecturer**: Léo Miolane – *lm4271nyu.edu*

  `leomiolane.github.io/linalg-for-ds.html`

- **Sections leaders**:

Alex



In person

Irina

Remote

Carles

Remote

Course components

Three main components:

1. Videos

   2-3 short videos to watch **before** each lecture

2. Lectures

   Deepens the concepts introduced in the videos

3. Recitations

   Practice!

Course components

Three main components:

1. Videos

   2-3 short videos to watch **before** each lecture

2. Lectures

   Deepens the concepts introduced in the videos

3. Recitations

   Practice!

Grades:

1. Weekly quizzes (5%)
2. Weekly homeworks (40%)
3. Exams: Midterm (20%) + Final (35%)

Weekly timeline

Mon	Tue	Wed	Thu	Fri	Sat	Sun
14	15	16	17	18	19	20
21	22	23	24	25	26	27

Grading

- Quizzes have to be answered on **Gradescope**.
- Homeworks questions are available on the **course's webpage** and have to be submitted on **Gradescope**.

Grading

- Quizzes have to be answered on **Gradescope**.
- Homeworks questions are available on the **course's webpage** and have to be submitted on **Gradescope**.
- I encourage you to type your homeworks using LaTeX.

  Some instructions and template available on the course's webpage.
- Otherwise, you can scan (using dedicated app) your handwritten work. **It has to be legible!!!**

Grading

- Quizzes have to be answered on **Gradescope**.
- Homeworks questions are available on the **course's webpage** and have to be submitted on **Gradescope**.
- I encourage you to type your homeworks using LaTeX.

  Some instructions and template available on the course's webpage.
- Otherwise, you can scan (using dedicated app) your handwritten work. **It has to be legible!!!**
- Midterm ($\sim$ mid-October) and Final will be «take-home exams».
- Limited time: after downloading the Midterm/Final questions, you will have to upload your work within few hours.

Check out the syllabus on the course webpage!

Questions on logistics ?

Vector spaces and subspaces

Quick recap of video 1.2

A **vector space** is a set V endowed with two 'nice and compatible' operations $+$ and $\cdot$ that verify:

- For all $u, v \in V$, $u + v \in V$.
- For all $u \in V$ and all $\lambda \in \mathbb{R}$, $\lambda \cdot u \in V$.

Example: $V = \mathbb{R}^n$, with the usual vector addition $+$ and scalar multiplication $\cdot$ is a vector space.

Quick recap of video 1.2

A subset S of a vector space V is called a **subspace** if it is closed under addition and multiplication by a scalar.

Example: For all $v \in \mathbb{R}^n$,

$$\mathrm{Span}(v) = \{\lambda v \,|\, \lambda \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^n$.

Question ?

Review of Span and linear dependency

Span

The *linear span* of vectors $x_1, \ldots, x_k$ as the set of all linear combinations of these vectors.

Linear dependency

- Vectors $x_1, \dots x_k$ are *linearly dependent* if one of them can be expressed as a linear combination of the others.
- They are said to be *linearly independent* otherwise.

Abuse of language: Instead of saying «$x_1, \dots, x_k$ are linearly dependent», we should say «the family $(x_1, \dots, x_k)$ is linearly dependent».

Basis

A family $(x_1, \ldots, x_n)$ of vectors of V is a basis of V if

1. $x_1, \ldots, x_n$ are linearly independent,
2. $\mathrm{Span}(x_1, \ldots, x_n) = V$.

The dimension

A useful lemma

Lemma

Let $v_1, \ldots, v_n \in V$ and let $x_1, \ldots, x_k \in \mathrm{Span}(v_1, \ldots, v_n)$.
Then, if $k > n$, $x_1, \ldots, x_k$ are linearly dependent.

Definition of the dimension

Definition

We say that a vector space V has dimension n if it admits a basis $(v_1, \ldots, v_n)$ with n vectors.

The dimension is well defined!

Theorem

If V admits a basis $(v_1, \ldots, v_n)$, then every basis of V has also n vectors. We say that V has dimension n and write $\dim(V) = n$.

Proof.

$\square$

Properties of the dimension

Proposition

Let V be a vector space that has dimension $\dim(V) = n$. Then

- Any family of vectors of V that are linearly independent contains at most n vectors.

  i.e. if $x_1, \dots, x_k \in V$ are linearly independent, then $k \leq n$.

- Any family of vectors of V that spans V contains at least n vectors.

  i.e. if $x_1, \dots, x_k \in V$ are such that $\mathrm{Span}(x_1, \dots, x_k) = V$, then $k \geq n$.

Proof.

$\square$

Properties of the dimension

Proposition

Let V be a vector space that has dimension $\dim(V) = n$. Then

- Any family of vectors of V that are linearly independent contains at most n vectors.

  i.e. if $x_1, \dots, x_k \in V$ are linearly independent, then $k \leq n$.

- Any family of vectors of V that spans V contains at least n vectors.

  i.e. if $x_1, \dots, x_k \in V$ are such that $\mathrm{Span}(x_1, \dots, x_k) = V$, then $k \geq n$.

Proof.

$\square$

Properties of the dimension

Proposition

Let V be a vector space of dimension n and let $x_1, \dots, x_n \in V$.

1. If $x_1, \dots, x_n$ are linearly independent, then $(x_1, \dots, x_n)$ is a basis of V.
2. If $\mathrm{Span}(x_1, \dots, x_n) = V$, then $(x_1, \dots, x_n)$ is a basis of V.

Very useful to show that a family of vector forms a basis:

Example: $x_1 = (12, 37)$ and $x_2 = (-9, 17)$ form a basis of $\mathbb{R}^2$.

