Session 1: Vector spaces

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

1. Recap of the videos
2. More about the dimension
3. Coordinates
4. Why do we care about all these things ?

   Application to data science: image compression

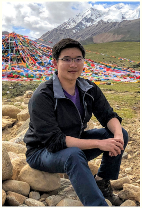
Logistics

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

Do the quiz!

Mon	Tue	Wed	Thu	Fri	Sat	Sun
14 Watch	15 the videos	16	17 Lecture	18 Homework	19	20
21	22 Recitations	23	24 Lecture	25	26	27 HW due date

Weekly Quizzes and Homeworks

- Quizzes have to be answered on **Gradescope**, after viewing the videos, but before the associated lecture.
- Homeworks questions are available on the **course's webpage** and have to be submitted on **Gradescope**.

Weekly Quizzes and Homeworks

- Quizzes have to be answered on **Gradescope**, after viewing the videos, but before the associated lecture.
- 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!!!**

Gradescope

DS-GA 1014 | Fall 2020

Entry Code: **M2ND83**

DESCRIPTION

Edit your course description on the Course Settings page.

THINGS TO DO

Review and publish grades for Quiz 1 now that you're all done grading.

ACTIVE ASSIGNMENTS	RELEASED	DUE (EDT)	SUBMISSIONS	% GRADED	PUBLISHED	REGRADES
Homework 1	SEP 02	SEP 20 AT 11:00PM	0	0%		ON
Quiz 2	SEP 03	SEP 10 AT 2:00PM	0	0%		ON
Quiz 1	AUG 23	SEP 10 AT 2:00PM	4	100%		ON

Midterm and Final

- **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!

Office hours + feedback

- I will have 2 office hours slots (+appointments):
  - One during New York 'standard hours'.
  - One early morning or late evening for students with a big time difference.
- Please fill the Google form with you preferences.
- Feedback, remarks about the lectures / videos / recitations / homeworks ... :
  - email me!
  - link for anonymous feedback on the course's website.

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.

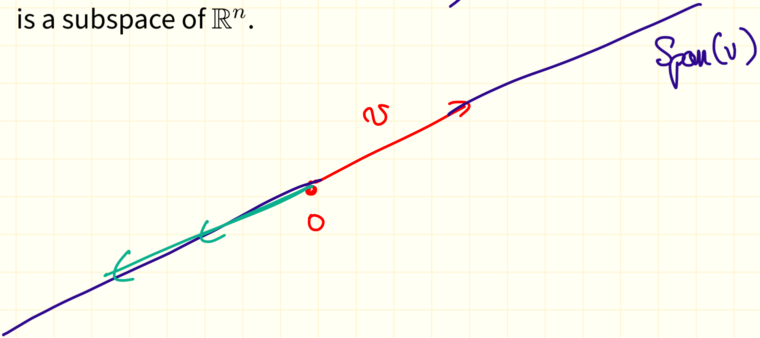
Quick recap of video 1.2

A non-empty 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 $\mathbb{R}^n$.



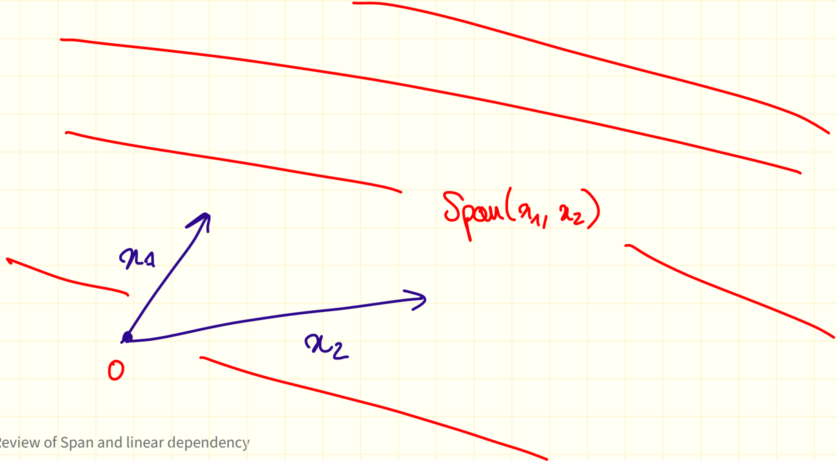
Remarks, questions ?

