Linear Algebra 2 Course Notes

Yehor Korotenko

November 15, 2025

Abstract

These are my notes taken for the Linear Algebra 2 course at Paris-Saclay University given by Professor Johannes Anschütz. The essential part of these notes is referenced from the book "Linear Algebra" written by Joseph Grifone [2].

My notes for other subjects are available on my website: dobbikov.com

These notes are translated into Ukrainian and English using the tool `sci-trans-git` [3]

Contents

CHAPTER 1

Euclidean Spaces

"Classical" linear algebra deals with vector spaces, where we only talk about linear combinations, subspaces, bases, matrices, etc. At some point, this is no longer sufficient. To be able to explore stronger, more complex, and useful notions, we will need to calculate the length of a vector, the angles between two vectors, the relative positioning between vectors, etc. To be able to study these concepts, we introduce the notion of a dot product (bilinear form) and then the vector spaces equipped with this product.

This chapter is devoted to the study of the two main notions:

- scalar products
- Euclidean spaces

1.1 Introduction

The vector spaces considered in this chapter are real. We assume that E is an $\mathbb{R}$-vector space.

Scalar product:

Definition 1.1. A bilinear form on E is a map

$$\begin{aligned} B : E \times E &\longrightarrow \mathbb{R} \\ (u,v) &\longmapsto B((u,v)) \end{aligned}$$

that satisfies the following conditions $\forall u, v, w \in E\ \forall \lambda \in \mathbb{R}$:

1. $B(u + \lambda v, w) = B(u, w) + \lambda B(v, w)$
2. $B(u, v + \lambda w) = B(u, v) + \lambda B(v, w)$

B is said to be

1. symmetric if $B(u, v) = B(v, u)\ \forall u, v \in E$
2. positive if $B(., u) \geq 0\, \forall u \in E$
3. defined if $B(u, u) = 0 \Leftrightarrow u = 0$

Notation. Scalar product is denoted: $< u, v >$

Example 1.2. .

1. $E = \mathbb{R}^n$, $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n) \in E$

$$< X, Y > := \sum_{n=1}^{n} x_i y_i$$

It is called the "canonical (or usual) scalar product".

2. $E = \mathbb{R}^2$ and $< X, Y >= 2x_1y_1 + x_2y_2$
3. $E = \mathcal{C}^0([-1, 1], \mathbb{R}) \ni f, g$ (a space of continuous functions)

$$< f, g > := \int_{-1}^{1} f(t) \cdot g(t)\, dt$$

4. $E = \mathcal{M}_n(\mathbb{R}) \ni A, B$

$$< A, B > := Tr(A^t B)$$

Proposition 1.3. A non-zero vector space has an infinite number of different scalar products.

Definition 1.4. A Euclidean space is a pair $(E, < . >)$ where E is a $\mathbb{R}$-vector space <u>of finite dimension</u> and $< . >$ is an inner product on E.

Property. Let $(E, < . >)$ be a Euclidean space. We define:

$$\|X\| := \sqrt{< X, X >} \qquad X \in E$$

the norm (or length) of X. (It is well defined because $< ., . >$ is always positive)

Property. Let $X, Y \in E$ be, then:

$$\|X + Y\|^2 = \|X\|^2 + 2\langle X, Y\rangle + \|Y\|^2$$

Proof.

$$\begin{aligned} \|X + Y\|^2 = \sqrt{\langle X + Y, X + Y\rangle}^2 &= \langle X + Y, X + Y\rangle \\ &= \langle X, X + Y\rangle + \langle Y, X + Y\rangle \\ &= \langle X, X\rangle + \langle X, Y\rangle + \langle Y, X\rangle + \langle Y, Y\rangle \\ &= \|X\|^2 + 2\langle X, Y\rangle + \|Y\| \end{aligned}$$

□

Lemma 1.5. Cauchy-Schwarz inequality We have

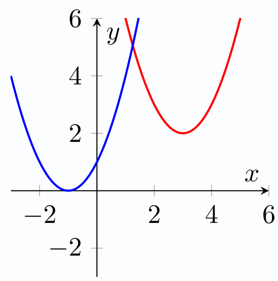
$$| < u, v > | \le \|u\| \cdot \|v\| \qquad \forall u, v \in E$$

with equality if and only if u and v are collinear, i.e $\exists t \in R$ such that $u = tv$ or $v = tu$

Proof. If $v = 0$, clear

If $v \neq 0$ we consider $\forall t \in \mathbb{R}$

$$\begin{aligned}\|u+tv\|^2 &=< u+tv, u+tv > \\ &=< u, u+tv > +t < v, u+tv > \\ &=< u,u > +t < u,v > +t < v,u > +t^2 < v,v > \\ &= \|u\|^2 + 2t < u,v > +t^2\|v\|^2 = f(t)\end{aligned}$$



Case 1: $f(t)$ has no distinct roots

$$\begin{aligned}&\Delta = 4 < u,v >^2 = 4\|u\|^2\|v\|^2 \leq 0 \\ &\Rightarrow < u,v >^2 \leq \|u\|^2 \cdot \|v\|^2 \\ &\Rightarrow | < u,v > | \leq \|u\|\|v\|\end{aligned}$$

Case 2: $f(t)$ has only one root:

$$\begin{aligned}&\Delta = 0 \\ &\Rightarrow \exists t \in \mathbb{R} \text{ tq } \|u+tv\|^2 = 0 \\ &\Rightarrow u+tv = 0 \Rightarrow u = -tv\end{aligned}$$

The following definition will be studied in the analysis course:

Definition 1.6. We say that $N : E \to \mathbb{R}_+$ is a norm if:

1. $N(\lambda u) = |\lambda| \cdot N(u) \quad \forall \lambda \in \mathbb{R}, \forall u \in E$
2. $N(u) = 0 \Rightarrow u = 0$
3. $N(u+v) \leq N(u) + N(v) \quad \forall u,v \in E$

Lemma 1.7. The application

$$\sqrt{< .,. >} = \|.\| : E \to \mathbb{R}_+$$

is called a Euclidean norm.

Proof. 1) and 2) are done

- $\|u+v\|^2 = \|u\|^2 + 2 < u,v > +\|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$

$$\Rightarrow \|u+v\|^2 \leq \|u\|^2 + \|v\|^2$$

Proposition 1.8. We have the following identities $\forall u, v \in E$

1. Parallelogram identity:
$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u^2\| + \|v\|^2)$$
2. Polarization identity:
$$\langle u, v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$$

Proof. .

1.
$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2\end{aligned}$$
2. $\|u-v\|^2 = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2$

For a:

- $(1) + (2)$: $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$
- $(1) - (2)$: $\|u+v\|^2 - \|u-v\|^2 = 4\langle u, v\rangle$

1.2 Orthogonality

Let E be an $\mathbb{R}$-vector space and $\langle , \rangle$ an inner product on E.

Definition 1.9. $u, v \in E$ are said to be orthogonal if $< u, v >= 0$. We denote $u \perp v$

- Two subsets A, B of E are orthogonal if:
$$\forall u \in A, \forall v \in B, \quad < u, v >= 0$$
- If $A \subseteq E$ we call the **orthogonal of** A, denoted $A^\perp$, the set
$$A^\perp = \{u \in E \mid < u, v >= 0 \quad \forall v \in A\}$$
Also known as **orthogonal complement of** A
- A family $(v_1, \ldots, v_n)$ of vectors in E is said to be orthogonal if $\forall i \neq j, v_i \perp v_j$. It is said to be orthonormal if it is orthogonal and additionally $\|v_i\| = 1 \quad \forall i \in \{1, \ldots, n\}$

Example 1.10. $E = \mathbb{R}^n$, $<,>$ canonical scalar product
$$v_i = (0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0)$$
$$< v_i, v_j >= \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$
$(v_1, \ldots, v_n)$ is a canonical basis

Proposition 1.11.

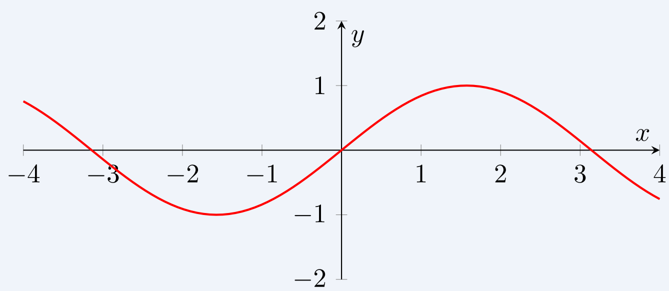
1. If $A \subseteq E$ then $A^\perp$ is a vector subspace of E
2. If $A \subseteq B$ then $B^\perp \subseteq A^\perp$
3. $A^\perp = Vect(A)^\perp$
4. $A \subset (A^\perp)^\perp$

Proof. Exercise

Example 1.12.

1. $E = \mathcal{C}^0([-1,1], \mathbb{R})$

$$< f, g >:= \int_{-1}^{1} f(t) \cdot g(t)\, dt$$



Then, $f(t) = \cos(t)$, $g(t) = \sin(t)$ are orthogonal: $2\cos(t)\sin(t) = \sin(2t)$

$$\int_{-1}^{1} \cos(t)\sin(t)\, dt = \frac{1}{2}\int_{-1}^{1} \sin(2t)\, dt = 0$$

Definition 1.13. If E is a Euclidean space, the set

$$L(E, \mathbb{R}) = \{f : E \to \mathbb{R} \mid f \text{ is linear}\}$$

is called the "dual of E". It is denoted E^*. An element $f \in E^*$ is called a linear form.

Recall:

Proposition 1.14. If F, F' are two finite-dimensional vector spaces, then $dim(L(F, F')) = dim(F) \cdot dim(F')$ In particular, $dim(F^*) = dim(F)$. Indeed, if $n = (e_1, \dots, e_p)$ is a basis of F and $n' = (e'_1, \dots, e'_q)$ is a basis of F', then the mapping

$$\begin{aligned} : L(F, F') &\longrightarrow Mat_{f \times p}(\mathbb{R}) \\ f &\longmapsto (f) = Mat_{n,n'}(f). \end{aligned}$$

is an isomorphism. Therefore $dim(F, F) = qp$

Theorem 1.15. Rank Theorem: If F is a finite-dimensional vector space and $f : F \to F'$ is linear, then

$dim(F) = dim(Ker(f)) + dim(Im(f))$

Proposition 1.16. If F, F' are two finite-dimensional vector spaces such that $dim(F) = dim(F')$ and $f : F \to F'$ is linear, then f is an isomorphism $\Leftrightarrow Ker(f) = 0$

Proof. Recall that if G, G' are finite-dimensional subspaces in the same vector space, then:

$$G = G' \Leftrightarrow G \subseteq G' \text{ and } dim(G) = dim(G')$$

$\Rightarrow$) f is injective $\Rightarrow Ker(f) = 0$
$\Leftarrow$) Let $Ker(f) = 0$.
Then, necessarily $dim(Ker(f)) = 0$ and by the rank theorem, we have $dim(F) = dim(Im(f))$, so $Im(f) = F'$

Lemma 1.17. Riesz's Lemma:
Let $(E, \langle ., . \rangle)$ be a finite-dimensional Euclidean space and $f \in E^*$. Then, $\exists! u \in E$ such that $f(x) = \langle u, x \rangle \quad \forall x \in E$. The linear form f is given by an inner product with a vector.

Notation. For any $v \in E$, we denote by f_v the mapping:

$$\begin{aligned} f_v : E &\longrightarrow \mathbb{R} \\ x &\longmapsto f_v(x) = < v, x > . \end{aligned}$$

f_v is linear $\forall v \in E$, i.e. E^*

Proof. Riesz Lemma
Consider the mapping

$$\begin{aligned} \phi : E &\longrightarrow E^* \\ v &\longmapsto \phi(v) = f_v. \end{aligned}$$

ϕ is linear (exercise). ϕ is injective:

$$v \in Ker(\phi) \Leftrightarrow f_v(x) = 0 \quad \forall x \in E$$

in particular for x = v, we have:

$$0 = f_v(v) = < v, v > \Rightarrow v = 0$$

$$\begin{aligned} dim(E) = dim(E^*) &\Rightarrow \phi \text{ is an isomorphism} \\ &\Rightarrow \phi \text{ bijective} \end{aligned}$$

$$\forall f \in E^*, \exists! n \in E \text{ such that } \phi(n) = f, \text{ i.e } f(x) = < n, x > \ \forall x \in E$$

In this case $E = \mathbb{R}^n$, the Riesz Lemma is very simple to understand:
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a linear form. If we denote $(e_1, \dots, e_n)$ the canonical basis of $\mathbb{R}^n$, any $x \in \mathbb{R}^n$ can be

written as

$$x = \sum_{n=1}^{n} \alpha_i e_i \qquad \alpha_i \in \mathbb{R}, \forall i \in \{1, \dots, n\}$$

$$\Rightarrow f(x) = \sum_{n=1}^{n} \alpha_i f(e_i) =< (\alpha_1, \dots, \alpha_n), (a_1, \dots, a_n) >=< (a_1, \dots, a_n), (\alpha_1, \dots, \alpha_n) >$$

1.3 Orthonormal bases

Let $(E, \langle , \rangle)$ be a Euclidean space and $F \subset E$ a vector subspace ($dim(F) < \infty$) because $dim(E) < \infty$.

