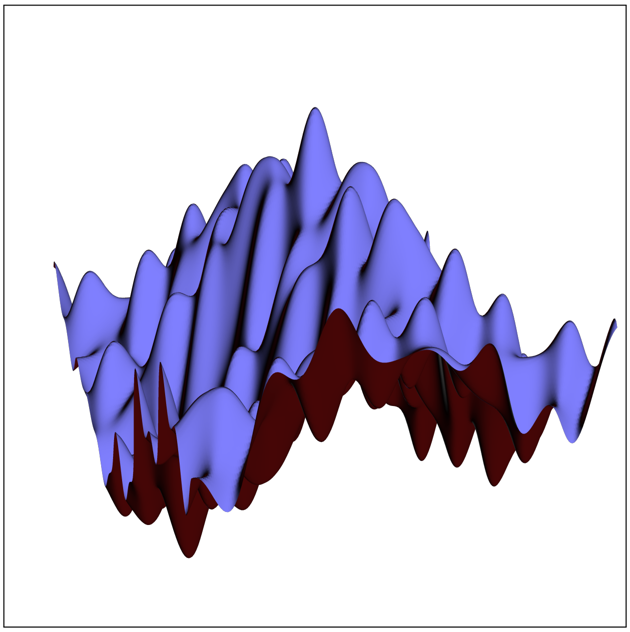
Université Grenoble Alpes

UE MAT305

Fonctions en plusieurs variables, calcul matriciel



Parcours Chimie

15 février 2024

TABLE DES MATIÈRES

AVERTISSEMENT AU LECTEUR

Ce polycopié correspond à l'Unité d'Enseignement MAT305. Cette unité d'enseignement est destinée aux étudiants de deuxième année du parcours chimie de l'Université Grenoble Alpes.

Ce polycopié est un outil pédagogique qui vient *s'ajouter* au cours. Le point de vue du cours et celui du polycopié peuvent différer offrant deux façons d'aborder une même notion mathématique.

Les chapitres de ce polycopié se décomposent de la façon suivante :

1. Le cours contient les notions à assimiler. Il convient d'en apprendre les définitions et les énoncés des résultats principaux. Les exercices corrigés sont des exemples ; essayez de rédiger une solution avant de lire celle proposée dans le texte.
2. La fiche de révision *n'*est *pas* la liste minimale des notions à connaître. Après avoir travaillé votre cours, lisez la fiche de révision : vous devez être capable de réciter chaque définition ou résultat de cette fiche sans la moindre hésitation (y compris l'énoncé des hypothèses éventuelles), sinon cela veut dire que vous devez retravailler plus soigneusement le cours.
3. L'index et le glossaire en fin de polycopié sont également des outils précieux pour la révision : ils contiennent la liste de toutes les notions définies et de toutes les notations introduites, en les parcourant vous pouvez tester l'acquisition des concepts du cours.

PROGRAMME

Prérequis pour cette UE : Les UE MAT102 et (MAT204 ou MAT208) de première année.

Programme résumé :

— Fonctions en plusieurs variables à valeur réelle, domaine de définition, lignes de niveau, extrema locaux et globaux, fonctions partielles, dérivées partielles, gradient, approximation linéaire, plan tangent ;

— Fonctions en n variables à valeurs dans $\mathbf{R}^m$, lien avec m fonctions en n variables, définition de la composition, dérivées partielles de la composée, changements de coordonnées, exemples : coordonnées polaires, cylindriques, sphériques.

— Intégrale multiple comme intégrale itérée, lien avec le volume.

— Matrices comme tableaux de nombres, somme, produit. Cas particulier du produit d'une matrice par un vecteur colonne, lien avec les systèmes linéaires, l'application $X \mapsto MX$ comme fonction de plusieurs variables.

— Sous-espaces vectoriels de $\mathbf{R}^n$, base d'un espace vectoriel, coordonnées dans une base. Application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$, matrice correspondante, application définie par n vecteurs, noyau, image, lien avec l'injectivité et la surjectivité.

— Exemples d'applications linéaires : rotation vectorielle, symétrie vectorielle en dimensions 2 et 3. matrice de la composée,

— Matrice jacobienne retour sur les dérivées partielles de la composée.

— Matrice inversible, calcul de l'inverse par un système linéaire. Déterminant pour les matrices 2x2 et 3x3, lien avec le volume, avec l'inversibilité.

— Formule de changement de variables pour les intégrales multiples en 2 ou 3 variables, exemple : coordonnées polaires.

— Changement de bases, diagonalisation, valeurs propres, vecteurs propres, cas d'une matrice symétrique, applications.

Fonctions en plusieurs variables

Emmanuel Peyre

Cours

1.1. Introduction. — Vous avez par le passé étudié le comportement des fonctions d'une variable réelle, utilisant leur dérivée pour en déduire leur tableau de variations et ainsi trouver leurs extrema ou dessiner leur graphe. L'idée de ce chapitre est d'étendre cette étude à des fonctions dépendant de plusieurs variables. Comme exemple, on peut penser à l'application qui à un point de coordonnées (x, y) sur une carte associe son altitude ou à celle qui à un point de coordonnées (x, y, z) associe la température en ce point ou la pression. La question qui se pose est de savoir de quels outils on dispose pour faire des études de fonctions dans ce cadre plus large.

1.2. Notions de base. — Rappelons d'abord un peu de vocabulaire mathématique sur les fonctions.

Définition 1.2.1

Soient E et F des ensembles. Une *fonction f de E dans F* est définie par les données suivantes :

(i) Une partie $\mathscr{D}_f$ de E appelée *domaine de définition de f* ;

(ii) Pour chaque élément x de $\mathscr{D}_f$, un élément noté $f(x)$ de l'ensemble F.

L'élément $f(x)$ s'appelle aussi la *valeur de f en x* ou *l'image de x par f*. L'ensemble E est appelé l'*ensemble de départ de f* et F l'*ensemble d'arrivée de f* .

Lorsque le domaine de définition est égal à l'ensemble de départ, autrement dit $\mathscr{D}_f = E$, on dit que f est une *application*.

Soit $n \in \mathbf{N}$ un entier naturel. Rappelons que $\mathbf{R}^n$ désigne l'ensemble des n-uplets $(x_1, \dots, x_n)$ où $x_1, \dots, x_n$ sont n nombres réels. Une fonction f en *n variables réelles* à valeurs réelles est une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$. On note $f(x_1, \dots, x_n)$ l'image du n-uplet $(x_1, \dots, x_n)$.

Représentation 1.2.2. — La figure 1.2.2 représente schématiquement ce qu'il faut pour définir une fonction, à savoir : l'ensemble de départ E, l'ensemble d'arrivée F, l'ensemble de définition

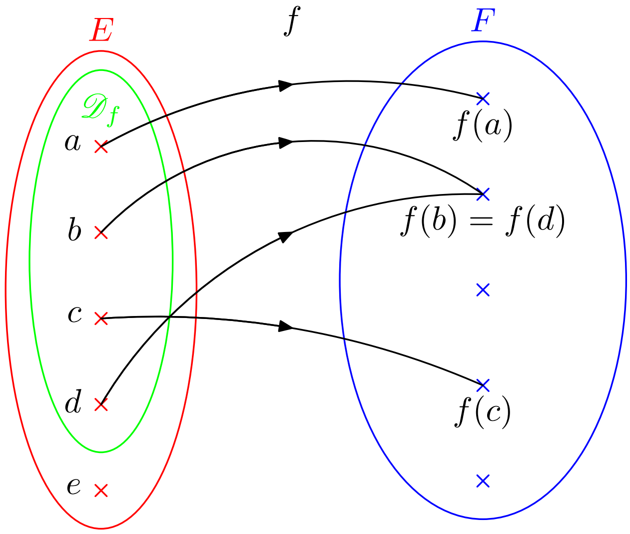


FIGURE 1. Une fonction f

$\mathscr{D}_f$ et pour chaque élément x de $\mathscr{D}_f$ (représenté par une croix sur la figure), son image $f(x)$ qui est un élément de l'ensemble d'arrivée F.

Remarques 1.2.3. — a) Il convient de ne pas confondre ensemble de départ et domaine de définition. Ainsi l'application racine carrée $x \mapsto \sqrt{x}$ est une fonction de $\mathbf{R}$ dans $\mathbf{R}$ dont le domaine de définition est $\mathbf{R}_+$, l'ensemble des nombres réels positifs ou nuls. Ainsi $\sqrt{x}$ n'est bien défini que si $x \geqslant 0$.

b) Pour définir une fonction en plusieurs variable on donne souvent une expression explicite qui indique comment calculer la fonction en un point. Par contre, le domaine de définition peut être implicite : dans ce cas, c'est la plus grande partie de l'ensemble de départ sur laquelle l'expression donnée a un sens. La phrase « Soit f la fonction de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par $(x,y) \mapsto \sqrt{x-y}$ » signifie que f est une fonction en deux variables réelles, dont le domaine de définition est l'ensemble des couples (x,y) tels que $x-y \geqslant 0$, c'est-à-dire $x \geqslant y$ et, si (x,y) vérifie cette condition, sa valeur en (x,y) est $\sqrt{x-y}$. Par exemple, $f(3,2)$ est bien défini et vaut $\sqrt{3-2} = 1$, mais $f(2,3)$ n'est pas défini.

1.3. Graphe, lignes de niveau. — Commençons par la définition mathématique du graphe d'une fonction :

Définition 1.3.1

Si f est une fonction d'un ensemble E dans un ensemble F, *Le graphe de f* est défini comme l'ensemble des couples $(x, f(x))$ lorsque x parcourt le domaine de définition de f. Autrement dit c'est l'ensemble des couples (x, y) où $x \in \mathscr{D}_f$ et $y = f(x)$. On le note souvent Γ_f.

Remarque 1.3.2. — Ainsi, si f est une fonction en n variables réelles à valeurs réelles, son graphe est une partie de $\mathbf{R}^n \times \mathbf{R}$ qu'on peut identifier avec $\mathbf{R}^{n+1}$ en assimilant un couple $((x_1, \ldots, x_n), x)$ avec le $n+1$-uplet $(x_1, \ldots, x_n, x)$. Par exemble, le graphe d'une fonction de $\mathbf{R}$ dans $\mathbf{R}$ est une partie du plan $\mathbf{R}^2$ et le graphe d'une fonction de $\mathbf{R}^2$ dans $\mathbf{R}$ est une surface contenue dans $\mathbf{R}^3$.

Exemples 1.3.3. — a) Si on considère la fonction f de $\mathbf{R}^2$ dans $\mathbf{R}$ définie par

$$(x, y) \mapsto \frac{\cos\left(x^2 + y^2\right)}{1 + x^2 + y^2}$$

en se restreignant aux couples $(x, y) \in [-12, 12] \times [-12, 12]$ le graphe à l'allure représentée dans la figure 1.3.3 a). Notons que comme $f(x, y)$ ne dépend que de la valeur de $x^2 + y^2$, qui est le carré

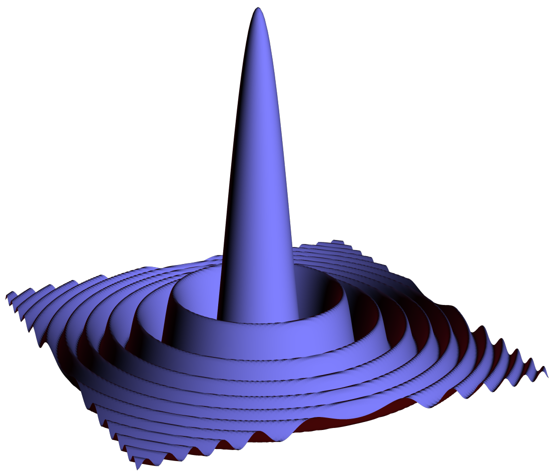


FIGURE 2. Graphe de f

de la distance du point (x, y) à l'origine $(0, 0)$, le graphe de f est invariant par rotation autour de l'axe Oz

b) Considérons maintenant la fonction g de $\mathbf{R}^2$ dans $\mathbf{R}$ définie par

$$(x, y) \mapsto \exp\left(-x^2 - y^2\right).$$

Une partie de son graphe est représenté dans la figure 1.3.3 b). Là encore, $g(x, y)$ ne dépend que

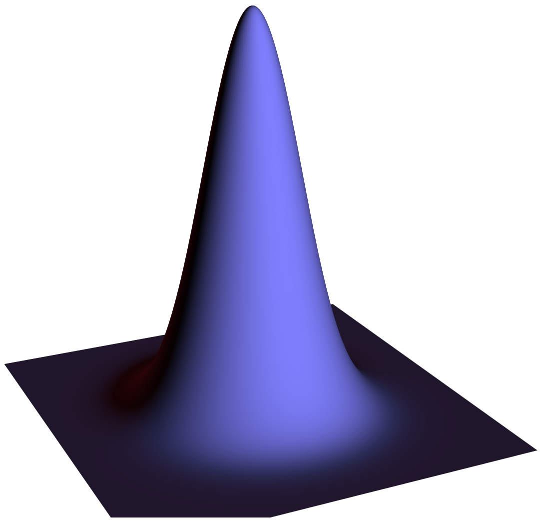


FIGURE 3. Graphe d'une distribution gaussienne

de la distance du point (x, y) à l'origine, si bien que l'axe Oz est à nouveau un axe de symétrie du graphe. Notons que cette fonction g est très utile en probabilité (*loi normale gaussienne*).

Définition 1.3.4

Soit f une fonction en n variables réelles à valeurs réelles. Par abus de langage, on appelle *ligne de niveau de* f un ensemble de la forme

$$\{(x_1, \ldots, x_n) \in \mathscr{D}_f | f(x_1, \ldots, x_n) = c\}$$

pour un nombre réel fixé c. Cet ensemble est également noté $f^{-1}(\{c\})$. Il s'agit donc de l'ensemble des points du domaine de définition dont l'image par f est c.

Remarque 1.3.5. — La terminologie s'explique par le fait que dans le cas d'une fonction en deux variables, une ligne de niveau est *souvent* –mais pas toujours– une courbe. Si vous avez utilisé des cartes IGN ou similaires, vous êtes habitués aux courbes de niveau : sur une carte IGN sont dessinées les courbes de niveau pour la fonction altitude (figure 1.3.5). Notons que si

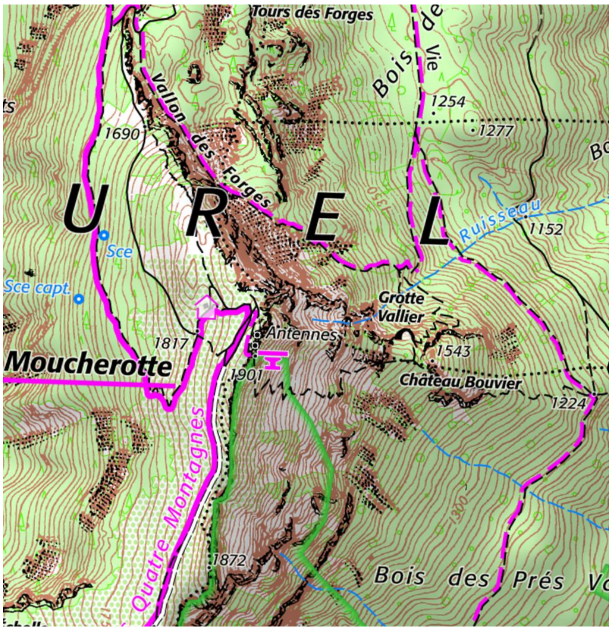


FIGURE 4. Lignes de niveau sur une carte

on regarde l'altitude d'un sommet, la ligne de niveau au sommet est réduite à un point. Si on considère une altitude trop élevée ou trop basse, la ligne de niveau est vide. Enfin si une zone est parfaitement nivelée, tous les points de la zone ont la même altitude et donc la ligne de niveau correspondante contiendra toute la zone en question.

1.4. Composée de fonctions. — La composée de fonctions est un opération qui nous sera très utile pour la suite. Nous allons d'abord en donner la définition mathématique abstraite avant de passer à quelques exemples.

Définition 1.4.1

Soient E, F et G des ensembles, soit f une fonction de E dans F et g une fonction de F dans G, alors on définit la composée des applications g et f comme la fonction de E dans G donnée par

$$g \circ f : x \mapsto g(f(x)).$$

Son domaine de définition est

$$\mathscr{D}_{g \circ f} = \{x \in \mathscr{D}_f \mid f(x) \in \mathscr{D}_g\}.$$

Notons que si f et g sont des applications, alors $g \circ f$ est une application de E dans G.

Remarque 1.4.2. — Avec les notations de la définition, si $x \in \mathscr{D}_{g \circ f}$ alors on sait déjà que x appartient au domaine de définition de f, donc $f(x)$ est bien défini. Par la description de $\mathscr{D}_{g \circ f}$, on a que $f(x)$ est un élément du domaine de définition de g donc $g(f(x))$ est bien défini.

Dans une composée, il faut faire attention au fait que l'ensemble de départ de la première application est l'ensemble d'arrivée de la seconde. Sinon la composée n'a pas de sens.

Représentation 1.4.3. — La figure 1.4.3 représente graphiquement la construction de la composée de deux fonctions. Sur ce graphique, les flèches pour $g \circ f$ sont obtenues en mettant bout

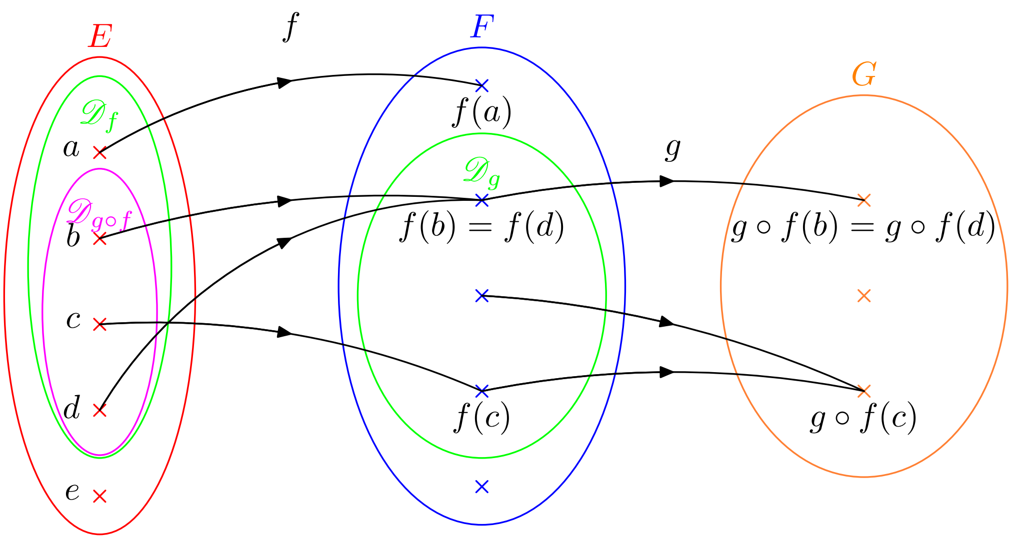


FIGURE 5. Composée de fonctions

à bout, quand cela est possible, une des flèches de f puis une des flèches de g.

Exemples 1.4.4. — a) Reprenons l'application f définie dans l'exemple 1.3.3 a) et considérons la fonctions en deux variables définie par $r : (x, y) \mapsto \sqrt{x^2 + y^2}$. Comme $x^2 + y^2$ est positif pour tout couple $(x, y) \in \mathbf{R}^2$, la fonction r est en fait une application de $\mathbf{R}^2$ dans $\mathbf{R}$. Soit $(x, y) \in \mathbf{R}^2$, les définitions de f et r donnent les égalités

$$f(x,y) = \frac{\cos(x^2 + y^2)}{1 + x^2 + y^2} = \frac{\cos(r(x,y)^2)}{1 + r(x,y)^2}$$

Donc $f = f_1 \circ r$ où f_1 est l'application de $\mathbf{R}$ dans $\mathbf{R}$ définie par

$$f_1 : t \mapsto \frac{\cos(t^2)}{1 + t^2}.$$

b) De même, l'application g de l'exemple 1.3.3 b) est l'application composée $g = g_1 \circ r$ où

$$g_1 : t \mapsto \exp(-t^2).$$

1.5. Distance euclidienne, extrema locaux et globaux. — Dans la suite de ce cours, nous utiliserons fréquemment la notion de distance entre deux points, rappelons rapidement la définition de la distance euclidienne usuelle dans l'espace $\mathbf{R}^n$.

Définition 1.5.1

Soient $A = (x_1, \dots, x_n)$ et $B = (y_1, \dots, y_n)$ des points de $\mathbf{R}^n$. La *distance de A à B* pour la distance euclidienne usuelle est le nombre réel

$$AB = d(A,B) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \dots + (y_n - x_n)^2}.$$

Exemple 1.5.2. — Dans le plan euclidien, la distance de $A = (1, 2)$ à $B = (4, 6)$ est

$$AB = \sqrt{(4-1)^2 + (6-2)^2} = \sqrt{9 + 16} = \sqrt{25} = 5.$$

Dans l'analyse vue au lycée, la notion d'intervelle est cruciale dans l'étude des fonctions. Par exemple, si a est un nombre réel, une assertion est dite vraie au voisinage de a si elle est vraie sur un intervalle ouvert contenant a. Nous allons généraliser cela à $\mathbf{R}^n$ en remplaçant les intervalles ouverts par des boules ouvertes.

Définition 1.5.3

Soit r un nombre réel strictement positif et soit $A \in \mathbf{R}^n$, alors la *boule ouverte centrée en A et de rayon r est l'ensemble*

$$B(A, r) = \{M \in \mathbf{R}^n | AM < r\}.$$

En dimension 2, une boule ouverte est donc un disque ouvert centré en le point choisi.

Terminologie 1.5.4

Soit $A \in \mathbf{R}^n$ On dit que quelque chose est vrai *au voisinage de A* si c'est vrai sur une boule ouverte centrée en A (et d'un rayon aussi petit qu'on veut).

La définition qui suit donne un exemple plus concret d'une propriété de ce type :

Définition 1.5.5

Soit f une fonction en n variables réelles à valeurs réelles. On suppose que $\mathscr{D}_f$, le domaine de définition de f, n'est pas vide. Soit $A \in \mathscr{D}_f$.

a) On dit que la fonction f admet un *maximum* (resp. un *minimum*) *global* en A si pour tout $x \in \mathscr{D}_f$, on a l'inégalité $f(x) \leqslant f(A)$ (resp. $f(x) \geqslant f(A)$).

b) On dit la fonction f admet un *maximum* (resp. un *minimum*) *local* en A s'il existe un nombre réel strictement positif r de sorte que pour $x \in \mathscr{D}_f \cap B(A, r)$, on a l'inégalité $f(x) \leqslant f(A)$ (resp. $f(x) \geqslant f(A)$).

