UNIVERSITE de HAUTE ALSACE - FACULTE DES SCIENCES et TECHNIQUES

Licence 1 Mathématiques

Mathématiques : ALGEBRE LINEAIRE

Elisabeth REMM

Chapitre 7

L'espace vectoriel euclidien $\mathbb{R}^n$

Table des matières

1. L'espace vectoriel réel $\mathbb{R}^n$

Dans tout ce chapitre nous allons nous intéresser à la géométrie vectorielle dans le plan et dans l'espace, c'est-à-dire à l'étude vectorielle des plans et des droites dans $\mathbb{R}^3$. Nous affinerons notre étude à l'aide d'un outil non plus linéaire mais bilinéaire en un sens que nous préciserons, le produit scalaire. Cet outil permet de faire des mesures de longueur, d'angle dans des sous-espaces vectoriels de $\mathbb{R}^n$.

1.1. **La base canonique de $\mathbb{R}^n$.** Par définition l'ensemble $\mathbb{R}^n$ est l'ensemble des n-uples de nombres réels :

$$\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n);\ x_i \in \mathbb{R}, i = 1, \cdots, n\}.$$

Nous identifierons naturellement $\mathbb{R}^1$ et $\mathbb{R}$. Dans le premier chapitre nous avons rappelé que $\mathbb{R}^n$ était muni d'une structure d'espace vectoriel réel, les opérations concernées étant

(1) l'addition : si $\overrightarrow{X} = (x_1, x_2, \cdots, x_n)$ et $\overrightarrow{Y} = (y_1, y_2, \cdots, y_n)$ sont deux éléments de $\mathbb{R}^n$ alors
$$\overrightarrow{X} + \overrightarrow{Y} = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n),$$

(2) la multiplication externe par un scalaire : si $\alpha \in \mathbb{R}$, alors
$$\alpha\overrightarrow{X} = \alpha(x_1, x_2, \cdots, x_n) = (\alpha x_1, \alpha x_2, \cdots, \alpha x_n).$$

Le vecteur nul $(0, 0, \cdots, 0)$ sera noté $\overrightarrow{0}$.

Considérons dans $\mathbb{R}^n$ les vecteurs
$$\overrightarrow{e_1} = (1, 0, \cdots, 0),\ \overrightarrow{e_2} = (0, 1, 0, \cdots, 0),\ \cdots, \overrightarrow{e_n} = (0, 0, \cdots, 0, 1).$$

Si $\overrightarrow{X} = (x_1, x_2, \cdots, x_n)$ il se décompose de manière unique
$$\overrightarrow{X} = x_1\overrightarrow{e_1} + x_2\overrightarrow{e_2} + \cdots + x_n\overrightarrow{e_n}$$
ce qui montre que $\mathcal{B} = \{\overrightarrow{e_1}, \overrightarrow{e_2}, \cdots, \overrightarrow{e_n}\}$ est une base de $\mathbb{R}^n$ et donc que $\mathbb{R}^n$ est bien de dimension n. Cette base "privilégiée" sera appelée base canonique de $\mathbb{R}^n$. Ainsi, si $\overrightarrow{X} = (x_1, x_2, \cdots, x_n)$, les scalaires x_i sont les composantes de $\overrightarrow{X}$ dans la base canonique.

1.2. **Les applications linéaires $f : \mathbb{R}^n \to \mathbb{R}^p$.** Considérons les deux espaces vectoriels réels $\mathbb{R}^n$ et $\mathbb{R}^p$. Si $f : \mathbb{R}^n \to \mathbb{R}^p$ est une application linéaire, elle se présente sous la forme
$$f(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_p).$$

Posons $f_i(x_1, x_2, \cdots, x_n) = y_i$ pour $i = 1, 2, \cdots, p$. Rappelons que l'on appelle forme linéaire, toute application linéaire à valeurs dans le corps de base. Ainsi, chacune des applications f_i est une forme linéaire $f_i : \mathbb{R}^n \to \mathbb{R}$. Ainsi f apparaît comme la composition des p formes linéaires :
$$f = (f_1, f_2, \cdots, f_p)$$
et donc l'étude de f se ramène à l'étude des applications linéaires de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

Etudions ces formes linéaires. Soit
$$f : \mathbb{R}^n \to \mathbb{R}$$
une forme linéaire et posons $f(x_1, x_2, \cdots, x_n) = y$. Comme f est linéaire, nous pouvons écrire
$$f(x_1, x_2, \cdots, x_n) = x_1 f(1, 0, \cdots, 0) + x_2 f(0, 1, , 0, \cdots, 0) + x_n f(0, 0, \cdots, 1)$$
soit
$$f(x_1, x_2, \cdots, x_n) = x_1 f(\overrightarrow{e_1}) + x_2 f(\overrightarrow{e_2}) + \cdots + x_n f(\overrightarrow{e_n}).$$

Nous retrouvons là le fait que f est entièrement définie par les images des vecteurs de base, ici la base canonique. Posons
$$f(\overrightarrow{e_1}) = a_1, f(\overrightarrow{e_2}) = a_2, \cdots, f(\overrightarrow{e_n}) = a_n.$$

Alors
$$f(x_1, x_2, \cdots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n.$$

Proposition 1. *Toute forme linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}$ s'écrit sous la forme*

$$f(x_1, x_2, \cdots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

avec $a_i = f(\overrightarrow{e_i}) \in \mathbb{R}$.

