# Bribes de géométrie

**Résumé**

Cette première série de séances pédagogiques se focalise sur les éléments de géométrie sous-jacents aux différents contextes qu'on aborde dans la suite. C'est une feuille de travail simple dont l'objectif est de vous mettre le pied à l'étrier.

## Table des matières

*Convention* 1. On fixe dans la suite un entier naturel $n$ strictement positif. Le cas nul, même s'il garderait un sens dans les énoncés a suivre, nous est pas d'un grand intérêt. Afin de ne pas alourdir le texte, la $i$-ème coordonnée d'un vecteur $x \in \mathbb{R}^n$ sera systématiquement notée $x_i$. La transposée d'une matrice $A$ sera notée $A^T$.

# 1 Structure euclidienne usuelle sur $\mathbb{R}^n$

Un espace euclidien est un espace vectoriel muni d'un produit scalaire, c'est-à-dire d'une forme bilinéaire symétrique définie positive. On vous renvoie à la section (C) pour un rafraîchissement de mémoire. L'espace vectoriel $\mathbb{R}^n$ vient muni d'une structure euclidienne, dite usuelle, définie par le produit scalaire suivant : à tout couple de vecteurs $(x, y)$ dans $\mathbb{R}^n$ on associe le scalaire

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i.$$

Cette quantité est souvent exprimée de manière plus compacte par

$$\langle x, y\rangle = x^T y.$$

La longueur au sens intuitif du vecteur $x$, appelée norme euclidienne de $x$, s'exprime alors sous la forme

$$\|x\|_2 = \langle x, x\rangle^{1/2} = \sum_{i=1}^{n} x_i^2.$$

En ce sens la norme $\|\cdot\|_2$ est dite ***norme associée au produit scalaire*** $\langle\cdot\,,\cdot\rangle$. On rappelle que celle-ci permet d'écrire l'inégalité de Cauchy-Schwarz sous la forme

$$\big|\langle x, y\rangle\big| \leq \|x\|_2\|y\|_2$$

où $x$ et $y$ sont des vecteurs quelconques de $\mathbb{R}^n$.

## 1.1 Produit scalaire de deux vecteurs

Le signe du produit scalaire de deux vecteurs porte une signification géométrique dont la compréhension est essentielle à l'étude qualitative des problèmes d'optimisation, ainsi qu'aux méthodes itératives qui permettent d'en approcher une solution.
On se place en un premier temps dans le cas de dimension 2, du plan euclidien. Soient $x$ et $y$ deux vecteurs de $\mathbb{R}^2$, on désigne par $\theta$ l'angle orienté (dans le sens direct) entre $x$ et $y$ et par $\phi$ (resp. $\psi$) celui entre $x$ (resp. $y$) et la partie positive de l'axe des abscisses.

**Question 1-1**

1. Exprimer les coordonnées de $x$ et $y$ en fonction de leurs normes respectives et des angles $\phi$ et $\psi$.
2. En déduire une expression du produit scalaire de $\langle x, y\rangle$ en fonction de $\theta$ et des normes de $x$ et $y$.

L'étude précédente s'étend relativement facilement au cas $n \geq 2$. Soit $P$ un plan contenant les vecteurs $x$ et $y$ de $\mathbb{R}^n$, si ceux-ci sont indépendants $P$ est bien entendu l'espace vectoriel engendre par $x$ et $y$. La restriction du produit scalaire $\langle\cdot\,,\cdot\rangle$ à $P$ prend la forme usuelle dans toute base orthonormée de $P$. Pour en trouver une dans $P$ il suffit de normaliser un vecteur non nul $f_1$ de $P$ puis de normaliser un vecteur $f_2$ de son orthogonal dans $P$. C'est la première étape de l'algorithme de Gram-Schmidt. La base $(f_1, f_2)$ nous ramène au cas de $\mathbb{R}^2$ précédent.

Le travail précédent permet d'établir en particulier le fait que l'orthogonalité de deux vecteurs, au sens intuitif que vous avez « *l'angle entre les deux vecteur est égal à* $\pi/2[\pi]$ », correspond à l'annulation de leur produit scalaire. C'est cette définition, plus algébrique et n'impliquant pas de définir ce qu'est un angle, qui est utilisée en pratique.

**Définition 1.1.** Deux vecteurs $x$, $y$ d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$ sont dits ***orthogonaux*** si leur produit scalaire est nul.

Cette notion d'orthogonalité de deux vecteurs va également permettre de décrire certains sous-espace d'un espace euclidien.

**Définition 1.2.** Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien. L'***orthogonal*** d'un sous-espace vectoriel $F$ de $E$ est le sous-espace vectoriel défini par

$$F^{\perp} = \{x \in E \mid \forall y \in F, \langle x, y\rangle = 0\}.$$

Par abus on parlera d'orthogonal à un vecteur (ou une famille de vecteur) quand on sous-entend l'orthogonal au sous-espace vectoriel engendré. À noter que lorsque un vecteur est orthogonal à une famille de vecteurs, il est également orthogonal à toute combinaison linéaire d'éléments de celle-ci.
On revient à la signification géométrique du produit scalaire. On est désormais en mesure de voir qu'étant donné un vecteur non nul $\nu \in \mathbb{R}^2$, le lieu des vecteurs $x$ sujets à la contrainte $\langle \nu, x\rangle = \nu^T x \geq 0$ est celui des vecteurs $x$ tels que $\measuredangle x\nu$ est aigu (ou orthogonal). C'est précisément le demi-espace de $\mathbb{R}^2$ délimité par $\{\nu\}^{\perp}$ et vers lequel pointe $\nu$. La contrainte $\nu^T x \leq 0$ correspond de son côté au demi-espace délimité par $\{\nu\}^{\perp}$ et hors duquel pointe $\nu$. C'est l'ensemble des vecteurs $x$ qui décrivent un angle $\measuredangle x\nu$ obtus.

**Question 1-2** Décrire le lieu de $\mathbb{R}^2$ donné par la relation matricielle :

$$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dans $\mathbb{R}^n$ un demi-espace n'est plus délimité par une droite mais par un sous-espace vectoriel de dimension $n-1$, c'est ce qu'on appelle un ***hyperplan*** de $\mathbb{R}^n$. Si $\nu$ est un vecteur non nul de $\mathbb{R}^n$, la condition de $\nu^T x \geq 0$ décrit les vecteurs $x$ dans le demi-espace délimité par l'orthogonal de $\nu$ et vers lequel pointe $\nu$. Inverser l'inégalité revient à étudier le demi-espace hors duquel pointe $\nu$.