Note.

$$F^\perp := \{x \in E \mid \langle X, Z \rangle = 0 \, \forall z \in F\}$$

the orthogonal of F.

Theorem 1.18. On a $E = F \oplus F^\perp$.
In particular, $dim(F^\perp) = dim(E) - dim(F)$ and $F = (F^\perp)^\perp$

Proof. We must show that:

1. $F \cap F^\perp = \emptyset$
2. $E = F + F^\perp$ i.e. $\forall x \in E, \exists x' \in F,\ x'' \in F^\perp$ such that $x = x' + x''$

1. Let $x \in F \cap F^\perp$
$\Rightarrow \langle X, Z \rangle = 0 \, \forall Z \in F$ because $x \in F \Rightarrow \langle X, X \rangle = 0 \Rightarrow x = 0(\langle , \rangle$ is defined)
2. Let $x \in E$. Consider $f_x \in E^*$, i.e. $f_x : E \to \mathbb{R}, y \mapsto \langle x, y \rangle$ and $f := f_{x|F} : F \to \mathbb{R} \Rightarrow f \in E^*$ Riesz Lemma $\Rightarrow \exists! x' \in F$ such that $f = f_{x'} : F \to \mathbb{R}, z \mapsto \langle x', z \rangle$
$\Rightarrow f_x(z) = f_{x'}(z) = f(z) \, \forall z \in F$ (Attention: not the equality for all z in E)
Let $x'' := x - x'$, i.e. $x = x' + x'' \in F$. Let's show $x'' \in F^\perp$.
If $z \in F$, $\langle x'', z \rangle = \langle x - x', z \rangle = \langle x, z \rangle - \langle x', z \rangle = 0$. Therefore, $x'' \in F^\perp$ and $E = F \oplus F^\perp$ ($dim(E) = dim(F) + dim(F^\perp)$)
$F \subseteq (F^\perp)^\perp$ because $\langle x, z \rangle = 0 \, \forall x \in F \, \forall z \in F^\perp$

$$dim(F) = dim(E) - dim(F^\perp)$$

because $E = G \oplus G^\perp$, therefore $dim(G) = dim(E) - dim(G^\perp)$ for $G = F^\perp$, $dim(F^\perp) = dim(G)$

□

Definition 1.19. Let E be a vector space equipped with a scalar product $\langle , \rangle$

- A family $(v_i)_{i \geq 0}$ of vectors in E is said to be orthogonal if for $i \neq j$ we have $\langle v_i, v_j \rangle = 0$ i.e. $v_i \perp v_j$
- An orthonormal family of E is an orthogonal family $(v_i)_{i \geq 0}$ such that, furthermore, $\|v_i\| = 1$ for $i \geq 0$

Example 1.20. 1. $E = \mathbb{R}^n$ equipped with the standard dot product. The standard basis $(e_1, \dots, e_n)$ is orthogonal because

$$\langle e_i, e_j \rangle = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

2. In $E = \mathcal{C}^0([-1,1], \mathbb{R})$ equipped with $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\,dt$. The family $(\cos(t), \sin(t))$ is orthogonal. The family $(1, t^2)$ is not orthogonal:

$$\left\langle 1, t^2 \right\rangle = \int_{-1}^{1} 1t^2\,dt = \frac{2}{3} \neq 0$$

Proposition 1.21. An orthogonal family consisting of <u>non-zero</u> vectors is linearly independent. In particular, an orthonormal family is linearly independent.

Proof. Suppose $(v_1, \ldots, v_n)$ is orthogonal with $v_i \neq 0\, \forall i = 1, \ldots, n$
if $\sum_{j=1}^{n} \underset{\in \mathbb{R}}{\alpha_i} v_i = 0$, then

$$\forall i \in \{1, \ldots, n\} 0 = \left\langle v_i, \sum_{j=1}^{n} \alpha_j v_j \right\rangle = \sum_{j=1}^{n} \alpha_j \langle v_i, v_j \rangle = \alpha_i \underset{\neq 0}{\|v_i\|^2}$$

Therefore, $\alpha_i = 0\, \forall i = 1, \ldots, n$.
If $(v_1, \ldots, v_n)$ is orthonormal, then $\|v_i\| = 1$. Therefore, $v_i \neq 0$, $\forall i = 1, \ldots, n$. ☐

Intuition. Orthogonal (perpendicular) vectors are never contained within each other (i.e. $e_i = \lambda e_j$ is not possible). If the vectors are related, then the angle is < 90 (thus the vectors are not orthogonal, absurd), (they are contained within each other, they are not orthogonal, absurd). Therefore, they are indeed linearly independent.

Definition 1.22. $(E, \langle , \rangle)$ Euclidean space. A family $B = (e_1, \ldots, e_n)$ is an orthonormal basis (or ONB) if it is a basis and an orthonormal family.

Theorem 1.23. $(E, \langle , \rangle)$ Euclidean space. Then, it admits an ONB.

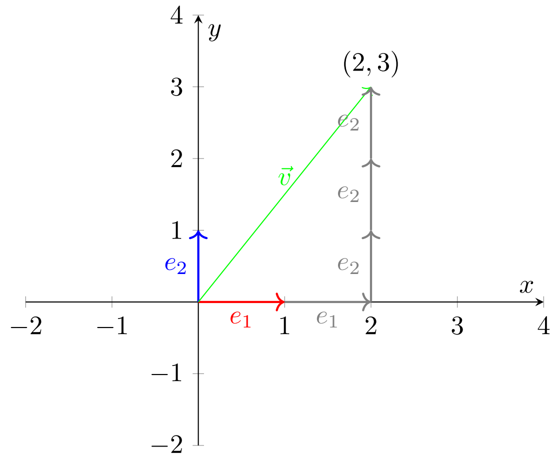
Proof. Let $n := dim(E)$. Let $(e_1, \ldots, e_p)$ be an orthogonal family (from the point of view of the cardinal p) such that $e_i \neq 0\, \forall i = 1, \ldots, p$.
Suppose, for the sake of contradiction, that $p < n$. Let $F = Vect(e_1, \ldots, e_p)$. Then, $E = F \oplus F^\perp$ and $dim(F) \leq p < n$. Therefore, $F^\perp \neq \{0\}$. Let $x \in F^\perp$, $x \neq 0$. Then, $(e_1, \ldots, e_p, x)$ is orthogonal of cardinal $> p$. Therefore, $p = n$ and $(e_1, \ldots, e_n)$ is a basis of E. To have an orthonormal family $(e'_1, \ldots, e'_n)$ it suffices to take $e'_i = \frac{1}{\|e_i\|} e_i\, \forall i = \{1, \ldots, n\}$. ☐

Proposition 1.24. Let $(E, \langle , \rangle)$ be a Euclidean space and let $(e_1, \ldots, e_n)$ be an orthonormal basis of E. If $x \in E$, we have:

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

In other words, the real number $\langle x, e_i \rangle$ is the $i^{\text{ème}}$ coordinate of x in the basis $(e_1, \ldots, e_n)$.

Intuition. The orthonormality of the base simplifies our lives. But first, a small introduction. Let an e.v $E = \mathbb{R}^2$ and the base $(e_1, e_2) = (\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix})$. Let a vector $\vec{v} = (2, 3)$:



So, we can write $\vec{v} = \vec{(2,3)} = 2 \cdot \vec{e_1} + 3 \cdot \vec{e_2}$. The x and y (the coordinates of v) give us how many parts of each base vector (the number can be $\in \mathbb{R}$) and take their sums, to obtain $\vec{v}$. (The simplest: how much we have to go left and up).

In the orthonormal base $\langle v, e_i\rangle$ gives us how much we take of a vector e_i to make the vector $\vec{v}$ and $\vec{e_i}$ gives the direction. Hence $\langle v, e_1\rangle$ is equivalent to 2, and $\langle v, e_2\rangle$ to 3, then:

$$\vec{v} = \underbrace{\langle v, e_1\rangle}_{=2} \cdot \vec{e_1} + \underbrace{\langle v, e_2\rangle}_{=3} \cdot \vec{e_2}$$

Usually, to find the coordinates in a base, we would have to solve a linear system, while an orthonormal base allows us to obtain them by calculating the scalar product with each vector of the base, which is much simpler.

Proof. Let $y := \sum_{i=1}^{n} \langle x, e_i\rangle e_i$. Then,

$$\begin{aligned}
&\forall j = 1, \ldots, n,\\
&\langle x - y, e_j\rangle\\
=&\langle x, e_j\rangle - \langle y, e_j\rangle\\
=&\langle x, e_j\rangle - \langle \sum_{i=1}^{n} \langle x, e_i\rangle e_i, e_j\rangle\\
=&\langle x, e_j\rangle - \underbrace{\sum_{i=1}^{n} \langle x, e_i\rangle}_{\substack{\text{moved out}\\ \text{like constant}}} \langle e_i, e_j\rangle\\
=&\langle x, e_j\rangle\\
-&\left(\langle x, e_1\rangle \underbrace{\langle e_1, e_j\rangle}_{=0} + \ldots + \langle x, e_{j-1}\rangle \underbrace{\langle e_{j-1}, e_j\rangle}_{=0} + \langle x, e_j\rangle \underbrace{\langle e_j, e_j\rangle}_{=1} + \langle x, e_{j+1}\rangle \underbrace{\langle e_{j+1}, e_j\rangle}_{=0} + \ldots + \langle x, e_n\rangle \underbrace{\langle e_n, e_j\rangle}_{=0} \right)\\
&(\forall i \neq j,\ \langle e_i, e_j\rangle = 0 \text{ because a scalar product of orthogonal vectors})\\
&(\forall j\ \langle e_j, e_j\rangle = 1 \text{ because a scalar product of the same vector})\\
=&\langle x, e_j\rangle - \langle x, e_j\rangle \underbrace{\langle e_j, e_j\rangle}_{=1} = 0
\end{aligned}$$

Therefore, $x - y \in Vect(e_1, \dots, e_n)^\perp = E^\perp = \{0\}$. Thus $x = y$ $\square$

Corollary 1.25. $\forall x \in E$, $\|x\|^2 = \sum_{i=1}^n \langle x, e_i\rangle^2$

Proof. If $x = \sum_{i=1}^n \langle x, e_i\rangle e_i = \sum_{i=1}^n x_i e_i$ then

$$\|x\|^2 = \langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j e_j\rangle = \sum_{i,j=1}^n x_i x_j \langle e_i, e_j\rangle = \sum_{i=1}^n x_i^2$$

$\square$

1.4 Matrices and scalar products

Proposition 1.26. Let $(E, \langle,\rangle)$ be a Euclidean space and $\varepsilon = (e_1, \dots, e_n)$ an orthonormal basis. Let $f \in \mathcal{L}(E, E)$ and $A = (a_{i,j})_{1\leq i,j\leq n}$ be the representative matrix of f in ε, i.e., $A = Mat_\varepsilon(f)$

$$a_{i,j} = \langle f(e_i), e_j\rangle \ \forall i, j = 1, \dots, n$$

Proof. A is the matrix whose columns are the vectors $f(e_j)$ written in the basis ε:

$$A = (f(e_1)| \dots |f(e_n)) \quad f(e_j) = \begin{pmatrix} a_{1,j} \\ \dots \\ a_{n,j} \end{pmatrix}$$

Because $\forall v \in E$, $v = c_1 e_1 + \dots c_n e_n$ then $f(v) = c_1 f(e_1) + \dots c_n f(e_n)$ by linearity, so we only have to study each $f(e_j)$

$$f(e_j) = a_{1,j} e_1 + \dots a_{n,j} e_n \Rightarrow$$

$$\langle f(e_j), e_i\rangle = \left\langle \sum_{k=1}^n a_{k,j} e_k, e_i \right\rangle = \sum_{k=1}^n a_{k,j} \langle e_k, e_i\rangle = a_{k,j}$$

because $\langle e_k, e_j\rangle = \begin{cases} 0 \text{ si } k \neq j \\ 1 \text{ si } k = j \end{cases}$ Therefore:

$$a_{i,j} = \langle f(e_j), e_i\rangle$$

$\square$

The matrix of a cross product is very useful in linear algebra. Before giving a definition:

Let E be a vector space of finite dimension n, a space K and a bilinear form $b : E \times E \longrightarrow K$. If $\{e_1, \dots, e_n\}$ is a basis of E, then: $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{j=1}^n y_j e_j$, then we have:

$$b(x, y) = \sum_{i,j=1}^n x_i y_j b(e_i, e_j)$$

b is therefore determined by the knowledge of the values $b(e_i, e_j)$ on a basis.

Definition 1.27. We call **the matrix of** b in the basis $\{e_i\}$ the matrix:

$$M(b)_{e_i} = \begin{pmatrix} b(e_1, e_1) & b(e_1, e_2) & \dots & b(e_1, e_n) \\ b(e_2, e_1) & b(e_2, e_2) & \dots & b(e_2, e_n) \\ \dots & \dots & \dots & \dots \\ b(e_n, e_1) & \dots & \dots & b(e_n, e_n) \end{pmatrix}$$

Thus, the element of the i$^{\text{ème}}$ row and j$^{\text{ème}}$ column is the coefficient of $x_i y_j$.