Nous en déduisons

Corollaire 1. *Toute application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ s'écrit sous la forme $f = (f_1, \cdots, f_p)$ où les f_i sont des formes linéaires $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ qui s'écrivent donc sous la forme*

$$\left\{\begin{array}{l} f_1(x_1, x_2, \cdots, x_n) = a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ f_2(x_1, x_2, \cdots, x_n) = a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \cdots \\ f_p(x_1, x_2, \cdots, x_n) = a_{p,1}x_1 + a_{p,2}x_2 + \cdots + a_{p,n}x_n \end{array}\right.$$

Nous déduisons que la matrice de l'application linéaire f relative aux bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,n} \end{pmatrix}$$

L'expression analytique de f s'écrit ainsi

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_p \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

1.3. **Les sous-espaces vectoriels de $\mathbb{R}^n$.** Soit F un sous-espace vectoriel de $\mathbb{R}^n$. Sa dimension p est inférieure à n et rappelons que si $p = n$ alors $F = \mathbb{R}^n$. Du théorème de la base incomplète nous déduisons directement le résultat suivant :

Proposition 2. *Tout sous-espace vectoriel F de $\mathbb{R}^n$ est défini comme l'ensemble des solutions d'un système linéaire à n variables*

$$\left\{\begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = 0 \\ \cdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = 0 \end{array}\right.$$

La résolution de tels systèmes linéaires a déjà été vue. Rappelons toutefois, qu'une manière un peu longue mais algorithmique de résoudre un tel système est la méthode du pivot. A ce niveau là, il serait intéressant de concevoir un programme sur PYTHON de résolution des systèmes linéaires.

Notons que F n'est certainement pas défini par un unique système linéaire. Ainsi au système précédent, on peut rajouter autant de combinaisons linéaires de lignes de ce système, la solution sera toujours F. Autrement dit, l'entier m apparaissant dans ce système n'est pas défini. Nous allons voir comment trouver le plus petit entier, qui lui aura un sens auprès de F.

Considérons le système linéaire précédent définissant le sous-espace vectoriel F. La matrice de ce système

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{p,2} & \cdots & a_{m,n} \end{pmatrix}$$

peut être vue comme la matrice d'une application linéaire $f : \mathbb{R}^n \to \mathbb{R}^m$ écrite dans les bases canoniques. Si p est la dimension de F, alors par définition de cette matrice, comme F est le noyau de f, nous avons $p = \dim \ker f$ et le théorème noyau image implique

$$\mathrm{rg}(f) = n - p.$$

Ce rang correspond au nombre d'équations linéaires indépendantes dans le système linéaire de définition de f.

Exemples : les sous-espaces vectoriels de $\mathbb{R}3$: Droites et plans vectoriels dans $\mathbb{R}^3$

Un plan vectoriel dans $\mathbb{R}^3$ est par définition un sous-espace vectoriel $\mathcal{P}$ de $\mathbb{R}^3$ de dimension 2. D'après la remarque ci-dessus, il est défini comme le noyau d'une application linéaire de rang 1 et donc par une équation linéaire

$$a_1x_1 + a_2x_2 + a_3x_3 = 0$$

dont au moins un des coefficients est non nul. Supposons par exemple que ce soit a_1. Nous pouvons écrire

$$x_1 = -\frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3$$

et le plan $\mathcal{P}$ est l'ensemble des vecteurs qui s'écrivent

$$\overrightarrow{X} = (-\frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3, x_2, x_3).$$

Une base de $\mathcal{P}$ est donnée par la famille

$$\{\overrightarrow{v_1} = (-\frac{a_2}{a_1}, 1, 0), \overrightarrow{v_2} = (-\frac{a_3}{a_1}, 0, 1)\}.$$

Une droite vectorielle de $\mathbb{R}^3$ est par définition un sous-espace vectoriel $\mathcal{D}$ de dimension 1. Elle est donc définie par un système linéaire de rang 2 à trois variables :

$$\left\{ \begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 = 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 = 0 \end{array} \right.$$

les deux équations étant indépendantes. La résolution de ce système permet de trouver une base de $\mathcal{D}$. Pour vérifier que ces deux équations sont bien indépendantes, plusieurs approches sont à notre disposition. La plus simple est de résoudre ce système, La solution doit s'écrire qu'avec un seul paramètre. La deuxième, plus sophistiquée mais plus facile à généraliser consiste à regarder tous les déterminants d'ordre 2 de la matrice

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}$$

du système, à savoir

$$\begin{cases} \det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}, \\ \det \begin{pmatrix} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{pmatrix} = a_{1,1}a_{2,3} - a_{1,3}a_{2,1}, \\ \det \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{pmatrix} = a_{1,2}a_{2,3} - a_{1,3}a_{3,1}, \end{cases}$$

et le rang du système est égal à 2 si l'un de ces déterminants est non nul. Supposons par exemple que ce soit le premier, soit $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$. La résolution du système donne alors

$$x_1 = -\frac{a_{1,3}a_{2,2} - a_{2,3}a_{1,2}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}x_3, \; x_2 = -\frac{a_{1,3}a_{2,1} - a_{2,3}a_{1,1}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}x_3$$

et tout vecteur de $\mathcal{D}$ s'écrit

$$\left(-\frac{a_{1,3}a_{2,2} - a_{2,3}a_{1,2}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}, -\frac{a_{1,3}a_{2,1} - a_{2,3}a_{1,1}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}, 1\right) x_3.$$