**Question 1-3** Représenter le lieu de $\mathbb{R}^3$ décrit par la reations $x_1 + x_2 + x_3 \geq 0$.

## 1.2 Dualité

Si $E$ est un espace vectoriel on note $E^{\vee}$ l'espace vectoriel $\mathcal{L}(E, \mathbb{R})$ des formes linéaires sur $E$. C'est ce qu'on appelle l'espace ***dual*** de $E$, il est de même dimension que $E$. Quand $E$ est euclidien, c'est-à-dire muni d'un produit scalaire $\langle\cdot\,,\cdot\rangle$ et de dimension finie, on a un isomorphisme d'espace vectoriel donné par l'application linéaire naturelle

$$\begin{array}{rcl} \Psi : E & \longrightarrow & E^{\vee} \\ x & \longmapsto & \left[\Phi_x : y \longmapsto \langle x, y\rangle\right] \end{array}$$

Pour s'en convaincre il suffit d'étudier le noyau de $\Psi$ : si celui-ci est trivial $\Psi$ est injective et donc bijective car $E$ et $E^{\vee}$ ont même dimension, finie. Un vecteur $x$ est dans le noyau de $\Psi$ si

$$\forall y \in \mathbb{R}^2, \quad \langle x, y\rangle = 0.$$

En particulier, on doit avoir $\langle x, x\rangle = 0$ donc $x = 0$.
Le résultat précédent signifie qu'étudier, par exemple, l'image d'une sous-partie $A$ de $\mathbb{R}^n$ par toutes les formes linéaires de son dual revient à regarder l'image de $A$ par les applications linéaires $\Phi_x = \langle x\,,\cdot\rangle$

pour $x$ qui varie dans $\mathbb{R}^n$. Cette seconde approche est plus explicite, elle est donc plus évidente à implémenter (avec toutes les précautions d'usage quant aux questions de discrétisation). Elle est également loin d'être dénuée d'intérêt applicatif, on pourrait par exemple s'intéresser à la trace que laisse un objet 3D par projections sur des droites vectorielles de $\mathbb{R}^3$.

## 2 Sous-espaces affines de $\mathbb{R}^n$

Afin de rendre compte des situations géométriques auxquelles on fera face, il va nous falloir élargir le panel des sous-ensembles de $\mathbb{R}^n$ que nous sommes capables de manipuler. La première étape est d'être en mesure de travailler avec des translatés d'espaces vectoriels : les sous-espaces affines de $\mathbb{R}^n$. On se limitera a une approche utilitariste qui ne s'attarde pas sur la définition d'espace affine en général (on vous renvoie vers l'annexe pour plus de formalisme). De ce point de vue, de nombreux abus de terminologie seront fait.

### 2.1 Définition

**Définition 2.1.** Un sous-espace affine de $\mathbb{R}^n$ est un translaté de sous-espace vectoriel de $\mathbb{R}^n$. Plus précisément, une partie $A$ de $\mathbb{R}^n$ est un sous-espace affine de $\mathbb{R}^n$ s'il existe un point $x$ dans $\mathbb{R}^n$ et un sous-espace vectoriel $F$ de $\mathbb{R}^n$ tels que

$$A = x + F$$

ou la notation $x + F$ désigne l'ensemble $\{x + y \mid y \in \mathbb{R}^n\}$.
Quand cette écriture a lieu, l'espace vectoriel $F$ est unique, il est appelé la direction de $A$. Le point $x$ ne l'est cependant pas, voir pour un exemple (2-4).

La dimension d'un sous-espace affine de $\mathbb{R}^n$ est la dimension de sa direction. En particulier, un hyperplan affine est le translaté d'un hyperplan vectoriel de $\mathbb{R}^n$. Dans la pratique on omet les adjectifs, le fait d'être vectoriel ou non est clair du contexte.

**Exemple 2.1.** Tout sous-espace vectoriel de $\mathbb{R}^n$ est un sous-espace affine de $\mathbb{R}^n$. En effet, on a toujours $F = 0 + F$.

**Exemple 2.2.** Tout singleton $\{x\}$ de $\mathbb{R}^n$ définit un sous-espace affine de $\mathbb{R}^n$. Pour tout $x \in \mathbb{R}^n$ on peut toujours écrire $\{x\} = x + \{0\}$.

**Exemple 2.3.** La partie $A = \{(1, t) \in \mathbb{R}^2 \mid t \in \mathbb{R}\}$ est un sous-espace affine de $\mathbb{R}^2$. On constate effectivement que

$$A = (1, 0) + \{(0, t) \mid t \in \mathbb{R}\}$$

le second terme étant évidemment un sous-espace vectoriel de $\mathbb{R}^2$. On peut par exemple remarquer que c'est le noyau de la projection $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ sur la première coordonnée (ou encore l'espace vectoriel engendré par $(0, 1)$). Donc

$$A = (1, 0) + \mathrm{Ker}(\pi_1)$$

**Question 2-4**

1. Donner une écriture différente de l'espace affine $A$ dans l'exemple (2.3).
2. Décrire toutes les écritures possibles de l'espace affine $A$ ; que pouvez vous en déduire en général ?

**Question 2-5** Décrire les sous-espaces affines de $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$.

## 2.2 Se donner un sous-espace affine de $\mathbb{R}^3$

Un sous-espace vectoriel de $\mathbb{R}^n$ est ou bien décrit par les équations que vérifient ses éléments ou alors explicitement à l'aide d'un certains nombres de paramètres. Dans le premier cas on dit que le sous-espace vectoriel est donné implicitement, dans le second qu'il est donné paramétriquement. Cela correspond à se donner un sous-espace vectoriel ou bien comme le noyau ou alors comme l'image d'une application linéaire. On peut encore dire qu'il est donné par contrainte ou par le biais d'une base par laquelle l'espace vectoriel est généré. La situation des sous-espaces affines est similaire.