Example 1.28. The matrix of the canonical scalar product in $\mathbb{R}^3$ is:

$$\langle X, Y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

$$Mat(\langle , \rangle)_{e_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Proposition 1.29. Scalar product represented by a matrix.
Let us note:

$$\underbrace{A = M(b)_{e_i}}_{\text{matrice de produit scalair}} \qquad \underbrace{X = M(x)_{e_i}}_{\substack{\text{coordonnées de } x \\ \text{dans la base } e_i}} \qquad \underbrace{Y = M(y)_{e_i}}_{\substack{\text{coordonnées de } y \\ \text{dans la base } e_i}} \qquad (x, y \in E)$$

Then, we have:

$$b(x, y) = X^t A Y$$

Example 1.30. Let's take the example with $b = \langle , \rangle$ the canonical scalar product in $\mathbb{R}^3$. Let $X = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ and $Y = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$ be in the canonical basis of $\mathbb{R}^3$. Therefore:

$$\begin{aligned} \langle x, y \rangle = X^t A Y &= \overbrace{(1, 2, -1)}^{X^t} \times \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{A} \times \overbrace{\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}}^{Y} \\ &= \underbrace{(1, 2, -1)}_{X} \times \underbrace{\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}}_{A \times Y} \\ &= 1 \cdot 2 + 2 \cdot 3 + (-1) \cdot 1 = 2 + 6 - 1 = 7 \end{aligned}$$

TODO. change of basis of the matrix of a bilinear form

1.5 Orthogonal Projections

Let $(E, \langle , \rangle)$ be a Euclidean space, $F \subseteq E$ a vector subspace. Then, $E = F \oplus F^{\perp}$. Therefore, $\forall x \in E$ can be written as

$$x = \underset{\in F}{x_F} + \underset{\in F^{\perp}}{x_{F^{\perp}}}$$

Definition 1.31. The **orthogonal projection** of E into F is the projection p_F of E onto F parallel to $F^{\perp}$, i.e.

$$\begin{aligned} p_F : E = F \oplus F^{\perp} &\longrightarrow F \\ x = x_F + x_{F^{\perp}} &\longmapsto p_F(x = x_F + x_{F^{\perp}}) = x_F. \end{aligned}$$

Remark 1.32. 1. p_F is linear

2. $\forall x \in E\, p_F(x)$ is completely characterized by the following property:
   Let $y \in E$, then

   $$y = p_F(x) \Leftrightarrow \left(\underset{\Rightarrow y = x_F}{y \in F} \text{ et } x - y \in F^{\perp} \right)$$

   In particular, $\langle p_F(x), x - p_F(x) \rangle = 0$. Then, if $(v_1, \ldots, v_R)$ is an ONB of F, we have:

   $$\forall x \in E,\, p_F(x) = \sum_{i=1}^{k} \langle x, v_i \rangle \, v_i$$

   Indeed, it suffices to verify that the vector $y = \sum_{i=1}^{k} \langle x, v_i \rangle \, v_i$ satisfies:

   $$y \in F \text{ and } x - y \in F^{\perp}$$

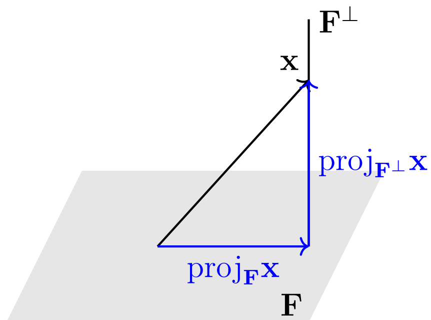


Figure 1.1: Projection

Proposition 1.33. Let $x \in E$. Then,

$$\|x - p_F(x)\| = inf\{\|x - y\| \mid y \in F\}$$

i.e. $\|x - p_F(x)\|$ is the distance from x to F.
See Figure 1.1

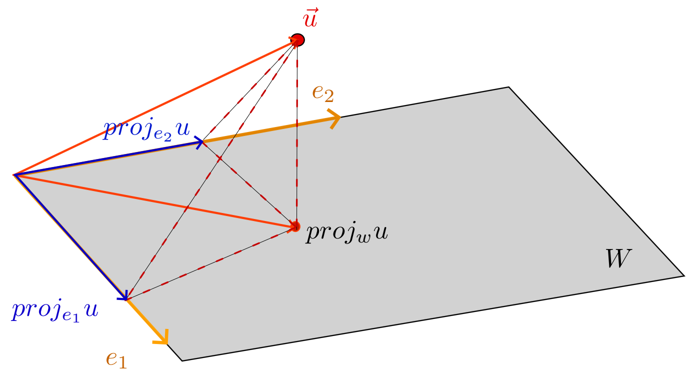


Figure 1.2: Projection with BON

Proof. Since $p_F(x) \in F$ it suffices to prove that, if $y \in F$, then

$$\|x - p_F(x)\| \leq \|x - y\|$$

But, $$\underbrace{\|x-y\|^2}_{(x-p_F(x))+(p_F(x)-y)} = \|x-p_F(x)\|^2 + 2\overbrace{\left\langle \overset{\in F^\perp}{x - p_F(x)}, \overset{\in F}{p_F(x) - y} \right\rangle} = 0 + \underbrace{\|p_F(x)-y\|^2}_{\geq 0} \geq \|x-p_F(x)\|^2 \quad \square$$

Theorem 1.34. Gram-Schmidt
Let E be a vector space equipped with a scalar product $\langle , \rangle$. Let $(v_1, \ldots, v_n)$ be a free family of elements $\in E$. Then, there exists an orthogonal family $(w_1, \ldots, w_n)$ such that

$$\forall i = 1, \ldots, n \quad Vect(v_1, \ldots, v_i) = Vect(w_1, \ldots, w_i)$$

Moreover, this theorem gives us a method for constructing an orthonormal basis from an arbitrary basis.

Proof. of Theorem 1.34 Let's construct the orthogonal basis: $\{w_1, \ldots, w_p\}$. First, let's set:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 + \lambda w_1, \qquad \text{avec } \lambda \text{ tel que } w_1 \perp w_2 \end{cases}$$

By imposing this condition, we find:

$$0 = \langle v_2 + \lambda w_1, w_1 \rangle = \langle v_2, w_1 \rangle + \lambda \|w_1\|^2$$

Since $w_1 \neq 0$, we obtain $\lambda = -\frac{\langle v_2, w_1 \rangle}{\|w_1\|^2}$. We notice that:

$$\begin{cases} v_1 = w_1 \\ v_2 = w_2 - \lambda w_1 \end{cases}$$

therefore $Vect\{v_1, v_2\} = Vect\{w_1, w_2\}$.

Once w_2 is constructed, we construct w_3 by setting:

$$w_3 = v_3 + \mu w_1 + \nu w_2$$
$$\text{with } \mu \text{ and } \nu \text{ such that: } w_3 \perp w_1 \text{ and } w_3 \perp w_2$$

We can see $w_3 = v_3 - \lambda' w_1 - \lambda'' w_2$ as $w_3 = v_3 - proj_{F_2} v_3$ where $F_i = Vect\{w_1, \ldots, w_i\}$

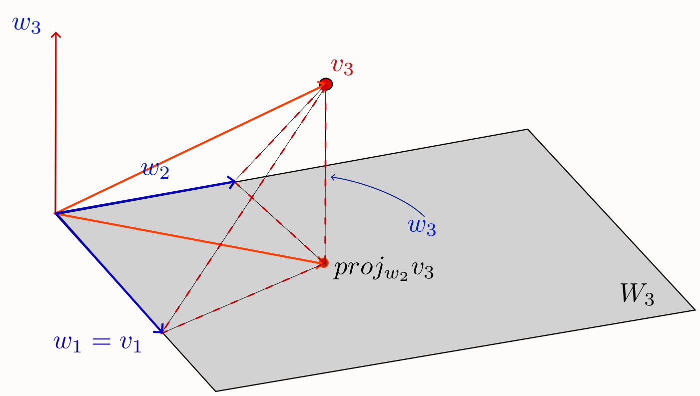


Figure 1.3: Vecteur par projection

This gives

$$0 = \langle v_3 + \mu w_1 + \nu w_2, w_1 \rangle = \langle v_3, w_1 \rangle + \mu \underbrace{\langle w_1, w_1 \rangle}_{= \|w_1\|^2} + \nu \underbrace{\langle w_2, w_1 \rangle}_{=0}$$
$$= \langle v_3, w_1 \rangle + \mu \|w_1\|^2$$

hence $\mu = -\frac{\langle v_3, w_1 \rangle}{\|w_1\|^2}$. Similarly, by imposing that $w_3 \perp w_2$, we find $\nu = -\frac{\langle v_3, w_2 \rangle}{\|w_2\|^2}$. As

$$\begin{cases} v_1 = w_1 \\ v_2 = w_2 - \lambda w_1 \\ v_3 = w_3 - \mu w_1 - \nu w_2 \end{cases}$$

we can see that $Vect\{w_1, w_2, w_3\} = Vect\{v_1, v_2, v_3\}$. That is, $\{w_1, w_2, w_3\}$ is an orthogonal basis of the space spanned by v_1, v_2, v_3. We can now clearly see the recurrence process.

Suppose we have constructed $w_1, \ldots, w_{k-1}$ for $k \leq p$. Let's set:

$$w_k = v_k + \text{ combinaison linéaire des vecteurs déjà trouvés}$$
$$= v_k + \lambda_1 w_1 + \ldots + \lambda_{k-1} w_{k-1}$$

The conditions $w_k \perp w_i$ (for $i \in \{1, \ldots, k-1\}$) are equivalent to:

$$\lambda_i = -\frac{\langle v_k, w_i \rangle}{\|w_i\|^2}$$

as can be verified immediately. Since $v_k = w_k - \lambda_1 - \ldots - \lambda_{k-1} w_{k-1}$, we see by recurrence that $Vect\{w_1, \ldots, w_k\} = Vect\{v_1, \ldots, v_k\} \Leftrightarrow \{w_1, \ldots, w_k\}$ is an orthogonal basis of $Vect\{v_1, \ldots, v_k\}$.

What remains for us is to normalize it, i.e. $\forall i \in \{1, \ldots, k\}$ $e_i = \frac{w_i}{\|w_i\|}$, hence $\{e_1, \ldots, e_k\}$ is an orthonormal basis of $F = Vect\{v_1, \ldots, v_k\}$. □

Proposition 1.35. To understand this proposition, I advise you to read section 1.6

Every orthogonal projection is self-adjoint, i.e. if p is an orthogonal projection, then:

$$p^* = p$$

In matrix notation: let A be a matrix of the projection p, then:

$$A^T = A$$

1.6 Isometries and Adjoints

1.6.1 Isometries

Definition 1.36. An **isometry** of E (or **orthogonal transformation**) is an endomorphism $f \in \mathcal{L}(E) := \mathcal{L}(E, E)$ preserving the dot product, i.e.:

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$$

Definition 1.37. Let $x, y \in E$ be two non-zero vectors. We have, according to the Cauchy-Schwarz inequality (see lemma 1.5):

$$\frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \leq 1$$

Then, there exists one and only one $\theta \in [0, \pi]$ such that:

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \tag{1.1}$$

θ is called the **angle** (non-oriented) between the vectors x and y.

Proposition 1.38. If f is an isometry of E, then we have:

$$\|f(x)\| = \|x\| \quad \forall x \in E$$

Proof. Suppose that f is an isometry of E. Let $x, y \in E$. By definition: $\langle f(x), f(y) \rangle = \langle x, y \rangle$, therefore, let $y := x$, then, we have:

$$\begin{aligned}
&\underbrace{\langle f(x), f(x) \rangle}_{\|f(x)\|^2} = \underbrace{\langle x, x \rangle}_{\|x\|^2} \\
\Leftrightarrow &\|f(x)\|^2 = \|x\|^2 \\
\Leftrightarrow &\|f(x)\| = \|x\|
\end{aligned}$$

□

Proposition 1.39. Let f be an isometry in E, then:

1. f is bijective

2. f preserves the Euclidean distance and angles

Proof. Let f be an isometry in E and two vectors $u, v \in E$

1. $$\|f(u) - f(v)\| = \sqrt{\langle f(u), f(v)\rangle} = \sqrt{\langle u, v\rangle} = \|u - v\|$$

2. Let θ_1 be the angle between $f(u)$ and $f(v)$ and θ_2 be the angle between u and v, so:

$$\cos\theta_1 := \frac{\langle f(u), f(v)\rangle}{\|f(u)\| \cdot \|f(v)\|}$$

$$\cos\theta_2 := \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}$$

By definition, $\langle f(u), f(v)\rangle = \langle u, v\rangle$, according to proposition 1.38, $\forall x, \|f(x)\| = \|x\|$, so:

$$\cos\theta_1 := \frac{\langle f(u), f(v)\rangle}{\|f(u)\| \cdot \|f(v)\|} = \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} = \cos\theta_2$$

□

Definition 1.40. Let F be a vector subspace of E, therefore $E = F \oplus F^\perp$ where $\forall v \in E, \exists v_1 \in F, v_2 \in F^\perp$ such that $v = v_1 + v_2$. We set:

$$s_F(v) = v_1 - v_2$$

and we call s_F an orthogonal symmetry with axis F.