Une base de $\mathcal{D}$ est donnée par

$$\left\{\left(-\frac{a_{1,3}a_{2,2} - a_{2,3}a_{1,2}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}, -\frac{a_{1,3}a_{2,1} - a_{2,3}a_{1,1}}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}, 1\right)\right\}.$$

Remarque. La donnée d'une base d'un plan vectoriel de $\mathbb{R}^3$ permet inversement de retrouver une équation linéaire de ce plan. Il est bon de noter que cette équation n'est pas unique : si $a_1x_1 + a_2x_2 + a_3x_3 = 0$ est une équation définissant le plan, alors pour tout $\lambda \neq 0$, $\lambda a_1x_1 + \lambda a_2x_2 + \lambda a_3x_3 = 0$ est également une équation linéaire définissant ce plan. Il en est de même pour une droite vectorielle de $\mathbb{R}^3$. La donnée d'une base, constituée d'un vecteur non nul de $\mathbb{R}^3$ permet de retrouver le système de rang 2 définissant cette droite. Là aussi les équations formant le système ne sont pas uniques, tout système équivalent est également un système linéaire définissant cette droite.

2. Projection sur un sous-espace

Nous avons étudié précédemment la notion générale de projection dans un espace vectoriel (de dimension finie) et défini une projection dans cette espace comme un endomorphisme p vérifiant $p \circ p = p$. Rappelons que cette identité est équivalente à $E = \ker(p) \oplus \Im(p)$ et précisé que la projection était alors la projection de E "parallèlement" à $\ker(p)$. Nous allons illustrer plus géométriquement ces notions dans le cadre classique de $\mathbb{R}^n$.

2.1. **Projection le long d'un supplémentaire.** Soit F_1 un sous-espace vectoriel de $\mathbb{R}^n$. Choisissons un sous-espace supplémentaire F_2 (celui ci n'est évidemment pas unique, on en choisit un). Nous avons donc

$$\mathbb{R}^n = F_1 \oplus F_2.$$

Tout vecteur $\overrightarrow{X}$ s'écrit de manière unique

$$\overrightarrow{X} = \overrightarrow{X_1} + \overrightarrow{X_2}$$

avec $\overrightarrow{X_1} \in F_1$ et $\overrightarrow{X_2} \in F_2$.

Définition 1. *Le vecteur $\overrightarrow{X_1} \in F_1$ est appelé le projeté du vecteur $\overrightarrow{X}$ sur le sous-espace F_1 le long du supplémentaire F_2.*

Il est clair que ce vecteur F_1 dépend du choix de F_2.

Exemple : Projection d'un vecteur sur un plan

Considérons dans l'espace $\mathbb{R}^3$ un plan vectoriel $\mathcal{P}$ d'équation

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

Supposons que les vecteurs $\{\overrightarrow{U_1}, \overrightarrow{U_2}\}$ forment une base de $\mathcal{P}$. D'après le théorème de la base incomplète, nous pouvons trouver un vecteur $\overrightarrow{U_3}$ tel que $\{\overrightarrow{U_1}, \overrightarrow{U_2}, \overrightarrow{U_3}\}$ soit une base de $\mathbb{R}^3$. Soit $\mathcal{D}$ le sous-espace vectoriel de $\mathbb{R}^3$ ayant pour base $\overrightarrow{U_3}$. Ce sous-espace vectoriel est donc de dimension 1, c'est une droite vectorielle dans $\mathbb{R}^3$. Soit $\overrightarrow{X}$ un vecteur non nul de $\mathbb{R}^3$. Il admet une décomposition unique dans la base $\{\overrightarrow{U_1}, \overrightarrow{U_2}, \overrightarrow{U_3}\}$:

$$\overrightarrow{X} = x'_1\overrightarrow{U_1} + x'_2\overrightarrow{U_2} + x'_3\overrightarrow{U_3}$$

et nous avons

$$\overrightarrow{X_1} = x'_1\overrightarrow{U_1} + x'_2\overrightarrow{U_2} \in \mathcal{P}, \overrightarrow{X_2} = x'_3\overrightarrow{U_3} \in \mathcal{D}.$$

Le vecteur $\overrightarrow{X_1}$ est le projeté du vecteur $\overrightarrow{X}$ sur le plan $\mathcal{P}$ le long de la droite $\mathcal{D}$.

2.2. **L'application linéaire projection.** Revenons au cas général : $\mathbb{R}^n = F_1 \oplus F_2$. Soit $\overrightarrow{X} \in \mathbb{R}^n$ et $\overrightarrow{X} = \overrightarrow{X_1} + \overrightarrow{X_2}$ la décomposition de $\overrightarrow{X}$ associée. Par définition $\overrightarrow{X_1}$ est de projeté de $\overrightarrow{X}$ sur F_1 le long de F_2. Considérons l'application

$$p : \mathbb{R}^n \to \mathbb{R}^n$$

donnée par

$$p(\overrightarrow{X}) = \overrightarrow{X_1}.$$

Cette application est bien linéaire. Elle vérifie l'identité

$$p \circ p = p.$$

En effet si $\overrightarrow{X} \in F_1$ sa décomposition s'écrit $\overrightarrow{X} = \overrightarrow{X} + \overrightarrow{0}$ et dans ce cas $p(\overrightarrow{X}) = \overrightarrow{X}$. Nous avons donc

$$\mathrm{Im}(p) = F_1, \ \ker(p) = F_2.$$

Définition 2. *Dans l'espace vectoriel $\mathbb{R}^n$, une projection est une application linéaire $p : \mathbb{R}^n \to \mathbb{R}^n$ vérifiant l'identité*

$$p \circ p = p.$$

Une telle application définit donc une projection sur le sous-espace vectoriel $F_1 = \mathrm{Im}(p)$ le long du sous-espace $F_2 = Ker(p)$.