### 2.2.1 Paramétriquement

Soit $P$ le sous-espace vectoriel de $\mathbb{R}^3$ donné parametriquement par

$$P = \{(t, 3t + u, -u) \mid (t, u) \in \mathbb{R}^2\}$$

C'est l'espace vectoriel engendré par la famille libre de $\mathbb{R}^3$, $\{(1, 3, 0), (0, 1, -1)\}$. C'est encore l'image de $\mathbb{R}^2$ par l'application linéaire $(t, u) \mapsto (t, 3t + u, -u)$. Tout translaté de $P$ est un sous-espace affine de $\mathbb{R}^3$, par exemple

$$A = (1, 2, -1) + P = \{(1 + t, 3t + u + 2, -u - 1) \mid (t, u) \in \mathbb{R}^2\}$$

est un sous-espace affine de $\mathbb{R}^3$ donné paramétriquement. On voit que la forme des coordonnées de $A$ est affine au sens que vous entendiez plus jeunes. Une application affine est une application linéaire une fois qu'on se débarrasse des constantes, plus formellement

**Définition 2.2.** Une application $f : \mathbb{R}^n \to \mathbb{R}^p$ est affine s'il existe $x_0 \in \mathbb{R}^p$ tel que $x \mapsto f(x) - x_0$ soit linéaire.

Dans notre cas $x_0 = (1, 2, -1)$. Afin d'éviter un détour qui nous écarterait de nos objectifs premiers nous garderons l'approche intuitive de la notion d'application affines : une applications dont l'expression des coordonnées est affines dans les variables de départ. Pour plus détail on vous renvoie à (B.1).

**Question 2-6** Écrire paramétriquement :

1. la droite de $\mathbb{R}^2$ de vecteur directeur $(1, -1)$ et passant par $(2, 3)$ ;
2. le plan de $\mathbb{R}^3$ donné par l'équation $x_1 + x_2 + x_3 = 2$.

### 2.2.2 Implicitement

Un sous-espace vectoriel $F$ de $\mathbb{R}^n$ est donne implicitement s'il est décrit comme le noyau d'une application linéaire. Soit $M$ une matrice de taille $(p, n)$ et de coefficients $(m_{i,j})$. Un vecteur $x \in \mathbb{R}^n$ est dans le noyau de l'application linéaire associée à $M$ si

$$Mx = 0$$

relation compacte pour designer le système linéaire

$$\left\{\begin{array}{ccccccc} m_{1,1}x_1 & + & \cdots & + & m_{1,n}x_n & = & 0 \\ m_{2,1}x_1 & + & \cdots & + & m_{2,n}x_n & = & 0 \\ \vdots & & & & \vdots & \vdots & \vdots \\ m_{p,1}x_1 & + & \cdots & + & m_{p,n}x_n & = & 0. \end{array}\right.$$

Un sous-espace affine de $\mathbb{R}^n$ est de manière similaire donné par un système linéaire avec un second membre non nul, qui prend la forme

$$Mx = R \Longleftrightarrow \left\{\begin{array}{ccccccc} m_{1,1}x_1 & + & \cdots & + & m_{1,n}x_n & = & r_1 \\ m_{2,1}x_1 & + & \cdots & + & m_{2,n}x_n & = & r_2 \\ \vdots & & & & \vdots & \vdots & \vdots \\ m_{p,1}x_1 & + & \cdots & + & m_{p,n}x_n & = & r_3. \end{array}\right.$$

où $R$ désigne un vecteur de $\mathbb{R}^n$. Pour s'en convaincre on se donne une solution particulière $x_0$ de $Mx = R$. L'équation précédente est donc équivalente a

$$Mx = Mx_0 \iff M(x - x_0) = 0.$$

L'ensemble des solutions de $Mx = R$ est donc ou bien vide ou alors $x_0 + \mathrm{Ker}(M)$ pour une solution particul'nière $x_0$. Une fois donné une base de $\mathrm{Ker}(M)$ on obtient une écriture paramétrique de $F$.

*Remarque* 1. Dans la pratique, on n'a pas à trouver d'abord une solution particulière de notre système afin d'en trouver toutes les solutions. Le pivot de Gauss s'en charge, vous êtes invités à vous rafraîchir la mémoire avec l'annexe (A).

**Question 2-7** Dessiner le lieu de $\mathbb{R}^2$ décrit par les contraintes

$$\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

— Décrire chacun des composants du lieu géométrique précédent paramétriquement.
— Que change le fait de rajouter la contrainte $x - 3y \leq 6$ ?
— Quel lieu correspond à la situation où l'on change le sens de toutes les inégalités ?

**Question 2-8** On désigne par $A$ la partie de $\mathbb{R}^2$ donnée par [a]

$$A = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; \begin{array}{ll} x + 2y & \leq 3 \\ x - y & \geq 2 \end{array} \right\}.$$

1. Représenter $A$ graphiquement en indiquant les éléments qui permettent de décomposer votre représentation.
2. Quel est le lieu qu'on obtient si on inverse les inégalités.
3. Donner l'équation d'un demi-espace dont l'intersection avec $A$ est un lieu non-vide borné de $\mathbb{R}^2$.

*a.* Question 1-1 Exam 2019.

### 2.3 Cas extrémaux

On a toujours le choix entre exprimer un sous-espace affine paramétriquement ou implicitement. Il est cependant plus simple d'exprimer une droite paramétriquement (on a besoin que d'un paramètre pour le faire) et un hyperplan implicitement (une seule équation y suffit).

— On pense une droite comme un point de départ auquel on ajoute les multiples scalaires d'un vecteur. On peut par exemple écrire

$$D = \{(0, 1) + \lambda(2, 3) \mid \lambda \in \mathbb{R}\}$$

Dans la pratique on verra souvent une écriture équivalente où l'on pense une droite comme l'unique droite passant par deux points donnés, dans le cas de $D$ on peut écrire

$$D = \{(1 - t)(0, 1) + t(2, 4) \mid t \in \mathbb{R}\}$$

— Un hyperplan affine $H$ est pour sa part donné par une équation de la forme $\nu^T x = r$. Si $x_0$ est une solution particulière de l'équation précédente celle-ci est équivalente à

$$\nu^T (x - x_0) = 0$$

Donc $H = x_0 + \{\nu\}^{\perp}$. C'est le translaté de l'orthogonal de $\nu$ par $x_0$.

