Mat305

Calcul matriciel et fonctions de plusieurs variables

1. Applications

La notion d'application est une généralisation de la notion de fonction réelle (de $\mathbb{R}$ dans $\mathbb{R}$).

Définition 1.1 (Application)**.** *Une application $f : E \to F$ est la donnée d'un ensemble de départ E, d'un ensemble d'arrivée F est d'un procédé qui associe à* **chaque** *élément de E un* **unique** *élément de F. Pour tout $x \in E$ on note $f(x)$ l'élément de F associé à x qu'on appelle image de x par f.*

Remarque 1.1. Certains éléments de F peuvent n'être l'image d'aucun élément de E. Par exemple si on considère l'application

$$f : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x(x-1) \end{cases}$$

alors, pour $y < -\dfrac{1}{4}$, il n'existe aucun $x \in \mathbb{R}$ tel que $f(x) = y$ et, pour $y > -\dfrac{1}{4}$, il en existe deux.

Définition 1.2 (Images et antécédents)**.**
Soit $f : E \to F$ une application.
- *Pour un élément y de F, on appelle* **antécédent** *de y tout $x \in E$ tel que $y = f(x)$. Il peut y en avoir plusieurs, un seul ou aucun.*
- *On définit* **l'image de l'application** *$f : E \to F$ par*

$$Im(f) = f(E) = \{f(x) \mid x \in E\} = \{y \in F \mid \exists x \in E, y = f(x)\}$$

  On a $Im(f) \subset F$.
- *Une application telle que $Im(f) = F$ est dite* **surjective** *: tout $y \in F$ a au moins un antécédent par f.*
- *Une application* **injective** *est une application telle que tout $y \in F$ a au plus un antécédent par f.*
- *Une application qui est à la fois surjective et injective est dite* **bijective**.

Définition 1.3 (Graphe)**.** *Soit $f : E \to F$ une application. On appelle graphe de f le sous-ensemble de $E \times F$ défini par*

$$Gr(f) = \{(x, f(x)) \mid x \in E\} = \{(x, y) \in F \mid \exists x \in E, y = f(x)\}$$

Remarque 1.2. Un graphe est donc un sous-ensemble de $E \times F$. dans votre scolarité vous avez donc représenté des graphes de fonctions réelles c'est-à-dire que vous avez représenté un sous ensemble de $\mathbb{R} \times \mathbb{R}$ (le plan) qui est le plus souvent une courbe.

Exercice 1.1. Dessiner les graphes des applications

$$f_1 : \begin{cases} \mathbb{R} & \to \mathbb{R} \\ x & \mapsto \quad x(x-1)(x-2) \end{cases}$$
$$f_2 : \begin{cases} \mathbb{R} & \to \mathbb{R} \\ x & \mapsto \quad \left|\frac{x}{2}+3\right| \end{cases}$$
$$f_3 : \begin{cases} \mathbb{N} & \to \mathbb{Z} \\ n & \mapsto \quad 10-n \end{cases}$$

Définition 1.4 (Composée d'applications)**.**
Étant données deux applications $f : E \to F_1$ et $g : F_2 \to G$, l'application

$$x \mapsto g(f(s))$$

est bien définie si une des deux conditions suivantes est réunie
- — $F_1 \subset F_2$,
- — $f(E) \subset F_2$.

La deuxième condition est moins contraignante que la première. Dans chacun de ces cas on définit

$$g \circ f : \begin{cases} E & \to & G \\ x & \mapsto & g(f(x)) \end{cases}$$

qu'on appelle composée de f par g.

Exercice 1.2. Pour les situations suivantes $g \circ f$ et $f \circ g$ sont-elles bien définies ? dans le cas d'une réponse positive, a-t-on $g \circ f = f \circ g$?
- — $f_1(x) = 2x$ et $g_1(x) = x + 1$,
- — $f_2(x) = \ln(x)$ et $g_2(x) = 5x$,
- — $f_3(x, y) = (x + y, xy, y\mathrm{e}^x)$ et $g_3(x) = (x^3, \mathrm{e}^x)$,
- — pour $n \in \mathbb{N}$, $f_4(n) = 2n - 5$ et $g_4(x) = E(x)$

où $E(x)$ est la partie entière de x. Attention, $E(1.2) = 1$ mais $E(-1.2) = -2$.