Remarque sur la matrice de l'application linéaire p

Considérons dans $\mathbb{R}^n = F_1 \oplus F_2$ la projection sur F_1 le long de F_2. Soit p l'application linéaire projection associée. Nous avons

$$F_1 = \mathrm{Im}(p), \ F_2 = \ker(p)$$

et pour tout vecteur $\overrightarrow{X_1} \in F_1$, $p(\overrightarrow{X_1}) = \overrightarrow{X_1}$ (la restriction de p à F_1 est l'application identité. Supposons que $\dim F_1 = m$ et considérons une base de $\mathbb{R}^n$ (ce n'est certainement pas la base canonique) :

$$\mathcal{B} = \{\overrightarrow{U_1}, \cdots, \overrightarrow{U_m}, \overrightarrow{U_{m+1}}, \cdots, \overrightarrow{U_n}\}$$

telle que $\{\overrightarrow{U_1}, \cdots, \overrightarrow{U_m}\}$ soit une base de F_1 et $\{\overrightarrow{U_{m+1}}, \cdots, \overrightarrow{U_n}\}$ est une base de F_2. La matrice de p relative à cette base est

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Nous voyons donc sur cet exemple que le choix de la base pour écrire la matrice d'une application linéaire est important : la matrice de p dans la base canonique est sûrement beaucoup plus compliquée.

Afin de pouvoir parler de longueur de vecteurs, d'angles de deux vecteurs, nous devons introduire une nouvel outil de mesure, à savoir le produit scalaire.

3. Le produit scalaire euclidien dans $\mathbb{R}^n$

3.1. **Définition.**

Définition 3. *On appelle produit scalaire euclidien dans l'espace vectoriel réel $\mathbb{R}^n$ l'application de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}$:*

$$\overrightarrow{X} \cdot \overrightarrow{Y} = (x_1, x_2, \cdots, x_n) \cdot (y_1, y_2, \cdots, y_n) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Cette application a les propriétés suivantes : Pour tous $\alpha_1, \alpha_2 \in \mathbb{R}$ et $\overrightarrow{X}, \overrightarrow{X_1}, \overrightarrow{X_2}, \overrightarrow{Y}, \overrightarrow{Y_1}, \overrightarrow{Y_2} \in \mathbb{R}^n$,

(1) $(\alpha_1\overrightarrow{X_1} + \alpha_2\overrightarrow{X_2}) \cdot \overrightarrow{Y} = \alpha_1(\overrightarrow{X_1} \cdot \overrightarrow{Y}) + \alpha_2(\overrightarrow{X_2} \cdot \overrightarrow{Y})$,

(2) $\overrightarrow{X} \cdot (\alpha_1\overrightarrow{Y_1} + \alpha_2\overrightarrow{Y_2}) = \alpha_1(\overrightarrow{X} \cdot \overrightarrow{Y_1}) + \alpha_2(\overrightarrow{X} \cdot \overrightarrow{Y_2})$

Une telle application, a deux variables de $\mathbb{R}^n$, vérifiant les deux propriétés ci-dessus est appelée une forme bilinéaire sur $\mathbb{R}^n$. L'étude générale de telles applications sur un espace vectoriel réel ou complexe seraétudiée plus tard.

Quelques propriétés du produit scalaire euclidien

(1) Soit $\overrightarrow{X} \in \mathbb{R}^n$. Alors l'identité

$$\forall \overrightarrow{Y} \in \mathbb{R}^n, \ \overrightarrow{X} \cdot \overrightarrow{Y} = 0$$

implique $\overrightarrow{X} = \overrightarrow{0}$.

Démonstration. En effet, posons $\overrightarrow{X} = (x_1, \cdots, x_n)$. Prenons dans un premier temps $\overrightarrow{Y} = \overrightarrow{e_1} = (1, 0, \cdots, 0)$. Alors $\overrightarrow{X} \cdot \overrightarrow{e_1} = 0$ implique $x_1 = 0$. De même, en prenant $\overrightarrow{Y} = \overrightarrow{e_i}$, nous obtenons $x_i = 0$ et en déduisons $\overrightarrow{X} = 0$.

(2) Pour tout $\overrightarrow{X} \in \mathbb{R}^n$ **non nul**

$$\overrightarrow{X} \cdot \overrightarrow{X} > 0.$$

En effet $\overrightarrow{X} \cdot \overrightarrow{X} = x_1^2 + x_2^2 + \cdots + x_n^2$.

(3) Le produit scalaire est commutatif :

$$\overrightarrow{X} \cdot \overrightarrow{Y} = \overrightarrow{Y} \cdot \overrightarrow{X}.$$

Lorsque nous considérons dans l'espace vectoriel $\mathbb{R}^n$ le produit scalaire euclidien, nous parlerons alors de $\mathbb{R}^n$ comme espace vectoriel euclidien.