**Question 2-9** Retrouver une écriture implicite de la droite $D$ précédente. Comment s'y prendre pour transformer une écriture paramétrique en une écriture implicite ? [a]

*a.* Vous êtes invités à étudier plus en profondeur le passage représentation implicite $\rightarrow$ représentation paramétriques et inversement

# 3 Sous-espaces de $\mathbb{R}^n$

On ne verra dans le cadre de ce cours qu'un contexte restreint de la géométrie des parties de $\mathbb{R}^n$. On se limite par exemple aux parties de $\mathbb{R}^n$ qu'on se donne à l'aide de fonctions de $\mathbb{R}^n$ vers $\mathbb{R}$. De plus, dans le cas des descriptions paramétriques on se limite aux graphes.

## 3.1 Se donner des parties de $\mathbb{R}^n$

Tout comme dans le cas des sous-espaces affines de $\mathbb{R}^n$, on peut se donner une partie de $\mathbb{R}^n$ de manière ***implicite*** ou ***paramétrique***.
Une partie de $\mathbb{R}^n$ est décrite ***implicitement*** lorsqu'elle est donnée comme zéros de fonctions. Dans notre cadre, une telle partie $A$ est de la forme

$$A = \{x \in \mathbb{R}^n \mid f(x) = 0\}$$

où $f$ est une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Notez que cette écriture regroupe les parties décrites par les équations de la forme $f(x) = \star$ pour une constante $\star \in \mathbb{R}$. Une telle partie est zéros de la fonction $g = f - \star$.

**Définition 3.1.** Étant donné une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on appelle ***courbe de niveau r*** de $f$ la partie de $\mathbb{R}^n$ définie par

$$\mathcal{C}_r = \{x \in \mathbb{R}^n \mid f(x) = r\}.$$

Dans la suite on s'intéresse également au lieu de sous-niveau $r$ d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ; c'est la partie de $\mathbb{R}^n$

$$\mathcal{C}_{\leq r} = \{x \in \mathbb{R}^n \mid f(x) \leq r\}.$$

**Question 3-10** Dessiner les courbes de niveaux 0, 1 et 2 des fonctions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ données par les expressions $f(x, y) = x^2 + y^2$ et $g(x, y) = x^2 + 4y^2$.

**Question 3-11** Comment décrire la surface délimitée par les deux branches d'une hyperbole ?

On dira qu'une partie de $\mathbb{R}^{n+1}$ est décrite ***paramétriquement*** si elle prend la forme du graphe d'une fonction

$$\Gamma(f) = \{\big(t, f(t)\big) \mid t \in \mathbb{R}^n\}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Une partie de $\mathbb{R}^{n+1}$ décrite paramétriquement au sens précédent s'écrit facilement sous forme implicite : c'est les zéros de la fonction

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & f(x) - y \end{array}$$

L'inverse n'est cependant pas vrai en général.

**Question 3-12** Trouver un exemple d'une partie de $\mathbb{R}^2$ qu'on peut décrire implicitement mais pas paramétriquement. Que faudrait-il modifier pour qu'il soit possible de parler d'écriture paramétrique dans votre exemple ?

*Remarque* 2. Même avec une notion plus fine d'écriture paramétrique le passage implicite - paramétrique n'est pas toujours possible. Je vous invite à vous intérsser au théorème des fonctions implicites pour plus de détails.

Les graphes de fonctions apparaissent régulièrement dans la suite ; parmi les parties définies à partir de ceux-ci on retrouve l'épigraphe d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{epi}(f) = \{(x,t) \mid t \geq f(x)\}.$$

**Question 3-13** Représenter l'épigraphe de la fonction sin sur $[0, \frac{\pi}{2}]$ puis sur $[\frac{\pi}{2}, \pi]$.

**Question 3-14** Représenter graphiquement l'intersection de l'épigraphe de la fonction $x \mapsto -\sqrt{x}$ sur $\mathbb{R}_+$ et de la partie

$$\left\{(x,y) \mid y \leq \sqrt{x}\right\}.$$

**Question 3-15** On considère les fonctions suivantes notées $f$, $g$ et $h$ respectivement données par les expressions

$$f(x,y) = x + 4y, \quad g(x,y) = xy, \quad h(x,y) = \frac{x^2}{2} + \frac{y^2}{4}.$$

Représenter les courbes de niveaux $\mathcal{C}_1(f)$, $\mathcal{C}_2(g)$ et $\mathcal{C}_4(h)$.

## 3.2 Convexité dans $\mathbb{R}^n$

Les techniques d'optimisation se déroulent au mieux dans des contextes convexes. C'est une notion qu'il est important de garder en mémoire ; des comportements pathologiques voient le jour hors de ce cadre.

### 3.2.1 Parties convexes de $\mathbb{R}^n$

**Définition 3.2.** Une partie $A \subset \mathbb{R}^n$ est dite ***convexe*** si l'on peut relier 2 points quelconques de $A$ par un segment intégralement inclus dans $A$. Plus formellement : $A$ est convexe si

$$\forall x, y \in A, \forall t \in [0,1] \qquad (1-t)x + ty \in A.$$

**Exemple 3.1.** Tout intervalle de $\mathbb{R}$ est convexe.

**Exemple 3.2.** Les sous-espaces affines de $\mathbb{R}^n$ sont convexes, il en va de même des demi-espaces de $\mathbb{R}^n$.

Dans la pratique reconnaître des parties convexes de $\mathbb{R}^n$ revient à les construire à partir de parties convexes à l'aide d'opérations préservant la convexité.

- L'intersection de parties convexes est convexe ; le lieu admissible d'un programme linéaire est donc convexe.
- L'image d'une partie convexe par une fonction affine est convexe ; les changements de variables affines dans les contraintes d'un porgramme linéaire ne changent donc pas le caractère convexe de celui-ci.

**Question 3-16** Donner deux arguments qui permettent de justifier le fait que la partie décrite question (3-14) est convexe.

**Question 3-17** Est-ce que l'union de parties convexes est convexe ?

**Question 3-18** Qu'est-ce que l'enveloppe convexe d'une partie de $\mathbb{R}^n$ ?

Les parties convexes de $\mathbb{R}^n$ jouissent de propriétés géométriques qui permettent de les isoler (séparer), lorsqu'elles ne s'intersectent pas [1]. Ces problématiques sont en dehors de la portée de ce cours, elles nous donnent cependant une information géométrique utile pour généraliser l'étude des programmes linéaires au cas (convexe) non-linéaire.