Remarques 1.5.6. — a) On notera qu'un maximum global est aussi un maximum local. Un sommet est un maximum local pour l'altitude. Ainsi, le sommet du Mont-Blanc est un maximum local, alors que le sommet de l'Éverest est un maximum global.

Il convient d'insister sur cet exemple : un maximum local n'est pas forcément un maximum global.

b) La notion de minimum local est cruciale en thermodynamique : un équilibre stable est atteint lorque l'état du système est un minimum local de l'énergie libre.

1.6. Rappel sur les dérivées des fonctions en une variable. — Vous connaissez la notion de dérivée depuis la première.

Définition 1.6.1

Soit f une fonction en une variable réelle et à valeurs réelles. Soit $a \in \mathscr{D}_f$. On suppose que pour tout intervalle ouvert I contenant a, l'intersection $I \cap \mathscr{D}_f$ contient un autre point que a (on dit que *a n'est pas un point isolé de $\mathscr{D}_f$*). Soit $\lambda \in \mathbf{R}$. Les deux conditions suivantes sont équivalentes

(i) Le *taux d'accroissement*

$$\frac{f(x)-f(a)}{x-a}$$

tend vers λ quand $x \in \mathscr{D}_f - \{a\}$ tend vers a.

(ii) Il existe une fonction ε en une variable réelle telle que

$$f(x) = f(a) + \lambda(x-a) + (x-a)\varepsilon(x-a) \text{ pour tout } x \in \mathscr{D}_f$$

et $\varepsilon(h)$ tend vers 0 quand h tend vers 0.

Si ces conditions sont vérifiées, on dit que l'application f est *dérivable en a*. Le nombre réel λ est appelé la dérivée de f en a et on le note $f'(a)$. On dit que f est *dérivable* si elle est dérivable en tout point de son domaine de définition. L'application $x \mapsto f'(x)$ est alors appelée *la dérivée de f*.

Remarque 1.6.2. — Notons que la condition que a n'est pas isolé est automatiquement vérifiée pour tout $a \in \mathscr{D}_f$ lorsque $\mathscr{D}_f$ est un intervalle de $\mathbf{R}$ de longueur non nulle.

Dessin 1.6.3. — Rappelons que la dérivée donne la pente de la tangente en un point : en effet, si a et b sont deux éléments distincts du domaine de définition de l'application f. La droite passant par les deux points $(a, f(a))$ et $(b, f(b))$, qui appartiennent tous deux au graphe de f, a pour équation

$$y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a).$$

En particulier la pente de cette droite est le taux d'accroissement $\frac{f(b)-f(a)}{b-a}$. Lorsque b tend vers a, la pente de cette droite tend, par définition de la dérivée, vers $f'(a)$ (figure 6). La *tangente T au graphe de f en le point* $(a, f(a))$ a pour équation

$$y = f(a) + f'(a)(x-a).$$

Le fait que $f(a) + f'(a)(x-a)$ soit la meilleure approximation affine de f près de a correspond géométriquement au fait que la tangente est la droite qui est la plus proche du graphe de f près du point $(a, f(a))$. Cette tangente est représentée en rouge sur la figure.

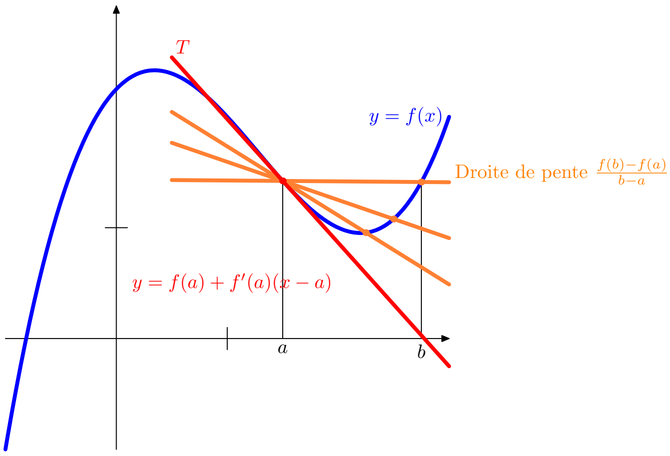


Figure 6. Dérivée et tangente

Les formules de calcul de dérivées rappelées dans le tableau qui suit font partie des formules à connaître.

Formules

$$
\begin{aligned}
(X^n)' &= (nX^{n-1}) \\
(u+v)' &= u' + v' \\
(\lambda u) &= \lambda u' \text{ pour tout } \lambda \in \mathbf{R} \\
(uv)' &= u'v + uv' \\
\left(\frac{u}{v}\right)' &= \frac{u'v - uv'}{v^2} \\
(v \circ u)' &= u'(v' \circ u) \\
\exp' &= \exp \\
\cos' &= -\sin \\
\sin' &= \cos
\end{aligned}
$$

Remarque 1.6.4. — Dans cette table X désigne l'application $t \mapsto t$, les lettres u et v désignent des fonctions dérivables. Les formules n'ont de sens que sur les domaines de définition des expressions données.

Justification de la dérivée d'une composée. — Soient I et J des intervalles ouverts de $\mathbf{R}$. Soit $u : I \to J$ et $v : J \to \mathbf{R}$ des applications. Soit $a \in I$, On pose $b = u(a)$. On suppose que l'application u est dérivable en a et que l'application v est dérivable en b. Par définition de la dérivabilité, il existe donc des fonctions en une variable ε_1 et ε_2, qui tendent toutes les deux vers 0 en 0 et telles que

$$u(x) = u(a) + (x-a)u'(a) + (x-a)\varepsilon_1(x-a) \tag{1}$$

pour tout $x \in I$ et

$$v(y) = v(b) + (y-b)v'(b) + (y-b)\varepsilon_2(y-b)$$

pour tout $y \in J$. Comme $b = u(a)$ et que cette dernière formule est valide si on prend $y = u(x)$ pour un $x \in I$, on obtient l'égalité

$$v(u(x)) = v(u(a)) + (u(x)-u(a))v'(u(a)) + (u(x)-u(a))\varepsilon_2(u(x)-u(a)).$$

Mais par la formule (1),

$$u(x)-u(a) = (x-a)(u'(a)+\varepsilon_1(x-a)).$$

On remplace donc $u(x)-u(a)$ par le terme de droite dans l'égalité précédente et on factorise $(x-a)$ ce qui donne la formule

$$\begin{aligned} v(u(x)) = {} & v(u(a)) + (x-a)u'(a)v'(u(a)) \\ & + (x-a)\Big(\varepsilon_1(x-a) + (u'(a)+\varepsilon_1(x-a))\varepsilon_2\big((x-a)(u'(a)+\varepsilon_1(x-a))\big)\Big). \end{aligned} \tag{2}$$

On définit alors la fonction

$$\varepsilon_3 : h \longmapsto \varepsilon_1(h) + (u'(a)+\varepsilon_1(h))\varepsilon_2\big(h(u'(a)+\varepsilon_1(h))\big)$$

Comme ε_1 et ε_2 tendent toutes les deux vers 0 quand h tend vers 0, on obtient que ε_3 tend également vers 0 quant h tend vers 0. D'autre part, avec l'expression donnée pour ε_3, la formule (2) se réécrit

$$v(u(x)) = v(u(a)) + (x-a)u'(a)v'(u(a)) + (x-a)\varepsilon_3(x-a).$$

ce qui prouve que la composée $v \circ u$ est dérivable au point a, de dérivée $u'(a) \times v' \circ u(a)$, comme annoncé. $\square$

Donnons une petite application mathématique au calcul de la dérivée d'une application composée. Pour cela soit I un intervalle de $\mathbf{R}$ et soit $f : I \to \mathbf{R}$ une application dérivable et strictement monotone (c'est-à-dire strictement croissante ou strictement décroissante) sur l'intervalle I. Dans ce cas, on peut déduire du théorème des valeurs intermédiaires qu'il existe un intervalle J de $\mathbf{R}$ et une application $g : J \to \mathbf{R}$ telle que

$$g(f(t)) = t$$

pour tout $t \in I$ et

$$f(g(t)) = t$$

pour tout $t \in J$. Les graphes des deux fonctions sont représentés sur la figure 12. Prenons $x \in I$ et

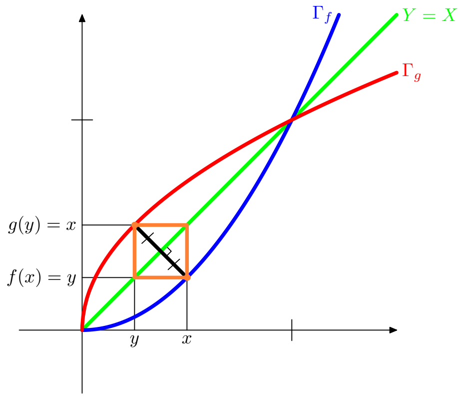


FIGURE 7. Fonction réciproque

posons $y = f(x)$. Alors $x = g(y)$. Sur la figure 12 les quatre points $(x, f(x))$, (y, y), $(y, g(y))$ et (x, x) forment alors un carré, ce qui montre que $(x, f(x))$ est le symétrique de $(y, g(y))$ par rappport à la droite diagonale Δ d'équation $Y = X$. Ceci montre que le graphe Γ_g est le symétrique de Γ_f par rapport à Δ. La fonction g s'appelle la *réciproque* de f, on la note f^{-1}. Nous reviendrons au chapitre 2 sur la définition générale de réciproque (définition 2.8.5)

Appliquons maintenant la formule de la dérivée pour la composée à $f \circ g$ en un point t où g est dérivable. Comme la dérivée de l'application identitée $\mathrm{Id}_I : t \mapsto t$ est la fonction constante de valeur 1, on obtient les égalités :

$$(f \circ g)' = g'(f' \circ g) = 1,$$

ce qui donne la formule

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}$$

qui n'a de sens qu'en les points $t \in J$ tels que $f'(f^{-1}(t)) \neq 0$.

Exemple 1.6.5. — L'application $f : x \mapsto x^2$ est strictement croissante sur $\mathbf{R}_+$ de réciproque l'application racine carrée $f^{-1} : x \mapsto \sqrt{x}$ dont la dérivée en tout réel $t \in \mathbf{R}_+^*$ est donnée par

$$(f^{-1})'(t) = \frac{1}{f'(f^{-1}(t))} = \frac{1}{2\sqrt{t}}.$$

Remarque 1.6.6. — Le lecteur attentif aura remarqué que nous n'avons pas démontré la dérivabilité de g. Pour faire cela, notons tout d'abord que la monotonie de f et donc de g permet de montrer que g est continue. Soit $b \in J$ et $a = g(b)$. On a donc $b = f(a)$. Pour tout $y \in J - \{b\}$, le taux d'accroissement pour g est donné par

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)}$$

où $x = g(y)$. C'est donc l'inverse de taux d'accroissement de f. Si on fait tendre y vers b, alors x tend vers a et, si $f'(a) \neq 0$, le taux d'accroissement tend vers $\frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}$.

1.7. Fonctions partielles. — En fixant la valeur de certaines des variables on peut en partie ramener l'étude des fonctions en plusieurs variables à celle des fonctions en une variable.

Définition 1.7.1

Soit f une fonction en n variables réelles. On se donne un entier $i_0 \in \{1, \ldots, n\}$ et pour $i \in \{1, \ldots, n\}$ tel que $i \neq i_0$, un nombre réel $a_i \in \mathbf{R}$. On peut alors considérer la fonction d'une variable réelle

$$f(a_1, \ldots, a_{i-1}, \bullet, a_{i+1}, \ldots, a_n) : x \mapsto f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n).$$

Une telle fonction est dite *fonction partielle* de la fonction f. Son domaine de définition est l'ensemble

$$\{x \in \mathbf{R} \mid (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) \in \mathscr{D}_f\}.$$

Remarques 1.7.2. — a) Ainsi si $f : (x, y) \mapsto f(x, y)$ est une fonction en deux variables réelles, et a un nombre réel on peut fixer la valeur de la première variable à a pour obtenir une fonction $f(a, \bullet) : t \mapsto f(a, t)$ ou la seconde variable pour obtenir une fonction $f(\bullet, a) : t \mapsto f(t, a)$. On notera que la place du symbole $\bullet$ dans la notation indique quelle variable n'est pas fixée.

b) Avec les notations de la remarque qui précède, le graphe de la fonction partielle $f(a, \bullet)$ (resp. $f(\bullet, a)$) s'obtient géométriquement en prenant l'intersection du graphe de f avec le plan vertical donné par l'équation $X = a$ (resp. $Y = a$) (cf. figure 8).

Donnons deux exemples de méthodes pour se ramener de l'étude d'une fonction en deux variables à celle de fonctions en une variable :

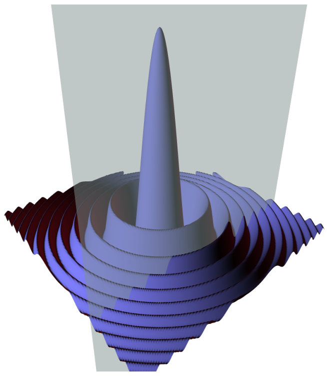


FIGURE 8. Fonction partielle

Exemples 1.7.3. — a) Commençons par la question du maximum. Soit f une fonction de $\mathbf{R}^2$ dans $\mathbf{R}$. On veut déterminer si f admet un maximum sur son domaine de définition. Pour cela on suppose que pour tout nombre réel a la fonction partielle $f(a,\bullet) : t \mapsto f(a,t)$ admet un maximum global. On peut alors définir une fonction de $\mathbf{R}$ dans $\mathbf{R}$ par

$$m : u \mapsto \max_{t\in\mathscr{D}_{f(u,\bullet)}} f(u,t).$$

Sous l'hypothèse faite, la fonction f admet un maximum global si et seulement si la fonction m en admet un et la valeur du maximum global de f est la valeur du maximum de m, ce qu'on peut résumer par la formule

$$\max_{(x,y)}(f(x,y)) = \max_x(\max_y(f(x,y)).$$

Notons que dans cette méthode, on peut parfaitement échanger les rôles de la première variable et de la seconde.

La difficulté de cette méthode est que pour obtenir le maximum de f, il faut étudier le maximum de *toutes* les fonctions partielles $f(u,\bullet)$, et donc d'une infinité de fonctions.

b) L'intégration en plusieurs variables peut se ramener aussi à l'intégration de fonctions en une variable Soit f une application de $\mathbf{R}^2$ dans $\mathbf{R}$. On suppose que pour $a \in \mathbf{R}$ la fonction partielle $f(a,\bullet) : t \mapsto f(a,t)$ est intégrable sur $\mathbf{R}$, on peut alors définir l'application

$$F : u \mapsto \int_{\mathbf{R}} f(u,t)\mathrm{d}t$$

Si F est également intégrable, on peut considérer l'intégrale double

$$\int_{\mathbf{R}}\left(\int_{\mathbf{R}} f(x,y)\mathrm{d}y\right)\mathrm{d}x = \int_{\mathbf{R}} F(u)\mathrm{d}u$$

Si l'application $|f| : (x,y) \mapsto |f(x,y)|$ vérifie également les hypothèses qui précèdent alors le théorème de Fubini affirme qu'on peut échanger le rôle des variables et on définit alors

$$\int_{\mathbf{R}^2} f(x,y)\mathrm{d}x\,\mathrm{d}y = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x,y)\mathrm{d}y \right) \mathrm{d}x.$$

Cette construction s'étend facilement à des applications en n variables.

Pour une fonction en n variables f, on peut également donner un sens à son intégrabilité en considérant son extension par 0 : on définit l'extension par 0 d'une fonction en n variables f comme l'application $\widetilde{f} : \mathbf{R}^n \to \mathbf{R}$ donnée par

$$(x_1,\ldots,x_n) \longmapsto \begin{cases} f(x_1,\ldots,x_n) \text{ si } (x_1,\ldots,x_n) \in \mathscr{D}_f \\ 0 \text{ sinon.} \end{cases}$$

1.8. Notion de continuité. — Comme nous allons maintenant le voir toutes les questions concernant les fonctions à plusieurs variables ne se réduisent pas si facilement à des questions sur des fonctions en une variable.

Pour aborder la question de la continuité, et la motiver, commençons par regarder ce que signifie une discontinuité pour la fonction altitude : imaginez que vous vous promenez de nuit et par temps de brouillard intense dans la région de Grenoble. Le fait de disposer d'un GPS dernier cri ne vous met pas à l'abri d'un accident, la région grenobloise étant riche en falaises. Ainsi si vous êtes, par exemple, proche du sommet du Moucherotte où une belle prairie se termine par une falaise de 200 mètres environ, la moindre erreur sur vos coordonnées fait que vous posez le pied sur la prairie ou dans le vide. Pour traduire cela en termes mathématiques : notons a la fonction altitude en deux variables correspondant à la position sur une carte, prenons un point A juste au bord de la falaise, il existe un nombre réel strictement positif H (la hauteur de la falaise) de sorte que quelque soit le nombre réel strictement positif r (la précision sur vos coordonnées GPS) et un point B tel que la distance AB est strictement inférieure à r (l'endroit où vous risquez de poser le pied) tel que $|a(B) - a(A)| \geqslant H$.

Une falaise c'est donc une discontinuité dans la fonction altitude (*cf.* figure 9). Le contraire, c'est la définition de la continuité qui signifie que si on se donne la précision souhaitée pour l'altitude, et qu'on connait suffisamment précisément notre position, l'erreur commise sur l'altitude est inférieure au seuil qu'on s'est fixé, et on ne risque pas de chute.

Définition 1.8.1

Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$. Soit $A \in \mathscr{D}_f$. On dit que f est continue en A si pour tout nombre réel $\varepsilon \in \mathbf{R}_+^*$, on peut trouver un nombre réel r strictement positif tel que, si

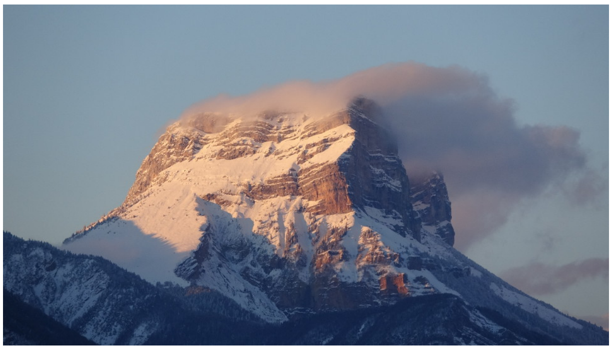


Figure 9. Les dangers de la discontinuité

$M \in B(a,r) \cap \mathscr{D}_f$, alors

$$|f(M) - f(A)| < \varepsilon.$$

Remarque 1.8.2. — En reprenant l'image de la falaise, la fin de la définition permet d'exclure qu'il y ait une falaise de hauteur $\geqslant \varepsilon$.

La continuité *ne* se teste *pas* avec les fonctions partielles !

Exemple 1.8.3. — On considère la fonction f en deux variables f définie par

$$f : (x,y) \longmapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Si on utilise des *coordonnée polaires*

$$x = r\cos(\theta) \quad \text{et} \quad y = r\sin(\theta)$$

alors, si $r \neq 0$,

$$f(r\cos(\theta), r\sin(\theta)) = \cos(\theta)\sin(\theta) = \frac{1}{2}\sin(2\theta).$$

En particulier $f(x,y)$ vaut $\frac{1}{2}$ si $x = y \neq 0$ et vaut $-\frac{1}{2}$ si $x = -y \neq 0$. Donc en prenant comme hauteur de falaise $\varepsilon = \frac{1}{2}$, pour tout $r > 0$, le point $\left(\frac{r}{2}, \frac{r}{2}\right) \in B((0,0), r)$ et

$$\left| f\left(\frac{r}{2}, \frac{r}{2}\right) - f(0,0) \right| \geqslant \frac{1}{2}.$$

Donc la fonction f n'est pas continue en $(0,0)$. Pourtant, si on regarde les fonctions partielles : si $a \in \mathbf{R}_+^*$, alors $f(a, \bullet) : t \mapsto \frac{at}{a^2+t^2}$ qui est une application continue de $\mathbf{R}$ dans $\mathbf{R}$, comme quotient de fonctions continues. D'autre part $f(0,\bullet)$ est la fonction constante de valeur 0, elle est donc également continue. Mais comme l'égalité $f(y,x)=f(x,y)$ pour tout $(x,y)\in\mathbf{R}^2$, on a également une égalité de fonctions $f(a,\bullet)=f(\bullet,a)$ pour tout $a\in\mathbf{R}$. Par conséquent, la fonction $f(\bullet,a)$ est également continue pour tout $a\in\mathbf{R}$.

Considérons les lignes de niveau de la fonction en deux variables f, comme $f(r\cos(\theta),r\sin(\theta))=\frac{1}{2}\sin(2\theta)$ pour $r\neq 0$. Soit $\alpha\in\mathbf{R}$. On déduit de la formule précédente que, pour $r\in\mathbf{R}_+^*$ et tout $\theta\in\mathbf{R}$, l'égalité $f(r\cos(\theta),r\sin(\theta))=\alpha$ équivaut à $\sin(2\theta)=2\alpha$ qui n'a pas de solution si $\alpha\notin[-\frac{1}{2},\frac{1}{2}]$ et équivaut à

$$2\theta\equiv\operatorname{asin}(2\alpha)[2\pi]\quad\text{ou}\quad 2\theta\equiv\pi-\operatorname{asin}(2\alpha)[2\pi]$$

Donc si $\alpha\in\left]-\frac{1}{2},\frac{1}{2}\right[-\{0\}$ la ligne de niveau $f^{-1}(\{\alpha\})$ est la réunion de deux droites dont on retire l'origine $(0,0)$. Une représentation graphique de ces lignes de niveau est faite dans la figure 10. Comme l'illustre la figure, les différentes lignes de niveau se « rencontrent » en l'origine

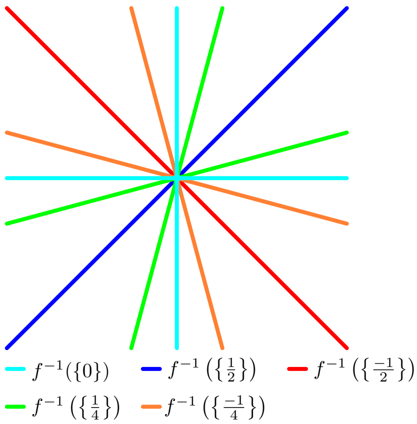


FIGURE 10. Lignes de niveaux

$(0,0)$ ce qui correspond au fait que pour tout nombre réel $\alpha\in\left[-\frac{1}{2},\frac{1}{2}\right]$ et tout nombre $r>0$ on peut trouver un point $A\in B((0,0),r)$ tel que $f(A)=\alpha$. De même, sur une carte IGN, une falaise corespond à une superposition des lignes des niveau (figure 1.3.5).