Définition 1.5.
- — *On appelle application identité sur E l'application*

$$id_E : \begin{cases} E & \to & E \\ x & \mapsto & x \end{cases}$$

- — *On dit que $f : E \to F$ est* **inversible** *s'il existe $g : F \to E$ telle que $g \circ f = id_E$ et $f \circ g = id_F$.*
- — *on note $g = f^{-1}$ et on l'appelle* **l'application réciproque** *de f.*

Attention, la notation f^{-1} n'a rien à voir avec $\frac{1}{f}$.

Exercice 1.3. Reprendre les fonction f_i et g_i définies plus haut et voir si elles sont inversibles.

Nous avons vu les notions d'application bijective et d'application inversible. Ce sont en fait la même propriété.

Proposition 1.1. *Une application f est inversible si et seulement si elle est bijective.*

Preuve.
• Supposons tout d'abord que f est inversible. Alors il existe g comme ci-dessus.

Montrons tout d'abord que f est surjective : en effet comme $f \circ g = id_F$ alors pour tout $y \in F$ on a $f \circ g(y) = y$ et donc $f(g(y)) = y$ ce qui implique que y a au moins $g(y)$ comme antécédent par f.

Montrons maintenant que f est injective : soit x_1 et x_2 deux éléments distincts de E. Comme $x_1 \neq x_2$ on obtient que $g \circ f(x_1) = x_1 \neq x_2 = g \circ f(x_2)$. Cela implique que $f(x_1) \neq f(x_2)$ et donc que x_1 et x_2 n'ont pas la même image par f.

L'application f est injective et surjective, elle est donc bijective.

• Supposons maintenant que f est bijective. Alors tout élément y de F a un unique antécédent par f. Ceci permet de définir l'application g qui a $y \in F$ lui associe son unique antécédent par f.

Alors pour x dans E on a $g(f(x)) = x$ car $f(x)$ a pour seul antécédent x donc son image par g est x.

Et pour y dans F on a $f(g(y)) = y$ car $g(y)$ est un antécédent de y par f donc $f(g(y)) = y$.

Ainsi $g \circ f = id_E$ et $f \circ g = id_F$. g est donc l'application réciproque de f et f est inversible.

$\blacksquare$

Si une application $f : E \to F$ est inversible alors le graphe de sa réciproque f^{-1} est

$$\{(f(x), x) \parallel x \in E\} \subset F \times E.$$

Dans les cas où $E = F = \mathbb{R}$, le graphe de f^{-1} est le symétrique orthogonal du graphe de f par rapport à la diagonal d'équation $y = x$ dans $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

2. Applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$

Rappels : $\mathbb{R}^n$ est l'ensemble des n-uplets de réels :

$$\{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ pour tout } 1 \leq i \leq n\}.$$

C'est un $\mathbb{R}$-espace vectoriel (on verra plus tard la définition) c'est-à-dire en gros qu'on a une addition dans $\mathbb{R}^n$:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

et on peut multiplier un élément de $\mathbb{R}^n$ (appelé vecteur) par un réel :

$$\lambda(x_1, \ldots, y_n) = (\lambda x_1, \ldots, \lambda y_n).$$

Définition 2.1.
Une application $f : \mathbb{R}^n \to \mathbb{R}^p$ est dite linéaire lorsque pour tout $\lambda \in \mathbb{R}$ et tout couple de vecteurs $\overrightarrow{v}, \overrightarrow{w} \in \mathbb{R}^n$ on a

$$\begin{aligned} f(\overrightarrow{v} + \overrightarrow{w}) &= f(\overrightarrow{v}) + f(\overrightarrow{w}) \\ f(\lambda \overrightarrow{v}) &= \lambda f(\overrightarrow{v}) \end{aligned}$$

Exemple 2.1. — les fonctions linéaires réelles $x \mapsto ax$ sont des applications linéaires de $\mathbb{R}$ dans $\mathbb{R}$.
- — l'application $(x, y) \mapsto (y, x)$, la symétrie orthogonal par rapport à l'axe $y = x$, est une application linéaire (c'est aussi une isométrie) de $\mathbb{R}^2$ dans $\mathbb{R}^2$.
- — l'application $(x, y) \mapsto (x, 0)$, la projection orthogonal sur l'axe $y = 0$, est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$.
- — Soient a, b, c, d quatre réels, alors $(x, y) \mapsto (ax + by, cx + dy)$ est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Toute application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ peut se mettre sous cette forme.