**Définition 3.3.** Un point $x$ d'une partie $A \subset \mathbb{R}^n$ est ***point du bord*** de $A$ si pour tout disque $D$ centré en $x$ et de rayon non nul $D \cap A \neq \emptyset$.
L'ensemble des point du bord d'une partir $A$ est appelée ***bord*** de $A$, elle est notée $\partial A$. L'union $\overline{A} = A \cup \partial A$ est appelée ***adhérence*** de $A$.

**Question 3-19** Dessiner la définition précédente.

**Définition 3.4.** Soit $A$ une partie de $\mathbb{R}^n$ et $x \in \partial A$. Un ***hyperplan d'appui*** à $A$ en $x$ est un hyperplan $H$

— passant par $x$
— defini par un vecteur normal $\nu$ tel que $\forall y \in A, \langle y - x, \nu \rangle \leq 0$.

Un hyperplan d'appui à $A$ cantonne donc la partie $A$ à l'un des demi-espaces qu'il definit.

**Question 3-20** Dessiner un exemple de partie
— n'ayant pas d'hyperplan d'appui en un point donné de son bord ;
— n'ayant aucun hyperplan d'appui ;
— ayant plus d'un hyperplan d'appui en un même point.

**Proposition 3.1.** *Une partie convexe de $\mathbb{R}^n$ admet un hyperlan d'appui en tout point de son bord* [2].

Cette propriété est une conséquence du théorème de séparations des convexes fermés par des hyperplans.

**Question 3-21** On reprend les notations de la question (3-15)
1. Donner un hyperplan d'appui à $\mathcal{C}_{\leq 4}(h)$.
2. Justifier le fait que $\mathcal{C}_{\leq 2}(g)$ n'a pas d'hyperplan d'appui.

### 3.2.2 Fonctions convexes

**Définition 3.5.** Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est ***convexe*** si
— son domaine de définition est une partie convexe de $\mathbb{R}^n$ ;
— $\forall x$, $y$ dans le domaine de définition de $f$,

$$\forall t \in [0,1], f(ty + (1-t)x) \leq tf(y) + (1-t)f(x).$$

1. Il faut un peu plus de rigueur, mais l'idée est là ...
2. Celui-ci pouvant être vide.

**Question 3-22** Dessiner la définition précédente et en donner une interprétation géométrique.

**Définition 3.6.** Une fonction $f$ est dite ***concave*** si son opposé $-f$ est convexe.

**Question 3-23** Donner des exemples de fonctions convexes.

**Question 3-24** Prouver, en n'utilisant que la définition, que la fonction $x \mapsto x^2$ définie sur $\mathbb{R}$ est convexe.

**Question 3-25** Justifier le fait

1. que la somme pondérée, de poids positifs de fonctions convexes est convexe ;
2. qu'une fonctoin sous-linéaire est convexe ? en donner des exemples ;
3. que le max de fonctions convexes est convexe ;
4. que la composition d'une fonction $f \circ g$ d'une fonction $f$ convexe avec une fonction affine $g$ est convexe.

**Question 3-26** Généraliser l'étude du max au contexte qui suit : soit $g : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ une fonction qui pour tout $x \in A \subset \mathbb{R}^n$ a une fonction partielle $y \mapsto g(x, y)$ qui est convexe, montrer que $f(y) = \sup_{x \in A} g(x, y)$ est également convexe.

**Question 3-27** Montrer que les sous-niveaux d'une fonction convexe sont convexes. La réciproque est-elle vraie ?

# A Pivot de Gauss

On se donne une matrice $M \in \mathcal{M}_{m,n}(\mathbb{R})$. Une opération élémentaire sur $M$ est une des deux opérations suivantes :

— pour $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$ interchanger les lignes $L_i$ et $L_j$

$$L_i \leftrightarrow L_j$$

— pour $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$, $i \neq j$, $\lambda_i \in \mathbb{R}^*$ et $\lambda_j \in \mathbb{R}$, remplacer la ligne $L_i$ par $\lambda_i L_i + \lambda_j L_j$ :

$$L_i \leftarrow \lambda_i L_i + \lambda_j L_j.$$

Le pivot de Gauss est un algorithme de transformation d'une matrice $M$ en matrice triangulaire supérieure et éventuellement en matrice identité, dans le but d'inverser une matrice, de calculer un déterminant ou de résoudre un système linéaire. Il est d'utilisation constante en calcul numérique. [3]

## A.1 Opérations élémentaires

On traduit par la suite les opérations élémentaires précédentes à l'aide du produit matriciel.

3. L'expression de l'inverse d'une matrice inversible à l'aide des cofacteurs est à prohiber ! Elle est théoriquement intéressante et particulièrement esthétique mais numériquement inéfficace.

**Question 1-28** Trouver des matrices $P(i,j)$ et $U(i,j,\lambda,\nu)$ dont la multiplications avec $M$ réalise les opérations élémentaires précédentes.

Même si ces opérations ne sont pas à utiliser lors de la résolution d'un système linéaire, elle le seront pour calculer un déterminant.

**Question 1-29** Trouver les matrices $P_t(i,j)$ et $U_t(i,j,\lambda,\nu)$ qui réalisent les opérations précédentes sur les colonnes.

### A.2 Systèmes linéaires

**Question 1-30** Déduire de la section A.1 que les opérations élémentaires (sur les lignes) du pivot de Gauss transforment tout système linéaire dont $M$ est une matrice (si appliqué des deux côtés de l'égalité) en un système linéaire ayant les mêmes solutions.

**Question 1-31** Comment procéder pour décrire toute les solutions d'un système linéaire ?

**Question 1-32** Discuter de l'efficacité de l'implémentation de l'algorithme

### A.3 Déterminant

**Question 1-33** Retrouver les relations standards sur le déterminant d'une matrice qui subit une opération élémentaire à l'aide de la formule

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

**Question 1-34** Comment procéder pour calculer le déterminant d'une matrice à l'aide du pivot de Gauss ?

### A.4 Matrices équivalentes

Deux matrices $A$ et $B$ sont dites équivalentes s'il existe deux matrices inversibles $P$ et $Q$ telle que

$$A = PBQ.$$

**Question 1-35** Justifier le fait que toute matrice est équivalente à une matrice ayant un bloc identité et des zéros partout ailleurs.