1.9. Dérivées partielles

La dérivabilité des fonctions partielles conduit aux dérivées partielles.

Définition 1.9.1

Soit f une fonction en n variables réelles, soit $i \in \{1,\ldots,n\}$ et soit $A = (a_1,\ldots,a_n) \in \mathbf{R}^n$. On suppose que le domaine de définition de la fonction partielle

$$f(a_0,\ldots,a_{i-1},\bullet,a_{i+1},\ldots,a_n)$$

contient un intervalle ouvert contenant a_i et que cette fonction est dérivable en a_i. On dit alors que la fonction f est dérivable relativement à la fonction X_i en A et on note

$$\frac{\partial f}{\partial X_i}(a_0,\ldots,a_{i-1},a_i,a_{i+1},\ldots,a_n) = f(a_0,\ldots,a_{i-1},\bullet,a_{i+1},\ldots,a_n)'(a_i)$$

Ce nombre s'appelle la *dérivée partielle de f par rapport à la i-ème variable en A.*

Remarques 1.9.2. — a) Dans la notation $\frac{\partial f}{\partial X_i}$, on utilise au dénominateur la notation choisie pour la variable considérée (éventuellement en majuscule). Par exemple, pour deux variables, si on considère une fonction $f : (x,y) \mapsto f(x,y)$ on pourra noter $\frac{\partial f}{\partial X}$ et $\frac{\partial f}{\partial Y}$ les dérivées partielles.

b) Si f est dérivable par rapport à la variable X_i en tous les points du domaine de définition $\mathscr{D}_f$, alors on peut définir *la dérivée partielle de f par rapport à X_i* qui est une fonction définie sur $\mathscr{D}_f$ par

$$\frac{\partial f}{\partial X_i} : (x_1,\ldots,x_n) \longmapsto \frac{\partial f}{\partial X_i}(x_1,\ldots,x_n).$$

Ainsi $\frac{\partial}{\partial X_i}$ devient un *opérateur* qui à la *fonction* f associe la *fonction* $\frac{\partial f}{\partial X_i}$. Cela permet en particulier d'itérer la construction et de considérer, là où cela est bien défini

$$\frac{\partial}{\partial X_{i_1}}\left(\frac{\partial}{\partial X_{i_2}}\left(\cdots\left(\frac{\partial}{\partial X_{i_m}} f\right)\cdots\right)\right).$$

Notation 1.9.3

La notation

$$\frac{\partial^m f}{\partial X_{i_1}\ldots\partial X_{i_m}}$$

est une abréviation pour

$$\frac{\partial}{\partial X_{i_1}}\left(\frac{\partial}{\partial X_{i_2}}\left(\cdots\left(\frac{\partial}{\partial X_{i_m}}f\right)\cdots\right)\right).$$

Terminologie 1.9.4

Soit f une fonction en n variables à valeurs réelles telle que pour tout point A de son domaine de définition $\mathscr{D}_f$, ce domaine $\mathscr{D}_f$ contient une boule ouverte centrée en A (on dit que $\mathscr{D}_f$ est *ouvert*). La fonction f sera dite de classe $\mathscr{C}^k$ si elle vérifie les deux conditions suivantes :

(i) Les dérivées partielles k-èmes $\frac{\partial^k f}{\partial X_{i_1} \ldots \partial X_{i_k}}(A)$ sont bien définies pour tout $A \in \mathscr{D}_f$ et tous $i_1, \ldots, i_k \in \{1, \ldots, n\}$;

(ii) Ces dérivées partielles k-èmes $\frac{\partial^k f}{\partial X_{i_1} \ldots \partial X_{i_k}}$ sont continues sur $\mathscr{D}_f$.

Énonçons un théorème au sujet dérivées partielles d'ordre supérieur :

Théorème 1.9.5 (Schwarz)

Si f est une fonction en n variables à valeurs réelle de classe $\mathscr{C}^2$, alors, on a l'égalité

$$\frac{\partial^2 f}{\partial X_i \partial X_j} = \frac{\partial^2 f}{\partial X_j \partial X_i}$$

pour tous $(i, j) \in \{1, n\}$.

Ce théorème est admis.

L'independance par rapport à l'ordre de dérivation peut-être fausse si on ne suppose pas que les dérivées partielles sont continues (*cf.* exercice 1.11).

Remarque 1.9.6. — Plus généralement, si f est de classe $\mathscr{C}^k$, alors les dérivées partielles k-émes ne dépendent pas de l'ordre dans lequel on calcule les dérivées partielles. Si σ est une permutation

de l'ensemble $\{1,\dots,k\}$,
$$\frac{\partial^k f}{\partial X_{i_1}\dots\partial X_{i_k}} = \frac{\partial^k f}{\partial X_{i_{\sigma(1)}}\dots\partial X_{i_{\sigma(k)}}}.$$

Définition 1.9.7

Soit f une fonction en n variables réelles et à valeur réelle et soit $A \in \mathscr{D}_f$. On suppose que $\mathscr{D}_f$ contient une boule ouverte centrée en A et que f admet des dérivées partielles par rapport à toutes les variables en A. Alors le *gradient de f en A* est le n-uplet de nombre réels
$$\operatorname{grad}_A(f) = \overrightarrow{\operatorname{grad}}_A(f) = \nabla f(A) = \left(\frac{\partial f}{\partial X_1}(A),\dots,\frac{\partial f}{\partial X_n}(A)\right)$$

La lettre ∇ se lit « nabla ».

***Exemples 1.9.8.* —** a) On peut appliquer toute les formules connues concernant les dérivées aux fonctions partielles pour en déduire des résultats sur les dérivées partielles : Soient f et g des fonctions de $\mathbf{R}^n$ dans $\mathbf{R}$, et soit $A \in \mathscr{D}_f \cap \mathscr{D}_g$. Soit $i \in \{1,\dots,n\}$. On suppose que f et g admettent toutes deux une dérivée partielle relativement à la variable X_i en A. Alors

a) La fonction somme $f+g : (x_1,\dots,x_n) \mapsto f(x_1,\dots,x_n)+g(x_1,\dots,x_n)$ aussi et
$$\frac{\partial(f+g)}{\partial X_i}(A) = \frac{\partial f}{\partial X_i}(A) + \frac{\partial g}{\partial X_i}(A);$$

b) Si $\lambda \in \mathbf{R}$, la fonction $\lambda f : (x_1,\dots,x_n) \mapsto \lambda f(x_1,\dots,x_n)$ admet également une dérivée partielle relativement à X_i et
$$\frac{\partial(\lambda f)}{\partial X_i}(A) = \lambda\frac{\partial f}{\partial X_i}(A);$$

c) La fonction produit $fg : (x_1,\dots,x_n) \mapsto f(x_1,\dots,x_n)g(x_1,\dots,x_n)$ aussi et
$$\frac{\partial(fg)}{\partial X_i}(A) = \frac{\partial f}{\partial X_i}(A)g(A) + f(A)\frac{\partial g}{\partial X_i}(A);$$

d) Si $g(A) \neq 0$, la fonction quotient $f/g : (x_1,\dots,x_n) \mapsto f(x_1,\dots,x_n)/g(x_1,\dots,x_n)$ admet également une dérivée partielle relativement à X_i en A et
$$\frac{\partial(f/g)}{\partial X_i}(A) = \frac{1}{g(A)^2}\left(g(A)\frac{\partial f}{\partial X_i}(A) - f(A)\frac{\partial g}{\partial X_i}(A)\right);$$

e) Si h est une fonction en une variable réelle dérivable en $f(A)$, alors
$$(h\circ f)(a_1,\ldots,a_{i-1},\bullet,a_{i+1},\ldots,a_n)=h\circ f(a_1,\ldots,a_{i-1},\bullet,a_{i+1},\ldots,a_n)$$
donc $h\circ f$ admet une dérivée partielle relativement à X_i en A et
$$\frac{\partial(h\circ f)}{\partial X_i}(A)=\frac{\partial f}{\partial X_i}(A)\times h'(f(A)).$$

b) Toutes ces formules donnent directement des formules analogues pour le gradient : avec les notations qui précèdent, on suppose que f et g admettent toutes deux des dérivées partielles relativement à toutes les variables en A. Alors

a) La fonction somme $f+g$ aussi et
$$\mathrm{grad}_A(f+g)=\mathrm{grad}_A(f)+\mathrm{grad}_A(g);$$

b) Si $\lambda\in\mathbf{R}$, la fonction λf admet également des dérivées partielles relativement à toutes les variables en A et et
$$\mathrm{grad}_A(\lambda f)=\lambda\,\mathrm{grad}_A(f);$$

c) La fonction produit fg aussi et
$$\mathrm{grad}_A(fg)=f(A)\,\mathrm{grad}_A(g)+g(A)\,\mathrm{grad}_A(f);$$

d) Si $g(A)\neq 0$, la fonction quotient f/g admet également des dérivées partielles relativement à toutes les variables en A et
$$\mathrm{grad}_A(f/g)=\frac{1}{g(A)^2}(g(A)\,\mathrm{grad}_A(f)-f(A)\,\mathrm{grad}_A(g));$$

e) Si h est une fonction en une variable réelle dérivable en $f(A)$, alors $h\circ f$ admet des dérivées partielles relativement à toutes les variables en A et
$$\mathrm{grad}_A(h\circ f)=h'(f(A))\,\mathrm{grad}_A(f).$$

1.10. Développement à l'ordre 1 d'une fonction en plusieurs variables, fonctions différentiables. — Commençons par un rappel sur l'approximation affine des fonctions dérivables.

Rappel 1.10.1. — Soient I un intervalle de $\mathbf{R}$ et $f: I\to\mathbf{R}$ une application. Soit $a\in I$. Par définition, si f est dérivable en a, on a une formule de la forme :
$$f(x)=f(a)+f'(a)(x-a)+(x-a)\varepsilon(x-a)$$
où ε est une fonction de $\mathbf{R}$ dans $\mathbf{R}$ telle que $\varepsilon(h)$ tend vers 0 lorsque h tend vers 0. Autrement dit $(x-a)\varepsilon(x-a)$ peut être considéré comme « négligeable » par rapport à la valeur absolue

$|x-a|$. Mais $|x-a| = \sqrt{(x-a)^2}$ est la distance du nombre réel x au nombre réel a au sens de la définition de la distance donnée dans la définition 1.5.1. Notons que la tangente au graphe de f en le point $(a, f(a))$ n'est rien d'autre que le graphe de la fonction

$$x \mapsto f(a) + f'(a)(x-a)$$

qui est l'*approximation affine* de f en a (figure 11).

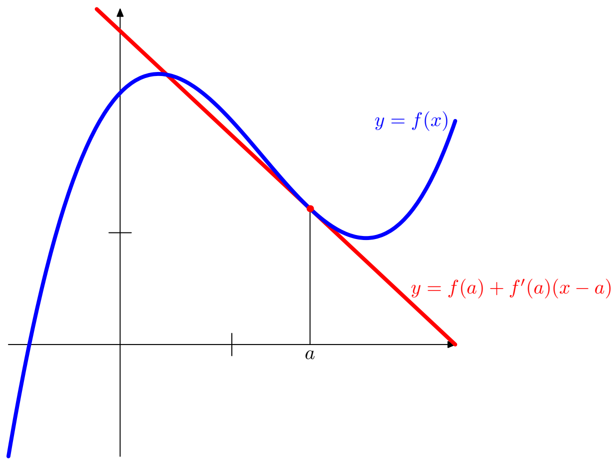


FIGURE 11. Approximation affine et tangente

Notre but est maintenant de généraliser cela aux fonctions en plusieurs variables. Une des difficultés de cette généralisation est que l'existence en un point des dérivées partielles ne garantit pas qu'il y ait une formule simple donnant une valeur approchée de la fonction au voisinage du point, nous reviendrons sur cette difficulté un peu plus tard.

Pour commencer, il convient de définir mathématiquement ce que veut dire « négligeable ».

Définition 1.10.2

Soient f et g des fonctions en n variables réelles à valeurs réelles. On suppose que le domaine de définition de f est contenu dans celui de g, c'est-à-dire $\mathscr{D}_f \subset \mathscr{D}_g$ et que g est à valeurs positives, c'est-à-dire qu'elle vérifie $g(t) \geqslant 0$ pour tout $t \in \mathscr{D}_g$. Soit $A \in \mathbf{R}^n$. On suppose que pour tout $\eta \in \mathbf{R}_+^*$ l'intersection $B(a,\eta) \cap \partial D_f \neq \emptyset$ (on dit que le point A est *adhérent* à l'ensemble $\mathscr{D}_f$).

Sous hypothèse, on dit que f est *négligeable relativement à g en A* si pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $M \in B(A, \eta) \cap \mathscr{D}_f$

$$|f(M)| \leqslant \varepsilon g(M).$$

Une fonction négligeable relativement à g en A est souvent notée $o_A(g)$ (voire $o(g)$ si le point A est clair par le contexte).

Dessin 1.10.3. — Prenons $n = 1$, $A = 0 \in \mathbf{R}$ et $g : x \mapsto |x|$. Soit $f : \mathbf{R} \to \mathbf{R}$ une application. Alors dire que $f = o(g)$ signifie que, près de 0, le graphe est contenu dans un cône arbitrairement étroit contenant l'axe des abscisses. Autrement dit, le graphe de f doit être tangent à l'axe des abscisses (figure 12).

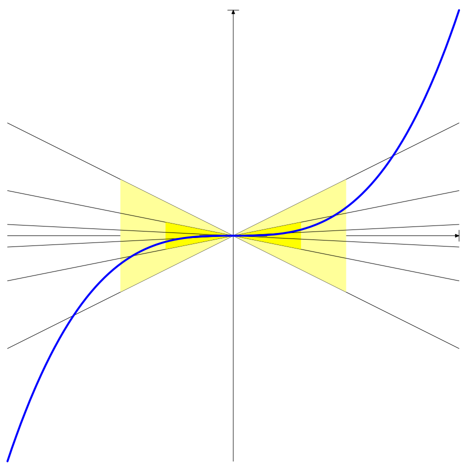


Figure 12. Petit o

Dans la figure 12, la zone colorée en jaune clair est donnée par les conditions

$$|x| < \frac{1}{2} \quad \text{et} \quad |y| \leqslant \frac{1}{2}|x|$$

et celle en jaune légèrement plus sombre par les conditions

$$|x| < \frac{1}{3} \quad \text{et} \quad |y| \leqslant \frac{1}{5}|x|.$$

Exemple 1.10.4. — Ainsi on écrit

$$x^2 = o(|x|) \quad \text{ou} \quad x^3 = o(|x|)$$

ce qui signifie que les applications $X^2 : x \mapsto x^2$ et $X^3 : x \mapsto x^3$ sont toutes les deux négligeables relativement à l'application $x \mapsto |x|$ en 0.

On prendra garde au fait que ces deux égalités *n'impliquent pas* l'égalité de fonctions $X^2 = X^3$! La notation $o(g)$ est une notation pratique pour désigner *n'importe quelle* fonction négligeable relativement à g en le point considéré.

Passons maintenant à la généralisation de l'expression de l'approximation affine d'une fonction f en une variable dérivable en un point a :

$$f(a) + f'(a)(x - a).$$

Définition 1.10.5

Soit f une fonction en n variables réelles à valeurs réelles. Soit A un point du domaine de définition de f de sorte qu'il existe une boule ouverte centrée en A contenu dans ce domaine de définition. On suppose que f admet des dérivées partielles par rapport à toutes les variables $X_1, \ldots, X_n$ en A. Alors on note $\mathrm{d}f_A$ (ou $\mathrm{d}f(A)$) l'application de $\mathbf{R}^n$ dans $\mathbf{R}$ définie par

$$\mathrm{d}f_A : (u_1, \ldots, u_n) \mapsto \frac{\partial f}{\partial X_1}(A)u_1 + \cdots + \frac{\partial f}{\partial X_n}(A)u_n.$$

Faisons le lien avec le gradient en utilisant la notion de produit scalaire :

Définition 1.10.6

Soient $\vec{u} = (u_1, \ldots, u_n)$ et $\vec{v} = (v_1, \ldots, v_n)$ des éléments de $\mathbf{R}^n$. On définit le *produit scalaire de $\vec{u}$ par $\vec{v}$* (pour le produit scalaire usuel) par

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots u_n v_n.$$

On définit également la norme de $\vec{u}$ par

$$||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + \ldots u_n^2}.$$

Remarques 1.10.7. — a) Pour $n = 2$ c'est la formule que vous apprise pour le produit scalaire dans le secondaire lorsqu'on utilise une base orthonormée du plan euclidien.

b) Notons que si $A = (a_1, \dots, a_n)$ et $B = (b_1, \dots, b_n)$ alors la notation $\overrightarrow{AB}$ désigne le n-uplet $(b_1 - a_1, \dots, b_n - a_n)$ et la distance de A à B est donnée par

$$AB = d(A, B) = ||\overrightarrow{AB}||.$$

Par définition, on obtient la formule utile suivante :

Proposition 1.10.8

Avec les notations de la définition 1.10.5, Si $\vec{u} = (u_1, \dots, u_n)$, alors

$$\mathrm{d}f_A(\vec{u}) = \mathrm{grad}_A(f) \cdot \vec{u}.$$

Exemple 1.10.9. — Pour $i \in \{1, \dots, n\}$, notons X_i la fonction « i-ème composante » définie par $X_i : (x_1, \dots, x_n) \mapsto x_i$. Si $A = (a_1, \dots, a_n)$ et $j \in \{1, \dots, n\}$, la fonction partielle $X_i(a_1, \dots, a_{j-1}, \bullet, a_{j+1}, \dots, a_n)$ est la fonction constante de valeur a_i si $j \neq i$ et la fonction identité $x \mapsto x$ si $j = i$. Par conséquent, on obtient que les dérivées partielles de la fonction X_i sont données par

$$\frac{\partial X_i}{\partial X_j} = \begin{cases} 1 & \text{si } j = i, \\ 0 & \text{sinon.} \end{cases}$$

Par conséquent

$$\mathrm{d}X_{iA} : (u_1, \dots, u_n) \longmapsto u_i$$

(Autrement dit, on a l'égalité de fonction $\mathrm{d}X_{iA} = X_i$). Notons que cela donne l'égalité de fonctions

$$\mathrm{d}f_A = \frac{\partial f}{\partial X_1}(A)\mathrm{d}X_{1A} + \dots + \frac{\partial f}{\partial X_n}(A)\mathrm{d}X_{nA} \tag{3}$$

qu'on abrège parfois en [1]

$$\mathrm{d}f = \frac{\partial f}{\partial X_1}\mathrm{d}X_1 + \dots + \frac{\partial f}{\partial X_n}\mathrm{d}X_n.$$

1. Formellement, cette égalité est une égalité entre fonctions de $\mathbf{R}^n$ dans un ensemble d'applications de $\mathbf{R}^n$ dans $\mathbf{R}$.

Terminologie 1.10.10

Une *forme affine* sur $\mathbf{R}^n$ est une application $f : \mathbf{R}^n \to \mathbf{R}$ telle qu'il existe $n+1$ nombre réels $\alpha_1, \dots, \alpha_n$ et β de sorte que

$$f(x_1, \dots, x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n + \beta$$

pour tous $(x_1, \dots, x_n) \in \mathbf{R}^n$. (On dit aussi que f est polynomiale de degré $\leqslant 1$).

Remarque 1.10.11. — Donnons-nous une fonction f en n variables définie sur une boule ouverte $B(A, r)$ où $A = (a_1, \dots, a_n) \in \mathbf{R}^n$. L'idée de l'approximation affine en A est de trouver une forme linéaire g de sorte que la différence $|f(M) - g(M)|$ soit aussi petite que possible quand M tend vers A. Mais si $i \in \{1, \dots, n\}$ et qu'on se restreint aux points M de la forme $M = (a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$ pour $t \in \mathbf{R}$, on obtient que la fonction partielle $g(a_1, \dots, a_{i-1}, \bullet, a_{i+1}, \dots, a_n)$ doit être l'approximation affine en a_i de la fonction partielle $f(a_1, \dots, a_{i-1}, \bullet, a_{i+1}, \dots, a_n)$. Or, par définition, la dérivée en a_i de cette dernière fonction est $\frac{\partial f}{\partial X_i}(A)$. Donc la fonction partielle $g(a_1, \dots, a_{i-1}, \bullet, a_{i+1}, \dots, a_n)$ est la fonction

$$t \mapsto f(A) + \frac{\partial f}{\partial X_i}(A)(t - a_i)$$

Si on introduit les nombres réels $\alpha_1, \dots, \alpha_n$ et β de sorte que

$$g(x_1, \dots, x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n + \beta$$

pour tous $(x_1, \dots, x_n) \in \mathbf{R}^n$, la fonction partielle qui précède s'écrit également

$$t \longmapsto \alpha_i t + c$$

où c est donné par

$$c = \beta + \alpha_1 a_1 + \dots \alpha_{i-1} a_{i-1} + \alpha_{i+1} a_{i+1} + \cdots + \alpha_n a_n.$$

En regardant les valeurs de cette fonction partielle en a_i et $a_i + 1$ on obtient donc que, nécessairement, $\alpha_i = \frac{\partial f}{\partial X_i}(A)$ et $g(A) = f(A)$. Donc la seule possibilité pour g est la fonction affine

$$M \longmapsto f(A) + \mathrm{d}f_A(\overrightarrow{AM}).$$

Cette remarque explique la définition suivante :

Définition 1.10.12

Soit f une fonction en n variables réelles à valeurs réelles et soit $A \in \mathscr{D}_f$. On suppose que $\mathscr{D}_f$ contient une boule ouverte centrée en A et que f admet des dérivées partielles par rapport à toutes les variables en A. On dit que f est *différentiable en* A si

$$f(M) = f(A) + \mathrm{d}f_A(\overrightarrow{AM}) + o_A(AM). \tag{4}$$

L'application $\mathrm{d}f_A$ s'appelle alors la *différentielle de* f *en* A. On dit que f est différentiable si elle l'est en tout point de son domaine de définition $\mathscr{D}_f$.