Proposition 2.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Alors pour tous $\lambda_1, \ldots, \lambda_k$ dans $\mathbb{R}$ et tous $\overrightarrow{v_1}, \ldots, \overrightarrow{v_k}$ vecteurs de $\mathbb{R}_n$ on a :*

$$f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_k \overrightarrow{v_k}) = \lambda_1 f(\overrightarrow{v_1}) + \cdots + \lambda_k f(\overrightarrow{v_k}).$$

Preuve. Nous allons faire une preuve par récurrence. La propriété que nous voulons montrer pour $k \in \mathbb{N}$ est notée H_k : $\forall\ \lambda_1, \ldots, \lambda_k \in \mathbb{R},\ \ \forall\ \overrightarrow{v_1}, \ldots, \overrightarrow{v_k} \in \mathbb{R}^n$:

$$f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_k \overrightarrow{v_k}) = \lambda_1 f(\overrightarrow{v_1}) + \cdots + \lambda_k f(\overrightarrow{v_k}).$$

Nous initions le processus pour $k = 1$: H_1 est la propriété 2 de la définition de linéarité. Nous prouvons ensuite l'hérédité : nous supposons queH_k est vraie et nous essayons de montrer que H_{k+1} est vraie. On a

$$f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_{k+1} \overrightarrow{v_{k+1}}) = f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_k \overrightarrow{v_k}) + f(\lambda_{k+1} \overrightarrow{v_{k+1}})$$

grâce à la première propriété de la définition de linéarité. Puis en appliquant la deuxième propriété au dernier terme on obtient

$$f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_{k+1} \overrightarrow{v_{k+1}}) = f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_k \overrightarrow{v_k}) + \lambda_{k+1} f(\overrightarrow{v_{k+1}})$$

et enfin en appliquant la propriété H_k on obtient H_{k+1} :

$$f(\lambda_1 \overrightarrow{v_1} + \cdots + \lambda_{k+1}\overrightarrow{v_{k+1}}) = \lambda_1 f(\overrightarrow{v_1}) + \cdots + \lambda_k f(\overrightarrow{v_k}) + \lambda_{k+1} f(\overrightarrow{v_{k+1}})$$

■

Dans la suite, on notera les éléments de $\mathbb{R}^n$ et de $\mathbb{R}^p$ en colonne :

$$\overrightarrow{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \qquad \overrightarrow{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \in \mathbb{R}^p$$

Définition 2.2 (Base canonique)**.**
On appelle **base canonique** *de* $\mathbb{R}^n$ *le n-uplet de vecteurs* $(\overrightarrow{e_1}, \dots, \overrightarrow{e_n})$ *ordonné où*

$$\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \overrightarrow{e_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \overrightarrow{e_i} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \textit{avec 1 à la ième coordonnée} \ , \ \overrightarrow{e_n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Tout vecteur $\overrightarrow{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ s'écrit de façon unique dans la base canonique

$$\overrightarrow{x} = x_1 \overrightarrow{e_1} + \cdots + x_n \overrightarrow{e_n}.$$

et pour toute application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ on a

$$f(\overrightarrow{x}) : x_1 f(\overrightarrow{e_1}) + \cdots + x_n f(\overrightarrow{e_n}).$$

Ainsi, si on note pour chaque $\overrightarrow{e_i}$:

$$f(\overrightarrow{e_i}) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{pi} \end{pmatrix}$$

alors

$$
\begin{aligned}
f(x_1,\dots,x_n) &= f(x_1\overrightarrow{e_1}+\cdots+x_n\overrightarrow{e_n})\\
&= x_1 f(\overrightarrow{e_1})+\cdots+x_n f(\overrightarrow{e_n})\\
&= x_1\begin{pmatrix} a_{11}\\ \vdots \\ a_{p1}\end{pmatrix}+\cdots+x_n\begin{pmatrix} a_{1n}\\ \vdots \\ a_{pn}\end{pmatrix}\\
&= \begin{pmatrix} x_1a_{11}+\cdots+x_ia_{1i}+\cdots+x_na_{1n}\\ \dots \\ x_1a_{p1}+\cdots+x_ia_{pi}+\cdots+x_na_{pn}\end{pmatrix}\\
&= A\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}
\end{aligned}
$$

où

$$
A=\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n}\\ \vdots & & \vdots & & \vdots\\ a_{p1} & \dots & a_{pi} & \dots + & a_{pn}\end{pmatrix}.
$$