## B Géométrie affine

Cette section est centrée autour des applications affines de $\mathbb{R}^n$ dans $\mathbb{R}^m$, plus particulièrement autour des projections et transformations de l'espace affine euclidien. L'objectif est de vous apporter les éléments de langage et les résultats de structure qui vous permettent de les expliciter et les utiliser dans les contextes de ML. Pour rappel, la majeure partie des algorithmes de ML sont de nature géométrique. Notre démarche dans la suite est particulièrement pragmatique [4].

4. Elle pourrait heurter certaines âmes sensibles et quelques matheux qui ont gardés un peu de sens esthétique.

### B.1 Applications affines

**Définition B.1.** Une application $f : \mathbb{R}^n \to \mathbb{R}^m$ est dite ***affine*** s'il existe un point $\boldsymbol{a} \in \mathbb{R}^m$ tel que $f - \boldsymbol{a}$ soit une application linéaire.

Cette définition formelle nécessite une réalisation plus concrète afin qu'on puisse l'utiliser. Pour cela on va introduire une généralisation de la notion de base dans le cas de $\mathbb{R}^n$ qui nous permettra de se placer en un point quelconque pour y travailler.

**Définition B.2.** Un repère de l'espace affine $\mathbb{R}^n$ est la donnée d'un point $\boldsymbol{o}$ appelé origine du repère et d'une base $\boldsymbol{v} = (v_1, \dots, v_n)$ de $\mathbb{R}^n$. On note $(\boldsymbol{o}, \boldsymbol{v})$ une telle donnée.

Le repère canonique de $\mathbb{R}^n$ est le repère $(\underline{0}, e_1, \dots, e_n)$, que vous aviez eu l'habitude de manipuler dans le secondaire et en physique. Du moins dans les cas de dimensions 1 ou 2.

**Définition B.3.** Soient $(\boldsymbol{o}, \boldsymbol{v})$ et $(\boldsymbol{l}, \boldsymbol{w})$ deux repères respetivement de $\mathbb{R}^n$ et $\mathbb{R}^m$. Une application affine $f$ dans les repères $(\boldsymbol{o}, \boldsymbol{v})$ et $(\boldsymbol{l}, \boldsymbol{w})$ prend la forme

$$\forall x \in \mathbb{R}^n, \quad f(x) = M(x - \boldsymbol{o})_{\boldsymbol{v}} + f(\boldsymbol{o}). \tag{1}$$

où $M$ est une matrice de taille $(m, n)$ et $\bullet_{\boldsymbol{v}}$ désigne l'écriture du vecteur $\bullet$ dans la base $\boldsymbol{v}$. Le membre de droite de l'équation 1 est sous-entendu être décrit dans le repère $(\boldsymbol{l}, \boldsymbol{w})$. Cette équation est appelée équation de $f$ dans les repères $(\boldsymbol{o}, \boldsymbol{v})$ et $(\boldsymbol{l}, \boldsymbol{w})$.

**Exemple B.1.** Une application affine peut être donnée par une experssion dans le repère canonique

$$x \mapsto Mx + b$$

où $x$ et $b$ sont des vecteurs respectivement dans $\mathbb{R}^n$ et $\mathbb{R}^m$ et $M \in \mathcal{M}_{m,n}(\mathbb{R})$. Par exemple

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

Ce qui nous donne une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ affine en $x$ et $y$ au sens qu'on entend intuitivement.

Dans le but de trouver des expressions agréables ou standardisées d'une application affine on est amené à écrire ces applications affines dans des repères adaptés. Pour pouvoir exprimer un tel contexte il nous faut être en mesure de formaliser la notion de changement de repères.

**Définition B.4.** Le changement de repères de $\mathbb{R}^n$ de $(\boldsymbol{o}, \boldsymbol{v})$ vers $(\boldsymbol{l}, \boldsymbol{w})$ est l'application affine donnée dans les repères précédents par

$$x \longmapsto P_{\boldsymbol{v}}^{\boldsymbol{w}}(x - \boldsymbol{o})_{\boldsymbol{v}} + \boldsymbol{l}$$

où $P_{\boldsymbol{v}}^{\boldsymbol{w}}$ est la matrice de passage de la base $\boldsymbol{v}$ vers la base $\boldsymbol{w}$ ; c'est-à-dire la matrice des vecteurs de $\boldsymbol{v}$ écrits dans la base $\boldsymbol{w}$.

Voici comment formaliser le passage de l'écriture d'une application dans un repère vers un autre.

**Proposition B.1.** *Soient $(\boldsymbol{o}, \boldsymbol{v})$ et $(\boldsymbol{o'}, \boldsymbol{v'})$ deux bases de $\mathbb{R}^n$, $(\boldsymbol{l}, \boldsymbol{w})$ et $(\boldsymbol{l'}, \boldsymbol{w'})$ deux bases de $\mathbb{R}^m$. On note $M$ la matrice d'une application affine $f$ dans les repères $(\boldsymbol{o}, \boldsymbol{v})$ et $(\boldsymbol{l}, \boldsymbol{w})$, et $M'$ celle de $f$ dans les repères $(\boldsymbol{o'}, \boldsymbol{v'})$ et $(\boldsymbol{l'}, \boldsymbol{w'})$. Si $P_{\boldsymbol{v'}}^{\boldsymbol{v}}$ désigne la matrice de passage de $\boldsymbol{v}$ vers $\boldsymbol{v'}$ et $P_{\boldsymbol{w'}}^{\boldsymbol{w}}$ la matrice correspondante dans le cas de $\boldsymbol{w}$ et $\boldsymbol{w'}$ alors pour tout $x \in \mathbb{R}^n$*

$$M'(x - \boldsymbol{o'})_{\boldsymbol{v'}} + f(\boldsymbol{o'}) = f(x) = P_{\boldsymbol{w'}}^{\boldsymbol{w}} M \left(P_{\boldsymbol{v'}}^{\boldsymbol{v}}\right)^{-1} (x - \boldsymbol{o})_{\boldsymbol{v}} + f(\boldsymbol{o})$$

**Question 2-36** Écrire la fonction de l'exemple (B.1) dans le repère de $\mathbb{R}^2$ donné au départ et à l'arrivée par $\big((0,0), \{(1,1), (1,-1)\}\big)$.