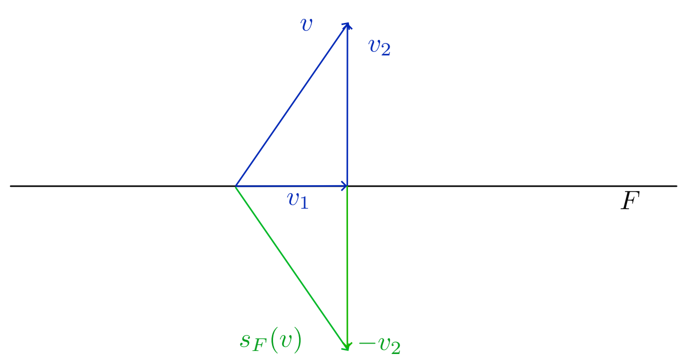


Figure 1.4: Orthogonal symmetry with axis F

Proposition 1.41. Orthogonal symmetry is an isometry.

Proof. TODO or not needed □

Proposition 1.42. f is an isometry if and only if it transforms every orthonormal basis into an orthonormal basis.

Proof. Let f be an isometry, then it transforms any basis into a basis because f is bijective by prop. 1.39.

- ($\Rightarrow$) Suppose that f is an isometry. Let $\{e_i\}$ be an orthonormal basis, then we have:

$$\langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{i,j}$$

  Therefore, $\{f(e_i)\}$ is an orthonormal basis.

- ($\Leftarrow$) Suppose that there exists an orthonormal basis $\{e_i\}$ such that $\{f(e_i)\}$ is also an orthonormal basis. Moreover, let $x = x_1e_1 + \ldots x_ne_n$ and $y = y_1e_1 + \ldots + y_ne_n$ with $x_i, y_i \in \mathbb{R}$
  Since $\{e_i\}$ is orthonormal, then we have:

$$\langle x, y \rangle = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i \tag{1.2}$$

  On the other hand:

$$\begin{aligned}\langle f(x), f(y) \rangle &= \left\langle \sum_{i=1}^{n} x_i f(e_i), \sum_{i=1}^{n} y_i f(e_i) \right\rangle = \sum_{i,j=1}^{n} x_iy_j \langle f(e_i), f(e_j) \rangle \\ &= \sum_{i,j=1}^{n} x_iy_j \langle e_i, e_j \rangle \underset{\text{car } \{e_i\} \text{ orthonormée}}{=} = \sum_{i=1}^{n} x_iy_i \underset{\text{D'apres 1.2}}{=} \langle x, y \rangle\end{aligned}$$

  Therefore f is an isometry.

□

Proposition 1.43. If $\{e_i\}$ is an orthonormal basis, f an isometry and $A = M(f)_{e_i}$, then $A^TA = I = AA^T$.

Proof. To prove this, we will use proposition 1.29.
By definition of isometry, we have:

$$\begin{aligned}&\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E \\ \Leftrightarrow &\underbrace{(AX)^T(AY)}_{\langle f(x), f(y) \rangle} = X^TA^TAY = \underbrace{X^TY}_{\langle x, y \rangle} \\ \Leftrightarrow &A^TA = I\end{aligned}$$

□

Proposition 1.44. If A is a matrix of isometry in an orthonormal basis, then $det(A) = \pm 1$

Proof. By proposition 1.43, we have: $A^TA = I$, hence:

$$\begin{aligned} det(A^TA) = det(I) = 1 \Rightarrow & det(A)^2 = 1 \quad \text{(because } det(A^T) = det(A)) \\ \Rightarrow & det(A) = \pm 1 \end{aligned}$$

□

Intuition. An isometry performs a rotation or a reflection; it preserves distances, and therefore the area (or volume) of a figure constructed by the base of this transformation is equal to 1.

1.6.2 Adjoint endomorphism

Proposition 1.45. Let E be a Euclidean space and $f \in End(E)$. There exists one and only one endomorphism $f^* \in E$ such that

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle, \quad \forall x, y \in E$$

f^* is called the **adjoint** of f.

If $\{e_i\}$ is an orthonormal basis and $A = M(f)_{e_i}$, then the matrix $A^* = M(f^*)_{e_i}$ is the transpose of A, i.e. $A^* = A^T$

Proof. Again, for the proof, we will use proposition 1.29 which is very useful, so I advise you to master this concept.

Let $\{e_i\}$ be an orthonormal basis of E and let us denote

$$A = M(f)_{e_i} \quad A^* = M(f^*)_{e_i} \quad X = M(x)_{e_i} \quad Y = M(y)_{e_i}$$

Since we are in an orthonormal basis, the statement is written:

$$\underbrace{(AX)^TY}_{\langle f(x), y \rangle} = X^TA^TY = \underbrace{X^T(A^*Y)}_{\langle x, f^*(y) \rangle} \quad \forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$$

which implies that $A^* = A$ and, furthermore, demonstrates the uniqueness of such adjoint. □

1.7 Orthogonal Groups

Reminder:

Definition 1.46. A general linear group:

$$GL(n, \mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid det(A) \neq 0\}$$

is a group of all linear transformations (square matrices) that are invertible (because $det(A) \neq 0$).

Definition 1.47. **Orthogonal Group**: The set:

$$O(n, \mathbb{R}) := \{A \in \mathcal{M}_n(\mathbb{R}) \mid A^TA = I\} = \{A \in \mathcal{M}_n(\mathbb{R}) \mid AA^T = I\}$$

satisfies the following properties:

1. if $A, B \in O(n, \mathbb{R})$, then $AB \in O(n, \mathbb{R})$
2. $I \in O(n, \mathbb{R})$

3. if $A \in O(n, \mathbb{R})$ then $A^{-1} \in O(n, \mathbb{R})$

In particular, $O(n, \mathbb{R})$ is a subgroup of $GL(n, \mathbb{R})$ (group of invertible matrices) (see definition 1.46).

Intuition. The meaning of orthogonal matrices is clear: they represent the matrices of orthogonal transformations (isometries) in **an orthonormal basis** (see defn 1.9).

We can notice that if $det(A) = 1$, this isometry represents a rotation; furthermore, we have the following definition:

Definition 1.48. The set of direct orthogonal matrices (i.e. such that $det(A) = 1$)

$$SO(n, \mathbb{R}) = \{A \in O(n, \mathbb{R}) \mid det(A) = 1\}$$

is a group, called the **special orthogonal group**.

Example 1.49. The matrix

$$A = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$

is orthogonal. We can verify that $A^T A = I$, or, it is sufficient to show that c_1, c_2, c_3 is an orthonormal family, i.e.:

$$\|c_i\|^2 = 1 \quad \text{and} \quad \langle c_i, c_j \rangle = 0 \quad \text{if } i \neq j$$

We can interpret A as the matrix of a transformation f in the canonical basis $\{e_i\}$, so we have: $c_i = f(e_i)$, according to proposition 1.42 f is orthogonal. Moreover, we see that $det(A) = +1$. Consequently, f is a direct orthogonal transformation.

Proposition 1.50. The change-of-basis matrix from an orthonormal basis to an orthonormal basis is an orthogonal matrix.

Proof. I'm providing intuition. A transition matrix transforms one basis into another; it transforms the vectors of the basis, so it transforms the basis of the O.N.B. into vectors of the basis of the O.N.B. Therefore, according to proposition 1.42, this matrix is orthogonal. □

CHAPTER 2

Determinants

This chapter is more of a cheat sheet for determinants because I'm not going to give proofs but the useful properties, examples, and intuition.

Definition 2.1. Let $A = [a_{i,j}] \in \mathcal{M}_n(\mathbb{R})$ be a square matrix $n \times n$, then:

$$\det(A) = \sum_{\sigma \in S_n} \text{signe}(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)}$$

where

- S_n is a group of all permutations of $\{1, \dots, n\}$
- $\text{signe}(\sigma)$ is a sign of permutation

This definition is very formal, so at the end of this chapter we will reformulate it. First, we will study the properties of determinants:

2.1 Most Important Properties

Proposition 2.2. the properties of the determinant. For this proposition, we denote $\det(c_1, \dots, c_n)$ a determinant where $\forall i$, r_i and $\forall i$, y_i represent a column (or a column vector). And $\forall i, \lambda_i \in \mathbb{R}$.

1. **Determinant of the identity matrix is 1:**

$$\det(I_n) = 1$$

2. **Determinant of the rank 1 matrix is its only element**:

$$\det(\begin{bmatrix} a_{1,1} \end{bmatrix}) = a_{1,1} \qquad \text{où } a_{1,1} \in \mathbb{R}$$

3. **Linearity 1**:

$$\det(r_1, \dots, r_i + y_i, \dots, r_n) = \det(r_1, \dots, r_i, \dots, r_n) + \det(r_1, \dots, y_i, \dots, r_n)$$

4. **Linearity 2**:

$$\det(r_1, \dots, \lambda_i r_i, \dots, r_n) = \lambda_i \det(r_1, \dots, r_i, \dots, r_n)$$

Note. That's why:

$$\det(\lambda A) = \lambda^n \det(A)$$

5. **Same columns**: Suppose that $i \neq j$ and $c_i = c_j$ then:

$$\det(c_1, \dots, c_i, \dots, c_j, \dots, c_n) = 0$$

If there are two identical columns, then det is equal to 0.

6. **Column swaps**:

$$\det(c_1, \dots, c_i, \dots, c_j, \dots, c_n) = -\det(c_1, \dots, \underbrace{c_j, \dots, c_i}_{\text{permutation}}, \dots, c_n)$$

In other words, a column permutation changes the sign.

7. **Determinant of multiplied matrices**: Let $A, B \in \mathcal{M}_n(\mathbb{R})$

$$\det(AB) = \det(A)\det(B)$$

8. **Determinant of a transposed matrix**: Let $A \in \mathcal{M}_n(\mathbb{R})$

$$\det(A^T) = \det(A)$$

2.2 Expansion along a row/column

Definition 2.3. Let $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ be a square matrix, i.e.:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,i-1} & a_{1,i} & a_{1,i+1} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,i-1} & a_{2,i} & a_{2,i+1} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,i-1} & a_{j-1,i} & a_{j-1,i+1} & \dots & a_{j-1,n} \\ a_{j,1} & a_{j,2} & \dots & a_{j,i-1} & a_{j,i} & a_{j,i+1} & \dots & a_{j,n} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,i-1} & a_{j+1,i} & a_{j+1,i+1} & \dots & a_{j+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,i-1} & a_{n,i} & a_{n,i+1} & \dots & a_{n,n} \end{bmatrix}$$

Then, $A_{j,i}$ is a matrix where row j and column i are deleted, i.e.:

$$A_{j,i} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,i-1} & & a_{1,i+1} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,i-1} & & a_{2,i+1} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,i-1} & & a_{j-1,i+1} & \dots & a_{j-1,n} \\ & & & & & & & \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,i-1} & & a_{j+1,i+1} & \dots & a_{j+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,i-1} & & a_{n,i+1} & \dots & a_{n,n} \end{bmatrix} \in \mathcal{M}_{n-1}(\mathbb{R})$$

This allows us to expand the determinant with respect to a row or column:

Proposition 2.4. Let $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ be a square matrix and let $1 \leq k \leq n$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{k,i} \det(A_{k,i})$$

be the calculation of the determinant with respect to the $k^{\text{ième}}$ row.

Example 2.5. Let

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$$

$$A_{2,1} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 5 \\ 7 & 6 \end{pmatrix}$$

Ce qui est au centre des lignes est le $a_{i,j}$. Ici: $a_{2,1}$

$$A_{2,2} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix}$$

$$A_{2,3} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 4 \\ 3 & 7 \end{pmatrix}$$

Figure 2.1: Development with respect to the second row

So:

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i+2} a_{2,i} \det(A_{2,i}) \\
&= (-1)^{1+2} \cdot a_{2,1} \cdot \det(A_{2,1}) + (-1)^{2+2} \cdot a_{2,2} \cdot \det(A_{2,2}) + (-1)^{3+2} \cdot a_{2,3} \cdot \det(A_{2,3}) \\
&= (-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 6 \end{vmatrix} + (-1)^{2+2} \cdot 9 \cdot \begin{vmatrix} 1 & 5 \\ 3 & 6 \end{vmatrix} + (-1)^{3+2} \cdot 8 \cdot \begin{vmatrix} 1 & 4 \\ 3 & 7 \end{vmatrix} \\
&= (-1) \cdot 2 \cdot (-11) + 1 \cdot 9 \cdot (-9) + (-1) \cdot 8 \cdot (-5) \\
&= 22 - 81 + 40 \\
&= -19
\end{aligned}$$

Proposition 2.6. Let $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ be a square matrix and let $1 \leq k \leq n$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{i,k} \det(A_{i,k})$$

be the calculation of the determinant with respect to the $k^{\text{ième}}$ column.