Même si f admet des dérivée partielles en tout point de $\mathscr{D}_f$ et qu'on peut définir l'application $\mathrm{d}f_A$, la formule d'approximation (4) peut être fausse !

Exemple 1.10.13. — On reprend l'exemple 1.8.3.

$$f : (x,y) \longmapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Rappelons que $f(r\cos(\theta), r\sin(\theta)) = \frac{1}{2}\sin(2\theta)$ et que cette fonction n'est pas continue en $(0,0)$. La proposition 1.10.16 qui suit montre donc que f n'est pas différentiable en $(0,0)$. Pourtant toutes les fonctions partielles sont dérivables : si $a \in \mathbf{R}_+^*$, alors $f(a,\bullet) : t \mapsto \frac{at}{a^2+t^2}$ est dérivable et $f(0,\bullet)$ est la fonction constante de valeur 0 également dérivable. L'égalité de fonctions $f(a,\bullet) = f(\bullet,a)$ pour tout $a \in \mathbf{R}$ implique que la fonction partielle $f(\bullet,a)$ est également dérivable pour tout $a \in \mathbf{R}$.

On a néanmoins le résultat suivant, qui sera admis :

Théorème 1.10.14

Soit f une fonction en n variables à valeurs réelles dont le domaine de définition est un ouvert. Si la fonction est de classe $\mathscr{C}^1$ (c'est-à-dire que les dérivées partielles $\frac{\partial f}{\partial X_i}$ sont bien définies sur $\mathscr{D}_f$ et *continues*) alors la fonction f est différentiable.

Remarques 1.10.15. — a) L'idée derrière les équations (3) et (4) est que si on modifie chaque paramètre X_i d'une « petite » quantité ΔX_i alors le changement pour la valeur de f est

$$\Delta f = \frac{\partial f}{\partial X_1}(A)\Delta X_1 + \cdots + \frac{\partial f}{\partial X_n}(A)\Delta X_n + \varepsilon$$

où ε est un terme d'erreur négligeable par rapport à l'ordre de grandeur de la modification qu'on peut définir comme $\max_{1\leqslant i\leqslant n}(|\Delta X_i|)$. La définition 1.10.12 donne un sens mathématique précis à cette formule d'approximation que vous avez pu rencontrer dans les autres unités d'enseignement.

b) De cette description, on peut déduire que le gradient de f donne la direction dans laquelle la fonction f augmente le plus rapidement et $\|\mathrm{grad}_A(f)\|$ donne la « vitesse » à laquelle f change dans cette direction. Nous donnerons plus tard un sens précis à cette remarque. Comme un écoulement de fluide suit naturellement la direction dans laquelle l'altitude diminue le plus rapidement possible, on peut lire dans la nature le gradient de l'altitude : il pointe dans la direction opposée au sens des cours d'eau.

Proposition 1.10.16

Si une fonction en n variables est différentiable en un point A, alors elle est continue en A.

On prendra également garde au fait qu'une fonction peut être continue en un point sans être différentiable en ce point. Nous en verrons un exemple un peu plus tard.

Preuve de la proposition. — Soit f une fonction différentiable en un point $A=(a_1,\dots,a_n)\in\mathbf{R}^n$. Posons

$$C = 1+\sum_{i=1}^{n}\left|\frac{\partial f}{\partial X_i}(A)\right|.$$

Pour tout point $M=(x_1,\dots,x_n)\in\mathbf{R}^n$, on a

$$|x_i-a_i| = \sqrt{(x_i-a_i)^2}\leqslant\sqrt{\sum_{i=1}^{n}(x_i-a_i)^2} = AM.$$

En utilisant l'inégalité triangulaire $|a+b|\leqslant|a|+|b|$, valable pour tous nombres réels a et b, on en déduit la majoration

$$\left|\mathrm{d}f_A\left(\overrightarrow{AM}\right)\right|\leqslant\sum_{i=1}^{n}\left|\frac{\partial f}{\partial X_i}(A)\right||x_i-a_i|\leqslant(C-1)AM.$$

Donc, par définition de $o_A(AM)$, la formule (4) montre qu'on peut choisir $r_0>0$ de sorte que $B(A,r_0)\subset\mathscr{D}_f$ et que, pour tout $M\in B(A,r_0)$,

$$|f(M)-f(A)|\leqslant C\,AM.$$

Soit ε un nombre réel strictement positif arbitraire. Comme C est positif, on peut poser $r = \min\left(r_0, \frac{\varepsilon}{C+1}\right) \in \mathbf{R}_+^*$. Alors pour tout $M \in B(A, r)$, on obtient

$$|f(M) - f(A)| \leqslant C\, AM \leqslant C\frac{\varepsilon}{C+1} < \varepsilon.$$

La définition de la continuité en A (définition 1.8.1) est donc bien satisfaite. □

Remarque 1.10.17. — Notons qu'en utilisant la notation o, une fonction en n variables f est continue en un point $A \in \mathscr{D}_f$ si et seulement si on a la relation

$$f(M) = f(A) + o_A(1). \tag{5}$$

Intuitivement, la relation (4) est plus précise que la relation (5) ; c'est exactement ce qui est démontré dans la preuve précédente.

Terminologie 1.10.18

Sous les hypothèses de la définition 1.10.12, lorsque la fonction f est différentiable en un point $A = (a_1, \ldots, a_n)$, de son domaine de définition, le graphe de l'application affine

$$M \longmapsto f(A) + \mathrm{d}f_A(\overrightarrow{AM})$$

est un hyperplan affine appelé *hyperplan tangent au graphe* Γ_f au point A (figure 13).

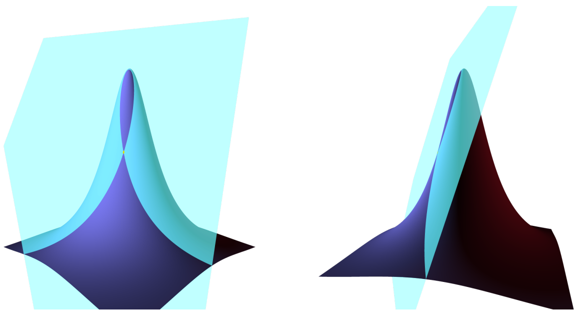


FIGURE 13. Plan tangent

Exemple 1.10.19. — Comme pour les fonctions dérivables, il est rare qu'il soit nécessaire d'utiliser la définition pour prouver qu'une fonction est différentiable. Dans la plupart des cas, il suffit d'utiliser les règles du calcul différentiel pour obtenir le résultat : on suppose que f et g

sont toutes les deux des fonctions en n variables réelles et à valeurs réelles différentiables en un point A. Alors

a) La fonction somme $f+g : (x_1,\ldots,x_n) \mapsto f(x_1,\ldots,x_n)+g(x_1,\ldots,x_n)$ est aussi différentiable en A et

$$\mathrm{d}(f+g)_A = \mathrm{d}f_A + \mathrm{d}g_A\,;$$

b) Si $\lambda \in \mathbf{R}$, la fonction $\lambda f : (x_1,\ldots,x_n) \mapsto \lambda f(x_1,\ldots,x_n)$ est également différentiable en A et

$$\mathrm{d}(\lambda f)_A = \lambda \mathrm{d}f_A\,;$$

c) La fonction produit $fg(x_1,\ldots,x_n) \mapsto f(x_1,\ldots,x_n)g(x_1,\ldots,x_n)$ est différentiable en A et

$$\mathrm{d}(fg)_A = f(A)\mathrm{d}g_A + g(A)\mathrm{d}f_A\,;$$

d) Si $g(A) \neq 0$, la fonction quotient $f/g : (x_1,\ldots,x_n) \mapsto f(x_1,\ldots,x_n)/g(x_1,\ldots,x_n)$ est également différentiable en A et

$$\mathrm{d}(f/g)_A = \frac{1}{g(A)^2}(g(A)\mathrm{d}f_A - f(A)\mathrm{d}g_A);$$

e) Si h est une fonction en une variable réelle dérivable en $f(A)$, alors la composée $h \circ f$ est différentiable en A et

$$\mathrm{d}(h \circ f)_A = h'(f(A))\mathrm{d}f_A.$$

Proposition 1.10.20

Soit f une fonction en n variables à valeurs réelles qui est différentiable en un point $A \in \mathscr{D}_f$. Si f admet un minimum ou un maximum local en A, alors $\mathrm{grad}_A(f)$ est nul.

Démonstration. — On écrit $A = (a_1,\ldots,a_n)$. Soit $i \in \{1,\ldots,n\}$. La fonction partielle $f(a_1,\ldots,a_{i-1},\bullet,a_{i+1},\ldots,a_n)$ est dérivable en a_i et admet un minimum ou un maximum local en ce point; donc sa dérivée s'annule en a_i. Par définition des dérivées partielles cela implique que

$$\frac{\partial f}{\partial X_i}(A) = 0$$

pour tout $i \in \{1,\ldots,n\}$. Par conséquent, $\mathrm{grad}_A(f) = 0$. □

On prendra garde au fait que la réciproque est *fausse* en général ! La condition $\operatorname{grad}_A(f) = 0$ *n'implique pas* que f admet un minimum ou un maximum local en A (*cf.* l'exercice corrigé 1.2). Notons que vous avez déjà vu des contre-exemples pour les fonctions en une variable réelle : L'application ⧉ : $t \mapsto t^3$ est dérivable et ⧉$'(0) = 0$. Néanmoins ⧉ est strictement croissante sur $\mathbf{R}$ et n'a donc ni minimum ni maximum en 0 (figure 14).

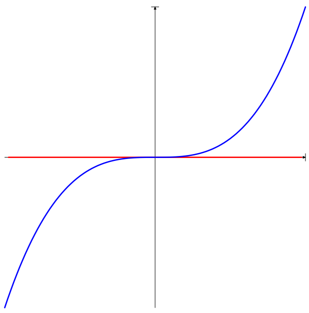


FIGURE 14. Graphe de ⧉

1.11. Fonctions à valeurs dans $\mathbf{R}^m$. — Dans ce paragraphe m et n sont des entiers naturels. Nous allons considérer les fonctions de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Une telle fonction f associe donc à tout élément $A = (x_1, \ldots, x_n) \in \mathscr{D}_f$ un m-uplet de nombre réels

$$f(A) = f(x_1, \ldots, x_n) \in \mathbf{R}^m.$$

Remarques 1.11.1. — a) Soit $f_1, \ldots, f_m$ des fonctions en n variables réelles à valeurs réelles ayant le même ensemble de définition $\mathscr{D} \subset \mathbf{R}^n$. On peut alors considérer la fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$, définie sur $\mathscr{D}$ et donnée par

$$f : (x_1, \ldots, x_n) \longmapsto (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)).$$

b) Inversement, soit f est une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ définie sur un domaine de définition $\mathscr{D}_f$. Pour tout $i \in \{1, \ldots, m\}$ notons Y_i l'application de $\mathbf{R}^m$ dans $\mathbf{R}$ correspondant à la i-ème composante d'un m-uplet. Autrement dit $Y_i : (y_1, \ldots, y_m) \mapsto y_i$. Alors la composée $f_i = Y_i \circ f$ est une fonction en n variables réelles à valeurs réelles définie sur $\mathscr{D}_f$ qui à un n-uplet $(x_1, \ldots, x_n)$ associe la i-ème composante du m-uplet $f(x_1, \ldots, x_n)$. Cela définit donc m fonctions de $\mathbf{R}^n$ dans $\mathbf{R}$ et

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$$

pour tout $(x_1, \ldots, x_n) \in \mathscr{D}_f$.

On peut résumer le contenu de cette remarque de la façon suivante :

Notation 1.11.2

Se donner une fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ définie sur $\mathscr{D}$ revient exactement à se donner m fonctions $f_1, \ldots, f_m$ de $\mathbf{R}^m$ dans $\mathbf{R}$ définies sur $\mathscr{D}$, via la relation

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$$

pour tout $(x_1, \ldots, x_n) \in \mathscr{D}_f$. Par abus de langage, on note alors $f = (f_1, \ldots, f_m)$ et on dit que f_i est la *i-ème composante de f*.

Le cas où $n = 1$ est particulièrement utile :

Définition 1.11.3

Une *courbe paramétrée* est une application de $\mathbf{R}$ dans $\mathbf{R}^m$.

Exemple 1.11.4. — Considérons l'application

$$f : t \longmapsto (\cos(t), \sin(t))$$

(figure 15). Elle est à valeur dans le cercle unité donné par

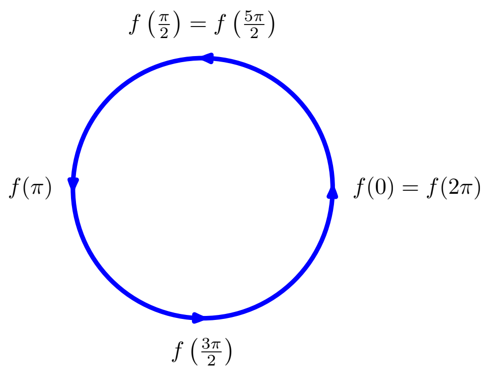


Figure 15. Courbe circulaire

$$\mathbf{S}^1 = \{ (x, y) \in \mathbf{R}^2 \mid x^1 + y^2 = 1 \}$$

Cette application est 2π-périodique, c'est-à-dire qu'elle vérifie
$$f(t+2\pi) = f(t)$$
pour tout $t \in \mathbf{R}$.

Remarque 1.11.5. — Les courbes paramétrées sont notamment utiles pour décrire l'évolution d'un système au cours du temps. Pensez par exemple à un système formé de N particules dans l'espace, chacune ayant trois coordonnées. La configuration à l'instant t est donné par un point $M(t)$ de $\mathbf{R}^{3N}$, ce qui donne une courbe paramétrée à valeurs dans $\mathbf{R}^{3N}$.

Remarque 1.11.6. — Considérons une fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Supposons que chacune de ses composantes f_i admet des dérivées partielles en un points $A \in \mathbf{R}^n$. En ce point, ces dérivées nous donnent donc $n \times m$ nombres réels $\frac{\partial f_i}{\partial X_j}$ où $1 \leqslant i \leqslant m$ et $1 \leqslant j \leqslant n$.

Définition 1.11.7

Pour une courbe paramétrée f à valeurs dans $\mathbf{R}^m$ si toutes les composantes f_i admettent une dérivée en un point $t \in \mathscr{D}_f$. Alors le *vecteur vitesse* en t est le m-uplet
$$f'(t) = (f_1'(t), \ldots, f_m'(t)).$$
Si les composantes sont dérivable deux fois en t, le *vecteur accélération* est défini par :
$$f''(t) = (f_1''(t), \ldots, f_m''(t)).$$
Plus géralement, on peut définier les dérivées k-èmes $f^{(k)}(t)$ à partir des dérivées k-èmes des composantes.

Remarque 1.11.8. — Notons que les composantes d'une courbe paramétrée sont des fonctions de $\mathbf{R}$ dans $\mathbf{R}$, et que cela fait donc sens de parler de leur dérivée.

Exemple 1.11.9. — Reprenons l'exemple 1.11.4. La vitesse à l'instant t est donnée par
$$f'(t) = (-\sin(t), \cos(t))$$
et l'accélération par
$$f''(t) = (-\cos(t), -\sin(t))$$
comme illustré par la figure 16

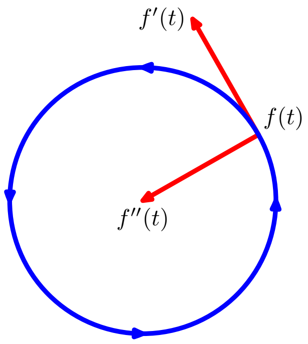


FIGURE 16. Vitesse et accélération

Définition 1.11.10

On dite qu'une fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est différentiable en $A \in \mathscr{D}_f$ si pour tout $j \in \{1, \dots, m\}$ la j-ème composante f_j de f est différentiable et l'application

$$\begin{array}{rcl} \mathrm{d}f_A : \mathbf{R}^n & \longrightarrow & \mathbf{R}^m \\ \vec{u} & \longmapsto & (\mathrm{d}(f_1)_A(\vec{u}), \dots, \mathrm{d}(f_m)_A(\vec{u})) \end{array}$$

s'appelle la différentielle de f en A. (Autrement dit, la j-ème composante de la différentielle de f en A est la différentielle de la j-ème composante de f en A, $(\mathrm{d}f_A)_j = \mathrm{d}(f_j)_A$.

Exemple 1.11.11. — Par définition, pour une courbe paramétrée φ, on a les relations

$$\mathrm{d}\varphi_t(u) = u\varphi'(t) \text{ et } \varphi'(t) = \mathrm{d}\varphi_t(1).$$

Définition 1.11.12

Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit g une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$. On suppose que le domaine de définition de f est contenu dans celui de g : $\mathscr{D}_f \subset \mathscr{D}_g$. Soit $A \in \mathbf{R}^n$ un point tel que pour tout nombre réel $\eta > 0$, $B(A, \eta) \cap \mathscr{D}_f \neq \emptyset$. Alors on dit que f est *négligeable par rapport à g en A* si elle vérifie une des conditions équivalentes suivantes :

(i) Pour tout $j \in \{1, \dots, m\}$, on a $f_j = o_A(g)$;

(ii) La fonction $\|f\| : (x_1, \dots, x_n) \mapsto \|f(x_1, \dots, x_n)\|$ est négligeable par rapport à g en A.

Si c'est conditions sont vérifiées, on note, comme pour les fonctions à valeurs réelles, $f = o_A(g)$ ou $f = o(g)$.

Proposition 1.11.13

Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ différentiable en un point $A \in \mathscr{D}_f$, alors

$$f(M) = f(A) + \mathrm{d}f_A(\overrightarrow{AM}) + o(AM).$$

Cette formule découle de la formule valide pour chacune des composantes.

Proposition 1.11.14

Soient $m, n, p \in \mathbf{N}$. Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit g une fonction de $\mathbf{R}^m$ dans $\mathbf{R}^p$. On suppose que f est différentiable en $A \in \mathscr{D}_f$, que $f(A) \in \mathscr{D}_g$ et que g est différentiable en $f(A)$. Alors la composée $g \circ f$ est différentiable en A et sa différentielle en ce point est donnée par la formule :

$$\mathrm{d}(g \circ f)_A = \mathrm{d}g_{f(A)} \circ \mathrm{d}f_A.$$

Cela se démontre comme pour la composée de fonctions en une variable.

Exemple 1.11.15. — Appliquons cela à un changement de système de coordonnées : on considère la fonction de $\mathbf{R}^2$ dans $\mathbf{R}^2$ donnée par $(r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, dont la restriction à $\mathbf{R}_+ \times [0, 2\pi[$ définit les *coordonnées polaires*. Soit f une fonction de $\mathbf{R}^2$ dans $\mathbf{R}$. Calculons la différentielle de φ :

$$\mathrm{d}\varphi_{r,\theta}(u, v) = (\cos(\theta)u - r\sin(\theta)v, \sin(\theta)u + r\cos(\theta)v).$$

On obtient

$$\begin{aligned}\mathrm{d}(f \circ \varphi)_{r,\theta}(u, v) &= \mathrm{d}f_{\varphi(r,\theta)} \circ \mathrm{d}\varphi_{(r,\theta)}(u, v)\\ &= \frac{\partial f}{\partial X}(\varphi(r, \theta)) \times (\cos(\theta)u - r\sin(\theta)v)\\ &\quad + \frac{\partial f}{\partial Y}(\varphi(r, \theta)) \times (\sin(\theta)u + r\cos(\theta)v)\end{aligned}$$

Ce qui nous permet d'écrire

$$\frac{\partial f \circ \varphi}{\partial r} = \cos(\theta)\frac{\partial f}{\partial X} + \sin(\theta)\frac{\partial f}{\partial Y}$$
$$\frac{\partial f \circ \varphi}{\partial \theta} = r\left(-\sin(\theta)\frac{\partial f}{\partial X} + \cos(\theta)\frac{\partial f}{\partial Y}\right)$$

et

$$\frac{\partial f}{\partial X} = \cos(\theta)\frac{\partial f \circ \varphi}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial f \circ \varphi}{\partial \theta}$$
$$\frac{\partial f}{\partial Y} = \sin(\theta)\frac{\partial f \circ \varphi}{\partial r} + \frac{\cos(\theta)}{r}\frac{\partial f \circ \varphi}{\partial \theta}$$

Une autre application du calcul de la composée de la différentielle est donné par la proposition suivante :

Proposition 1.11.16

Soit f une fonction en n variables réelles à valeurs réelles et soit φ une courbe paramétrée *à valeur dans une des lignes de niveau de* f (autrement dit la composée $f \circ \varphi$ est constante). Alors pour tout $t \in \mathscr{D}_\varphi$, la vitesse $\varphi'(t)$ est orthogonale au gradient $\mathrm{grad}_{\varphi(t)}(f)$:

$$\varphi'(t) \cdot \mathrm{grad}_{\varphi(t)}(f) = 0.$$

Démonstration. — Comme l'application composée $f \circ \varphi$ qui va de $\mathbf{R}$ dans $\mathbf{R}$ est supposée constante, la dérivée s'annule : $(f \circ \varphi)' = 0$. Calculons cette dérivée à l'aide de la différentielle de la composée :

$$(f \circ \varphi)'(t) = \mathrm{d}(f \circ \varphi)_t(1) = \mathrm{d}f_{\varphi(t)} \circ \mathrm{d}\varphi_t(1) = \mathrm{d}f_{\varphi(t)}(\varphi'(t)) = \mathrm{grad}_{\varphi(t)}(f) \cdot \varphi'(t)$$

où la dernière égalité résulte de la proposition 1.10.8. □

Remarque 1.11.17. — Sur une carte IGN, cela se traduit par le fait que les cours d'eaux naturels croisent à angle droit les courbes de niveau (figure 1.3.5).

Terminons ce chapitre avec la formule donnant les dérivées partielles d'une composée.

Formules

Soient $m, n, p \in \mathbf{N}$. Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit g une fonction de $\mathbf{R}^m$ dans $\mathbf{R}^p$. On suppose que f est différentiable en $A \in \mathscr{D}_f$ et que g est différentiable en $f(A)$. Si on notes $X_1, \ldots, X_n$ les variables pour f et $Y_1, \ldots, Y_m$ les variables pour g, alors

$$\frac{\partial (g \circ f)_k}{\partial X_i}(A) = \sum_{j=1}^{m} \frac{\partial g_k}{\partial Y_j}(f(A)) \frac{\partial f_j}{\partial X_i}(A).$$

Nous reviendrons sur cette formule dans le chapitre sur le calcul matriciel.