*Remarque* 3. Dans les faits, on ne fera pas souvent ce type de changement de repères. À partir des propriétés des applications à l'étude on sera à même de trouver un bon repère dans lesquels les décrire. Il reste que de tels changements de repères doivent être compris et faits quand nécessaire.

**Exemple B.2.** Les applications affines constantes sont celles dont l'expression dans tous les repères a une matrice qui est nulle.

**Exemple B.3.** Avec la définition qu'on a donné toute application linéaire va décrire une application affine.

### B.1.1 Translations et homothéties

On s'attarde un instant sur les plus simples des applications affines non triviales.

**Exemple B.4.** La translation de vecteur $b$ est l'application affine donnée dans le repère canonique de $\mathbb{R}^n$ par $x \mapsto x + b$.

**Exemple B.5.** L'***homothétie*** de rapport $\lambda$ et centre $\boldsymbol{o}$ est l'application affine de $\mathbb{R}^n$ dans lui-même dont l'écriture dans le repère $(\boldsymbol{o}, \boldsymbol{v})$ prend la forme

$$x \mapsto \lambda(x - \boldsymbol{o})_{\boldsymbol{v}} + \boldsymbol{o}$$

L'homothétie de rapport 1 est simplement l'identité.

**Question 2-37** Donner une particularité que partagent les homothéties et les translations.

**Question 2-38** Que donne la composition de deux translations ? Celle de deux homothéties ?

**Question 2-39** Dessiner l'image du carré de sommets $(\pm 1, \pm 1)$ dans $\mathbb{R}^2$ par les homothéties suivantes :
- de rapport 2 et de centre l'origine
- de rapport 2 et de centre $(1, 0)$.

### B.1.2 Projections affines

Les projections affines sont la généralisation des projections linéaires. Pour rappel une projection $p : E \to E$ de l'espace vectoriel $E$ sur lui même est un application linéaire qui satisfait la relation $p^2 = p$. Cette relation garanti le fait que $G = \mathrm{Ker}(p)$ et $F = \mathrm{Ker}(p - \mathrm{id})$ sont des espaces vectoriels supplémentaires dans $E$. Ainsi, pour tout $x = x_F + x_G$, la projection $p$ est définie par $p(x) = x_F$. L'application $p$ est appelée projection sur $F$ parallèlement à $G$.

**Définition B.5.** Une application affine $f$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$ est une projection affine si elle a un point fixe et sa matrice dans un repère est la matrice d'une projection linéaire.

Une projection affine, dans un repère $(\boldsymbol{o}, \boldsymbol{v})$ de $\mathbb{R}^n$, s'écrit donc sous la forme

$$x \mapsto M(x - \boldsymbol{o})_{\boldsymbol{v}} + \boldsymbol{o}$$

On voit dans cette écriture que $\boldsymbol{o}$ est fixé par la projection. On vient d'écrire la projection de $\boldsymbol{o} + \mathrm{Ker}(M - I_n)$ parallèlement à $\boldsymbol{o} + \mathrm{Ker}(M)$.

**Question 2-40** Étudier des exemples de projections affines dans $\mathbb{R}^2$. Dessiner systématiquement les droites affines qui caractérisent cette projection.

## B.2 Transformations orthogonales

### B.2.1 Aparté sur l'orthogonalité

On se donne un produit scalaire $\langle \cdot, \cdot \rangle$ sur $\mathbb{R}^n$ [5]. On rappelle que c'est une forme bilinéaire symétrique définie positive. Elle vient avec une norme associée dite norme euclidienne définie par

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

**Définition B.6.** Deux vecteurs $x$, $y$ in $\mathbb{R}^n$ sont dit orthogonaux pour $\langle \cdot, \cdot \rangle$ si $\langle x, y \rangle = 0$.

Nous avons vu que cette notion correspond dans le cas du produit scalaire usuel au fait que l'angle entre $x$ et $y$ est $\pi/2$ $[\pi]$.

**Définition B.7.** Une base $\boldsymbol{v}$ est dite ***orthognale*** si les vecteurs qui la constituent sont orthogonaux deux à deux. Elle est ***orthonormale*** si ses vecteurs sont de norme 1.

L'exemple le plus simple d'une base orthonormée est celui de la base canonique de $\mathbb{R}^n$ pour le produit scalaire usuel.

**Question 2-41** Donner d'autres bases orthonormales de $\mathbb{R}^3$.

**Question 2-42** Comment construire une base orthonormée de $\mathbb{R}^3$ à partir de la base

$$\{(1, 0, 1), (0, 1, 1), (1, 1, 0)\}?$$

**Définition B.8.** Étant donné une partie $K$ de $\mathbb{R}^n$ on appelle orthogonal de $K$ pour $\langle \cdot, \cdot \rangle$ le sous-espace vectoriel

$$K^{\perp} = \{x \in \mathbb{R}^n \mid \forall y \in K,\ \langle x, y \rangle = 0\}.$$

Ainsi un hyperplan $H$ de $\mathbb{R}^n$ quand il est donné implicitement est décrit comme l'orthogonal d'un vecteur de $\mathbb{R}^n$ qu'on qualifie de vecteur ***normal***.

**Question 2-43** Décrire dans $\mathbb{R}^3$ l'orthogonal à la partie $\{(1, 1, 1), (1, 0, 1)\}$.

### B.2.2 Transformations vectorielles

La notion de produit scalaire permet à la fois de formaliser la notion d'angle ainsi que celle de distance [6]. Les transformations affines de $\mathbb{R}^n$, c'est-à-dire les applications affines bijectives, qui présèrvent les angles et les distances sont d'un grand intérêt géométrique ; elles ne touchent pas aux dispositions relatives de configurations de parties dans $\mathbb{R}^n$ et donc présèrvent les propriétés géométriques de celles-ci. Elles présèrvent par exemple la perpendicularité ou le parallélisme. Pour l'instant on se limite au cas vectoriel avant d'exposer la situation affine, plus générale.
Avant de poursuivre notre travail on prend un instant pour étudier de plus prêt une écriture matricielle de $\langle \cdot, \cdot \rangle$. Étant donné une base $\boldsymbol{v}$ de $\mathbb{R}^n$, on appelle matrice de $\langle \cdot, \cdot \rangle$ dans la base $\boldsymbol{v}$ la matrice

$$\boldsymbol{A} = \big(\langle v_i, v_j \rangle\big)_{i,j}.$$

5. Pensez au produit scalaire usuel si ça vous facilite la vie.
6. On étudiera plus en détail la seconde question un peu plus tard.