Example 2.7. Let

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$$

$$A_{1,2} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 8 \\ 3 & 6 \end{pmatrix}$$

$$A_{2,2} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix}$$

$$A_{3,2} = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 5 \\ 2 & 8 \end{pmatrix}$$

Figure 2.2: Development with respect to the second column

So:

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i+2} a_{i,2} \det(A_{i,2}) \\
&= (-1)^{1+2} \cdot a_{1,2} \cdot \det(A_{1,2}) + (-1)^{2+2} \cdot a_{2,2} \cdot \det(A_{2,2}) + (-1)^{3+2} \cdot a_{3,2} \cdot \det(A_{3,2}) \\
&= (-1)^{1+2} \cdot 4 \cdot \begin{vmatrix} 2 & 8 \\ 3 & 6 \end{vmatrix} + (-1)^{2+2} \cdot 9 \cdot \begin{vmatrix} 1 & 5 \\ 3 & 6 \end{vmatrix} + (-1)^{3+2} \cdot 7 \cdot \begin{vmatrix} 1 & 5 \\ 2 & 8 \end{vmatrix} \\
&= (-1) \cdot 4 \cdot (-12) + 1 \cdot 9 \cdot (-9) + (-1) \cdot 7 \cdot (-2) \\
&= 48 - 81 + 14 \\
&= -19
\end{aligned}$$

2.3 Determinant of a triangular matrix

Corollary 2.8. The determinant of a triangular matrix is the product of its diagonal elements. I.e., let a triangular matrix be

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{n,n} \end{bmatrix}$$

then

$$\det(A) = a_{1,1} \cdot a_{2,2} \cdot \ldots \cdot a_{n,n}$$

Example 2.9. Let

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 9 & 8 \\ 0 & 0 & 6 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Let's expand this determinant with respect to the first column:

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i+2} a_{i,2} \det(A_{i,2}) \\
&= (-1)^{1+1} \cdot a_{1,1} \cdot \det(A_{1,1}) + (-1)^{2+1} \cdot a_{2,1} \cdot \det(A_{2,1}) + (-1)^{3+1} \cdot a_{3,1} \cdot \det(A_{3,1}) \\
&= (-1)^2 \cdot 1 \cdot \begin{vmatrix} 9 & 8 \\ 0 & 6 \end{vmatrix} + \underbrace{(-1)^3 \cdot 0 \cdot \begin{vmatrix} 4 & 5 \\ 0 & 6 \end{vmatrix}}_{=0} + \underbrace{(-1)^4 \cdot 0 \cdot \begin{vmatrix} 4 & 5 \\ 9 & 8 \end{vmatrix}}_{=0} \\
&= \underbrace{1}_{=a_{1,1}} \cdot \begin{vmatrix} 9 & 8 \\ 0 & 6 \end{vmatrix} \\
&= \det(\begin{bmatrix} 9 & 8 \\ 0 & 6 \end{bmatrix} =: B) \\
&= (-1)^{1+1} \cdot b_{1,1} \cdot \det(B_{1,1}) + (-1)^{2+1} \cdot b_{2,1} \cdot \det(B_{2,1}) \qquad \text{développement par rapport à la premiere colonne} \\
&= 1 \cdot \underbrace{9}_{a_{2,2}} \cdot |6| + \underbrace{(-1) \cdot 0 \cdot |8|}_{=0} \\
&= \underbrace{1}_{=a_{1,1}} \cdot \underbrace{9}_{=a_{2,2}} \cdot \underbrace{6}_{=a_{3,3}}
\end{aligned}$$

2.4 Cofactor matrix and adjoint matrix

First, let's recall the definition of $A_{i,j}$. It is a square matrix where the $i^{\text{ième}}$ row and the $j^{\text{ième}}$ column are removed. (See definition 2.3).

Definition 2.10. Let $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ be a square matrix. We denote

$$b_{i,j} = (-1)^{i+j} \det(A_{i,j})$$

Next, we denote the matrix

$$N = \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{bmatrix} = \operatorname{Com}(A)$$

The matrix N is called the **cofactor matrix** of A. Then, the **adjoint matrix** of A is defined as the transposed cofactor matrix:

$$A^* = N^T = \begin{bmatrix} b_{1,1} & \dots & b_{n,1} \\ \vdots & \ddots & \vdots \\ b_{1,n} & \dots & b_{n,n} \end{bmatrix}$$

Theorem 2.11. Let $A \in \mathcal{M}_n\mathbb{R}$ be a square matrix and A^* its adjoint matrix, then we have:

$$A^*A = AA^* = \det(A)I_n = \begin{bmatrix} \det(A) & 0 & 0 & \dots & 0 & 0 \\ 0 & \det(A) & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \det(A) \end{bmatrix}$$

What is the use of such a matrix?

2.5 Inverse Matrix

Theorem 2.12. Let $A \in \mathcal{M}_n(\mathbb{R})$ be a square matrix such that $\det(A) \neq 0$, then:

$$A^{-1} = \frac{1}{\det(A)} \cdot A^*$$

is the inverse matrix of A.

Corollary 2.13. If $A \in \mathcal{M}_n(\mathbb{R})$ is an invertible square matrix, then:

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

CHAPTER 3

Reduction of Endomorphisms

While writing this chapter, I was inspired by the videos of the *3blue1brown* channel which I advise you to watch, at least the playlist concerning linear algebra. The second source of inspiration was Joseph Grifone's book [2].

3.1 Introduction

In the previous chapter, we studied the notion of an orthonormal basis, the utilities of which are: simplification of coordinate calculations in a basis and calculation of a projection. This notion is one of the first steps towards the study of SVD[1] which is applied in several fields, e.g.: the reduction of image sizes.

In this chapter, we continue the study of bases in order to finally understand the SVD. We will study the reduction of endomorphisms, *to be more precise* diagonalization and triangularization. To begin: a small exercise:

Exercise. Calculate

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^{15} = \underbrace{\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \cdot \ldots \cdot \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}}_{\text{15 fois}}$$

This doesn't seem very easy, does it? At the end of this chapter, we will find a way to simplify the calculation and in the end we will solve this exercise.

We know from linear algebra that we can represent a matrix of a mapping in different bases, i.e. let $\{e_i\}$ be a basis of E and f a mapping. Then this mapping in the basis $\{e_i\}$ is represented:

$$A = M(f)_{e_i} = \|f(e_1), \ldots, f(e_n)\|$$

Let $\{e'_i\}$ be another basis of E, then we can represent the mapping f in this basis as well, let's denote: $P = P_{e_i \to e'_i}$ a change-of-basis matrix from the basisFrom $\{e_i\}$ to the base $\{e'_i\}$

$$A' = M(f)_{e'_i} = P^{-1}AP = \|f(e'_1), \ldots, f(e'_n)\|_{e'_i}$$

Definition 3.1. The matrix A is **diagonalizable** if there exists a similar matrix [a] A' that is diagonal:

$$A' = \begin{bmatrix} a_{1,1} & 0 & \ldots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & a_{n,n} \end{bmatrix}$$

[a] A is similar to A' if there exists a change-of-basis matrix P such that $A' = P^{-1}AP$

[1] Singular Value Decomposition

Definition 3.2. The matrix A is **triangulable** if there exists a similar triangular (upper/lower) matrix A'

$$A' = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \dots & 0 & a_{n,n} \end{bmatrix} \text{ or } A' = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 \\ a_{2,1} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

So the problems in this chapter that we are going to solve are:

1. Determine whether an endomorphism f is diagonalizable/triangulable, i.e., if there exists such a matrix A'.
2. Determine the change-of-basis matrix P and the matrix A'.

Throughout the chapter, we assume that the vector space E is of finite dimension.

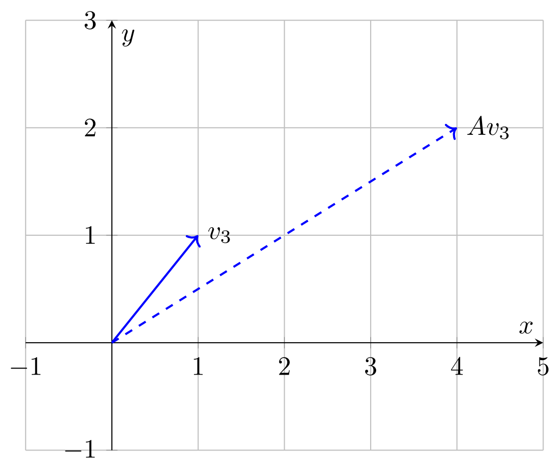
3.2 Eigenvectors - Eigenvectors

Let's start by clarifying the concept of a linear application and its matrix. Let's take for that the matrix from the exercise at the beginning of the chapter:

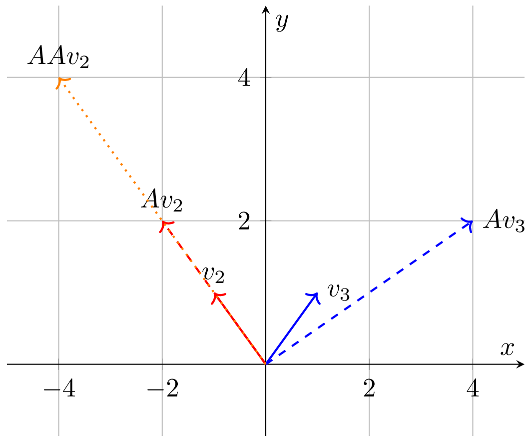
$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

This matrix transforms the vector space that we give it, or, to simplify, it transforms each vector of the vector space. Let's take a vector $v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, by applying A we obtain:

$$Av_3 = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$



We note that the vector Av_3 is no longer located on the same line as the vector v_3, which makes sense because if the vectors were on the same lines after a transformation, it would not make sense. On the other hand, sometimes there are cases when the vector applied to the matrix remains on the same line, for example the vector $v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, with $Av_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2v_2$



And this is not only the case for the vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$, by taking any vector generated by $v = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, we will obtain $Av = 2v$. Such vectors v and the scalars (here: 2) are called eigenvectors and eigenvalues respectively. So, we have the formal definition:

Definition 3.3. Let f be an endomorphism in E and a vector $v \in E$ is called an **eigenvector** of f if:

1. $v \neq 0$
2. There exists a real number λ such that $f(v) = \lambda v$

The scalar $\lambda \in \mathbb{R}$ is called the **eigenvalue** corresponding to v.

Intuition. Eigenvectors are vectors that, under the action of f, do not change direction, only length (not even always). This simplifies the calculation of such vectors. Can you calculate $A^3 v_3$? Not very easy, then the vector $A^3 v_2$?

$$Av_2 = 2v_2 \Rightarrow A^2 v_2 = 2 \cdot 2v_2 = 4v_2 \Rightarrow A^3 v_2 = 2 \cdot 4v_2 = 8v_2 = \begin{pmatrix} -8 \\ 8 \end{pmatrix}$$

That's cool, isn't it?

On the other hand, this is not the only use of eigenvectors, and we will come back to discuss it here, but first, how do we find such vectors?

3.3 Finding Eigenvalues

We are looking for vectors which, under the action of the endomorphism f, are scaled by a factor of $\lambda \in \mathbb{R}$, so we are supposed to solve this equation:

$$\begin{aligned}
& f(v) = \lambda v \\
\Leftrightarrow \quad & Av = \lambda v \quad \text{in matrix notation} \\
\Leftrightarrow \quad & Av = \lambda (Iv) \quad \text{where } I \text{ is an identity matrix} \\
\Leftrightarrow \quad & Av - \lambda I v = 0 \\
\Leftrightarrow \quad & (A - \lambda I)v = 0
\end{aligned}$$

Therefore, we must study the application $(A - \lambda I)$ and connect it to the notion of determinants. Recall: if the determinant of a matrix is non-zero, then that matrix (i.e., endomorphism) is injective. In our case, if $\det(A - \lambda I)$ were zero, the only vector v that would give $(A - \lambda I)v = 0$ was the zero vector $v = 0$ because $(A - \lambda I)$ is linear and (as we have supposed) injective.

On the other hand, according to the definition, eigenvectors are not zero, so the injective case is not suitable; therefore, to have eigenvectors, the application $(A - \lambda I)$must not be injective, which is equivalent to saying that

$\det(A - \lambda I) = 0$. So, we are supposed to calculate the following determinant:

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda \end{bmatrix}\right) = \begin{vmatrix} a_{1,1}-\lambda & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2}-\lambda & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n}-\lambda \end{vmatrix}$$

By expanding this determinant, we obtain an equation of the type:

$$(-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0$$

whose roots are the eigenvalues of f (remember: an eigenvalue is a factor λ). Don't focus too much on this equation for now, we'll come back to it.

Proposition 3.4. Let f be an endomorphism in a vector space E of finite dimension n and A the representative matrix of f in a basis of E. The eigenvalues of f are the roots of the polynomial:

$$P_f(\lambda) = \det(A - \lambda I)$$

This polynomial is called the **characteristic polynomial** of f.

Definition 3.5. The set of eigenvalues of f is called the **spectrum** of f and is denoted $\mathrm{Sp}_K(f)$ or $\mathrm{Sp}_K(A)$ if A is a matrix of f.