Remarks, questions ?

Review of Span and linear dependency

Span

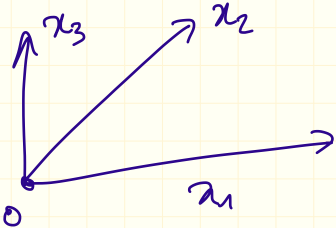
The *linear span* of vectors $x_1, \ldots, x_k$ as the set of all linear combinations of these vectors.



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».

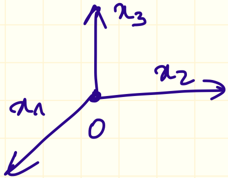


x_1, x_2, x_3 are lin dep
but x_1, x_2 are lin
independent.

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. $\mathrm{Span}(x_1, \ldots, x_n) = V$.



is a basis of $\mathbb{R}^3$

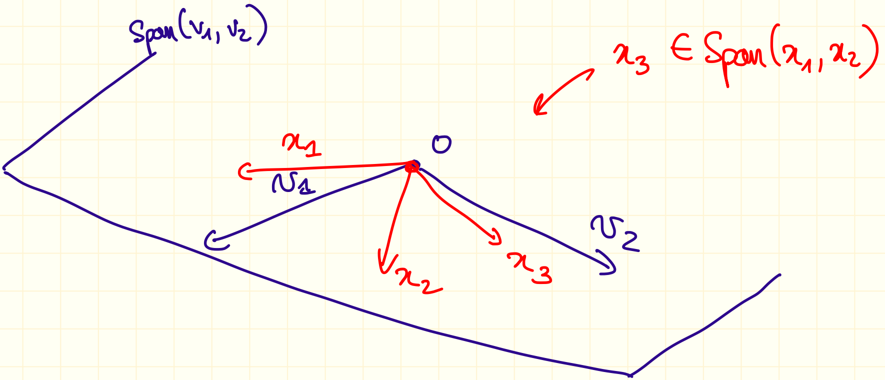
The dimension

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.

in $V = \mathbb{R}^3$

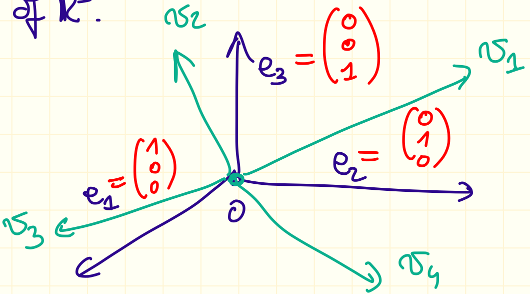


Definition of the dimension

Definition

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

$\mathbb{R}^3$ has dimension 3 because (e_1, e_2, e_3) is a basis of $\mathbb{R}^3$.



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.

Proof. By contradiction: let's assume that there exists two bases of V: $\left| \begin{array}{l} (v_1 \ldots v_n) \\ (x_1 \ldots x_k) \end{array} \right.$ with $k > n$

We have $x_1, \ldots x_k \in V = \mathrm{Span}(v_1 \ldots v_n)$

because $(v_1 - v_n)$ basis of V.

By "the lemma": we get that $x_1 \ldots x_k$ are lin. dependent.

CONTRADICTION!

$\rightarrow$ $(x_1 \ldots x_k)$ is supposed to be a basis

□

Properties of the dimension

Proposition

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

1. Any family of vectors of V that spans V contains at least n vectors.
2. Any family of vectors of V that are linearly independent contains at most n vectors.

Proof. ① Let consider $(u_1 \dots u_k)$ such that $\mathrm{Span}(u_1 \_ u_k) = V$. Then, if I consider a basis $(v_1 \dots v_n)$ of V I have:

$$\underline{v_1 \dots v_n} \in \underbrace{\mathrm{Span}(u_1 \dots u_k)}_{= V}$$

$v_1 \dots v_n$ are lin. indep: $n \leq k$. (from the lemma □)

Properties of the dimension

Proposition

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

1. Any family of vectors of V that spans V contains at least n vectors.
2. Any family of vectors of V that are linearly independent contains at most n vectors.