**Question 2-44** Montrer que dans la base $v$, on a pour tout $x$, $y \in \mathbb{R}^n$, $\langle x, y\rangle = x^T \boldsymbol{A} y$.

Si $\boldsymbol{v}$ est une base orthonormale dans ce cas $\boldsymbol{A} = I$ et on a, pour tout $x$, $y \in \mathbb{R}^n$ écrits dans cette base

$$\langle x, y\rangle = x^T y.$$

Ce qui nous ramène à l'écriture usuelle du produit scalaire.

**Hypothèse B.2.** Quitte à se ramener à une base orthonormée appropriée[7] on peut supposer que notre produit scalaire a la forme usuelle.

Soit $M$ une matrice inversible dans une base orthonormée de $\mathbb{R}^n$. Dire que $M$ préserve le produit scalaire $\langle\cdot,\cdot\rangle$ signifie que pour tout $x$, $y \in \mathbb{R}^n$ on a

$$(Mx)^T(My) = x^T y.$$

D'où l'on obtient, pour tout $x$, $y \in \mathbb{R}^n$

$$x^T\left(M^T M - I_n\right) y = 0.$$

En prenant $x$ et $y$ ayant des coefficients nuls sauf un on peut montrer que les coefficients de $M^T M - I_n$ sont tous nuls. Donc

$$M^T M = I_n.$$

**Proposition B.3.** *Une endomorphisme de $\mathbb{R}^n$ préserve un produit scalaire si et seulement sa matrice dans une base orthonormée satisfait*

$$M^T M = I_n.$$

*Une matrice qui satisfait la propriété précédente est dite* ***orthogonale****.*

**Question 2-45** Quel peut être le déterminant d'une matrice orthogonale ?

**Question 2-46** Traduire la proposition précédente en fonction des vecteurs colonnes de la matrice $M$.

On étudie un premier exemple de matrices orthogonales. On se fixe une base orthonormée de $\mathbb{R}^n$ dans laquelle on travaille. Soit $\theta \in \mathbb{R}$ on note $R(\theta)$ la matrice

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

**Question 2-47** Montrer que $R(\theta)$ est une matrice orthogonale pour tout $\theta$. Quel est son déterminant ?

**Question 2-48** Dessiner l'image du carré de points extrémaux $(\pm 1, \pm 1)$ par $R(\theta)$ pour $\theta = 0$, $\pi/4$.

Un exemple de matrice orthogonale qui n'est pas de déterminant 1 est celui-ci

$$s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

7. Ça existe toujours. Voir l'algorithme de Gram-Schmidt.

**Question 2-49** Interpréter géométriquement cette matrice.

**Question 2-50** Dessiner l'image du carré de points extrémaux $\{(0,0),(0,1),(1,0),(1,1)\}$ par $R(\theta)$ pour $\theta = 0$, $\pi/4$.

**Question 2-51** Déterminer toutes les matrices orthogonales de $\mathbb{R}^2$ muni de son produit scalaire usuel. Vous pourrez les étudier suivant leurs spectres.

**Question 2-52** Donner des exemples de matrices orthogonales en toute dimension. Essayez d'être le plus général possible.

**Question 2-53** Étendre la classification des matrices orthogonales au cas de dimension 3.

#### B.2.3 Transformations affines

***TODO***

# C Les formes bilinéaires symétriques

Les formes bilinéaires symétriques apparaissent sous différents aspects en mathématiques parmi ceux-ci, en réalité dès que l'on travaille avec une matrice symétrique. On peut par exemple évoquer le produit scalaire sur $\mathbb{R}^n$ qui nous permet de décrire un objet qui donne à la fois une information sur la norme d'un vecteur et l'angle que deux vecteurs forment. On peut encore évoquer le cas de la matrice de covariance, d'usage fréquent et symétrique. Ou enfin les polynôme de degré 2 ayant un nombre quelconque de variables.

## C.1 Définition

**Définition C.1.** Soit $E$ un $\mathbb{R}$-espace vectoriel. Une ***forme bilinéaire*** est une application $\phi : E \times E \to \mathbb{R}$ pour laquelle les applications partielles en tout point $x \in E$

— $t \mapsto \phi(x,t)$
— $t \mapsto \phi(t,x)$

sont linéaires.

**Exemple C.1.** La multiplication matricielle $\times : \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est une application bilinéaire sur l'espace vectoriel des matrices carrées de taille $n$.

**Définition C.2.** Une forme bilinéaire $\phi$ sur un $\mathbb{R}$-espace vectoriel $E$ est dite

— ***symétrique*** si $\forall x,\, y \in E,\, \phi(x,y) = \phi(y,x)$ ;
— ***antysmétrique*** si $\forall x,\, y \in E,\, \phi(x,y) = -\phi(y,x)$ ;
— ***positive*** si $\forall x,\, \phi(x) \geq 0$ ;
— ***négative*** si $\forall x,\, phi(x) \leq 0$ ;
— ***définie*** si $\forall x,\, \phi(x) = 0 \Rightarrow x = 0$.

**Question 3-54** Justifier le fait que le produit scalaire usuel sur $\mathbb{R}^n$ est une forme bilinéaire symétrique définie positive.

Dans la pratique, et quand on travaille sur $\mathbb{R}^n$, une forme bilinéaire symétrique est donnée par une matrice symétrique $P$ de taille $(n, n)$. L'application bilinéaire $\Phi$ associée s'écrit

$$\forall x, y \in \mathbb{R}^n, \quad \Phi(x, y) = x^T P y.$$

L'écriture à droite de l'égalité donne effectivement une forme bilinéaire symétrique. Si l'on s'est placés dans la base $b_1, \ldots, b_n$ pour écrire les vecteurs $x$ et $y$, la matrice $P$ est la matrice de coefficients $P_{ij} = b_i^T P b_j$.