On a donc le théorème

Theorème 2.1. *Soit $f:\mathbb{R}^n\to\mathbb{R}^p$ une application linéaire et A la matrice dont les colonnes sont les vecteurs images de la base canonique alors, pour tout $(x_1,\dots,x_n)\in\mathbb{R}^n$ on a*

$$
f(x_1,\dots,x_n)=A\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}.
$$

Theorème 2.2. *Soit $f:\mathbb{R}^n\to\mathbb{R}^p$ et $g:\mathbb{R}^p\to\mathbb{R}^q$ deux applications linéaires. Soit A la matrice de f et B la matrice de g alors $g\circ f$ est linéaire et a pour matrice $B.A$.*

Preuve. Il s'agit simplement de montrer que pour tout $X=\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}\in\mathbb{R}^n$ on a

$$
B(AX)=(BA)X.
$$

En effet, le terme de gauche étant l'image de X par $g\circ f$ et le terme de droite étant l'image de X par l'application dont la matrice dans la base canonique est BA.

La preuve est principalement calculatoire. La matrice $C=BA$ est une matrice $q\times n$ dont les coefficients valent

$$
c_{ij}=\sum_{k=1}^{p} b_{ik}a_{kj},\quad \forall i\in[\![1,\dots,q]\!],\ \forall j\in[\![1,\dots,n]\!].
$$

Donc la i-ème coordonnée de $(BA).\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ vaut $\sum_{j=1}^{n}\left(x_j \sum_{k=1}^{p} b_{ik}a_{kj}\right)$

D'autre part la k-ème coordonnée de $Y = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ vaut $Y_j = \sum_{j=1}^{n} a_{kj}x_j$ et donc la i-ème coordonnée de $B\left(A.\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)$ vaut $\sum_{k=1}^{p} b_{ik} \sum_{j=1}^{n} a_{kj}x_j$.

Les deux double-sommes étant égales on en déduit que $(BA)X = B(AX)$. Ce qu'il fallait démontrer. ■

Theorème 2.3. *Si $f : \mathbb{R}^n \to \mathbb{R}^p$ est une application linéaire inversible alors $n = p$ et son inverse f^{-1} est également linéaire, de matrice B l'inverse de A : $AB = BA = I_n$.*

Preuve. L'application identité sur $\mathbb{R}^n$ est linéaire et sa matrice représentative est I_n la matrice diagonale avec des 1 sur la diagonale. Donc, si une application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ de matrice A est inversible, et si son inverse g est aussi linéaire et a pour matrice B, on doit avoir $BA = I_n$ et $AB = I_p$.

D'autre part, on verra plus tard que si une application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ est inversible alors $p = n$.

Il ne nous reste plus qu'à prouver que g est linéaire. Soit $\lambda \in \mathbb{R}$ et $y_1, y_2 \in \mathbb{R}^p$. Comme g est l'inverse de f qui est bijective, il existe x_1 et x_2 tels que $f(x_1) = y_1$ et $f(x_2) = y_2$. On a aussi $g(y_1) = x_1$ et $g(y_2) = x_2$.

On a $f(g(y_1+y_2)) = y_1+y_2 = f(g(y_1))+f(g(y_2)) = f(g(y_1)+g(y_2))$ car $g = f^{-1}$ et pour la dernière égalité parce que f est linéaire. On peut alors appliquer g et trouver

$$g(y_1 + y_2) = g(f(g(y_1 + y_2))) = g(f(g(y_1) + g(y_2))) = g(y_1) + g(y_2),$$

et on a montré la première identité de la définition de linéarité pour g.

On a $f(g(\lambda y_1)) = \lambda y_1 = \lambda f(g(y_1)) = f(\lambda g(y_1))$ bet donc en appliquant s

$$g(\lambda y_1) = g(f(g(\lambda y_1))) = g(f(\lambda g(y_1))) = \lambda g(y_1),$$

ce qui prouve la deuxième identité de la définition de linéarité pour g.

On a démontré que g est linéaire. ■