Proof of the Proposition. $\square$

An inequality

Proposition

Let U and V be two subspaces of $\mathbb{R}^n$. Assume that $U \subset V$. Then

$$\dim(U) \leq \dim(V) \leq n.$$

If **moreover** $\dim(U) = \dim(V)$, then $U = V$.

An inequality

Proposition

Let U and V be two subspaces of $\mathbb{R}^n$. Assume that $U \subset V$. Then

$$\dim(U) \leq \dim(V) \leq n.$$

If **moreover** $\dim(U) = \dim(V)$, then $U = V$.

Proof.

$\square$

A bit of vocabulary

Definition

Let S be a subspace of $\mathbb{R}^n$.

- We call S a *line* if $\dim(S) = 1$.
- We call S an *hyperplane* if $\dim(S) = n - 1$.

Coordinates

Coordinates of a vector in a basis

Definition

If $(v_1, \dots, v_n)$ is a basis of V, then for every $x \in V$ there exists a unique vector $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

We say that $(\alpha_1, \dots, \alpha_n)$ are the coordinates of x in the basis $(v_1, \dots, v_n)$.

Coordinates of a vector in a basis

Definition

If $(v_1, \dots, v_n)$ is a basis of V, then for every $x \in V$ there exists a unique vector $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

We say that $(\alpha_1, \dots, \alpha_n)$ are the coordinates of x in the basis $(v_1, \dots, v_n)$.

Proof.

$\square$

Exercise

1. Show that the vectors $v_1 = (1, 1)$ and $v_2 = (1, -1)$ form a basis of $\mathbb{R}^2$.
2. Express the coordinates of $u = (x, y)$ in the basis (v_1, v_2) in terms of x and y.

Exercise

1. Show that the vectors $v_1 = (1, 1)$ and $v_2 = (1, -1)$ form a basis of $\mathbb{R}^2$.
2. Express the coordinates of $u = (x, y)$ in the basis (v_1, v_2) in terms of x and y.

Why do we care about this ?

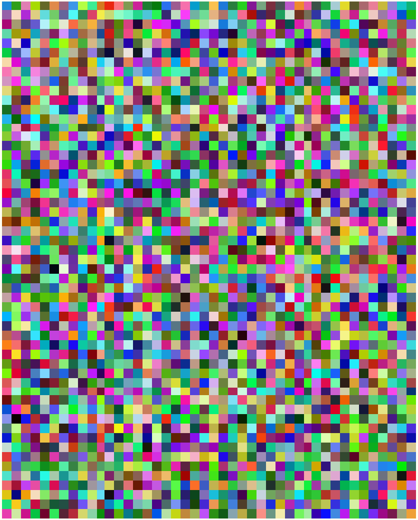
Application to image compression

- Image = Grid of pixels
- Represented as a vector $v \in \mathbb{R}^n$, for some large n.
- One need to store n numbers.

$$n = 44 \times 55 = 2420$$

Can we do better?

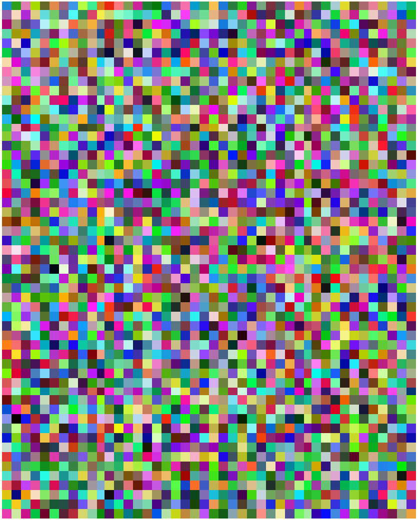
- If we want to store an arbitrary image, NO!



«Random» image

Can we do better?

- If we want to store an arbitrary image, NO!
- However, we are mainly storing images coming from the « real world »
- These images have some *structure*.



«Random» image

Can we do better?

- If we want to store an arbitrary image, NO!
- However, we are mainly storing images coming from the « real world »
- These images have some *structure*.

«Real» image

What do we mean by « structure » ?

Neighboring pixels are very likely to have similar colors.

What do we mean by « structure » ?

Neighboring pixels are very likely to have similar colors.

- There exists a basis $(w_1, \dots, w_n)$ of $\mathbb{R}^n$ in which «real» images $v \in \mathbb{R}^n$ are (approximately) **sparse**.
- This means that the coordinates $(\alpha_1, \dots, \alpha_n)$ of v in the basis $(w_1, \dots, w_n)$ contains a lot of zeros.

What do we mean by « structure » ?

Neighboring pixels are very likely to have similar colors.

- There exists a basis $(w_1, \dots, w_n)$ of $\mathbb{R}^n$ in which «real» images $v \in \mathbb{R}^n$ are (approximately) **sparse**.
- This means that the coordinates $(\alpha_1, \dots, \alpha_n)$ of v in the basis $(w_1, \dots, w_n)$ contains a lot of zeros.

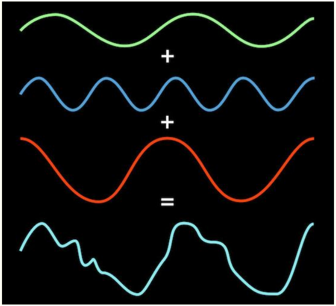
Store only the $k \ll n$ non-zero coordinates of v (in the w_i's basis') !

A toy example

Consider $n = 2$, that is images $v \in \mathbb{R}^2$ with only 2 pixels.

Examples of good bases

- **Fourier bases** (used in `.jpeg`, `.mp3`)



- `JPEG2000` uses **wavelet bases**, and achieves better performance than `JPEG`.
- In **Homework 4**, you will use wavelets to compress/denoise images.
- The course **DS-GA 1013** deepens these concepts!

Questions?