Explication. — Posons $A = (x_1, \ldots, x_n)$. Alors $\frac{\partial (g\circ f)_k}{\partial X_i}(A)$ est, par définition, la dérivée de l'application partielle $(g \circ f)_k(x_1, \ldots, x_{i-1}, \bullet, x_{i+1}, \ldots, x_n)$ en x_i, c'est-à-dire

$$\begin{aligned}
&(g_k \circ f(x_1, \ldots, x_{i-1}, \bullet, x_{i+1}, \ldots, x_n))'(x_i) \\
&= \mathrm{d}(g_k \circ f(x_1, \ldots, x_{i-1}, \bullet, x_{i+1}, \ldots, x_n))_{x_i}(1) \\
&= (\mathrm{d}g_k)_{f(A)} \circ \mathrm{d}(f(x_1, \ldots, x_{i-1}, \bullet, x_{i+1}, \ldots, x_n))_{x_i}(1) \\
&= (\mathrm{d}g_k)_{f(A)} \left(\frac{\partial f_1}{\partial X_i}, \ldots, \frac{\partial f_m}{\partial X_i} \right) \\
&= \sum_{j=1}^{m} \frac{\partial g_k}{\partial Y_j}(f(A)) \frac{\partial f_j}{\partial X_i}(A).
\end{aligned}$$

□

Exercices corrigés

Donnons tout d'abord un exemple d'étude de fonctions en deux variables :

Énoncé 1.1

On considère les fonctions f et g en deux variables définies par

$$f : (x, y) \longmapsto x^2 + y^2 \quad \text{et} \quad g : (x, y) \longmapsto \sqrt{x^2 + y^2}.$$

1. Décrire géométriquement les lignes de niveau de f et de g.
2. Justifier que f et g admettent un minimum global en un point que l'on précisera.
3. Justifier que f est différentiable et calculer $\operatorname{grad}_A(f)$ pour tout $A \in \mathbf{R}^2$.
4. Déterminer l'ensemble des points $A \in \mathbf{R}^2$ tels que $\operatorname{grad}_A(f) = (0, 0)$.
5. Écrire g comme une composée et en déduire que g est différentiable en tout point $A \in \mathbf{R}^2 - \{(0, 0)\}$ et calculer $\operatorname{grad}_A(g)$. Peut-on avoir $\operatorname{grad}_A(g) = (0, 0)$?
6. L'application g est-elle différentiable en $(0, 0)$?
7. Dessiner l'allure du graphe de g.

Solution. — 1. Soit $a \in \mathbf{R}$, la ligne de niveau correspondante pour f s'écrit

$$f^{-1}(\{a\}) = \{\, (x, y) \in \mathbf{R}^2 \mid f(x, y) = a \,\} = \{\, (x, y) \in \mathbf{R}^2 \mid x^2 + y^2 = a \,\}.$$

Notons O le point $(0, 0) \in \mathbf{R}^2$. Si $M = (x, y) \in \mathbf{R}^2$, alors la distance OM vaut $\sqrt{x^2 + y^2}$ et $x^2 + y^2 = OM^2$. Comme $x^2 + y^2 \geqslant 0$ pour tout $(x, y) \in \mathbf{R}^2$, l'ensemble $f^{-1}(\{a\})$ est vide si $a < 0$ et il ne contient que $(0, 0)$ si $a = 0$. Si $a > 0$, alors $x^2 + y^2 = a$ équivaut à $OM^2 = a$ et donc à $OM = \sqrt{a}$. L'ensemble $f^{-1}(\{a\})$ est donc le cercle de centre O et de rayon $\sqrt{a}$.

De même si $M = (x, y) \in \mathbf{R}^2$, alors la condition $g(x, y) = a$ équivaut à $OM = a$. L'ensemble $g^{-1}(\{a\})$ est donc vide si $a < 0$, réduit à $\{(0, 0)\}$ si $a = 0$ et un cercle de centre O et de rayon a si $a > 0$.

2. Pour tout point $(x, y) \in \mathbf{R}^2$, on a

$$f(x, y) = x^2 + y^2 \geqslant 0 = f(0, 0).$$

Donc f admet un minimum global en $(0, 0)$ qui vaut 0.

De même, pour tout $(x,y) \in \mathbf{R}^2$
$$g(x,y) = \sqrt{x^2+y^2} \geqslant 0 = g(0,0)$$
et donc la fonction g admet un minimum global en $(0,0)$ qui vaut également 0.

3. On a l'égalité de fonctions $f = X^2 + Y^2$ où $X : (x,y) \mapsto x$ et $Y : (x,y) \mapsto y$ sont les fonctions composantes qui sont différentiables. Donc la fonction en deux variables f est différentiable. En outre en utilisant les formules pour le gradient, si $A = (x,y) \in \mathbf{R}^2$,
$$\begin{aligned}\mathrm{grad}_A(f) &= \mathrm{grad}_A(X^2+Y^2)\\ &= \mathrm{grad}_A(X\times X) + \mathrm{grad}_A(Y\times Y)\\ &= 2X(A)\times \mathrm{grad}_A(X) + 2Y(A)\times \mathrm{grad}_A(Y)\\ &= 2x(1,0)+2y(0,1)\\ &= (2x,2y).\end{aligned}$$
Autre solution pour le calcul du gradient : Considérons les fonctions partielles
$$f(\bullet,y) : t \longmapsto t^2+y^2 \text{ et } f(x,\bullet) : t \longmapsto x^2+t^2$$
elles sont toutes les deux dérivables et leur dérivées respectives sont données par
$$\frac{\partial f}{\partial X}(A) = f(\bullet,y)'(x) = 2x$$
et
$$\frac{\partial f}{\partial Y}(A) = f(x,\bullet)'(y) = 2y$$
Donc
$$\mathrm{grad}_A(f) = \left(\frac{\partial f}{\partial X}(A), \frac{\partial f}{\partial Y}(A)\right) = (2x,2y).$$

4. Soit $A = (x,y) \in \mathbf{R}^2$, alors $\mathrm{grad}_A(f) = (0,0)$ équivaut à $(2x,2y) = (0,0)$ et donc à $2x = 0$ et $2y = 0$. Donc $\mathrm{grad}_A(f) = (0,0)$ uniquement si $A = (0,0)$. Notons que f admet un minimum en $(0,0)$ et est différentiable en ce point donc on savait que son gradient est nul en $(0,0)$.

5. On peut écrire $g = h\circ f$ où $h : t \mapsto \sqrt{t}$. Comme la fonction h est dérivable sur $\mathbf{R}_+^*$ et que l'application f ne s'annule qu'en $(0,0)$, il en résulte que g est différentiable en tout point de $\mathbf{R}^2 - \{(0,0)\}$ et
$$\mathrm{grad}_{(x,y)}(g) = h'(f(x,y))\times \mathrm{grad}_{(x,y)}(f) = \frac{1}{2\sqrt{x^2+y^2}}(2x,2y).$$
Il en résulte que pour tout $(x,y) \neq (0,0)$, $\mathrm{grad}_{(x,y)}(g) \neq (0,0)$.

6. La fonction partielle $g(0,\bullet) : t \mapsto \sqrt{t^2} = |t|$ n'est pas dérivable en 0, la fonction g ne peut donc pas être différentiable en $(0,0)$.
7. L'allure du graphe de g est représenté de deux manières différentes sur la figure 17. Il s'agit d'un cône dont le sommet correspond au point en lequel la fonction n'est pas différentiable.

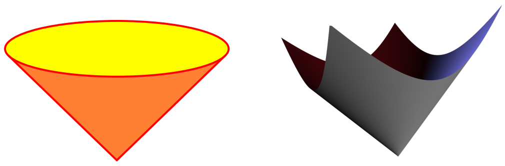


FIGURE 17. Graphe de la distance à O

□

Passons à un exemple où le gradient de la fonction s'annule en un point mais la fonction n'admet ni minimum ni maximum en ce point

Énoncé 1.2

On considère l'applications en deux variables

$$f : (x,y) \longmapsto xy.$$

1. Représenter sur un unique dessin
   (a) L'ensemble des $(x,y) \in \mathbf{R}^2$ tels que $f(x,y) = 0$;
   (b) L'ensemble des $(x,y) \in \mathbf{R}^2$ tels que $f(x,y) > 0$;
   (c) L'ensemble des $(x,y) \in \mathbf{R}^2$ tels que $f(x,y) < 0$;
2. Prouver que f est différentiable et calculer $\mathrm{grad}_A(f)$ pour $A \in \mathbf{R}^2$.
3. Pour quel A a-t-on $\mathrm{grad}_A(f) = 0$? Dire pour chacun d'entre eux s'il s'agit d'un minimum local ou d'un maximum local.

Solution. — 1. L'inégalité $xy > 0$ équivaut à

$$(x > 0 \text{ et } y > 0) \text{ ou } (x < 0 \text{ et } y < 0)$$

et l'inégalité $xy < 0$ équivaut à

$$(x > 0 \text{ et } y < 0) \text{ ou } (x < 0 \text{ et } y > 0).$$

les ensembles demandés sont représentés sur la figure 18

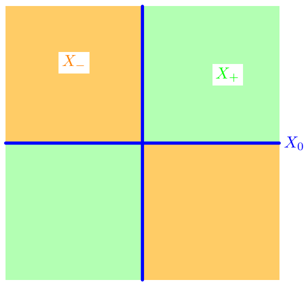


FIGURE 18. Signe du produit

2. L'application f est différentiable comme produit des fonctions différentiables $X : (x, y) \mapsto x$ et $Y : (x, y) \mapsto y$ et
$$\mathrm{grad}_A(XY) = Y(A)\,\mathrm{grad}_A(X) + X(A)\,\mathrm{grad}_A(Y) = y(1,0) + x(0,1) = (y, x)$$
pour tout point $A = (x, y) \in \mathbf{R}^2$.
3. Le gradient $\mathrm{grad}_A(f)$ est nul si et seulement si $A = (0,0)$. Mais pour tout nombre réel $r > 0$, on peut trouver dans la boule centrée en $(0,0)$ et de rayon r les deux points $A = (\frac{r}{2}, \frac{r}{2})$ pour lequel $f(A) = \frac{r^2}{4} > 0 = f(0,0)$ et $B = (\frac{r}{2}, -\frac{r}{2})$ pour lequel $f(B) = -\frac{r^2}{4} < 0 = f(0,0)$ donc f n'a ni maximum local ni minimum local en $(0,0)$. (Sur la figure 18, cela correspond au fait que *n'importe quel* disque ouvert centré en $(0,0)$ contient à la fois des point coloriés en orange et des points coloriés en vert.)

□

Fiche de révision

1.1. Formules de dérivées

Formules

$$
\begin{aligned}
(u+v)' &= u'+v' \\
(\lambda u) &= \lambda u' \text{ pour tout } \lambda \in \mathbf{R} \\
(uv)' &= u'v+uv' \\
\left(\frac{u}{v}\right)' &= \frac{u'v-uv'}{v^2} \\
(v\circ u)' &= u'(v'\circ u) \\
\exp' &= \exp \\
\cos' &= -\sin \\
\sin' &= \cos
\end{aligned}
$$

1.2. Gradient

Formules

$$
\begin{aligned}
\operatorname{grad}_A(f+g) &= \operatorname{grad}_A(f)+\operatorname{grad}_A(g) \\
\operatorname{grad}_A(\lambda f) &= \lambda \operatorname{grad}_A(f) \\
\operatorname{grad}_A(fg) &= f(A)\operatorname{grad}_A(g)+g(A)\operatorname{grad}_A(f) \\
\operatorname{grad}_A(f/g) &= \frac{1}{g(A)^2}\left(g(A)\operatorname{grad}_A(f)-f(A)\operatorname{grad}_A(g)\right) \\
\operatorname{grad}_A(h\circ f) &= h'(f(A))\operatorname{grad}_A(f).
\end{aligned}
$$

1.3. Différentielle

Définition 1.3.1

Soit f une fonction en n variables réelles à valeurs réelles et soit $A \in \mathscr{D}_f$. On suppose que $\mathscr{D}_f$ contient une boule ouverte centrée en A et que f admet des dérivées partielles par rapport à toutes les variables en A. On dit que f est *différentiableen A* si

$$f(M) = f(A) + \mathrm{d}f_A(\overrightarrow{AM}) + o_A(AM). \tag{6}$$

L'application $\mathrm{d}f_A$ s'appelle alors la *différentielle de f en A*. On dit que f est différentiable si elle l'est en tout point de son domaine de définition $\mathscr{D}_f$.

Proposition 1.3.2

Soient $m, n, p \in \mathbf{N}$. Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit g une fonction de $\mathbf{R}^m$ dans $\mathbf{R}^p$. On suppose que f est différentiable en $A \in \mathscr{D}_f$, que $f(A) \in \mathscr{D}_g$ et que g est différentiable en $f(A)$. Alors la composée $g \circ f$ est différentiable en A et sa différentielle en ce point est donnée par la formule :

$$\mathrm{d}(g \circ f)_A = \mathrm{d}g_{f(A)} \circ \mathrm{d}f_A.$$

Entraînement

1.1. Exercices

1.1.1. *Fonctions d'une variable*

Exercice 1.1. Pour chacune des fonctions d'une variable réelle suivantes, déterminer le domaine de définition et tracer le graphe de la fonction

1. $f_1 : x \mapsto \frac{x}{3}$;
2. $f_2 : x \mapsto 2x + 3$;
3. $f_3 : x \mapsto 1 + \frac{1}{x+2}$;
4. $f_4 : x \mapsto |2x - 3|$.

Exercice 1.2. Pour chacune des fonctions d'une variable réelle suivantes, déterminer le domaine de définition de la fonction et sa dérivée en tout point où elle est dérivable. Faire alors le tableau de variations de la fonction et tracer le graphe de la fonction.

1. $f_1 : x \mapsto x^2 - 2x + 1$;
2. $f_2 : x \mapsto \sqrt{3 - 2x}$;
3. $f_3 : x \mapsto \frac{1}{x^2+1}$;
4. $f_4 : x \mapsto \frac{x-1}{x+1}$;
5. $f_5 : x \mapsto \cos(x)$;
6. $f_6 : x \mapsto e^{-x^2}$
7. $f_7 : x \mapsto \begin{cases} \cos\left(\frac{1}{x}\right) & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$

1.1.2. *Fonctions en plusieurs variables*

Exercice 1.3. Trouver pour les applications suivantes le sous-ensemble de $\mathbf{R}^2$ sur lequel elles sont définies.

1. $f_1 : (x, y) \mapsto \frac{1}{x^2 - y}$,
2. $f_2 : (x, y) \mapsto \sqrt{x^2 - y}$,
3. $f_3 : (x, y) \mapsto \ln(x + \ln(y))$,
4. $f_4 : (x, y) \mapsto \mathrm{e}^{x^2 - y^2}$,
5. $f_5 : (x, y) \mapsto \frac{1}{y \sin(x)}$

Exercice 1.4. Pour chacune des fonctions en deux variables suivantes, déterminer leur domaine de définition et représenter l'allure du graphe en en précisant les éventuelles symétries.

1. $f_1 : (x,y) \mapsto \sqrt{x^2+y^2}$;
2. $f_2 : (x,y) \mapsto (x+2)^2+3$;
3. $f_3 : (x,y) \mapsto x+2y+3$
4. $f_4 : (x,y) \mapsto x+2y$
5. $f_5 : (x,y) \mapsto -x+y+4$
6. $f_6 : (x,y) \mapsto -x+y$

Exercice 1.5 (Équation d'état des gaz parfaits). L'équation d'état de n moles d'un gaz parfait est donnée par l'équation :

$$pV = nRT$$

où p désigne la pression dont on suppose qu'elle est uniforme sur le volume considéré, V est le volume de gaz, R est la constante universelle des gaz parfaits et T la température absolue (en Kelvin) également supposée uniforme.

1. (a) Exprimer T comme une fonction en deux variables p et V.
   (b) Quel est le domaine de définition de cette fonction du point de vue mathématique ?
   (c) Quelle partie de ce domaine a un sens du point de vue physique ?
   (d) Tracer l'allure du graphe de cette fonction.
   (e) Tracer des courbes de niveau de cette fonction.
2. (a) L'ensemble des solutions à coordonnées strictement positives de l'équation forme-t-il également le graphe d'une fonction g en les variables p et T ?
   (b) Quel est son domaine de définition ?
   (c) Tracer des courbes de niveau de la fonction g.

Exercice 1.6. On considère la fonction en deux variables f définie par la formule

$$f : (x,y) \mapsto \cos\left(\sqrt{x^2+y^2}\right).$$

1. Quelle est le domaine de définition de cette application ?
2. Exprimer f comme la composée de deux applications.
3. Rapeller l'équation cartésienne dans un repère othornomé d'un cercle $\mathscr{C}$ de centre un point A de coordonnées (x_A, y_A) et de rayon un nombre réel $R > 0$.
4. Justifier que les lignes de niveaux de f sont des réunions de cercles centrés en l'origine. Tracer la ligne de niveau $f^{-1}(\{0\})$.

Dans la suite, on fixe un point $\Omega = (x_\Omega, y_\Omega) \in \mathbf{R}^2$ et on considère l'application

$$f_\Omega : (x,y) \mapsto \cos\left(\sqrt{(x-x_\Omega)^2 + (y-y_\Omega)^2}\right).$$

5. Écrire f_Ω comme la composée de *trois* applications.
6. Décrire les lignes de niveau de f_Ω.
7. Donner un maximum et un minimum global pour la fonction f_Ω.

1.1.3. *Application réciproque*

Exercice 1.7. Pour chacune des fonctions d'une variable réelle suivantes, déterminer le domaine de définition de la fonction et sa dérivée en tout point où elle est dérivable. Faire alors le tableau de variations de la fonction et tracer le graphe de la fonction.

1. $f_1 : x \mapsto \arcsin(x)$;
2. $f_2 : x \mapsto \arccos(x)$;
3. $f_3 : x \mapsto \arctan(x)$.

Quelle est la valeur de l'intégrale

$$\int_{-\infty}^{+\infty} \frac{1}{1+x^2} \mathrm{d}x\,?$$

1.1.4. *Dérivées partielles, gradient*

Exercice 1.8. Soit n un entier strictement positif et soit $k \in \{1, \ldots, n\}$. On note X_k l'application k-ème composante de $\mathbf{R}^n$ dans $\mathbf{R}$ définie par

$$X_k : (x_1, \ldots, x_n) \mapsto x_k.$$

1. Dans cette question uniquement on suppose que $n = 2$.
   (a) Décrire les lignes de niveau de X_1 et X_2.
   (b) Donner une description géométrique des graphes de X_1 et X_2.
2. Soit $a_1, \ldots, a_n$ des nombres réels ; on pose $A = (a_1, \ldots, a_n)$.
   (a) Décrire en des termes simples la fonction partielle $X_k(a_1, \ldots, a_{i-1}, \bullet, a_{i+1}, \ldots, a_n)$ (on distinguera les cas $i \neq k$ et $i = k$).
   (b) Calculer les dérivées partielles $\frac{\partial X_k}{\partial X_i}(A)$ pour $i \in \{1, \ldots, n\}$.
   (c) Que vaut le gradient $\mathrm{grad}_A(X_k)$?
   (d) Décrire l'application $\mathrm{d}X_{k\,A}$.

Exercice 1.9. Calculer les dérivées partielles des fonctions suivantes en précisant leur domaine de définition.

1. $f_1 : (x, y) \mapsto x \ln(y^2 + 1)$;
2. $f_2 : (x, y) \mapsto (x - y) \ln(x^2 - y^2)$;
3. $f_3 : (x, y) \mapsto e^{(x^2+y^3)} - \cos(xy)$;
4. $f_4 : (x, y) \mapsto \dfrac{x^2}{x + 2y}$.

Pour chacune des fonctions précédentes calculer la différence $\dfrac{\partial^2 f}{\partial x \partial y}(x, y) - \dfrac{\partial^2 f}{\partial y \partial x}(x, y)$ pour tout (x, y) appartenant au domaine de définition de la fonction. Que constatez-vous ?

Exercice 1.10. On définit les applications de $\mathbf{R}^2$ dans $\mathbf{R}$ suivantes :

$$\begin{aligned} f_1 &: (x, y) \mapsto \sqrt{x^2 + y^2} \\ f_2 &: (x, y) \mapsto (x + 2)^2 + 3 \\ f_3 &: (x, y) \mapsto x + 2y + 3 \\ f_4 &: (x, y) \mapsto x^2 - 9y^2 \end{aligned}$$

Dans la suite de l'énoncé l'entier i parcourt $\{1, 2, 3, 4\}$.

1. Pour toute valeur c de l'application f_i, trouver une fonction $g_{i,c}$ de $\mathbf{R}$ dans $\mathbf{R}^2$ dont l'image, c'est-à-dire l'ensemble des valeurs, est la ligne de niveau
$$f_i^{-1}(\{c\}) = \{ (x, y) \in \mathbf{R}^2 | f_i(x, y) = c \}$$
2. Calculer le gradient de f_i.
3. Tracer quelques lignes de niveau et représenter le gradient en quelques points de ces lignes de niveau. Que constatez-vous ?
4. Calculer le vecteur dérivé (ou vitesse) $g'_{i,c}(t)$ et faire le produit scalaire de $g'_{i,c}(t)$ et du gradient de f_i en $g_{i,c}(t)$. Qu'obtenez-vous ?

Exercice 1.11. On considère la fonction $f : \mathbf{R}^2 \to \mathbf{R}$ définie par

$$f : (x, y) \longmapsto \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

1. Expliquer pourquoi f admet des dérivées partielles en tout point (x, y) de $\mathbf{R}^2$ et les calculer (on traitera à part le cas $(x, y) = (0, 0)$).

2. Démontrer que $\dfrac{\partial^2 f}{\partial x \partial y}(0,0)$ et $\dfrac{\partial^2 f}{\partial y \partial x}(0,0)$ existent et les comparer.

1.1.5. *Équations aux dérivées partielles*

Problème 1.1 (Équation des ondes). On note Δ l'opérateur *laplacien*

$$\Delta = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2}.$$

L'*équation des ondes* ou *équation de* D'ALEMBERT pour une fonction f en trois variables notées T, X et Y s'écrit

$$\Delta f = \frac{1}{c^2}\frac{\partial^2 f}{\partial T^2}. \tag{7}$$

On se donne un point $A \in \mathbf{R}^2$, un vecteur $u \in \mathbf{R}^2$ de norme égale à 1 et on considère l'application de $\mathbf{R} \times \mathbf{R}^2$ dans $\mathbf{R}$ donnée par

$$f_{A,u} : (t, M) \mapsto \cos\left(u \cdot \overrightarrow{AM} - ct\right)$$

où $u \cdot v$ désigne le produit scalaire des vecteurs u et v.