To clarify:

Example 3.6. Let f be an endomorphism in $\mathbb{R}^2$ whose representative matrix in the canonical basis is:

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Let's calculate its eigenvalues:

$$\begin{aligned} & & \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} v &= \lambda v \\ &\Leftrightarrow & \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} v - \lambda I v &= 0 \\ &\Leftrightarrow & \left(\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - \lambda I\right) v &= 0 \\ &\Rightarrow & \det\left(\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - \lambda I\right) &= 0 \\ &\Rightarrow & \det\left(\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - \lambda \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) &= 0 \\ &\Rightarrow & \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) &= 0 \\ & & &= (3-\lambda)(2-\lambda) = 0 \end{aligned}$$

We can clearly see that the solutions are: $\lambda_1 = 3$ and $\lambda_2 = 2$

We can find eigenvalues, nevertheless, we were looking for the eigenvectors. And we are there:

3.4 Finding Eigenvectors

Suppose that for $q \in \mathbb{N}^*$ we have already found q eigenvalues of a matrix $\{\lambda_1, \ldots, \lambda_q\}$, to find the eigenvectors, we still need to find the basis of:

$$\ker(A - \lambda_i I) \quad \forall i \in \{1, \ldots, q\}$$

which is equivalent to:

$$(A - \lambda_i I)\, v = 0 \quad \forall i \in \{1, \ldots, q\}$$

Example 3.7. Again, the matrix

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

in the canonical basis of $\mathbb{R}^2$. We have already found its eigenvalues: $\lambda_1 = 3$ and $\lambda_2 = 2$. So, let's find the vectors:

$$\begin{bmatrix} 3-\lambda_1 & 1 \\ 0 & 2-\lambda_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3-3 & 1 \\ 0 & 2-3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{cases} y = 0 \\ -y = 0 \\ x \in \mathbb{R} \end{cases}$$

Therefore, $\ker(A - 3I) = \begin{pmatrix} x \\ 0 \end{pmatrix} = \mathrm{Vect}(\begin{pmatrix} 1 \\ 0 \end{pmatrix})$. Here is our first eigenvector: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. For the second one:

$$\begin{bmatrix} 3-\lambda_2 & 1 \\ 0 & 2-\lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3-2 & 1 \\ 0 & 2-2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{cases} x + y = 0 \end{cases} \quad \Rightarrow \begin{cases} x = -y \end{cases}$$

Therefore, $\ker(A - 2I) = \begin{pmatrix} -y \\ y \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \mathrm{Vect}(\begin{pmatrix} -1 \\ 1 \end{pmatrix})$ and here is the second eigenvector: $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (it was our vector v_2 at the beginning of the chapter).

Finally, the useful property:

Proposition 3.8. Let $A \in \mathcal{M}_n(\mathbb{R})$ be a matrix with its eigenvectors: $\{\lambda_1, \ldots, \lambda_n\}$, then:

$$\mathrm{Tr}(A) = \lambda_1 + \ldots + \lambda_n$$
$$\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$$

3.5 Diagonalizable endomorphisms

Let's revisit the utility of eigenvectors. Let f be an endomorphism of E whose base is $\{e_1, \ldots, e_n\}$ and $\mathrm{Mat}_{e_i}(f) = A$ and the matrix of f in this base. Let's take the following example again:

Example 3.9. We have: $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ in the canonical basis $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We recall that we found two eigenvectors:

$$\begin{cases} v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{cases}$$

We notice that these two vectors are linearly independent and thus form a basis for $\mathbb{R}^2$. Let's try to change the basis of A using two methods:

1. We can calculate the coordinates of $f(v_1)$ and $f(v_2)$ in the basis $\{v_1, v_2\}$, we have:

$$f(v_1) = 3v_1 = 3 \cdot v_1 + 0 \cdot v_2$$
$$f(v_2) = 2v_2 = 0 \cdot v_1 + 2 \cdot v_2$$

   And thus $\mathrm{Mat}_{v_i}(f) = \|f(v_1), f(v_2)\|_{v_i} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$

2. We can calculate the change-of-basis matrix $P = P_{e_i \to v_i}$ from the basis $\{e_i\}$ to the basis $\{v_i\}$ and deduce the matrix of f in the new basis. We have:

$$\begin{cases} v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot e_1 + 0 \cdot e_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{e_i} \\ v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \cdot e_1 + 1 \cdot e_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}_{e_i} \end{cases}$$

   thus $P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (you can verify the calculation). And so:

$$A' = P^{-1}AP = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 3 & -2 \\ 0 & 2 \end{bmatrix}}_{AP} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

And there you have it, the magic: we found the diagonal matrix.

Next, let's generalize what we have done.

Definition 3.10. Let $\lambda \in K$, we denote:

$$E_\lambda := \{v \in E \mid f(v) = \lambda v\}$$

E_λ is a vector space of E called **eigen-space** corresponding to λ.

Remark 3.11.

1. If λ is not an eigenvalue of f, then $E_\lambda = \{0\}$
2. If λ is an eigenvalue, then:

$$E_\lambda = \{ \text{ eigenvectors associated with } \lambda\} \cup \{0\} \text{ and } \dim E_\lambda \geq 1$$

Proposition 3.12. Let $\lambda_1, \ldots, \lambda_p$ be pairwise distinct scalars. Then the eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_p}$ form a direct sum. In other words, if $\mathcal{B}_1, \ldots, \mathcal{B}_p$ are bases for $E_{\lambda_1}, \ldots, E_{\lambda_p}$, the family $\{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$ is linearly independent (but not necessarily a spanning set for E).

Proof. Let $E_{\lambda_1}, \ldots, E_{\lambda_p}$ be the eigenspaces associated with the eigenvalues $\lambda_1, \ldots, \lambda_p$ of an endomorphism f of a vector space E. We must show that these subspaces are in direct sum, meaning that if a vector belongs to their intersection, then it is zero.

Let's take an element v belonging to their sum, meaning it can be written in the form:

$$v = v_1 + v_2 + \cdots + v_p$$

with $v_i \in E_{\lambda_i}$ for all i.

Since each v_i is an eigenvector for f associated with λ_i, we have:

$$f(v_i) = \lambda_i v_i.$$

Let's apply f to the sum:

$$f(v) = f(v_1 + v_2 + \cdots + v_p) = f(v_1) + f(v_2) + \cdots + f(v_p).$$

Using the linearity of f, this gives:

$$f(v) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p.$$

However, v is also a combination of these same vectors:

$$v = v_1 + v_2 + \cdots + v_p.$$

Therefore, by rearranging:

$$(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p) - (v_1 + v_2 + \cdots + v_p) = 0.$$

Which gives:

$$(\lambda_1 - 1)v_1 + (\lambda_2 - 1)v_2 + \cdots + (\lambda_p - 1)v_p = 0.$$

Let's factor each term:

$$(\lambda_1 - \lambda)v_1 + (\lambda_2 - \lambda)v_2 + \cdots + (\lambda_p - \lambda)v_p = 0.$$

However, the λ_i are assumed to be distinct. We deduce that the coefficients are different, and that the sum is zero only if all v_i are zero (since eigenspaces are generally in direct sum).

Thus, $v = 0$, which proves that the eigenspaces are in direct sum. □

Thus, the eigenspaces are always in direct sum, but not necessarily equal to E:

$$E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p} \underset{\neq}{\subset} E$$

which holds if:

$$\dim E_{\lambda_1} + \ldots + \dim E_{\lambda_p} < \dim E$$

Theorem 3.13. Let f be an endomorphism in E and $\lambda_1, \ldots, \lambda_p$ its eigenvalues, then the following properties are equivalent:

1. f is diagonalizable
2. E is a direct sum of its eigenspaces: $E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p}$
3. $\dim E_{\lambda_1} + \ldots + \dim E_{\lambda_p} = \dim E$

Corollary 3.14. If f is an endomorphism of E with $\dim E = n$ and f admits n pairwise distinct eigenvalues, then f is diagonalizable.

But since the eigenvalues are the roots of the characteristic polynomial (see prop 3.4) we have:

Proposition 3.15. Let f be an endomorphism in E and λ an eigenvalue of order α (i.e., α is a root of $P_f(\lambda)$

of order α, i.e., $P_f(\lambda) = (X-\lambda)^\alpha Q(X)$). Then:

$$\dim E_\lambda \leq \alpha$$

Theorem 3.16. Let f be an endomorphism in E with $\dim E = n$. Then f is diagonalizable if and only if:

1. $P_f(X)$ is split, i.e:
$$P_f(X) = (-1)^n (X-\lambda_1)^{\alpha_1} \cdot \ldots \cdot (X-\lambda_p)^{\alpha_p}$$
(λ_i are the roots, hence the eigenvalues) and $\alpha_1 + \ldots + \alpha_p = n$. Thus, if the sum of the multiplicities of the roots is equal to the dimension of the vector space.
2. The dimensions of the eigenspaces are maximal, i.e $\forall i \in \{1, \ldots, p\}$
$$\dim E_{\lambda_i} = \alpha_i$$

Intuition. It's not always easy to understand the idea through characteristic polynomials, so another way to look at it is:

1. We find the eigenvalues: $\lambda_1, \ldots, \lambda_p$
2. Then we find the eigenspaces: $E_{\lambda_i} = \ker(f - \lambda_i I)$
3. We sum the dimensions: $\dim E_{\lambda_1} + \ldots + \dim E_{\lambda_p} =: d$.
   - If $d = \dim E$ i.e., if the sum of the dimensions is equal to the dimension of the space E, the eigenspaces span E and thus f is diagonalizable (because its matrix can be written in the basis of these eigenvectors).
   - Otherwise, the number of linearly independent eigenvectors is not sufficient to span E.

3.6 Applications

3.6.1 Calculation of Power

So, we're back where we started; I remind you of the exercise from the beginning of the chapter:

Exercise. Calculate

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^{15} = \underbrace{\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \cdot \ldots \cdot \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}}_{15 \text{ fois}}$$

Recall that the eigenvectors of A are:

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

which are linearly independent and span $\mathbb{R}^2$, thus forming a basis for $\mathbb{R}^2$. Therefore, we can express A in this new basis, and as we have already found:

$$A' = P^{-1}AP = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

in the basis (v_1, v_2) with the change-of-basis matrix:

$$P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Furthermore, by multiplying A' by A', we get:

$$A' \cdot A' = (P^{-1}AP)(P^{-1}AP) = P^{-1}A^2P = A'^2$$

hence

$$A'^n = P^{-1}A^nP \Rightarrow PA'^nP^{-1} = PP^{-1}A^nPP^{-1} = A^n$$

This allows us to first calculate the power of A':

$$A'^{15} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{15} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{13} = \begin{bmatrix} 3^2 & 0 \\ 0 & 2^2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{13} = \begin{bmatrix} 3^{15} & 0 \\ 0 & 2^{15} \end{bmatrix}$$

This is much easier than calculating A^15 directly, so now we just need to convert back to the canonical basis:

$$P \begin{bmatrix} 3^{15} & 0 \\ 0 & 2^{15} \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3^{15} & 0 \\ 0 & 2^{15} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3^{15} & 3^{15} - 2^{15} \\ 0 & 2^{15} \end{bmatrix}$$

What is very useful about diagonal matrices is that the power of such a matrix is equal to the same matrix with its diagonal elements raised to the power, i.e:

$$A' = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \Rightarrow A'^n = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^n \end{bmatrix}$$

Let's generalize: If $A \in \mathcal{M}_n(K)$ is diagonalizable (i.e., there exists P and A' such that $A' = P^{-1}AP$), then:

$$A^n = P(A'^n)P^{-1} = P \begin{bmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^n \end{bmatrix} P^{-1}$$

3.6.2 Resolution of a System of Recurrent Sequences

Let $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be two sequences such that:

$$\begin{cases} u_{n+1} = u_n - v_n \\ v_{n+1} = 2u_n + 4v_n \end{cases} \tag{3.1}$$

with $u_0 = 2$ and $v_n = 1$. Let $X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$, then the system 3.1 is written:

$$X_{n+1} = AX_n \quad \text{with} \quad A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

by recurrence we obtain:

$$X_n = A^nX_0 \quad \text{with } X_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

So, we are reduced to calculating the power of a matrix: A^n which we saw in section 3.6.1. You can check that there exists $P \in GL_2(\mathbb{R})$ s.t.

$$P = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix} \quad \text{with} \quad A = P \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} P^{-1}$$

and then

$$A^n = P \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} P^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} -2 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 2^n - 3^n & 2^n - 3^n \\ -2 \cdot 2^n + 2 \cdot 3^n & -2^n + 2 \cdot 3^n \end{pmatrix}$$

Whence

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} 2 \cdot 2^n - 3^n & 2^n - 3^n \\ -2 \cdot 2^n + 2 \cdot 3^n & -2^n + 2 \cdot 3^n \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \cdot 2^n - 2 \cdot 3^n + 2^n - 3^n \\ -4 \cdot 2^n + 4 \cdot 3^n - 2^n + 2 \cdot 3^n \end{pmatrix}$$

that is to say:

$$\begin{cases} u_n = 5 \cdot 2^n - 3 \cdot 3^n \\ v_n = -5 \cdot 2^n + 6 \cdot 3^n \end{cases}$$

3.6.3 Solving Differential Equations

Consider solving the differential system

$$\begin{cases} \dfrac{dx_1}{dt} = a_{11}x_1 + \cdots + a_{1n}x_n \\ \quad\vdots \\ \dfrac{dx_n}{dt} = a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases}$$

with $a_{ij} \in \mathbb{R}$ and $x_i : \mathbb{R} \to \mathbb{R}$ differentiable.