Proof. 2 Exercise

$\square$

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 $\mathbb{R}^2$.

x_1, x_2 are lin independent, since $\dim(\mathbb{R}^2) = 2$

I get that (x_1, x_2) is a basis of $\mathbb{R}^2$

An inequality

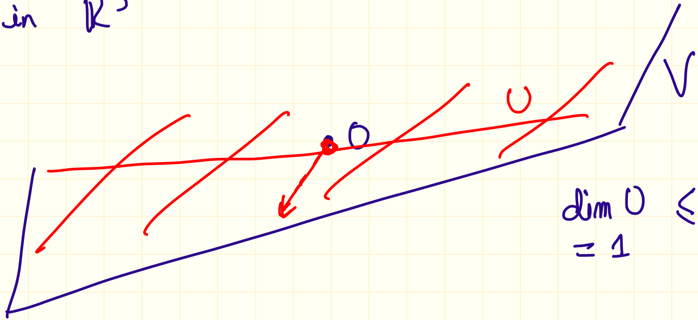
Proposition

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

$$\boxed{\dim(U) \leq \dim(V) \leq n.}$$

If **moreover** $\boxed{\dim(U) = \dim(V)\text{, then } U = V.}$

in $\mathbb{R}^3$



$$\underset{=1}{\dim U} \leq \underset{=2}{\dim V} \leq \underset{=3}{\dim \mathbb{R}^3}$$

Proof

Let $k = \dim U$

- Let $(u_1 \dots u_k)$ be a basis of U.

$u_1 \dots u_k$ are k lin. independent vectors that are inside U and therefore V

Hence (from Proposition 3 slides ago): $k \leq \dim V$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overset{\parallel}{\dim U}$.

- If moreover $\dim U = \dim V$

$u_1 \dots u_k$ are $k = \dim V$ lin. indep vectors of V $\longrightarrow$ $(u_1 \dots u_k)$ is a basis of V

Hence $U = \mathrm{Span}(u_1 \dots u_k) = V$

A bit of vocabulary

Definition

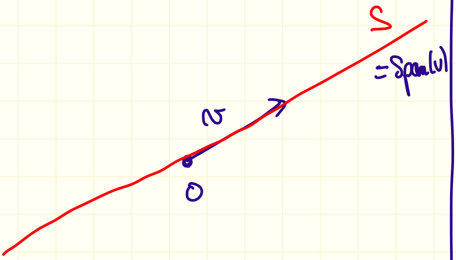
Let S be a subspace of $\mathbb{R}^n$.

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

$S = \{(x,y) \mid x+y=2\}$

$(0,0) \notin S$

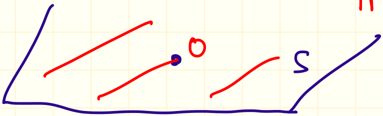
Line



Hyperplane.

in $\mathbb{R}^3$ $\dim(S) = 3-1$
$= 2$

hence S is an hyperplane



Coordinates

Coordinates of a vector in a basis

Definition & Theorem

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

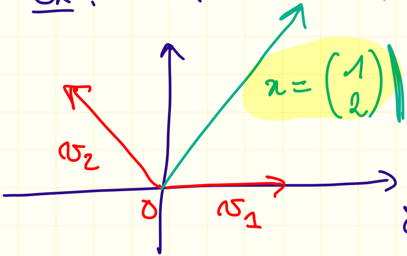
1

2

$$\boxed{x = \alpha_1 v_1 + \cdots + \alpha_n v_n.}$$

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

Ex: in $\mathbb{R}^2$ $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$



$x = 3v_1 + 2v_2$

therefore, the coordinates of x in the (v_1, v_2) basis are $(3, 2)$

Coordinates of a vector in a basis

Definition & Theorem

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

1

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

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

Proof. 1 Existence

$x \in V = \mathrm{Span}(v_1 \ldots v_n)$ (because $(v_1 \ldots v_n)$ basis of V)

hence there exists $\alpha_1 \ldots \alpha_n \in \mathbb{R}$ such that

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

Coordinates of a vector in a basis

2 Uniqueness.

Let $\beta_1 \dots \beta_n \in \mathbb{R}$ such that

$$x = \beta_1 v_1 + \dots + \beta_n v_n$$

We will show that $\beta_1 = \alpha_1, \dots, \beta_n = \alpha_n$.

$$0 = x - x = \underline{(\alpha_1 - \beta_1)} v_1 + \dots + \underline{(\alpha_n - \beta_n)} v_n$$

$v_1 \dots v_n$ are lin indep, from the def. we get $\alpha_1 - \beta_1 = 0 = \dots = \alpha_n - \beta_n$

Hence $\alpha_1 = \beta_1 \dots \alpha_n = \beta_n$.