1. Soit $t_0 \in \mathbf{R}$. Décrire les lignes de niveau de l'application $M \mapsto f_{A,u}(t_0, M)$.
2. Exprimer $f_{A,u}$ en termes des coordonnées (t, x, y).
3. Vérifier que $f_{A,u}$ est solution de l'équation (7).
4. On pose $A = (-1, 0)$, $B = (1, 0)$, $u = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ et $v = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. On pose $g_1 = f_{A,u} + f_{B,v}$.
   (a) Justifier que g_1 est solution de (7).
   (b) Soit M un point de $\mathbf{R}^2$. Donner la valeur minimale α et la valeur maximale β de la fonction $t \mapsto g_1(t, M)$. La différence $\beta - \alpha$ est appelée *l'amplitude* de g_1 en M.
   (c) Pour quels points M l'amplitude de g_1 est-elle maximale ?
   (d) Existe-t-il des points en lesquels l'amplitude est nulle ?
5. Soient $f_1, \ldots, f_m$ des solutions de (7) et soient $\alpha_1, \ldots, \alpha_m$ des nombres réels. Que peut-on dire de l'application

$$\alpha_1 f_1 + \cdots + \alpha_m f_m ?$$

Algèbre linéaire

Emmanuel Peyre

Cours

2.1. Introduction. — Dans le secondaire vous avez vu, dans le plan ou dans l'espace, comment à un couple (A, B) de points est associé un vecteur, noté $\overrightarrow{AB}$. On a $\overrightarrow{AB} = \overrightarrow{CD}$ si les segments $[A, D]$ et $[B, C]$ ont le même milieu (ou si ABDC est un parallélogramme, figure 1). On peut définir

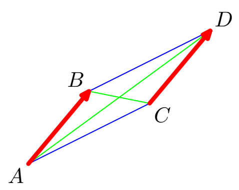


FIGURE 1. Égalité de vecteurs

sur l'ensemble E des vecteurs une addition grâce à la *relation de Chasles* (figure 2), pour tous points A, B, C de l'espace affine $\mathscr{E}$,

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

et la multiplication d'un vecteur par un nombre réel : $\lambda\vec{0} = \vec{0}$, $0\vec{u} = \vec{0}$ et, si $\vec{u} \neq \overrightarrow{0}$ et $\lambda \neq 0$, le

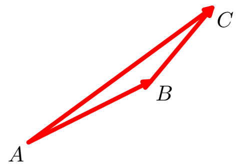


FIGURE 2. Somme de vecteurs

vecteur $\lambda\vec{u}$ est le vecteur qui est sur la droite dirigée par $\vec{u}$, de même sens que $\vec{u}$ si $\lambda > 0$, de sens opposé si $\lambda < 0$ et de longueur $||\lambda\vec{u}|| = |\lambda|\,||\vec{u}||$.

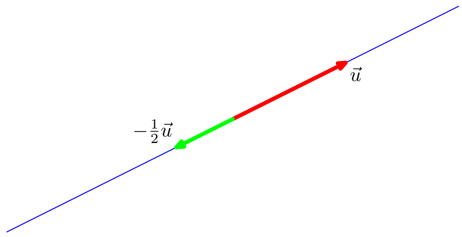


FIGURE 3. Multiplication par un nombre réel

2.2. Opérations vectorielles sur $\mathbf{R}^n$. — On va définir sur $\mathbf{R}^n$ des opérations d'addition et de multiplication par un nombre réel qui vont vérifier les règles du calcul vectoriel

Définition 2.2.1

Dans le suite n désigne un entier positif ou nul. Sur $\mathbf{R}^n$ on définit

a) Une *addition* :

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

pour tous nombres $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n \in \mathbf{R}$;

b) Une *multiplication par un nombre réel* :

$$\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$$

pour tous nombres $\lambda, x_1, \dots, x_n \in \mathbf{R}$.

***Exemples 2.2.2.* —** Prenons $n = 3$ dans cet exemple

$$(1, 2, 3) + (4, 5, 6) = (1 + 4, 2 + 5, 3 + 6) = (5, 7, 9)$$

et

$$4(1, 2, 3) = (4 \times 1, 4 \times 2, 4 \times 3) = (4, 8, 12).$$

Terminologie 2.2.3

Dans ce contexte, on dit que $\vec{u} \in \mathbf{R}^n$ est un *vecteur* et $\lambda \in \mathbf{R}$ un *scalaire*.

Formules

On note $\overrightarrow{0}$ ou $\vec{O}_{\mathbf{R}^n}$ le n-uplet dont toutes les composantes sont nulles :

$$\overrightarrow{0} = \overrightarrow{0}_{\mathbf{R}^n} = (0, \ldots, 0) \in \mathbf{R}^n.$$

Les opérations ainsi définies vérifient les règles de calcul suivantes : [a]

a) $\forall \vec{u}, \vec{v}, \vec{w} \in \mathbf{R}^n, \quad \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$;

b) $\forall \vec{u} \in \mathbf{R}^n, \quad \vec{u} + \overrightarrow{0} = \overrightarrow{0} + \vec{u} = \vec{u}$;

c) $\forall \vec{u} \in \mathbf{R}^n, \quad \vec{u} + (-1)\vec{u} = \overrightarrow{0}$;

d) $\forall \vec{u}, \vec{v} \in \mathbf{R}^n, \quad \vec{u} + \vec{v} = \vec{v} + \vec{u}$;

e) $\forall \lambda, \mu \in \mathbf{R}, \forall \vec{u} \in \mathbf{R}^n, \quad \lambda(\mu\vec{u}) = (\lambda\mu)\vec{u}$;

f) $\forall \lambda \in \mathbf{R}, \forall \vec{u}, \vec{v} \in \mathbf{R}^n, \quad \lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$;

g) $\forall \lambda, \mu \in \mathbf{R}, \forall \vec{u} \in \mathbf{R}^n, \quad (\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}$;

h) $\forall \vec{u} \in \mathbf{R}^n, \quad 1\vec{u} = \vec{u}$;

a. Rappelons que le symbole $\forall$ se lit « pour tout »

Remarque 2.2.4. — Dans les contructions et les formules précèdent, on peut parfaitement remplacer les nombres réels par les nombres complexes : l'ensemble $\mathbf{C}^n$ des n-uplets de nombres complexes peut être muni des opérations

a) Une addition :

$$(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n);$$

Une multiplication par un nombre complexe :

$$\lambda(u_1, u_2, \ldots, u_n) = (\lambda u_1, \lambda u_2, \ldots, \lambda u_n).$$

Toutes les règles de calculs restent valides à condition de remplacer à chaque fois **R** par **C**.

Exemple 2.2.5. — À titre d'exemple :

$$(1 + 2\mathrm{i}, 3\mathrm{i}) + (4, 5\mathrm{i}) = (1 + 2\mathrm{i} + 4, 3\mathrm{i} + 5\mathrm{i}) = (5 + 2\mathrm{i}, 8\mathrm{i})$$

et

$$\mathrm{i}(2\mathrm{i}, 3) = (\mathrm{i} \times 2\mathrm{i}, \mathrm{i} \times 3) = (-2, 3\mathrm{i}).$$

2.3. Matrices

Définition 2.3.1

Etant donnés deux entiers naturels m et n, une *matrice à m lignes et n colonnes à coefficients réels* est un famille de $m \times n$ réels $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ où le premier indice i, appelé l'indice de *ligne* va de 1 à m, et le second indice j, appelé indice de *colonne* va de 1 à n. Une telle matrice est notée à l'aide d'un tableau rectangulaire

$$A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,j} & \cdots & a_{m,n} \end{pmatrix}.$$

Le nombre réel $a_{i,j}$ est appelé le *coefficient d'indice i,j* de la matrice A. L'ensemble des matrices à m lignes et n colonnes et à coefficients réels est noté $\mathscr{M}_{m,n}(\mathbf{R})$.

Remarque 2.3.2. — Suivant les circonstances, il nous arrivera de noter une matrice sous forme condensée $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ ou sous la forme d'un tableau.

Exemple 2.3.3. — La matrice 2×2 dont les coefficients sont $a_{1,1} = 1$, $a_{1,2} = 2$, $a_{2,1} = 3$ et $a_{2,2} = 4$ s'écrit

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Exemple 2.3.4. — Si $n = 1$, il n'y a qu'une seule colonne et on parle de vecteur colonne. On peut identifier $\mathbf{R}^m$ avec $\mathscr{M}_{m,1}(\mathbf{R})$, c'est à dire écrire un m-uplet $(x_1, \ldots, x_m)$ comme un vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

En calcul matriciel, un vecteur de $\mathbf{R}^n$ est presque toujours écrit en *colonne*.

Donnons un exemple d'utilisation des matrices faisant le lien avec le premier chapitre :

Définition 2.3.5

Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ différentiable en un point $A \in \mathscr{D}_f$. La *matrice jacobienne* de f en A est la matrice

$$\mathrm{Jac}_A(f) = \begin{pmatrix} \frac{\partial f_1}{\partial X_1}(A) & \frac{\partial f_1}{\partial X_2}(A) & \cdots & \frac{\partial f_1}{\partial X_n}(A) \\ \frac{\partial f_2}{\partial X_1}(A) & \frac{\partial f_2}{\partial X_2}(A) & \cdots & \frac{\partial f_2}{\partial X_n}(A) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial X_1}(A) & \frac{\partial f_m}{\partial X_2}(A) & \cdots & \frac{\partial f_m}{\partial X_n}(A) \end{pmatrix}$$

Exemples 2.3.6. — a) Revenons sur les coordonnées polaires $\varphi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, la matrice jacobienne est donnée par

$$\mathrm{Jac}_{(r,\theta)}(\varphi) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

b) Si $n = 1$, c'est-à-dire dans le cas d'une courbe paramétrée f de $\mathbf{R}$ dans $\mathbf{R}^n$,

$$\mathrm{Jac}_t(f) = \begin{pmatrix} f_1'(t) \\ \vdots \\ f_m'(t) \end{pmatrix}$$

qui est le vecteur colonne correspondant à la vitesse.

c) Si $m = 1$ c'est-à-dire dans le cas d'une fonction à valeurs réelles,

$$\mathrm{Jac}_A(f) = \left(\frac{\partial f}{\partial X_1}(A) \ldots \frac{\partial f}{\partial X_n}(A) \right)$$

correspond au gradient.

On peut donc dire que la matrice jacobienne généralise à la fois vitesse et gradient.

2.4. Opérations sur les matrices

2.4.1. *Addition, multiplication par un nombre réel*

Définition 2.4.1

On peut additionner des matrices de même forme et multiplier une matrice par un nombre réel :

- *Addition :* Si $A = (a_{i,j})$ et $B = (b_{i,j})$ sont des matrices de $\mathscr{M}_{m,n}(\mathbf{R})$, leur somme $A+B$ est la matrice $(a_{i,j}+b_{i,j})$.
- *Multiplication par un scalaire :* Si $A = (a_{i,j})$ est une matrice de $\mathscr{M}_{m,n}(\mathbf{R})$, et λ est un réel, le produit λA est la matrice $(\lambda a_{i,j})$.

La matrice nulle dans $\mathscr{M}_{m,n}(\mathbf{R})$ est la matrice dont les $m \times n$ coefficients sont tous nuls. On la note $0_{m,n}$ ou simplement 0 si les entiers m et n sont clairs par le contexte.

On prendra garde au fait qu'on additionne uniquement des matrices de même forme, c'est-à-dire ayant le même nombre de lignes et de colonnes.

Exemple 2.4.2. — Par exemple :

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} + \begin{pmatrix} -3 & 1 \\ 5 & -3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 7 & 0 \\ 1 & 1 \end{pmatrix}$$

et

$$-2\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ -4 & -6 \\ -2 & 2 \end{pmatrix}.$$

Ces opérations sur les matrices vérifient les mêmes règles de calcul que les vecteurs de $\mathbf{R}^n$.

Formules

a) $\forall A, B, C \in \mathscr{M}_{m,n}(\mathbf{R}), \quad A + (B + C) = (A + B) + C$;
b) $\forall A \in \mathscr{M}_{m,n}(\mathbf{R}), \quad A + 0_{m,n} = 0_{m,n} + A = A$;
c) $\forall A \in \mathscr{M}_{m,n}(\mathbf{R}), \quad A + (-1)A = 0$;
d) $\forall A, B \in \mathscr{M}_{m,n}(\mathbf{R}), \quad A + B = B + A$;
e) $\forall \lambda, \mu \in \mathbf{R}, \forall A \in \mathscr{M}_{m,n}(\mathbf{R}), \quad \lambda(\mu A) = (\lambda\mu)A$;
f) $\forall \lambda \in \mathbf{R}, \forall A, B \in \mathscr{M}_{m,n}(\mathbf{R}), \quad \lambda(A + B) = \lambda A + \lambda B$;
g) $\forall \lambda, \mu \in \mathbf{R}, \forall A \in \mathscr{M}_{m,n}(\mathbf{R}), \quad (\lambda + \mu)A = \lambda A + \mu A$;
h) $\forall A \in \mathscr{M}_{m,n}(\mathbf{R}), \quad 1A = A$;

2.4.2. *Produit matriciel.* — Nous allons maintenant définir une opération très importante pour les matrices, le *produit matriciel.*

Définition 2.4.3

Soient m, n, p trois entiers strictement positifs. Soit $A = (a_{i,j})$ une matrice de $\mathscr{M}_{m,n}(\mathbf{R})$ et soit $B = (b_{j,k})$ une matrice de $\mathscr{M}_{n,p}(\mathbf{R})$. On appelle produit matriciel de A par B la matrice $C \in \mathscr{M}_{m,p}(\mathbf{R})$ dont le terme général $c_{i,k}$ est défini, pour tout $i = 1, \ldots, m$ et pour tout $k \in 1, \ldots, p$ par :

$$c_{i,k} = \sum_{j=1}^{n} a_{i,j} b_{j,k}.$$

Nous insistons sur le fait que le produit AB de deux matrices n'est défini que si le nombre de colonnes de A et le nombre de lignes de B sont les mêmes.

Moyen mémnotechnique 2.4.4. — Au brouillon, pour effectuer ce produit, nous conseillons de placer B au-dessus du produit et A à sa gauche.

$$\begin{array}{cc} & \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} & \cdots & b_{1,n} \\ \vdots & & \vdots & & \vdots \\ & \cdots & b_{j,k} & \cdots & \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,p} \end{pmatrix} \\ \begin{pmatrix} a_{1,1} & \cdots & & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & & \vdots & & \vdots \\ a_{m,1} & \cdots & & \cdots & a_{m,n} \end{pmatrix} & \begin{pmatrix} c_{1,1} & & \vdots & & c_{1,p} \\ & & \vdots & & \\ \cdots & \cdots & c_{i,k} & & \\ & & & & \\ c_{m,1} & & & & c_{m,p} \end{pmatrix} \end{array}$$

Posons par exemple

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix}$$

La matrice A a 3 lignes et 2 colonnes, la matrice B a 2 lignes et 4 colonnes. Le produit AB a donc un sens : c'est une matrice à 3 lignes et 4 colonnes.

$$\begin{array}{cc} & \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} & \begin{pmatrix} -3 & -1 & -1 & -1 \\ -9 & -4 & -2 & -1 \\ 3 & 3 & -1 & -3 \end{pmatrix} \end{array}$$

Ainsi pour obtenir le coefficient -2 du produit le calcul est $-2 = 2 \times (-1) + 3 \times 0$. Cela donne l'égalité

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -1 & -1 & -1 \\ -9 & -4 & -2 & -1 \\ 3 & 3 & -1 & -3 \end{pmatrix}.$$

On prend donc la première *ligne* de la première matrice et on effectue le produit scalaire avec chacune des *colonnes* de la seconde matrice pour obtenir la première ligne de la matrice produit, et ainsi de suite.

Notation 2.4.5

La *matrice identité* (également appelée *matrice unité*) est la matrice

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = (\delta_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

où $\delta_{i,j}$ est le symbole de Kronecker qui vaut 1 si $i = j$ et 0 sinon.

Le produit matriciel a presque toutes les propriétés auxquelles vous êtes habitués d'un produit, sauf qu'il n'est pas commutatif.

Proposition 2.4.6

Le produit matriciel possède les propriétés suivantes.

a) *Associativité* : Soient m, n, p, q des entiers naturels. Soient $A \in \mathscr{M}_{m,n}(\mathbf{R})$, $B \in \mathscr{M}_{n,p}(\mathbf{R})$ et $C \in \mathscr{M}_{p,q}(\mathbf{R})$. Alors
$$A(BC) = (AB)C\,;$$

b) *Linéarité à droite* : Soient m, n, p des entiers. Soit $A \in \mathscr{M}_{m,n}(\mathbf{R})$ et soient B et C des matrices de $\mathscr{M}_{n,p}(\mathbf{R})$. Si λ et μ sont des nombres réels, alors
$$A(\lambda B + \mu C) = \lambda AB + \mu AC\,;$$

c) *Linéarité à gauche* : Soient m, n, p des entiers, soient A et B des matrices de $\mathscr{M}_{m,n}(\mathbf{R})$ et soit $C \in \mathscr{M}_{n,p}(\mathbf{R})$. Si λ et μ sont des nombres réels, alors
$$(\lambda A + \mu B)C = \lambda AC + \mu BC\,;$$

d) *Élement unité* : Soient m, n des entiers et soit $A \in \mathscr{M}_{m,n}(\mathbf{R})$
$$I_m A = A I_n = A;$$

Notons en outre que $0_{p,m}A = 0_{p,n}$ et $A0_{n,p} = 0_{m,p}$ pour toute matrice $A \in \mathscr{M}_{m,n}(\mathbf{R})$.

Ces propriétés se démontrent à partir de la définition 2.4.3.

Il faut garde au fait que ce produit n'est pas commutatif : en général si $n \neq 2$ et si $A, B \in \mathscr{M}_n(\mathbf{R})$, alors
$$AB \neq BA.$$

En particulier, il convient d'être prudent vis-à-vis des identités remarquables :
$$(A+B)^2 = A^2 + AB + BA + B^2$$
mais en général, cela n'est pas égal à $A^2 + 2AB + B^2$. Cette identité n'est valable que si A et B *commutent*, c'est-à-dire si ces matrices vérifient $AB = BA$ (*cf.* la formule du binôme pour les matrices, proposition 2.4.9).

Exemple 2.4.7. — Soient $a, b, c, d \in \mathbf{R}$. Posons $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ alors
$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$
et
$$BA = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix}.$$

Exemple 2.4.8. — Revenons à titre d'exemple sur la formule 1.11 donnant la dérivée partielle d'une fonction composée :

$$\frac{\partial(g\circ f)_k}{\partial X_i}(A)=\sum_{j=1}^{m}\frac{\partial g_k}{\partial Y_j}(f(A))\frac{\partial f_j}{\partial X_i}(A).$$

En utilisant la matrice jacobienne et le produit matriciel cette formule s'écrit simplement

$$\mathrm{Jac}_A(g\circ f)=\mathrm{Jac}_{f(A)}(g)\,\mathrm{Jac}_A(f).$$

Formule que nous justifierons plus loin dans l'exemple 2.5.11.

Proposition 2.4.9

Soient A et B des matrices carrées de $\mathcal{M}_n(\mathbf{R})$. On suppose qu'elles *commutent*, c'est-à-dire que $AB=BA$, alors

$$(A+B)^n=\sum_{k=0}^{n}\binom{n}{k}A^kB^{n-k}$$

où $\binom{n}{k}$ est le coefficient binomial $\frac{n!}{k!(n-k)!}$.

2.5. Applications linéaires de $\mathbf{R}^n$ dans $\mathbf{R}^m$

2.5.1. *Définition.* — Les applications linéaires sont celles qui sont compatibles avec les opérations vectorielles.

Définition 2.5.1

Une application $\varphi:\mathbf{R}^n\to\mathbf{R}^m$ est dite *linéaire* si elle vérifie les deux conditions suivantes :

(i) *Compatibilité avec l'addition :* Pour tous vecteurs $\vec{u},\vec{v}\in\mathbf{R}^n$,

$$\varphi(\vec{u}+\vec{v})=\varphi(\vec{u})+\varphi(\vec{v});$$

(ii) *Compatibilité avec la multiplication par un nombre réel :* Pour tout $\lambda\in\mathbf{R}$, pour tout vecteur $\vec{u}\in\mathbf{R}^n$,

$$\varphi(\lambda\vec{u})=\lambda\varphi(\vec{u}).$$

On notera $\mathscr{L}(\mathbf{R}^n,\mathbf{R}^m)$ l'ensemble des applications linéaires de $\mathbf{R}^n$ dans $\mathbf{R}^m$.

Exemples 2.5.2. — a) Soit φ l'application de $\mathbf{R}^2$ dans $\mathbf{R}^2$ donnée par $(x,y) \mapsto (x+y, x-y)$. Prouvons que φ est linéaire en vérifiant les deux conditions de la définition :

Condition (i) : Soient $\vec{u} = (x_u, y_u)$ et $\vec{v} = (x_v, y_v)$ des vecteurs quelconques de $\mathbf{R}^2$. Calculons $\varphi(\vec{u}+\vec{v})$.

$$\vec{u}+\vec{v} = (x_u, y_u) + (x_v, y_v) = (x_u + x_v, y_u + y_v).$$

Rappelons que $f(x_1, \dots, x_n)$ est une abbréviation pour $f((x_1, \dots, x_n))$, c'est-à-dire l'image par f du n-uplet $(x_1, \dots, x_n)$. Donc

$$\begin{aligned}
\varphi(\vec{u}+\vec{v}) &= \varphi(x_u + x_v, y_u + y_v) \\
&= \big((x_u + x_v) + (y_u + y_v), (x_u + x_v) - (y_u + y_v)\big) \\
&= (x_u + y_u + x_v + y_v, x_u - y_u + x_v - y_v)
\end{aligned}$$

Par ailleurs,

$$\begin{aligned}
\varphi(\vec{u}) + \varphi(\vec{v}) &= \varphi(x_u, y_u) + \varphi(x_v, y_v) \\
&= (x_u + y_u, x_u - y_u) + (x_v + y_v, x_v - y_v) \\
&= (x_u + y_u + x_v + y_v, x_u - y_u + x_v - y_v)
\end{aligned}$$

Par conséquent,

$$\varphi(\vec{u}+\vec{v}) = \varphi(\vec{u}) + (\vec{v})$$

ce qui prouve que φ est compatible avec l'addition des vecteurs.