Définition 4. *L'espace vectoriel $\mathbb{R}^n$ muni du produit scalaire euclidien est appelé espace vectoriel euclidien.*

3.2. **Norme associée au produit scalaire.**

Définition 5. *Dans l'espace vectoriel euclidien $\mathbb{R}^n$, on appelle norme du vecteur $\overrightarrow{X}$, le scalaire*

$$||\overrightarrow{X}|| = \sqrt{\overrightarrow{X} \cdot \overrightarrow{X}}.$$

Nous avons donc, si $\overrightarrow{X} = (x_1, x_2, \cdots, x_n)$,

$$||\overrightarrow{X}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Ceci montre que la norme d'un vecteur est bien définie. Nous avons aussi

$$||\overrightarrow{0}|| = 0.$$

L'interprétation géométrique de cette fonction norme, qui est bien une fonction de $\mathbb{R}^n$ à valeurs réelles mais bien entendu non linéaire, est d'associer à tout vecteur, sa longueur.

Proposition 3. *(Inégalité de Cauchy-Schwarz) Pour tout $\overrightarrow{X}$ et $\overrightarrow{Y}$ dans $\mathbb{R}^n$, nous avons*

$$|\overrightarrow{X} \cdot \overrightarrow{Y}| \leq ||\overrightarrow{X}||||\overrightarrow{Y}||.$$

Démonstration. En effet, si nous posons $\overrightarrow{X} = (x_1, \cdots, x_n)$ et $\overrightarrow{Y} = (y_1, \cdots, y_n)$, alors

$$(\overrightarrow{X} \cdot \overrightarrow{Y})^2 = (x_1y_1 + \cdots + x_ny_n)^2 = \sum_i x_i^2y_i^2 + 2\sum_{i \neq j} x_iy_ix_jy_j.$$

De même

$$(||\overrightarrow{X}||||\overrightarrow{Y}||)^2 = (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = \sum_{i,j} x_i^2y_j^2.$$

On en déduit

$$(||\overrightarrow{X}||||\overrightarrow{Y}||)^2 - (\overrightarrow{X} \cdot \overrightarrow{Y})^2 = \sum_{i<j}(x_iy_j - x_jy_i)^2.$$

L'inégalité de Cauchy-Schwarz s'en déduit.

Proposition 4. *(Inégalité de Minkowski.) Pour tout $\overrightarrow{X}$ et $\overrightarrow{Y}$ dans $\mathbb{R}^n$, nous avons*

$$||\overrightarrow{X} + \overrightarrow{Y}|| \leq ||\overrightarrow{X}|| + ||\overrightarrow{Y}||.$$

Démonstration. En effet, si nous posons $\overrightarrow{X} = (x_1, \cdots, x_n)$ et $\overrightarrow{Y} = (y_1, \cdots, y_n)$, alors

$$(||\overrightarrow{X}|| + ||\overrightarrow{Y}||)^2 = (x_1^2 + \cdots + x_n^2) + (y_1^2 + \cdots + y_n^2) + 2\sqrt{\sum_{i,j} x_i^2 y_j^2}.$$

De même

$$(||\overrightarrow{X} + \overrightarrow{Y}||)^2 = \sum_i (x_i + y_i)^2 = \sum_i (x_i^2 + y_i^2 + 2x_i y_i).$$

Ainsi

$$(||\overrightarrow{X}|| + ||\overrightarrow{Y}||)^2 - (||\overrightarrow{X} + \overrightarrow{Y}||)^2 = 2\sqrt{\sum_{i,j} x_i^2 y_j^2} - 2\sum_i x_i y_i.$$

Nous en déduisons que

$$(||\overrightarrow{X} + \overrightarrow{Y}||)^2 \leq (||\overrightarrow{X}|| + ||\overrightarrow{Y}||)^2$$

Nous remarquons que l'égalité pour ces deux relations n'a lieu que lorsque $\overrightarrow{X}$ et $\overrightarrow{Y}$ sont liés.

3.3. **Bases orthonormées.** Soit $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \cdots, \overrightarrow{e_n}\}$ la base canonique de $\mathbb{R}^n$. Les vecteurs de cette base vérifient

(1) $\overrightarrow{e_i} \cdot \overrightarrow{e_i} = 1$,
(2) $\overrightarrow{e_i} \cdot \overrightarrow{e_j} = 0$ dès que $i \neq j$.

La représentation habituelle de repère canonique dans la plan ou dans l'espace conduit à la définition naturelle de l'orthogonalité de deux vecteurs, orthogonalité qui est "vérifiée" par la représentation de la base canonique :

Définition 6. *Deux vecteurs $\overrightarrow{X_1}$ et $\overrightarrow{X_2}$ de $\mathbb{R}^n$ sont dits orthogonaux si leur produit scalaire est nul :*

$$\overrightarrow{X_1} \cdot \overrightarrow{X_2} = 0.$$

Ainsi les vecteurs de la base canonique sont deux à deux orthogonaux et sont de longueur 1. Nous dirons qu'une telle base est une base orthonormée. D'une manière générale,

Définition 7. *Soit $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \cdots, \overrightarrow{v_n}\}$ une base de $\mathbb{R}^n$. Nous dirons que c'est une base orthonormée si les vecteurs de cette base vérifient*

(1) pour tout $i = 1, \cdots, n$, $\overrightarrow{v_i} \cdot \overrightarrow{v_i} = 1$,
(2) pour tout $i \neq j$, $\overrightarrow{v_i} \cdot \overrightarrow{v_j} = 0$.