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.

① v_1 and v_2 are lin independent, $\dim \mathbb{R}^2 = 2$

Hence (from the Prop.) (v_1, v_2) is a basis of $\mathbb{R}^2$

② We have to find $\alpha_1, \alpha_2 \in \mathbb{R}$

$$\begin{pmatrix} x \\ y \end{pmatrix} = u = \alpha_1 \underbrace{v_1}_{\begin{pmatrix} 1 \\ 1 \end{pmatrix}} + \alpha_2 \underbrace{v_2}_{\begin{pmatrix} 1 \\ -1 \end{pmatrix}} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_1 - \alpha_2 \end{pmatrix}$$

Exercise

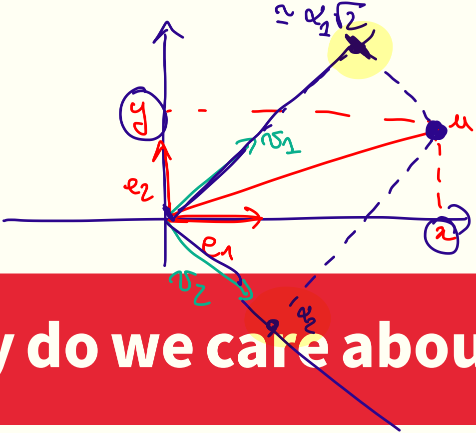
1. Show that the vectors $v_1 = (1, 1)$ and $v_2 = (1, -1)$ form a basis of $\mathbb{R}^2$.
2. Express the coordinates of $u = (x, y)$ in the basis (v_1, v_2) in terms of x and y.

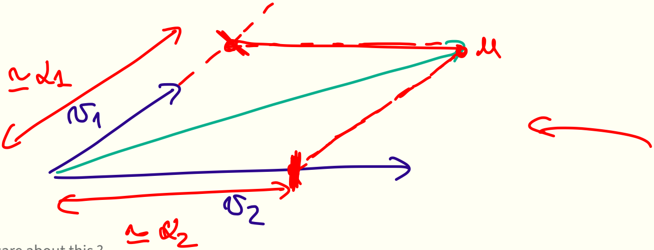
$$\begin{cases} \alpha_1 + \alpha_2 = x & (R_1) \\ \alpha_1 - \alpha_2 = y & (R_2) \end{cases} \qquad \begin{cases} 2\alpha_1 = x+y & (R_1)+(R_2) \\ \alpha_1 - \alpha_2 = y & (R_2) \end{cases}$$

$$\begin{cases} \alpha_1 = \dfrac{x+y}{2} \\ \alpha_2 = \dfrac{x-y}{2} \end{cases}$$

So the coordinates of $u = \begin{pmatrix} x \\ y \end{pmatrix}$ in the (v_1, v_2) basis are $\boxed{\left(\dfrac{x+y}{2}, \dfrac{x-y}{2} \right)}$

Why do we care about this ?





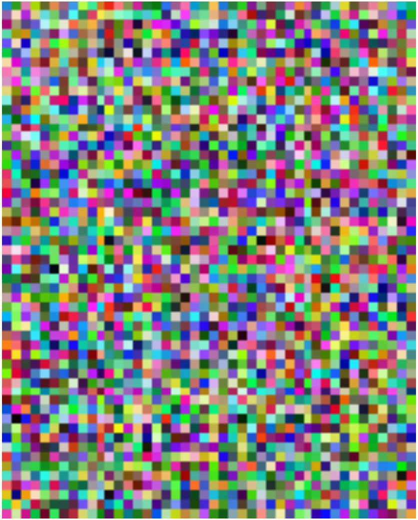
Application to image compression

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

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

Can we do better?