Condition (ii) : Soit $u = (x_u, y_u) \in \mathbf{R}^2$ et $\lambda \in \mathbf{R}$. Alors

$$\lambda\vec{u} = \lambda(x_u, y_u) = (\lambda x_u, \lambda y_u).$$

Donc

$$\begin{aligned}
\varphi(\lambda\vec{u}) &= \varphi(\lambda x_u, \lambda y_u) \\
&= ((\lambda x_u) + (\lambda y_u), (\lambda x_u) - (\lambda y_u)) \\
&= (\lambda(x_u + y_u), \lambda(x_u - y_u)) \\
&= \lambda(x_u + y_u, x_u - y_u) \\
&= \lambda\varphi(x_u, y_u) \\
&= \lambda\varphi(\vec{u})
\end{aligned}$$

Donc φ est également compatible avec la multiplication par un nombre réel.

Conclusion : L'application φ est linéaire.

b) Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ différentiable en un point A. La différentielle $\mathrm{d}f_A : \mathbf{R}^n \to \mathbf{R}^m$ est linéaire. En effet, $\mathrm{d}f_A$ est donnée par

$$(u_1,\dots,u_n) \longmapsto \left(\sum_{i=1}^{n} \frac{\partial f_1}{\partial X_i}(A)u_i,\dots,\sum_{i=1}^{n} \frac{\partial f_m}{\partial X_i}(A)u_i\right)$$

La vérification des deux conditions de la définition se fait comme dans l'exemple précédent et est laissée en exercice.

c) Terminons avec un exemple d'application qui *n'*est *pas* linéaire. L'application f de $\mathbf{R}^2$ dans $\mathbf{R}^2$ donnée par $(x,y) \mapsto x^2 + y^2$ vérifie

$$f(2,0) = 4 \text{ et } 2f(1,0) = 2 \neq 4$$

Donc pour $\vec{u} = (1,0)$ et $\lambda = 2$, on a $f(\lambda\vec{u}) \neq \lambda f(\vec{u})$ donc l'application f n'est pas compatible avec la multiplication par un nombre réel et elle n'est donc pas linéaire.

Notation 2.5.3

Pour tout ensemble X, on appelle *application identité de* X et on note Id_X l'application de X dans X donnée par $\mathrm{Id}_X : x \mapsto x$.

Donnons quelques propriétés des applications linéaires :

Proposition 2.5.4

Soit $\varphi : \mathbf{R}^m \to \mathbf{R}^m$ une application linéaire. Alors,

a) L'image du vecteur nul par φ est le vecteur nul :

$$\varphi\left(\overrightarrow{0}_{\mathbf{R}^n}\right) = \overrightarrow{0}_{\mathbf{R}^m};$$

b) Si $\vec{u},\vec{v} \in \mathbf{R}^n$ et $\lambda,\mu \in \mathbf{R}$, alors

$$\varphi(\lambda\vec{u} + \mu\vec{v}) = \lambda\varphi(\vec{u}) + \mu\varphi(\vec{v});$$

c) Plus généralement, si $\vec{u}_1,\dots,\vec{u}_r$ sont r vecteurs de $\mathbf{R}^n$ et $\lambda_1,\dots,\lambda_r \in \mathbf{R}$, on note

$$\sum_{i=1}^{r} \lambda_i\vec{u}_i = \lambda_1\vec{u}_1 + \cdots + \lambda_r\vec{u}_r$$

et on obtient la relation :
$$\varphi\left(\sum_{i=1}^{r}\lambda_i\vec{u}_i\right)=\sum_{i=1}^{r}\lambda\varphi(\vec{u}_i).$$

Idée de la preuve. — a) On a les égalités
$$\varphi\left(\overrightarrow{0}_{\mathbf{R}^n}\right)=\varphi\left(0\,\overrightarrow{0}_{\mathbf{R}^n}\right)=0\varphi\left(\overrightarrow{0}_{\mathbf{R}^n}\right)=\overrightarrow{0}_{\mathbf{R}^m}.$$

b) La définition d'une application linéaire donne
$$\varphi(\lambda\vec{u}+\mu\vec{v})=\varphi(\lambda\vec{u})+\varphi(\mu\vec{v})=\lambda\varphi(\vec{u})+\mu\varphi(\vec{v}).$$

L'assertion c) s'obtient par un petit raisonnement par récurrence. □

2.5.2. *Description matricielle des applications linéaires de* $\mathbf{R}^n$ *dans* $\mathbf{R}^m$. — Nous allons maintenant décrire de deux manières différentes les applications linéaires, ce qui conduira à des méthodes de calcul.

Remarque 2.5.5. — Si $j\in\{1,\dots,n\}$, on note $\vec{e}_j\in\mathbf{R}^n$ le vecteur dont les composantes sont nulles, sauf la j-ème qui vaut 1. Notons que si $(x_1,\dots,x_n)\in\mathbf{R}^n$, on a les égalités

$$\begin{aligned}\sum_{i=1}^{n}x_i\vec{e}_i&=(x_1,0,\dots,0)+(0,x_2,\dots,0)+\cdots+(0,0,\dots,x_n)\\&=(x_1,x_2,\dots,x_n).\end{aligned}\tag{8}$$

Soit $\varphi:\mathbf{R}^n\to\mathbf{R}^m$ une application linéaire. Notons $\vec{u}_j=\varphi(\vec{e}_j)\in\mathbf{R}^m$ pour tout $j\in\{1,\dots,n\}$. Alors pour tout $(x_1,\dots,x_n)\in\mathbf{R}^n$ on a les égalités
$$\begin{aligned}\varphi(x_1,\dots,x_n)&=\varphi(x_1\vec{e}_1+\cdots+x_n\vec{e}_n)\\&=x_1\varphi(\vec{e}_1)+\cdots+x_n\varphi(\vec{e}_n)\\&=x_1\vec{u}_1+\cdots+x_n\vec{u}_n.\end{aligned}$$

Donc, si on connaît les vecteurs $\vec{u}_i=\varphi(\vec{e}_i)$, on peut retrouver l'image de n'importe quel vecteur $(x_1,\dots,x_n)$ de $\mathbf{R}^n$ grâce à la formule
$$\varphi(x_1,\dots,x_n)=x_1\vec{u}_1+\cdots+x_n\vec{u}_n.$$

On peut dire que l'application φ est uniquement déterminée par les vecteurs $\vec{u}_i=\varphi(\vec{e}_i)$.

Inversement, donnons-nous n vecteurs $\vec{u}_1,\dots,\vec{u}_n$ de $\mathbf{R}^m$. L'application $\varphi:\mathbf{R}^n\to\mathbf{R}^m$ définie par
$$(x_1,\dots,x_n)\longmapsto x_1\vec{u}_1+\cdots+x_n\vec{u}_n$$
est linéaire et vérifie $\varphi(\vec{e}_j)=\vec{u}_j$ pour tout $j\in\{1,\dots,n\}$.

Conclusion : Étant donnés n vecteurs $\vec{u}_1, \dots, \vec{u}_n$ de $\mathbf{R}^m$, il existe une unique application linéaire φ de $\mathbf{R}^n$ dans $\mathbf{R}^m$ telle que $\varphi(\vec{e}_j) = \vec{u}_j$ pour $j \in \{1, \dots, n\}$, elle est donnée par

$$(x_1, \dots, x_n) \longmapsto x_1\vec{u}_1 + \cdots + x_n\vec{u}_n.$$

Se donner une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ revient donc à se donner n vecteurs dans $\mathbf{R}^m$, comme chacun de ces vecteurs est donné par m nombre réels, on obtient donc qu'une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est donnée par $m \times n$ nombres réels, ce qui amène naturellement à la description matricielle qui est l'objet des notations qui suivent.

Notation 2.5.6

Soit φ une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Posons à nouveau $\vec{e}_j = (\delta_{1,j}, \dots, \delta_{n,j})$ pour $j \in \{1, \dots, n\}$, où $\delta_{i,j}$ désigne le symbole de Kronecker et écrivons $\varphi(\vec{e}_j) = (a_{1,j}, \dots, a_{m,j})$. La *matrice de φ* (ou plus précisément *la matrice de φ dans les bases usuelles*) est la matrice

$$\operatorname{Mat}(\varphi) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathscr{M}_{m,n}(\mathbf{R})$$

c'est-à-dire la matrice dont la j-ème *colonne* correspond au m-uplet $\varphi(\vec{e}_j)$.

Remarque 2.5.7. — Inversement étant donné une matrice

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathscr{M}_{m,n}(\mathbf{R}).$$

On peut considérer l'application $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ définie par

$$(x_1, \dots, x_n) \longmapsto \left(\sum_{k=1}^n a_{1,k}x_k, \sum_{k=1}^n a_{2,k}x_k, \dots, \sum_{k=1}^n a_{m,k}x_k \right).$$

Cette application est linéaire. Soit $j \in \{1, \dots, n\}$. Calculons $\varphi(e_j)$ en utilisant le symbole de Kronecker et les symboles sommatoires

$$\varphi(\vec{e}_j) = \varphi(\delta_{1,j}, \dots, \delta_{n,j}) = \left(\sum_{k=1}^{n} a_{1,k}\delta_{k,j}, \sum_{k=1}^{n} a_{2,k}\delta_{k,j}, \dots, \sum_{k=1}^{n} a_{m,k}\delta_{k,j} \right)$$

Mais rappelons que $\delta_{k,j}$ vaut 1 si $k = j$ et 0 sinon. Par conséquent le seul terme non nul dans la somme $\sum_{k=1}^{n} a_{i,k}\delta_{k,j}$ est celui pour $k = j$ et il vaut $a_{i,j}$ ce qui donne l'égalité

$$\varphi(\vec{e}_j) = (a_{1,j}, \dots, a_{n,j})$$

Par conséquent $\mathrm{Mat}(\varphi) = A$.

Conclusion. L'application

$$\begin{array}{ccc} \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m) & \longrightarrow & \mathscr{M}_{m,n}(\mathbf{R}) \\ \varphi & \longmapsto & \mathrm{Mat}(\varphi) \end{array}$$

est *bijective* ; Autrement dit, pour tout $A \in \mathscr{M}_{m,n}(\mathbf{R})$, il existe une unique application linéaire $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ telle que $\mathrm{Mat}(\varphi) = A$.

Terminologie 2.5.8

Si $A \in \mathscr{M}_{m,n}(\mathbf{R})$, l'unique $\varphi \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ telle que $\mathrm{Mat}(\varphi) = A$ est appelée *l'application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ définie par A.*

Formules

Soit $A \in \mathscr{M}_{m,n}(\mathbf{R})$ et soit $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ l'application linéaire correspondante. Soit $\vec{u} \in \mathbf{R}^n$ et soit $U \in \mathscr{M}_{n,1}(\mathbf{R})$ le vecteur colonne correspondant. Alors $AU \in \mathscr{M}_{m,1}(\mathbf{R})$ est le vecteur colonne correspondant $\varphi(\vec{u})$.

Autrement dit si on identifie un n-uplet avec le vecteur colonne correspondant, une application linéaire φ de $\mathscr{M}_{n,1}(\mathbf{R})$ dans $\mathscr{M}_{m,1}(\mathbf{R})$ est donnée par le produit matriciel :

$$U \mapsto \mathrm{Mat}(\varphi)U.$$

Exemple 2.5.9. — Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ différentiable en un point $A \in \mathscr{D}_f$, la matrice de l'application linéaire $\mathrm{d}f_A : \mathbf{R}^n \to \mathbf{R}^m$ est la matrice jacobienne $\mathrm{Jac}_A(f)$:

$$\mathrm{Mat}(\mathrm{d}f_A) = \begin{pmatrix} \frac{\partial f_1}{\partial X_1}(A) & \frac{\partial f_1}{\partial X_2}(A) & \cdots & \frac{\partial f_1}{\partial X_n}(A) \\ \frac{\partial f_2}{\partial X_1}(A) & \frac{\partial f_2}{\partial X_2}(A) & \cdots & \frac{\partial f_2}{\partial X_n}(A) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial X_1}(A) & \frac{\partial f_m}{\partial X_2}(A) & \cdots & \frac{\partial f_m}{\partial X_n}(A) \end{pmatrix} = \mathrm{Jac}_A(f).$$

2.5.3. *Matrice de la composée.* — L'utilité du produit matriciel vient de la proposition suivante

Proposition 2.5.10

Soient $m, n, p \in \mathbf{N}$. Soir $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ et $\psi : \mathbf{R}^m \to \mathbf{R}^p$ des applications linéaires. Alors la matrice de la composée est le produit des matrices :

$$\mathrm{Mat}(\psi \circ \varphi) = \mathrm{Mat}(\psi)\,\mathrm{Mat}(\varphi).$$

Démonstration. — Notons tout d'abord que $\mathrm{Mat}(\psi) \in \mathscr{M}_{p,m}(\mathbf{R})$ et que $\mathrm{Mat}(\varphi) \in \mathscr{M}_{m,n}(\mathbf{R})$ si bien que le produit matriciel $\mathrm{Mat}(\psi)\,\mathrm{Mat}(\varphi)$ est bien défini. Posons à nouveau $\vec{e}_j = (\delta_{1,j}, \ldots, \delta_{n,j}) \in \mathbf{R}^n$ et calculons $\psi \circ \varphi(\vec{e}_j)$. Pour cela, on pose $\vec{f}_k = (\delta_{1,k}, \ldots, \delta_{m,k}) \in \mathbf{R}^m$ et on écrit les coefficients des deux matrices :

$$\mathrm{Mat}(\psi) = \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & & \vdots \\ b_{p,1} & \cdots & b_{p,m} \end{pmatrix} \text{ et } \mathrm{Mat}(\varphi) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}.$$

Cela signifie que

$$\varphi(\vec{e}_j) = (a_{1,j}, \ldots, a_{m,j}) = a_{1,j}\vec{f}_1 + \cdots + a_{m,j}\vec{f}_m$$

pour $j \in \{1, \ldots, n\}$ et

$$\psi(\vec{f}_k) = (b_{1,k}, \ldots, b_{p,k})$$

pour $k \in \{1, \dots, m\}$. Donc

$$\begin{aligned}
\psi \circ \varphi(\vec{e}_j) &= \psi(\varphi(\vec{e}_j)) \\
&= \psi(a_{1,j}\vec{f}_1 + \cdots + a_{m,j}\vec{f}_m) \\
&= a_{1,j}\psi(\vec{f}_1) + \cdots + a_{m,j}\psi(\vec{f}_m) \\
&= a_{1,j}(b_{1,1}, \dots, b_{p,1}) + \cdots + a_{m,j}(b_{1,m}, \dots, b_{p,m}) \\
&= (a_{1,j}b_{1,1} + \cdots + a_{m,j}b_{1,m}, \dots, a_{1,j}b_{p,1} + \cdots + a_{m,j}b_{p,m}) \\
&= \left(\sum_{k=1}^{m} b_{1,k}a_{k,j}, \dots, \sum_{k=1}^{m} b_{p,k}a_{k,j}\right)
\end{aligned}$$

Donc en mettant ce p-uplet en colonnes, on obtient

$$\mathrm{Mat}(\psi \circ \varphi) = \begin{pmatrix} \sum_{k=1}^{m} b_{1,k}a_{k,1} & \cdots & \sum_{k=1}^{m} b_{1,k}a_{k,n} \\ \vdots & & \vdots \\ \sum_{k=1}^{m} b_{p,k}a_{k,1} & \cdots & \sum_{k=1}^{m} b_{p,k}a_{k,n} \end{pmatrix} = \mathrm{Mat}(\psi)\,\mathrm{Mat}(\varphi). \qquad \square$$

Exemple 2.5.11. — Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit g une fonction de $\mathbf{R}^m$ dans $\mathbf{R}^p$. On suppose que f est différentiable en A et que g est différentiable en $f(A)$. Rappelons la formule de la proposition 1.11.14

$$\mathrm{d}(g \circ f)_A = \mathrm{d}g_{f(A)} \circ \mathrm{d}f_A.$$

En prenant les matrices cela donne directement la relation annoncée dans l'exemple 2.4.8

$$\mathrm{Jac}_A(g \circ f) = \mathrm{Jac}_{f(A)}(g)\,\mathrm{Jac}_A(f).$$

Comme expliqué dans cet exemple, cette formule redonne les formules pour les dérivées partielles d'une fonction composée

$$\frac{\partial (g \circ f)_k}{\partial X_i}(A) = \sum_{j=1}^{m} \frac{\partial g_k}{\partial Y_j}(f(A)) \frac{\partial f_j}{\partial X_i}(A).$$

Cette formule qui paraît compliquée, n'est donc qu'un avatar de la formule naturelle $\mathrm{d}(g \circ f)_A = \mathrm{d}g_{f(A)} \circ \mathrm{d}f_A$.

2.6. La transposition. — Une dernière opération sur les matrices est la transposition, qui intervient notamment dans l'écriture des produits scalaires en termes de coordonnées.

Définition 2.6.1

Etant donnée une matrice $A = (a_{i,j})$ de $\mathscr{M}_{m,n}(\mathbf{R})$, sa transposée est la matrice ${}^tA$ de $\mathscr{M}_{n,m}(\mathbf{R})$ dont le coefficient d'indice (i,j) est $a_{j,i}$. Sous forme de tableau :

$$ {}^t\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix} $$

Exemples 2.6.2. — a) Pour écrire la transposée d'une matrice, il suffit de transformer ses lignes en colonnes. Par exemple :

$$ A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} \quad , \quad {}^tA = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & -1 \end{pmatrix}. $$

b) La transposée d'une matrice avec une seule ligne est un vecteur colonne et vice-versa :

$$ {}^t\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} $$

c) Rappelons la définition du produit scalaire usuel sur $\mathbf{R}^n$ (définition 1.10.6) : Soient $\vec{u} = (u_1, \dots, u_n)$ et $\vec{v} = (v_1, \dots, v_n)$ des éléments de $\mathbf{R}^n$. Soit U la matrice colonne correspondant à $\vec{u}$ et V celle correspondant à $\vec{v}$. Alors

$$ \vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = {}^tUV. $$

Énonçons quelques propriétés de la transposition :

Formules

a) La transposée de la transposée est la matrice initiale.

$$^t(^tA) = A$$

pour toute matrice $A \in \mathscr{M}_{m,n}(\mathbf{R})$;

b) La transposée de la somme est la somme des transposée :

$$^t(A+B) = {}^tA + {}^tB$$

pour toutes matrices $A, B \in \mathscr{M}_{m,n}(\mathbf{R})$;

c) Soient m, n, p trois entiers strictement positifs. Soient $A = (a_{i,j})$ une matrice de $\mathscr{M}_{m,n}(\mathbf{R})$ et $B = (b_{j,k})$ une matrice de $\mathscr{M}_{n,p}(\mathbf{R})$. La transposée du produit de A par B est le produit de la transposée de B par la transposée de A.

$$^t(AB) = {}^tB\,{}^tA\,.$$

Attention : la transposée échange les deux facteurs d'un produit.

Exemple 2.6.3. — Par exemple, en reprenant les matrices A et B définies ci-dessus :

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & -1 \end{pmatrix} \\ \begin{pmatrix} 0 & -3 \\ 1 & -2 \\ -1 & 0 \\ -2 & 1 \end{pmatrix} & \begin{pmatrix} -3 & -9 & 3 \\ -1 & -4 & 3 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{pmatrix} \end{array}$$

Définition 2.6.4

Une matrice carrée $A \in \mathscr{M}_n(\mathbf{R})$ est dite *symétrique* si $^tA = A$.

Remarque 2.6.5. — Si

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

cela signifie que $a_{i,j} = a_{j,i}$, la matrice est symétrique par rapport à la diagonale.

Exemple 2.6.6. — Soit f une fonctions en n variables de classe $\mathscr{C}^2$ alors on peut considérer la *matrice hessienne* de f en un point A du domaine de définition de f, définie par

$$\mathrm{Hess}_A(f) = \left(\frac{\partial^2 f}{\partial X_i \partial X_j}(A)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}.$$

Par le théorème de Schwarz (théorème 1.9.5), cette matrice est symétrique.

Voyons maintenant une petite application de l'expression du produit scalaire à l'aide de la transposée.

Définition 2.6.7

Soit f un endomorphisme de $\mathbf{R}^n$. On dit que f est une *isométrie* si elle vérifie une des deux conditions équivalentes suivantes :

(i) Elle préserve la norme usuelle :

$$\forall \vec{u} \in \mathbf{R}^n, \quad ||f(\vec{u})|| = ||\vec{u}||;$$

(ii) Elle préserve le produit scalaire usuel :

$$\forall \vec{u}, \vec{v} \in \mathbf{R}^n, \quad f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v};$$

Remarque 2.6.8. — L'équivalence entre les deux assertions se prouve en utilisant la formule suivante :

$$||\vec{u}+\vec{v}||^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) = \vec{u}\cdot\vec{u} + \vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{u} + \vec{v}\cdot\vec{v} = ||\vec{u}||^2 + 2\vec{u}\cdot\vec{v} + ||\vec{v}||^2$$

qui donne la formule

$$\vec{u}\cdot\vec{v} = \frac{||\vec{u}+\vec{v}||^2 - ||\vec{u}||^2 - ||\vec{v}||^2}{2}.$$

Proposition 2.6.9

Un application linéaire $f : \mathbf{R}^n \to \mathbf{R}^n$ est une isométrie si et seulement si sa matrice $M = \mathrm{Mat}(f)$ vérifie l'égalité

$${}^tMM = I_n.$$

Terminologie 2.6.10

Une matrice M telle que ${}^tMM = I_n$ est dite *orthogonale*.