3.4. **Deuxième définition du produit scalaire euclidien de $\mathbb{R}^n$.** Commençons par interpréter le produit scalaire dans le plan, c'est-à-dire dans l'espace vectoriel $\mathbb{R}^2$. Soient $\overrightarrow{X_1}$ et $\overrightarrow{X_2}$ deux vecteurs du plan vectoriel

$$\overrightarrow{X_1} = x_1\overrightarrow{e_1} + y_1\overrightarrow{e_2}, \ \overrightarrow{X_2} = x_2\overrightarrow{e_1} + y_2\overrightarrow{e_2}.$$

Nous avons

$$\left| \frac{\overrightarrow{X_1} \cdot \overrightarrow{X_2}}{\left\|\overrightarrow{X_1}\right\| \left\|\overrightarrow{X_2}\right\|} \right| \leq 1.$$

En effet $(x_1x_2 + y_1y_2)^2 \leq (x_1 + y_1)^2(x_2 + y_2)^2$. Nous pouvons donc interpréter ce nombre comme le cosinus d'un angle, mais ceci nécessite que l'angle formé par les deux vecteurs de base soit un angle droit.

Ainsi, par définition

$$\cos(\theta) = \frac{\overrightarrow{X_1} \cdot \overrightarrow{X_2}}{||\overrightarrow{X_1}||\ ||\overrightarrow{X_2}||}$$

et θ désigne l'angle formé par les vecteurs $\overrightarrow{X_1}$ et $\overrightarrow{X_2}$ du plan.

Théorème 1. *Soient $\overrightarrow{X}$ et $\overrightarrow{Y}$ deux vecteurs de $\mathbb{R}^n$. Alors*

$$\overrightarrow{X} \cdot \overrightarrow{Y} = ||\overrightarrow{X}||\ ||\overrightarrow{Y}|| \cos(\overrightarrow{X}, \overrightarrow{Y})$$

où $(\overrightarrow{X}, \overrightarrow{Y})$ désigne l'angle des vecteurs $\overrightarrow{X}$ et $\overrightarrow{Y}$.

Ainsi, comme dans l'exemple ci-dessus, une représentation cartésienne d'un repère orthonormé est donnée par des vecteurs de longueur 1 et deux à deux perpendiculaires.

4. Bases orthonormales. Le procédé de Gram-Schmidt

4.1. **Bases orthonormales.** Nous avons vu que dans l'espace euclidien $\mathbb{R}^n$, la base canonique $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \cdots, \overrightarrow{e_n}\}$ est orthogonale, c'est-à-dire

$$\overrightarrow{e_i} \cdot \overrightarrow{e_j} = 0$$

dès que $i \neq j$. Cette base est même orthonormée (ou orthonormale) chacun des vecteurs de base est de longueur 1. Bien entendu ce n'est pas l'unique base de $\mathbb{R}^n$ qui soit orthonormée (ou orthonormale).

Définition 8. *Soit F un sous-espace vectoriel de $\mathbb{R}^n$. Une base $\mathcal{B}_F = \{\overrightarrow{v_1}, \cdots, \overrightarrow{v_p}\}$ de F est dite*

(1) orthogonale si pour tout i, j :

$$\overrightarrow{v_i} \cdot \overrightarrow{v_j} = 0.$$

(2) orthonormée si elle est orthogonale et pour tout i, $||v_i|| = 1$.

Remarquons que si la famille $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_p}\}$ était une famille génératrice de F, elle en serait une base. En effet

Théorème 2. *Toute famille de vecteurs non nuls deux à deux orthogonaux est libre.*

Démonstration. Soit $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_p}\}$ une famille de vecteurs non nuls deux à deux orthogonaux. Soit $\alpha_1 v_1 + \cdots + \alpha_p v_p = \overrightarrow{0}$ alors $(\alpha_1 v_1 + \cdots + \alpha_p v_p) \cdot v_1 = \overrightarrow{0} \cdot v_1$ d'où $\alpha||v_1||^2 = 0$. Mais puisque $v_1 \neq \overrightarrow{0}$ on a $||v_1|| \neq 0$ et donc l'égalité précédente implique $\alpha_1 = 0$. On montre de même que, pour tout $i \in \{1, \cdots, p\}$, $\alpha_i = 0$. Ainsi la famille $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_p}\}$ est libre.

4.2. **Un procédé d'orthogonalisation.**

Théorème 3. *Soit F un sous-espace vectoriel de $\mathbb{R}^n$. Alors F admet une base orthonormale.*

Démonstration. La démonstration qui suit donne le **principe de construction d'une telle base orthonormale à partir d'une base de F donnée**. Soit donc $\{\overrightarrow{v_1}, \overrightarrow{v_2},, \cdots, \overrightarrow{v_p}\}$ une base de F. Considérons les p vecteurs $\{\overrightarrow{w_1}, \overrightarrow{w_2},, \cdots, \overrightarrow{w_p}\}$ définis de proche en proche par les relations

$$\left\{\begin{array}{ll} \overrightarrow{w_1} & = \overrightarrow{v_1} \\ \overrightarrow{w_2} & = a_{2,1}\overrightarrow{w_1} + \overrightarrow{v_2} \\ \overrightarrow{w_3} & = a_{3,1}\overrightarrow{w_1} + a_{3,2}\overrightarrow{w_2} + \overrightarrow{v_3} \\ \cdots & \cdots \\ \overrightarrow{w_{p-1}} & = a_{p-1,1}\overrightarrow{w_1} + a_{p-2,2}\overrightarrow{w_2} + \cdots + a_{p-1,p-2}\overrightarrow{w_{p-2}} + \overrightarrow{v_{p-1}} \\ \overrightarrow{w_p} & = a_{p,1}\overrightarrow{w_1} + a_{p,2}\overrightarrow{w_2} + \cdots + a_{p,p-2}\overrightarrow{w_{p-2}} + a_{p,p-1}\overrightarrow{w_{p-1}} + \overrightarrow{v_p} \end{array}\right.$$