In matrix form, the system is written as:

$$\frac{dX}{dt} = AX, \quad \text{where} \quad A = (a_{ij}), \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{3.2}$$

Suppose A is diagonalizable. Then there exist A' a diagonal matrix and P an invertible matrix such that:

$$A' = P^{-1}AP.$$

If A is considered as the matrix of an endomorphism in the canonical basis, then A' is the matrix of f in the basis of eigenvectors $\{v_i\}$.
Likewise, X is the matrix of a vector $\vec{x}$ in the canonical basis and $X' = M(\vec{x})_{v_i}$ is related to X by

$$X' = P^{-1}X$$

> **Note.** Note! In this section X' does not describe the derivative, but a vector denoted X'!

By deriving this relation:

$$\frac{dX'}{dt} = P^{-1}\frac{dX}{dt}$$

(because A having constant coefficients, P will also have constant coefficients). Therefore:

$$\frac{dX'}{dt} = P^{-1}AX = \left(P^{-1}AP\right)X' = A'X'$$

The system 3.2 is therefore equivalent to the system

$$\frac{dX'}{dt} = A'X'$$

This system is easily integrated, because A' is diagonal.
Thus, we can solve the system $\frac{dX}{dt} = AX$ in the following way:

1. We diagonalize A. Let $A' = P^{-1}AP$ be a diagonal matrix similar to A;
2. we integrate the system $\frac{dX'}{dt} = A'X'$;
3. we return to X using $X = PX'$.

3.6.4 Example

Consider the system

$$\begin{cases} \dfrac{dx}{dt} = x - y \\ \dfrac{dy}{dt} = 2x + 4y \end{cases}$$

We have $A' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ and $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$

The system $\frac{dX'}{dt} = A'X'$ is written as:

$$\begin{cases} \dfrac{dx'}{dt} = 2x' \\ \dfrac{dy'}{dt} = 3y' \end{cases}$$

which immediately gives

$$\begin{cases} x' = C_1 e^{2t} \\ y' = C_2 e^{3t} \end{cases}$$

and thus, by reverting to X using $X = PX'$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} C_1 e^{2t} \\ C_2 e^{3t} \end{pmatrix} = \begin{pmatrix} C_1 e^{2t} + C_2 e^{3t} \\ -C_1 e^{2t} - 2C_2 e^{3t} \end{pmatrix}$$

that is to say:

$$\begin{cases} x = C_1 e^{2t} + C_2 e^{3t} \\ y = -C_1 e^{2t} - 2C_2 e^{3t} \end{cases}$$

3.7 Trigonalization

A matrix $A \in \mathcal{M}_n(K)$ is called upper triangular if it is of the form:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \dots & 0 & a_{n,n} \end{bmatrix}$$

respectively lower triangular:

$$A = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 \\ a_{2,1} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

Remark 3.17. Any upper triangular matrix A is similar to a lower triangular matrix.

Proof. Let A be an upper triangular matrix and f be the endomorphism of K^n which, in the basis $\{e_1, \dots, e_n\}$, is represented by the matrix A, then:

$$\begin{cases} f(e_1) = a_{1,1}e_1 \\ f(e_2) = a_{1,2}e_1 + a_{2,2}e_2 \\ \vdots \\ f(e_n) = a_{1,n}e_1 + a_{2,n}e_2 + \dots + a_{n,n}e_n \end{cases} \Leftrightarrow A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \dots & 0 & a_{n,n} \end{bmatrix}$$

Let's consider the basis

$$\varepsilon_1 = e_n, \quad \varepsilon_2 = e_{n-1}, \quad \ldots, \quad \varepsilon_n = e_1$$

then we have:

$$\begin{cases} f(\underbrace{\varepsilon_1}_{e_n}) == a_{1,n} \underbrace{\varepsilon_n}_{e_1} + a_{2,n} \underbrace{\varepsilon_{n-1}}_{e_2} + \ldots + a_{n,n} \underbrace{\varepsilon_1}_{e_n} \\ f(\underbrace{\varepsilon_2}_{e_{n-1}}) == a_{1,n-1} \underbrace{\varepsilon_n}_{e_1} + \ldots + a_{n-1,n-1} \underbrace{\varepsilon_2}_{e_{n-1}} \\ \vdots \\ f(\underbrace{\varepsilon_n}_{e_1}) = a_{1,1} \underbrace{\varepsilon_n}_{e_1} \end{cases}$$

thus

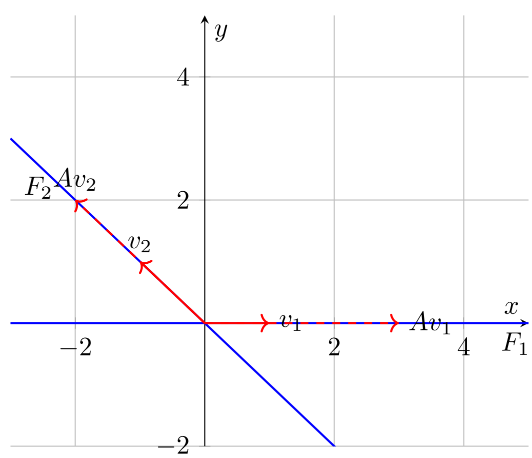
$$A' = M(f)_{\varepsilon_i} = \begin{bmatrix} a_{n,n} & & \ldots & 0 \\ a_{n-1,n} & a_{n-1,n-1} & \ldots & 0 \\ \vdots & & \ddots & \\ a_{1,n} & \ldots & & a_{1,1} \end{bmatrix}$$

□

3.7.1 The geometric intuition of diagonalization

Let's recall diagonalization. The matrix A representing the endomorphism f in $K^n = \text{Vect}(e_1, \ldots, e_n)$ is diagonalizable if there exist enough vector subspaces $\{F_1, \ldots, F_n\}$, each of dimension 1, such that $K^n = F_1 \oplus \ldots \oplus F_n$ and $\forall i \in \{1, \ldots, n\}, f(F_i) \subset F_i$ (a vector remains in the space after applying f). What can be seen geometrically:

Eigenvector Transformation



We already know that such an endomorphism is very useful, but it is not often that it can be diagonalized. Therefore, it would be useful to have something more general but still similar to diagonalization.

3.7.2 The Geometric Intuition of Trigonalization

The geometry of the trigonalizable endomorphism is similar yet still different. Let A be a representative matrix of the endomorphism f in K^n. It is trigonalizable if there exists a basis $\{v_1, \ldots, v_n\}$ of K^n, let's denote $F_1 = \text{Vect}(v_1), F_2 = \text{Vect}(v_1, v_2), \ldots, F_n = \text{Vect}(v_1, v_2, \ldots, v_n)$ such that

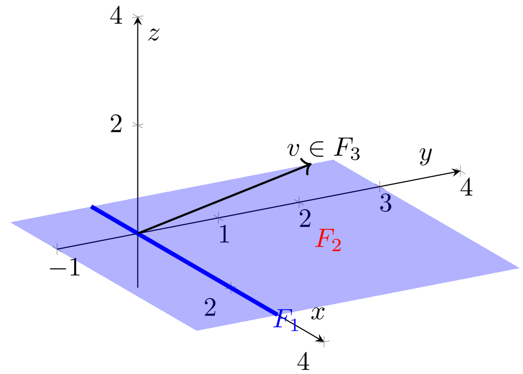
$$F_1 \subset F_2 \subset \ldots \subset F_n$$

and

$$\forall i \in \{1, \ldots, n\}, f(F_i) \subset F_i$$

Do you see the similarity? The endomorphism is stable by the subspace! The vector applied to f never leaves its subspace. Let's take the following matrix as an example:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \mathrm{Mat}(f)_{e_i}$$



As we have an intuition for trigonalizable endomorphisms, let's return to pure mathematics.

3.7.3 Theory

Theorem 3.18. An endomorphism is trigonalizable in K if and only if its characteristic polynomial splits in K.

This means that the characteristic polynomial has exactly n roots, where $n = \dim(E)$, and can be written as:

$$P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$$

with $\alpha_1 + \ldots + \alpha_p = n$

Proof. -

- Suppose the endomorphism f is trigonalizable and let $\{e_1, \ldots, e_n\}$ be a basis such that

$$M(f)_{e_i} = \begin{pmatrix} a_{1,1} & & & * \\ 0 & a_{2,2} & & \\ \vdots & & \ddots & \\ 0 & \ldots & 0 & a_{n,n} \end{pmatrix}$$

  We have:

$$P_f(X) = \det \begin{pmatrix} a_{1,1} - X & & & * \\ 0 & a_{2,2} - X & & \\ \vdots & & \ddots & \\ 0 & \ldots & 0 & a_{n,n} - X \end{pmatrix} = (a_{1,1} - X) \cdots (a_{n,n} - X)$$

  Thus, $P_f(X)$ is split (we can note that its roots are the eigenvalues of f).

- Suppose $P_f(X)$ is split and let us show by induction that f is trigonalizable.

  For $n = 1$, it is trivial.

  Suppose that the result holds for order $n - 1$. Since $P_f(X)$ is split, it admits at least one root $\lambda_1 \in K$

and thus an eigenvector $\varepsilon_1 \in E_{\lambda_1}$. Let's complete $\{\varepsilon_1\}$ to a basis $\{\varepsilon_1, \dots, \varepsilon_n\}$, so we have:

$$A = M(f)_{\varepsilon_i} = \begin{pmatrix} \lambda_1 & b_2 & \dots & b_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}, \quad \text{where: } B \in \mathcal{M}_{n-1}(K)$$

Let $F = \mathrm{Vect}(\varepsilon_2, \dots, \varepsilon_n)$ and $g : F \to F$ be the unique endomorphism of F such that $M(g)_{\varepsilon_2,\dots,\varepsilon_n} = B$, we have:

$$P_f(X) = \det(A - XI_n) = (\lambda_1 - X)\det(B - XI_{n-1}) = (\lambda_1 - X)P_g(X)$$

Since $P_f(X)$ is split, $P_g(X)$ is also split, and by the induction hypothesis, B is trigonalizable, so there exists a basis $\{v_2, \dots, v_n\}$ in which $M(g)_{v_2,\dots,v_n}$ is triangular, and thus the matrix of f in the basis $\{\varepsilon_1, v_2, \dots, v_n\}$ is triangular, so f is trigonalizable.

□

Corollary 3.19. Any matrix $A \in \mathcal{M}_n(\mathbb{C})$ is similar to a triangular matrix in $\mathcal{M}_n(\mathbb{C})$.

Intuition. According to the abstract algebra course, every polynomial in $\mathbb{C}$ is split.

Remark 3.20. -

1. If A is trigonalizable and A' is triangularly similar to A, then A' has its eigenvalues on the diagonals.
2. Any matrix $A \in \mathcal{M}_n(K)$ is trigonalizable over the closure K' of K. (e.g.: $A \in \mathcal{M}_n(\mathbb{R})$ is trigonalizable over $\mathbb{C}$).

Corollary 3.21. Let $A \in \mathcal{M}_n(K)$ have $\{\lambda_1, \dots, \lambda_n\}$ as its eigenvalues, then

$$\mathrm{Tr}(A) = \lambda_1 + \dots + \lambda_n$$
$$\det(A) = \lambda_1 \cdot \dots \cdot \lambda_n$$

Proof. $A' \in \mathcal{M}_n(K')$ is triangular and similar to A (recall: K' is the closure of K), so the eigenvalues are on the diagonal of A'. Now, similar matrices have the same trace and determinant, so $\mathrm{Tr}(A) = \mathrm{Tr}(A') = \lambda_1 + \dots + \lambda_n$ and $\det(A) = \det(A') = \lambda_1 \cdot \dots \cdot \lambda_n$. □

We will show the trigonalization process with the following example:

Example 3.22. Let the matrix

$$A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$$

. We have the characteristic polynomial $P_A(X) = -(X-1)^2(X+2)$ which is split over $\mathbb{R}$, so A is trigonalizable (according to Theorem 3.18). Therefore, if we consider A as an endomorphism in the canonical basis, we know that there exists a basis $\{v_i\}$ of $\mathbb{R}^3$ such that:

$$M(f)_{v_i} = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -2 \end{pmatrix}$$

I remind you that this means:

$$\begin{cases} f(v_1) = v_1 \\ f(v_2) = av_1 + v_2 \\ f(v_3) = bv_1 + cv_2 - 2v_3 \end{cases} \tag{3.3}$$

Let's start by finding v_1. We know that v_1 is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$, i.e., $(f - \mathrm{Id})v_1 = 0$. So, let's calculate $(A - I)v_1 = 0$ (in other words, we are looking for v_1 that spans $\ker(A - I)$):

$$(A - I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} -5x - 2z & = 0 \\ 5x + y + 2z & = 0 \end{cases}$$

Thus, we can choose $v_1 = \begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}$ (in other words, $\ker(A - I) = \mathrm{Vect}(\begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix})$).

Next, let's find v_2. According to 3.3,

$$\begin{aligned} & f(v_2) = av_1 + v_2 \\ \Rightarrow & f(v_2) - v_2 = av_1 \\ \Rightarrow & (f - I)v_2 = av_1 \\ \Rightarrow & (A - I)v_2 = av_1 \end{aligned}$$

Thus we have:

$$(A - I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = a\begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix} \Leftrightarrow \begin{cases} -5x - 2z = 2a \\ 5x + y + 2z = -5a \end{cases}$$

So, by taking $a = 1$, we get

$$\begin{cases} -5x - 2z = 2 \\ 5x + y + 2z = -5 \end{cases}$$

therefore $v_2 = \begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$ (just by solving the system).

