

# Modélisation de Données Complexes

Matthieu Boyer

Cours TalENS 1 2025-2026



## Table des matières

<b>1</b>	<b>Calcul vectoriel</b>	<b>1</b>
1.1	Espaces euclidiens . . . . .	1
1.2	Applications sur les espaces euclidiens . . . . .	3
<b>2</b>	<b>Immersion euclidiennes</b>	<b>3</b>
2.1	Immersion . . . . .	3
2.2	Métrique angulaire . . . . .	3
2.3	Métriques adaptées . . . . .	3
<b>3</b>	<b>Algorithmes sur les immersions</b>	<b>3</b>
3.1	PCA . . . . .	3
3.2	t-SNE . . . . .	3

Ce polycopié ne doit pas être vu comme un remplacement au cours, mais simplement comme un résumé du contenu qui permet d'être sûr de ne rien manquer. On ne parlera ici des espaces vectoriels que dans un cadre très restreint adapté à ce qu'on va vouloir en faire. En particulier, on ne fera pas d'algèbre linéaire, on ne parlera pas de base quelconque ni d'isomorphismes. Je vous renvoie au cours donné avec Clément Allard l'an dernier pour plus de détails et des preuves.

## 1 Calcul vectoriel

### 1.1 Espaces euclidiens

Autour de 300 avant notre ère, Euclide, mathématicien grec écrit les Éléments, un traité de géométrie qui construit le monde autour de 5 axiomes (postulats indémontrables par les autres), sous forme de constructions géométriques réalisables :

1. On peut tracer une ligne droite de tout point à tout autre point ;
2. On peut étendre une ligne droite finie continuellement en une ligne droite ;
3. On peut décrire un cercle de tout centre avec tout rayon ;
4. On suppose que tous les angles droits sont égaux ;

5. On suppose qu'étant donnée une droite et un point n'appartenant pas à la droite, il y a exactement une droite qui passe par ce point et qui ne croise pas la droite de départ.

Les espaces vectoriels euclidiens sont une manière algébrique de décrire un espace vérifiant les axiomes d'Euclide.

**Définition 1.1** Un espace vectoriel (réel) est un ensemble  $E$  de points (appelés vecteurs) munis d'une addition interne associative commutative unitale et d'une multiplication externe distributive sur l'addition.

Autrement dit, dans l'espace  $E$  :

- Il existe un vecteur origine ou vecteur nul, dénoté  $0_E$  ou simplement  $0$  ;
- L'addition de deux vecteurs se fait par la relation de Chasles<sup>1</sup> ;
- La multiplication par  $\lambda \in \mathbb{R}$  d'un vecteur  $x$  est l'agrandissement/le rétrécissement de  $x$  par un facteur  $\lambda$ .

Il n'y a pas de "points" dans  $E$ , mais on peut les construire en voyant un point  $P$  comme le vecteur partant de  $0$  et allant jusqu'à  $P$ .

**R** Il est important de noter que cette définition ne présume rien sur le nombre de vecteurs, ni sur ce qu'ils représentent : pour tout ensemble  $X$ , l'ensemble des fonctions de  $X$  dans  $\mathbb{R}$  est un espace vectoriel (muni de l'addition point à point et de la multiplication du résultat). On va beaucoup, dans la suite, faire appel à l'intuition géométrique des espaces vectoriels, mais il nous est utile d'avoir une définition très générale.

**Définition 1.2** Une base d'un espace vectoriel  $E$  est une famille  $(e_i)_{i \in I}$  d'éléments de  $E$  telle que tout vecteur  $x$  de  $E$  s'écrit de manière unique comme une somme finie de  $e_i$ . On dit que  $E$  est de dimension finie si  $E$  possède une base finie.

Tous les espaces vectoriels possèdent une base, si l'on accepte l'axiome du choix (ou plutôt le lemme de Zorn, équivalent). Tous les espaces vectoriels ne possèdent pas de base finie, comme c'est par exemple le cas de l'espace des fonctions. Toutes les bases d'un espace vectoriel ont même cardinal, appelé dimension de l'espace.

**Définition 1.3** L'espace engendré par une famille  $(e_i)$  est l'ensemble  $\text{Vect}(e_i)$  des combinaisons linéaires des  $(e_i)$ .

**R** Être une base, ce n'est pas juste engendrer  $E$  (dans ce cas, la famille est génératrice), c'est engendrer  $E$  sans avoir de degré de liberté superflu. Cela revient à dire qu'il n'existe pas d'indice  $k \in I$  tel que  $e_k \in \text{Vect}(e_i)_{i \in I \setminus \{k\}}$ , et dans ce cas là, on dit que la famille est libre. Dans la suite, on ne considèrera que des familles libres (sauf mention contraire).

L'espace engendré par un vecteur est une droite, celui engendré par 2 vecteurs est un plan et celui engendré par  $n - 1$  vecteurs dans un espace de dimension  $n$  est appelé un hyperplan.

**Définition 1.4** Une application  $\varphi : E \rightarrow F$  est dite linéaire si pour tout  $x, y \in E$ , pour tout  $\lambda \in \mathbb{R} : \varphi(\lambda x + y) = \lambda \varphi(x) + \varphi(y)$ . Une application  $\psi : E \times F \rightarrow G$  est dite bilinéaire si pour tout  $x \in E$  sa restriction à  $F$  est linéaire (et de même pour tout  $y \in F$ ).

**Définition 1.5** Un produit scalaire sur  $E$  est une application bilinéaire symétrique définie positive de  $E \times E$  dans  $\mathbb{R}$ . Sa norme engendrée est l'application  $\|\cdot\| : x \mapsto \sqrt{\langle x, x \rangle}$

Symétrique signifie que  $\langle x, y \rangle = \langle y, x \rangle$ , positive signifie que  $\langle x, x \rangle \geq 0$  et définie signifie que  $\langle x, x \rangle = 0 \Rightarrow x = 0$ .

**Définition 1.6** Deux vecteurs sont orthogonaux si leur produit scalaire est nul. Une base est orthonormale si tous ses vecteurs sont orthogonaux deux à deux et sont de norme 1.

Tout espace vectoriel admet une base orthonormale.

1. Qui est un théorème de géométrie affine et non vectorielle.

**Proposition 1.1** Dans le cas où  $E$  admet une base orthonormée  $(e_i)$ , si  $x = \sum x_i e_i$  et  $y = \sum y_i e_i$  alors :

$$\langle x, y \rangle = \sum x_i y_i$$

Toutes les définitions ci-dessus sont une extension simple de ce qui a été vu en cours pour le produit scalaire en dimension 2 et 3.

**Théorème 1.1 — Pythagore** Si la famille des  $u_i$  est orthogonale :

$$\left\| \sum u_i \right\|^2 = \sum \|u_i\|^2$$

## 1.2 Applications sur les espaces euclidiens

## 2 Immersions euclidiennes

### 2.1 Immersion

### 2.2 Métrique angulaire

### 2.3 Métriques adaptées

## 3 Algorithmes sur les immersions

### 3.1 PCA

### 3.2 t-SNE