Il est clair que ces vecteurs sont encore linéairement indépendants. En effet la matrice de ces vecteurs relatives à la base $\{\overrightarrow{v_1}, \overrightarrow{v_2},, \cdots, \overrightarrow{v_p}\}$ est du type

$$\begin{pmatrix} 1 & a_{2,1} & a_{3,1} + a_{3,2}a_{2,1} & \cdots & a_{p,1} + a \\ 0 & 1 & a_{3,2} & \cdots & b \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

et donc son déterminant vaut 1 (pour une matrice triangulaire, le déterminant est le produit des éléments de la diagonale et les vecteurs $\{w_1, \cdots, w_n\}$ forment donc une nouvelle base de F. Montrons que l'on peut choisir les coefficients $a_{i,j}$ de manière à ce que ces vecteurs soient orthogonaux deux à deux. La méthode est simple : On commence par déterminer $a_{2,1}$ pour que $\overrightarrow{w_2}$ soit orthogonal à $\overrightarrow{w_1}$. Une fois ce coefficient choisi, on détermine $a_{3,1}$ et $a_{3,2}$ pour que $\overrightarrow{w_3}$ soit orthogonal aux vecteurs $\overrightarrow{w_1}$ et $\overrightarrow{w_2}$. On procède ainsi jusqu'au rang p.

(1) Ecrivons que $\overrightarrow{w_2}$ est orthogonal à $\overrightarrow{w_1}$:

$$\overrightarrow{w_2} \cdot \overrightarrow{w_1} = a_{2,1}\overrightarrow{w_1} \cdot \overrightarrow{w_1} + \overrightarrow{v_2} \cdot \overrightarrow{w_1} = 0$$

ce qui donne

$$a_{2,1} = -\frac{\overrightarrow{v_2} \cdot \overrightarrow{w_1}}{\overrightarrow{w_1} \cdot \overrightarrow{w_1}} = -\frac{\overrightarrow{v_2} \cdot \overrightarrow{w_1}}{||\overrightarrow{w_1}||^2}.$$

On a donc $\overrightarrow{w_2} = -\dfrac{\overrightarrow{v_2} \cdot \overrightarrow{w_1}}{||\overrightarrow{w_1}||^2} w_1 + v_2$ avec $w_1 \cdot w_2 = 0$.

(2) Ecrivons que $\overrightarrow{w_3}$ est orthogonal à $\overrightarrow{w_1}$ et à $\overrightarrow{w_2}$:

$$\left\{\begin{array}{l} a_{3,1}\overrightarrow{w_1} \cdot \overrightarrow{w_1} + a_{3,2}\overrightarrow{w_2} \cdot \overrightarrow{w_1} + \overrightarrow{v_3} \cdot \overrightarrow{w_1} = 0 \\ a_{3,1}\overrightarrow{w_1} \cdot \overrightarrow{w_2} + a_{3,2}\overrightarrow{w_2} \cdot \overrightarrow{w_2} + \overrightarrow{v_3} \cdot \overrightarrow{w_2} = 0 \end{array}\right.$$

d'où

$$\left\{\begin{array}{l} a_{3,1} = -\dfrac{\overrightarrow{v_3} \cdot \overrightarrow{w_1}}{\overrightarrow{w_1} \cdot \overrightarrow{w_1}} \\ \\ a_{3,2} = -\dfrac{\overrightarrow{v_3} \cdot \overrightarrow{w_1}}{\overrightarrow{w_2} \cdot \overrightarrow{w_2}} \end{array}\right.$$

d'où $\overrightarrow{w_3} = v_3 - \dfrac{\overrightarrow{v_3} \cdot \overrightarrow{w_1}}{||\overrightarrow{w_2}||^2} w_2 - \dfrac{\overrightarrow{v_3} \cdot \overrightarrow{w_1}}{||\overrightarrow{w_1}||^2} w_1$.

(3) Ainsi de suite jusqu'à la détermination du vecteur $\overrightarrow{w_p}$. Nous déterminons les coefficients $a_{p,1}, \cdots, a_{p,p-1}$ en résolvant le système

$$\left\{\begin{array}{l} (a_{p,1}\overrightarrow{w_1} + a_{p,2}\overrightarrow{w_2} + \cdots + a_{p,p-2}\overrightarrow{w_{p-2}} + a_{p,p-1}\overrightarrow{w_{p-1}} + \overrightarrow{v_p}) \cdot \overrightarrow{w_1} = 0 \\ (a_{p,1}\overrightarrow{w_1} + a_{p,2}\overrightarrow{w_2} + \cdots + a_{p,p-2}\overrightarrow{w_{p-2}} + a_{p,p-1}\overrightarrow{w_{p-1}} + \overrightarrow{v_p}) \cdot \overrightarrow{w_2} = 0 \\ \cdots \\ (a_{p,1}\overrightarrow{w_1} + a_{p,2}\overrightarrow{w_2} + \cdots + a_{p,p-2}\overrightarrow{w_{p-2}} + a_{p,p-1}\overrightarrow{w_{p-1}} + \overrightarrow{v_p}) \cdot \overrightarrow{w_{p-1}} = 0 \end{array}\right.$$