For v_3, we have two choices:

1. either proceed similarly by solving the system,
2. or notice that there exists an eigenvector of A corresponding to the eigenvalue -2, i.e., $\exists v_3 \in \mathbb{R}^3$ such that $f(v_3) = -2v_3$. In this case, we can take this vector v_3 and thus set $b = c = 0$.

Remark 3.23. Why can we do this? Because for every eigenvalue of f, there always exists an eigenspace with multiplicity at least 1, and this applies to the eigenvalue -2 as well.

So, let's find v_3:

$$(A + 2I)v_3 = 0 \Leftrightarrow \begin{cases} -2x - 2z = 0 \\ 3y = 0 \end{cases}$$

thus we can take $v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

Consequently, the matrix A is similar to

$$A' = M(f)_{v_i} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

with the change-of-basis matrix:

$$P = \|v_1, v_2, v_3\| = \begin{pmatrix} 2 & -2 & 1 \\ 0 & -3 & 0 \\ -5 & 4 & -1 \end{pmatrix}$$

3.8 Annihilating Polynomials

In the previous sections, we learned that to know if a matrix is diagonalizable, we must study the eigenspaces, which is not always very easy and is not the fastest way. So, in this section we will see one of the other methods of studying diagonalizability, one of these methods is the study of annihilating polynomials.

Remark 3.24. In this section, I will not write out most of the proofs, but rather the intuition behind why it is true and why it works.

Definition 3.25. Let $f \in \mathbb{K}^n$ be an endomorphism. A polynomial $Q(X) \in K[X]$ is an **annihilating polynomial** of f if $Q(f) = 0$.

Example 3.26. Let f be a projection, then we know that $f^2 = f$, hence $f^2 - f = 0$, so $Q(X) = X^2 - X = X(X-1)$ is an annihilating polynomial of f.

What is important is that the annihilating polynomials are closely related to the eigenvalues:

Proposition 3.27. Let $Q(X)$ be an annihilating polynomial of f, then the eigenvalues of f appear among the roots of Q, i.e.:

$$\mathrm{Sp}(f) \subset \mathrm{Rac}(Q)$$

Proof. Let $Q(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0$ be an annihilating polynomial for f and λ an eigenvalue of f. Thus, $\exists v \neq 0 \in E$ such that $f(v) = \lambda v$, moreover:

$$Q(f) = a_n f^n + a_{n-1} f^{n-1} + \ldots + a_0 \,\mathrm{Id} = 0$$

Since $f(v) = \lambda v$, it follows that $f^2(v) = f(\lambda v) = \lambda^2 v$, whence $f^k(v) = \lambda^k v \ \forall k \in \mathbb{N}$. Then:

$$Q(f(v)) = 0 = (a_n f^n + a_{n-1} f^{n-1} + \ldots + a_0 \,\mathrm{Id})v = (a_n \lambda^n + a_{n-1} \lambda^{n-1} + \ldots + a_0 \,\mathrm{Id})v = 0$$

Since $v \neq 0$, it follows that $a_n \lambda^n + a_{n-1} \lambda^{n-1} + \ldots + a_0 \,\mathrm{Id} = 0$, whence λ is a root of Q. □

Note. However, the equality is not generally true; for example, $\mathrm{Id}^2 = \mathrm{Id}$, thus $Q(X) = X^2 - X = X(X-1)$ annuls Id with roots 0 and 1, but 0 is not an eigenvalue of Id.

Theorem 3.28. Cayley-Hamilton. Let $f \in K^n$ be an endomorphism and $P_f(X)$ its characteristic polynomial, then

$$P_f(f) = 0$$

In other words, the characteristic polynomial of an endomorphism is its annihilating polynomial.

Intuition. The characteristic polynomial describes the structure of f, i.e., what operations must be performed to lose at least one dimension; if factors of the form $(X - \lambda)^n$ are obtained, then one must apply $f(v) - \lambda v) = v_r$, and then to the result v_r again, i.e., $f(v_r) - \lambda v_r$, and we repeat n times (this occurs in the case of trigonalizable matrices).

The theorem remains true even in cases where the endomorphism is not trigonalizable, because we can choose the closure K' of the field K in which our endomorphism is defined, and it becomes trigonalizable (e.g., $\mathbb{C}$ for $\mathbb{R}$).

Furthermore, the characteristic polynomial gives us $\ker(P_f(X)) = E$, i.e., the vectors that become null under the action of $P_f(f)$. The interesting fact is that all vectors in E belong to this kernel, and thus $\forall v \in E$, $p_f(f)v = 0$, from which $p_f(f) = 0$.

Definition 3.29. Let Q be a split polynomial:

$$Q(X) = (X - a_1)^{\alpha_1} \cdots (X - a_r)^{\alpha_r}$$

The polynomial

$$Q_1 = (X - a_1) \cdots (X - a_r)$$

is called the **radical** of Q (i.e., a split polynomial (the same polynomial but without powers next to the parentheses)).

Furthermore, $Q_1 \mid Q$ i.e., the radical of a polynomial divides the polynomial itself.

Proposition 3.30. Let f be an endomorphism and

$$P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$$

its characteristic polynomial. Then, if f is diagonalizable, the radical Q_1 also annihilates f, i.e.

$$Q_1(f) = (f - \lambda_1) \cdots (f - \lambda_r) = 0$$

Intuition. I will provide the intuition behind the proof. If f is diagonalizable with a characteristic polynomial

$$P_f(X) = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_p)^{\alpha_p}$$

with $r := \alpha_i > 1$, this does <u>not mean</u> that one must apply $(f - \lambda_i \operatorname{Id})$ r times to reduce the dimension as in the case of trigonalizable matrices, but rather that E_{λ_i}, the eigenspace for the eigenvalue λ_i, has dimension $\alpha_i = r$, and therefore $\forall v \in E_{\lambda_i}, f(v) = \lambda_i v$.

Since $E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p}$, if $v \in E$, then $\exists i \in \{1, \ldots, p\}$ such that $v \in E_{\lambda_i}$, and thus $f(v) - \lambda_i v = 0$, i.e., $(f - \lambda_i \operatorname{Id})(v) = 0$. Hence, the radical of P_f annihilates f.

3.9 The Kernel Lemma

Lemma 3.31. of Kernels Let $f \in K^n$ be an endomorphism and

$$Q(X) = Q_1(X) \cdots Q_p(X)$$

a polynomial factored into a product of pairwise coprime polynomials. If $Q(f) = 0$ then:

$$E = \operatorname{Ker} Q_1(f) \oplus \ldots \oplus \operatorname{Ker} Q_p(f)$$

Intuition. Since $Q(f) = 0$, therefore $\forall v \in E, Q(f)(v) = 0$ which means $\operatorname{Ker}(Q(f)) = E$. $\exists v_1, \ldots, v_p$ such that $v = v_1 + \ldots + v_p$. However, all polynomials are pairwise coprime, so only one of them annihilates v_i, thus $v_i \in \operatorname{Ker} Q_i(f)$ and this remains true for all $v_1, \ldots, v_p$. And since the polynomials are coprime, if $k \neq j$ and $Q_k(v_i) = 0$, then $Q_j(v_i) \neq 0$ because Q_j and Q_k are different. Therefore, $\forall i, j \operatorname{Ker} Q_i \cap \operatorname{Ker} Q_j = \{0\}$.

Remark 3.32. Let's revisit the example of f which is a projection, thus $f^2 - f = 0$ and $Q(X) = X^2 - X = X(X-1)$ annihilates f. Now X and $X - 1$ are coprime, then

$$E = \operatorname{Ker} f \oplus \operatorname{Ker}(f - \operatorname{Id})$$

More generally, let f be an endomorphism and $Q(X) = (X - \lambda_1) \cdots (X - \lambda_p)$ such that $Q(f) = 0$, we have:

$$E = \underbrace{\operatorname{Ker}(f - \lambda_1 \operatorname{Id})}_{E_{\lambda_1}} \oplus \ldots \oplus \underbrace{\operatorname{Ker}(f - \lambda_p \operatorname{Id})}_{E_{\lambda_p}}$$

Of course, $\lambda_i \neq \lambda_j$. And thus f is diagonalizable because it is a direct sum of these eigenspaces.

Corollary 3.33. An endomorphism f is diagonalizable if and only if there exists an annihilating polynomial Q of f that is split and has only simple roots [a]

[a] scindé: $(X - \lambda_i)^{\alpha_i}$ - X est à la puissance 1! racines simples: si $\alpha_i = 1$ aussi i.e les facteurs $(X - \lambda)$ sont à la puissance 1!

3.10 Finding Annihilating Polynomials. Minimal Polynomial

Definition 3.34. A **minimal polynomial** of f, denoted $m_f(X)$, is defined as the monic polynomial [a] that annihilates f and has the smallest degree.

[a] i.e de coefficient 1 du terme du plus haut degré, i.e: $1 * X^n + a_{n-1}X^{n-1} + \ldots + a_0$

Proposition 3.35. The annihilating polynomials of f are of the form:

$$Q(X) = A(X) m_f(X) \quad \text{with} \quad A(X) \in K[X]$$

i.e., $m_f(X)$ divides $Q(X)$.

Proposition 3.36. The roots of the minimal polynomial $m_f(X)$ are exactly the roots of the characteristic polynomial $P_f(X)$, i.e., the eigenvalues.

Proof. We know that $P_f(X) = A(X) m_f(X)$ so if λ is a root of $m_f(X)$, then it is also a root of $P_f(X)$. Conversely, if λ is a root of $P_f(X)$ then it is an eigenvalue, however $m_f(X)$ annihilates f, therefore λ is also a root of $m_f(X)$. □

Theorem 3.37. An endomorphism f is diagonalizable if and only if its minimal polynomial is split and all its roots are simple.

Example 3.38. 1. $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. $P_A(X) = -(X-1)(X+2)^2$, so we have two possibilities:

- $m_A(X) = (X-1)(X+2)$ - so A is diagonalizable
- $m_A(X) = (X-1)(X+2)^2$ - so A is not diagonalizable

Let's calculate:

$$(A-I)(A+2I) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Therefore, $m_f(X) = (X-1)(X+2)$ and thus A is diagonalizable.

2. $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$. We have: $P_A(X) = -(X-1)(X-2)^2$, thus:

$$m_A(X) = \begin{cases} (X-1)(X-2) & \text{i.e } A \text{ diagonalisable} \\ (X-1)(X-2)^2 & \text{i.e } A \text{ pas diagonalisable} \end{cases}$$

Let's calculate:

$$(A-I)(A-2I) = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 0 & -2 & 2 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence $m_A(X) \neq (X-1)(X-2)$ and thus A is not diagonalizable.

APPENDIX A

Reminders of Linear Algebra concepts

A.1 Matrices

A.1.1 Multiplication of matrices

Definition A.1. Let $A \in \mathcal{M}_{p,n}(\mathbb{R})$ and $B \in \mathcal{M}_{n,q}(\mathbb{R})$ such that $A = (a_{j,i})$ and $B = (b_{m,k})$, then:

$$AB = C = (c_{j,k} = \sum_{i=1}^{n} a_{j,i} b_{i,k})$$

A.1.2 The trace

Definition A.2. The trace of the $n \times n$ square matrix A, denoted $\mathrm{tr}(A)$, is the sum of the diagonal elements

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

where a_{ii} are diagonal elements of the matrix A.

Property. of the trace.

- Linearity:

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$$

$$\mathrm{tr}(cA) = c\mathrm{tr}(A), \quad c \in \mathbb{R} \text{ (ou } \mathbb{C})$$

- Transposed:

$$\mathrm{tr}(A) = \mathrm{tr}(A^T)$$

- Multiplication of matrices:

$$\mathrm{tr}(AB) = \mathrm{tr}(BA), \quad (\text{si } A \text{ et } B \text{ are of size } n \times n)$$

  However, the trace is not distributive over multiplication:

$$\mathrm{tr}(ABC) \neq \mathrm{tr}(A)\mathrm{tr}(BC)$$

- Eigenvalues:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$$

  where λ_i are the eigenvalues of A. This makes the trace an important tool in spectral analysis.

- Trace of the Identity Matrix

$$\mathrm{tr}(I_n) = n$$

since all the diagonal elements are equal to 1.

Example A.3. For

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

the trace is:

$$\mathrm{tr}(A) = 3 + 5 + 9 = 17$$

Example A.4. If

$$B = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$$

then

$$\mathrm{tr}(B + C) = \mathrm{tr}\begin{bmatrix} 6 & 3 \\ 1 & 8 \end{bmatrix} = 6 + 8 = 14$$

which corresponds well to

$$\mathrm{tr}(B) + \mathrm{tr}(C) = (2 + 3) + (4 + 5) = 14$$

thus confirming linearity.

Bibliography

[1] Johannes Anschütz. *Algèbre linéaire 2 (OLMA252)*. 2024-2025.

[2] Grifone Joseph. *Algèbre linéaire*. fre. 4e édition. Toulouse: Cépaduès Éditions , DL 2011, 2011. ISBN: 978-2-85428-962-6.

[3] Yehor Korotenko. *sci-trans-git*. Version 0.2.0-alpha. 2025. DOI: 10.5281/zenodo.15775111. URL: https://github.com/DobbiKov/sci-trans-git.