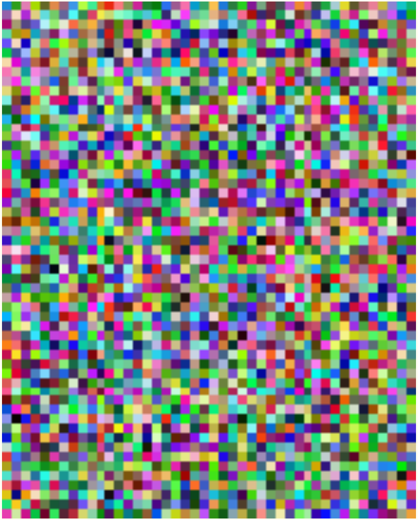
- If we want to store an arbitrary image, NO!



«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.

"set of all images"

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

image from "real world"

What do we mean by « structure » ?

Neighboring pixels are very likely to have similar colors.

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

$v = \boxed{x \mid y}$ ... "color of the pixel #2"

I take here $\omega_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\omega_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

then the coordinates of v in the (ω_1, ω_2) basis

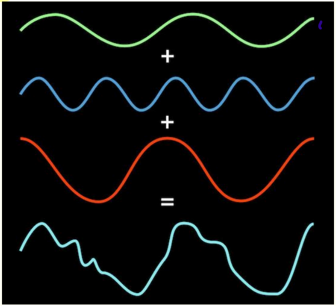
are
$$\begin{cases} \alpha_1 = \dfrac{x+y}{2} \simeq x \simeq y \\ \alpha_2 = \dfrac{x-y}{2} \simeq 0 \end{cases}$$

v is an image from the "real-world"

$\longrightarrow \boxed{x \simeq y}$ $\qquad (\alpha_1, \alpha_2) \simeq (x, 0)$

Examples of good bases

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



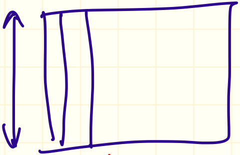
- `JPEG2000` uses **wavelet bases**, and achieves better performance than JPEG.
- In **Homework 4**, you will use wavelets to compress/denoise images.
- The course **DS-GA 1013** deepens these concepts!

Questions?

$\mathbb{R}^4$

$\mathbb{R}^{2048}$

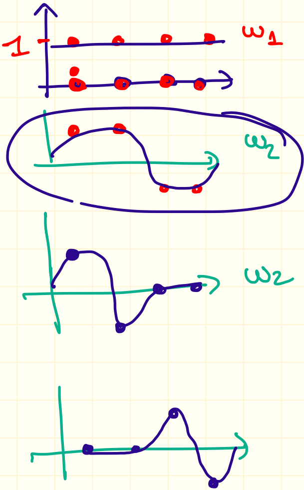
$$\begin{cases} \omega_1 = (1\ 1\ 1\ 1) \\ \omega_2 = (1\ 1\ -1\ -1) \\ \omega_3 = (1\ -1\ 0\ 0) \\ \omega_4 = (0\ 0\ 1\ -1) \end{cases}$$



20

$\alpha_1\quad \alpha_2\quad \alpha_3\quad \alpha_4$

$(1, \cancel{0{,}001}, 0, 2)$



Questions?

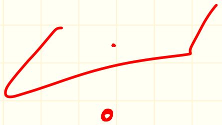
$$S = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} x + y = 1 \\ x + 2y = 0 \end{array} \right\}$$

define $\quad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

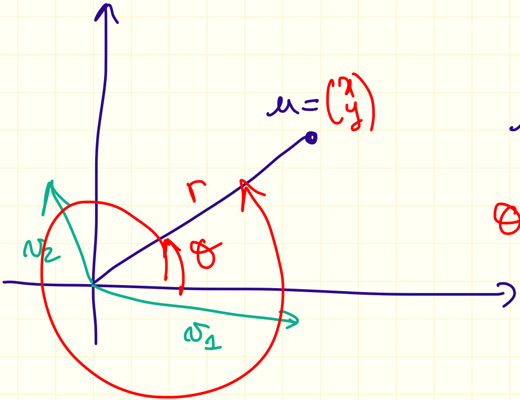
$b \neq 0$

$$\underline{S} = \left\{ \underline{(x,y)} \in \mathbb{R}^2 \;\middle|\; \underbrace{A \cdot \begin{pmatrix} x \\ y \end{pmatrix}}_{0 \,\neq\, b} = b \right\}$$

$$(0,0) \notin S$$



Questions?



$$u = \theta \, \underline{v_1} + r \, \underline{v_2}$$

$$\theta \to \theta + 2\pi$$

(r, θ) and $(r, \theta + 2\pi)$

Questions?