ce qui donne $\overrightarrow{w_p} = v_p - \dfrac{\overrightarrow{v_p} \cdot \overrightarrow{w_{p-1}}}{||\overrightarrow{w_{p-1}}||^2} w_{p-1} - \cdots - \dfrac{\overrightarrow{v_p} \cdot \overrightarrow{w_1}}{||\overrightarrow{w_1}||^2} w_1$. La base $\{\overrightarrow{w_1}, \overrightarrow{w_2}, \cdots, \overrightarrow{w_{p-2}}, \overrightarrow{w_{p-1}}, \overrightarrow{w_p}\}$ ainsi trouvée est orthonormale. Si nous voulons une base orthonormée, on considère alors les vecteurs

$$\overrightarrow{w_i'} = \frac{\overrightarrow{w_i}}{||\overrightarrow{w_i}||}.$$

On a bien

$$||\overrightarrow{w_i'}|| = \sqrt{\frac{\overrightarrow{w_i}}{||\overrightarrow{w_i}||} \cdot \frac{\overrightarrow{w_i}}{||\overrightarrow{w_i}||}} = \frac{\sqrt{\overrightarrow{w_i} \cdot \overrightarrow{w_i}}}{||\overrightarrow{w_i}||} = 1.$$

(On peut simplement utiliser la propriété d'une norme, i.e.$||\lambda \overrightarrow{u}|| = |\lambda| ||\overrightarrow{u}||$ pour tout scalaire λ de $\mathbb{R}$ et vecteur $\overrightarrow{u}$ de $\mathbb{R}^n$)

Remarque. Le procédé d'orthogonalisation d'une base d'un sous-espace de $\mathbb{R}^n$ reste bien entendu valable pour l'espace $\mathbb{R}^n$. A partir d'une base quelconque non orthogonale de $\mathbb{R}^n$, nous en déduisons une base orthogonale et donc orthonormée. Rappelons que la base canonique est déjà orthonormée.

4.3. **Matrices de passage d'une base orthonormée à une autre base orthonormée.** Les matrices de passage d'une base orthonormée à une autre base orthormée, en particulier de la base canonique à une base orthonormée ont une structure particulière. Elles s'appellent des **matrices orthogonales**. Nous allons étudier leur structure.

Soient $\mathcal{B} = \{\overrightarrow{v_1}, \overrightarrow{v_2},, \cdots, \overrightarrow{v_n}\}$ et $\mathcal{B}' = \{\overrightarrow{w_1}, \overrightarrow{w_2},, \cdots, \overrightarrow{w_n}\}$ deux bases orthonormées de $\mathbb{R}^n$. Considérons la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$. Rappelons qu'elle se construit en mettant en colonne les composantes des vecteurs de $\mathcal{B}'$ relatives à la base $\mathcal{B}$. Ainsi si

$$\overrightarrow{w_j} = \alpha_{1,j}\overrightarrow{v_1} + \alpha_{2,j}\overrightarrow{v_2} + \cdots + \alpha_{i,j}\overrightarrow{v_i} + \cdots + \alpha_{n,j}\overrightarrow{v_n}$$

alors

$$P = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,j} & \cdots \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,j} & \cdots \alpha_{2,n} \\ \cdots & & & & \\ \alpha_{i,1} & \alpha_{i,2} & \cdots & \alpha_{i,j} & \cdots \alpha_{i,n} \\ \cdots & & & & \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,j} & \cdots \alpha_{n,n} \end{pmatrix}$$

Comme les vecteurs $\overrightarrow{w}_i$ sont deux à deux orthogonaux, si $i \neq j$

$$\overrightarrow{w}_i \cdot \overrightarrow{w}_j = (\alpha_{1,i}\overrightarrow{v_1} + \alpha_{2,i}\overrightarrow{v_2} + \cdots + \alpha_{n,i}\overrightarrow{v_n}) \cdot (\alpha_{1,j}\overrightarrow{v_1} + \alpha_{2,j}\overrightarrow{v_2} + \cdots + \alpha_{n,j}\overrightarrow{v_n}).$$

En développant et compte tenu du fait que la famille $\mathcal{B}$ est orthonormée, il résulte

$$\overrightarrow{w}_i \cdot \overrightarrow{w}_j = \alpha_{1,i}\alpha_{1,j} + \alpha_{2,i}\alpha_{2,j} + \cdots + \alpha_{n,i}\alpha_{n,j} = 0.$$

Si $i = j$, cette relation devient

$$\overrightarrow{w}_i \cdot \overrightarrow{w}_i = \alpha_{1,i}^2 + \alpha_{2,i}^2 + \cdots + \alpha_{n,i}^2 = 0.$$

On en déduit que la matrice P vérifie

$$P \cdot^t P = Id.$$

Définition 9. *On appelle matrice orthogonale, toute matrice carrée A vérifiant l'identité*

$$A \cdot^t A = Id.$$

Une telle matrice est toujours inversible et son inverse est donc $A^{-1} =^t A$.

Théorème 4. *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées de $\mathbb{R}^n$. Alors la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice orthonormée. La matrice de passage de $\mathcal{B}'$ et $\mathcal{B}$ est ${}^tP$.*

Corollaire 2. *Soit $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ et soit $\mathcal{B}$ une base quelconque de $\mathbb{R}^n$. Si la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice orthonormée, alors la base $\mathcal{B}'$ une base orthonormée.*