Idée de la preuve de la proposition. — Soient $\vec{u}, \vec{v} \in \mathbf{R}$, soit U le vecteur colonne correspondant à $\vec{u}$ et V le vecteur colonne correspondant à $\vec{v}$. Alors le vecteur colonne pour $f(\vec{u})$ est MU et celui pour $f(\vec{v})$ est MV. Donc, par l'exemple 2.6.2 c),

$$f(\vec{u}) \cdot f(\vec{v}) = {}^t(MU)MV = {}^tU({}^tMM)V$$

Donc f est une isométrie si et seulement si ${}^tU({}^tMM)V = {}^tUV$ pour tous $U, V \in \mathscr{M}_{1,n}(\mathbf{R})$. Cela est vérifié lorsque ${}^tMM = I_n$. Inversement si ${}^tU({}^tMM)V = {}^tUV$ pour tous $U, V \in \mathscr{M}_{1,n}(\mathbf{R})$ alors, cela est en particulier vrai lorsqu'on prend $U = E_{i,1}$ et $V = E_{j,1}$ où $E_{i,1}$ est le vecteur colonne où toutes les lignes sont nulles sauf la i-ème qui vaut 1. Mais étant donné une matrice $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$, On a que ${}^tE_{i,1}AE_j = a_{i,j}$. Par conséquent, la condition « ${}^tU({}^tMM)V = {}^tUV$ pour tous $U, V \in \mathscr{M}_{1,n}(\mathbf{R})$ » implique que ${}^tMM = I_n$. □

Remarque 2.6.11. — On garde les notations de la proposition 2.6.9 Notons $\vec{u}_i$ le vecteur de $\mathbf{R}^n$ correspondant à la i-ème colonne de M. Autrement dit $\vec{u}_i = f(\vec{e}_i)$ où $\vec{e}_i$ est défini comme dans la notation 2.5.6. Alors la matrice M est orthogonale si et seulement si

$$\vec{u}_i \cdot \vec{u}_j = \delta_{i,j}$$

pour tous $i, j \in \{1, \dots, n\}$, où δ est le symbole de Kronecker.

Exemple 2.6.12. — Soit $\theta \in \mathbf{R}$. Alors la matrice

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

est une matrice orthogonale, puisque $\cos(\theta)^2 + \sin(\theta)^2 = 1$. Elle correspond à une *rotation de centre* $(0,0)$ *et d'angle* θ (figure 4).

2.7. Sous-espaces vectoriels

Définition 2.7.1

Un *sous-espace vectoriel* de $\mathbf{R}^n$ est une partie E de $\mathbf{R}^n$ qui satisfait les trois conditions suivantes :

(i) La partie E contient le vecteur nul : $\overrightarrow{0} \in E$;

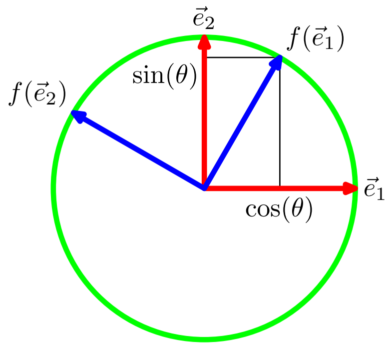


FIGURE 4. Rotation d'angle θ

(ii) La partie E est *stable par addition* :

$$\forall \vec{u}, \vec{v} \in E, \quad \vec{u} + \vec{v} \in E;$$

(iii) La partie E est *stable par multiplication par un nombre réel* :

$$\forall \lambda \in \mathbf{R}, \forall \vec{u} \in E, \quad \lambda \vec{u} \in E.$$

Exemples 2.7.2. — a) La partie $\{\overrightarrow{0}\}$ dont l'unique élément est le vecteur nul est un sous-espace vectoriel de $\mathbf{R}^n$. En effet

$$\overrightarrow{0} + \overrightarrow{0} = \overrightarrow{0}$$

et si $\lambda \in \mathbf{R}$,

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

b) L'ensemble $\mathbf{R}^n$ est un sous-espace vectoriel de $\mathbf{R}^n$.

c) Si F_1 et F_2 sont des sous-espaces vectoriels de $\mathbf{R}^n$, l'intersection $F_1 \cap F_2$ aussi. Vérifions cela à partir de la définition. Rappelons d'abord que si $\vec{u} \in \mathbf{R}^n$, alors on a une équivalence

$$\vec{u} \in F_1 \cap F_2 \Longleftrightarrow (\vec{u} \in F_1 \text{ et } \vec{u} \in F_2)$$

Condition (i). Comme $\overrightarrow{0} \in F_1$ et $\overrightarrow{0} \in F_2$, $\overrightarrow{0} \in F_1 \cap F_2$.

Condition (ii). Soient $\vec{u}_1$ et $\vec{u}_2$ des vecteurs arbitraires de $F_1 \cap F_2$. Comme $\vec{u}_1 \in F_1$, $\vec{u}_2 \in F_1$ et F_1 est stable par addition, on obtient que $\vec{u}_1 + \vec{u}_2 \in F_1$. De la même manière on montre que $\vec{u}_1 + \vec{u}_2 \in F_2$. Par conséquent $\vec{u}_1 + \vec{u}_2 \in F_1 \cap F_2$ ce qui prouve que l'intersection $F_1 \cap F_2$ est stable par addition.

Condition (iii). Soit $\lambda \in \mathbf{R}$ et soit $\vec{u}$ un vecteur arbitraire de $F_1 \cap F_2$. Comme $\vec{u} \in F_1$ et que F_1 est stable par multiplication par un nombre réel, on obtient que $\lambda\vec{u} \in F_1$. On démontre de même que $\lambda\vec{u} \in F_2$ Donc $\lambda\vec{u} \in F_1 \cap F_2$ ce qui prouve que $F_1 \cap F_2$ est stable par multiplication par un nombre réel.

Conclusion. L'intersection $F_1 \cap F_2$ est un sous-espace vectoriel de $\mathbf{R}^n$.

d) Plus généralement, si $F_1, \ldots, F_r$ sont des sous-espaces vectoriels alors l'intersection

$$F_1 \cap \cdots \cap F_r = \{\vec{u} \in \mathbf{R}^n \mid \forall i \in \{1, \ldots, r\}, \vec{u} \in F_i\}$$

est aussi un sous espace vectoriel de $\mathbf{R}^n$. La preuve peut-être faite par récurrence sur r en notant que, si $r \geqslant 2$

$$F_1 \cap \cdots \cap F_r = (F_1 \cap \cdots \cap F_{r-1}) \cap F_r.$$

Cela est faux, en général, pour la réunion. Prenons deux sous-espaces vectoriels F_1 et F_2 tels que F_1 n'est pas contenu dans F_2 et F_2 n'est pas contenu dans F_1. Choisissons alors deux vecteurs $\vec{u}_1$ et $\vec{u}_2$ tels que $\vec{u}_1 \in F_1$, $\vec{u}_1 \notin F_2$, $\vec{u}_2 \in F_2$ et $\vec{u}_2 \notin F_1$. Alors $\vec{u}_1$ et $\vec{u}_2$ appartiennent tous les deux à la réunion $F_1 \cup F_2$, mais $\vec{u}_1 + \vec{u}_2$ n'appartient pas à cette réunion (figure 5). La

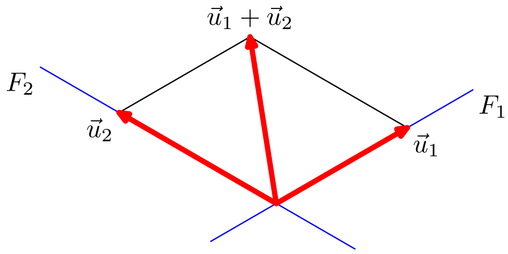


FIGURE 5. Réunion de sous-espaces vectoriels

réunion de F_1 et de F_2 n'est donc pas stable par addition.

Proposition 2.7.3

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application *linéaire*. Alors

a) L'ensemble

$$f^{-1}(\{0\}) = \{\vec{u} \in \mathbf{R}^n \mid f(\vec{u}) = \overrightarrow{0}\}$$

est un sous-espace vectoriel de $\mathbf{R}^n$ qu'on appelle le *noyau de f* et qu'on note $\mathrm{Ker}(f)$.

b) L'ensemble des valeurs de f c'est-à-dire des vecteurs de $\mathbf{R}^m$ de la forme $f(\vec{u})$ pour un $\vec{u} \in \mathbf{R}^n$ (avec des notations ensemblistes

$$f(\mathbf{R}^m) = \{\vec{v} \in \mathbf{R}^m \mid \exists \vec{u} \in \mathbf{R}^n, \vec{v} = f(\vec{u})\}$$

est un sous-espace vectoriel de $\mathbf{R}^m$ qu'on appelle l'*image de* f et qu'on note $\mathrm{Im}(f)$.

Démonstration. — Démontrons d'abord l'assertion a) en vérifiant les trois conditions de la définition d'un sous-espace vectoriel.

Condition (i). Comme f est linéaire, $f(\overrightarrow{0}_{\mathbf{R}^n}) = \overrightarrow{0}_{\mathbf{R}^m}$ donc $\overrightarrow{0}_{\mathbf{R}^n} \in \mathrm{Ker}(f)$.

Condition (ii). Soient $\vec{u}$ et $\vec{v}$ des vecteurs de $\mathrm{Ker}(f)$. alors par définition du noyau, $f(\vec{u}) = \overrightarrow{0}$ et $f(\vec{v}) = \overrightarrow{0}$. Donc

$$f(\vec{u}+\vec{v}) = f(\vec{u}) + f(\vec{v}) = \overrightarrow{0} + \overrightarrow{0} = \overrightarrow{0}$$

où la seconde égalité résulte du fait que f est compatible avec l'addition. Cela prouve que $\vec{u}+\vec{v} \in \mathrm{Ker}(f)$ et donc $\mathrm{Ker}(f)$ est stable par addition.

Condition (iii). Soient $\lambda \in \mathbf{R}$ et soit $\vec{u} \in \mathrm{Ker}(f)$. Alors $f(\vec{u}) = \overrightarrow{0}$. Mais la compatibilité de f avec la multiplication par un nombre réel fournit la première égalité dans

$$f(\lambda\vec{u}) = \lambda f(\vec{u}) = \lambda\overrightarrow{0} = \overrightarrow{0}.$$

Cela prouve que $\lambda\vec{u} \in \mathrm{Ker}(f)$ et que $\mathrm{Ker}(f)$ est stable par multiplication par un nombre réel.

Conclusion. L'ensemble $\mathrm{Ker}(f)$ est un sous-espace vectoriel de $\mathbf{R}^n$. □

Pour l'assertion b) on va prouver un résultat plus fort :

Proposition 2.7.4

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire. Soit F un sous-espace vectoriel de $\mathbf{R}^n$. Alors

$$f(F) = \{\vec{v} \in \mathbf{R}^m \mid \exists \vec{u} \in F, \vec{v} = f(\vec{u})\}$$

est un sous-espace vectoriel de $\mathbf{R}^m$.

Démonstration. — À nouveau nous allons vérifier les trois conditions de la définition.

Condition (i). Comme F est un sous-espace vectoriel de $\mathbf{R}^n$, il contient le vecteur nul $\overrightarrow{0}_{\mathbf{R}^n} \in F$. Comme l'application f est linéaire, on a la relation

$$\overrightarrow{0}_{\mathbf{R}^m} = f\left(\overrightarrow{0}_{\mathbf{R}^n}\right) \in f(F).$$

Condition (ii). Soient $\vec{v}_1$ et $\vec{v}_2$ des vecteurs de $f(F)$. Par définition de $f(F)$, il existe des vecteurs $\vec{u}_1, \vec{u}_2 \in F$ tels que $\vec{v}_1 = f(\vec{u}_1)$ et $\vec{v}_2 = f(\vec{u}_2)$. alors, comme l'application f est linéaire,

$$\vec{v}_1 + \vec{v}_2 = f(\vec{u}_1) + f(\vec{u}_2) = f(\vec{u}_1 + \vec{u}_2)$$

Mais comme F est stable par addition, la somme $\vec{u}_1 + \vec{u}_2$ appartient à F, si bien que $\vec{v}_1 + \vec{v}_2 = f(\vec{u}_1 + \vec{u}_2) \in F$. Cela prouve que $f(F)$ est stable par addition.

Condition (iii). Soient $\lambda \in \mathbf{R}$ et $\vec{v} \in f(F)$. Par définition, il existe $\vec{u} \in F$ tel que $\vec{v} = f(\vec{u})$. Comme l'application f est linéaire,

$$\lambda\vec{v} = \lambda f(\vec{u}) = f(\lambda\vec{u}).$$

Mais F est stable par multiplication par un nombre réel, donc $\lambda\vec{u} \in F$. Par conséquent, $\lambda\vec{v} = f(\lambda\vec{u}) \in f(V)$, ce qui prouve que $f(V)$ est également stable par multiplication par un nombre réel.

Conclusion. L'ensemble $f(F)$ est un sous-espace vectoriel de $\mathbf{R}^m$. □

Exemples 2.7.5. — a) Considérons une matrice avec une seule ligne

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \in \mathscr{M}_{1,n}(\mathbf{R}).$$

alors cette matrice correspond à une application linéaire φ de $\mathbf{R}^n$ dans $\mathbf{R}$ qui est donnée par

$$\varphi : (x_1, x_2, \dots, x_n) \longmapsto a_1x_2 + a_2x_2 + \cdots + a_nx_n.$$

Il en résulte que son noyau

$$\mathrm{Ker}(\varphi) = \{\, (x_1, \dots, x_n) \in \mathbf{R}^n \mid a_1x_1 + \cdots + a_nx_n = 0 \,\}$$

est un sous-espace vectoriel de $\mathbf{R}^n$.

b) Considérons maintenant une matrice à m lignes et n colonnes :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathscr{M}_{m,n}(\mathbf{R}).$$

Notons $\vec{u}_i = (a_{1,i}, \dots, a_{m,i}) \in \mathbf{R}^m$ le vecteur correspondant à la i-ème colonne de la matrice A. Alors l'applicarion linéaire correspondant à la matrice A est l'application linéaire φ donnée par

$$\varphi : (x_1, \dots, x_n) \longmapsto x_1\vec{u}_1 + \dots x_n\vec{u}_n = \left(\sum_{k=1}^{n} a_{1,k}x_k, \dots, \sum_{k=1}^{n} a_{m,k}x_k \right).$$

Le noyau de φ est donc la partie de $\mathbf{R}^n$ définie par le système de m équations linéaires sans second membre :

$$\begin{cases} a_{1,1}X_1 + \cdots + a_{1,n}X_n = 0 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n = 0 \\ \dots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n = 0 \end{cases}$$

L'image de φ est l'ensemble de vecteurs

$$\{\, \vec{v} \in \mathbf{R}^m \mid \exists (\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n, \vec{v} = \lambda_1\vec{u}_1 + \dots \lambda_n\vec{u}_n \,\}.$$

Définition 2.7.6

Soient $\vec{u}_1, \ldots, \vec{u}_n$ des vecteurs de $\mathbf{R}^m$. Une *combinaison linéaire de* $\vec{u}_1, \ldots, \vec{u}_n$ est un vecteur $\vec{w} \in \mathbf{R}^m$ tel qu'il existe des nombres réels $\lambda_1, \ldots, \lambda_n$ tels que

$$\vec{w} = \lambda_1 \vec{u}_1 + \cdots + \lambda_n \vec{u}_n.$$

On vient de démontrer le résultat suivant :

Proposition 2.7.7

L'ensemble des combinaisons linéaires de $\vec{u}_1, \ldots, \vec{u}_n$ est un sous-espace vectoriel de $\mathbf{R}^m$ appelé le *sous-espace vectoriel engendré par* $\vec{u}_1, \ldots, \vec{u}_n$. On le note $\mathrm{Vect}(\vec{u}_1, \ldots, \vec{u}_n)$.

Exemple 2.7.8. — Pour $n = 1$, on obtient l'ensemble des multiples du vecteur $\vec{u}_1$, si celui-ci est non nul cela donne donc une droite passant par $\overrightarrow{0}$ et contenant $\vec{u}_1$. Pour $n = 2$, si les deux vecteurs ne sont pas *colinéaires*, cela va donner un plan (figure 6 et 7)

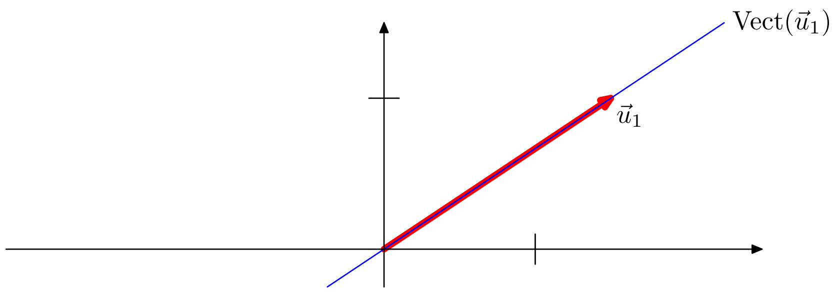


FIGURE 6. Sous-espace engendré par un vecteur

2.8. Compléments sur les fonctions. — Nous allons utiliser les notions d'injectivité, surjectivité et bijectivité pour les applications linéaires.

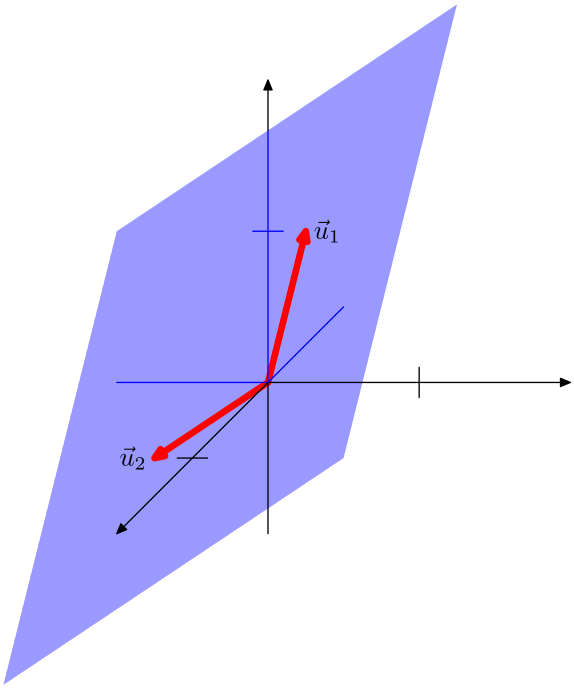


FIGURE 7. Sous-espace engendré par deux vecteurs

Définition 2.8.1

Soient X et Y des ensembles et soit $f : X \to Y$ une application. On dit que :

a) L'application f est *injective* si

$$\forall x_1, x_2 \in X, \quad f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

Autrement dit, pour tout $y \in Y$, il existe *au plus* un élément $x \in X$ tel que $f(x) = y$;

b) L'application f est *surjective* si pour tout $y \in Y$ il existe *au moins* un élément $x \in X$ tel que $f(x) = y$;

c) L'application f est *bijective* si pour tout $y \in Y$ il existe exactement un élément $x \in X$ tel que $f(x) = y$.

Remarque 2.8.2. — On notera que, par définition, une application est bijective si et seulement si elle est à la fois injective et surjective.

Représentation 2.8.3. — Représentons schématiquement ces différents types d'application la figure 8 représente une application qui n'est ni injective ni surjecyive, la figure 9 une application qui est injective sans être surjective, la figure 10 une application qui est surjective sans être injective et la figure 11 une application bijective.

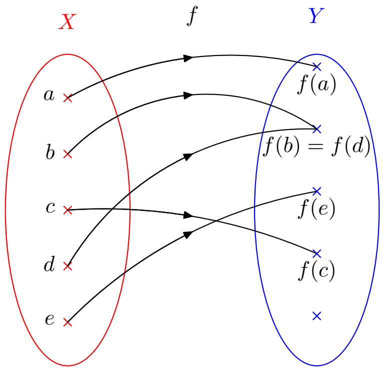


Figure 8. Application ni injective ni surjective

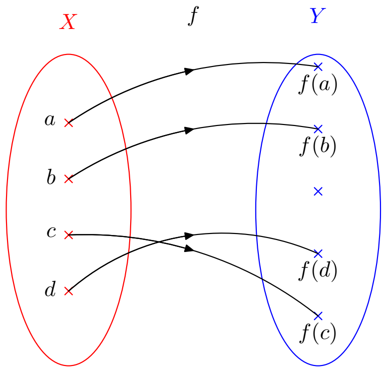


Figure 9. Application injective

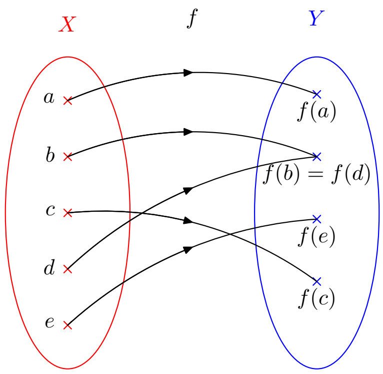


Figure 10. Application surjective

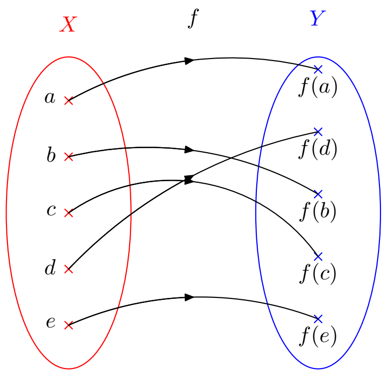


Figure 11. Application bijective

Proposition 2.8.4

Si X et Y sont des ensembles finis et $f : X \to Y$ une application, alors

a) Si f est injective, alors Y a plus d'éléments que X ;

b) Si f est surjective, alors X a plus d'éléments que Y ;

c) Si f est bijective, alors X et Y ont le même nombre d'éléments.

Si on considère cet énoncé en termes du schéma représentant f, le point est que, pour une application, le nombre de flèche est exactement le nombre d'éléments de X. Si l'application est injective, alors en chaque point de Y il arrive au plus une flèche, si bien que le nombre de flèche est inférieur ou égal au nombre d'éléments dans Y. De même, si f est surjective alors en tout point de Y il arrive au moins une flèche, si bien que le nombre de flèches est supérieur ou égal au nombre d'éléments dans Y.

Définition 2.8.5

Si $f : X \to Y$ est une application *bijective* alors l'applications f^{-1} qui à un élément $y \in Y$ associe l'unique élément $x \in X$ tel que $f(x) = y$ s'appelle la *réciproque de* f.

Remarques 2.8.6. — a) Avec les notations de la définition, on peut remarquer que pour $y \in Y$, $f(f^{-1}(y)) = y$ et pour tout $x \in X$, $f^{-1}(f(x)) = x$. Autrement dit $f \circ f^{-1} = \mathrm{Id}_Y$ et $f^{-1} \circ f = \mathrm{Id}_X$.

b) Notons que si $f : X \to Y$ est bijective, alors sa réciproque f^{-1} l'est également et $(f^{-1})^{-1} = f$.

Représentation 2.8.7. — La figure 12 donne une représentation schématique de l'application réciproque. Les flèches pour la réciproque f^{-1} s'obtiennent en renversant le sens des flèches de f.

Exemple 2.8.8. — Considérons l'élévation au carré (figure 13). L'application

$$\begin{array}{rcl} f_0 : \mathbf{R} & \longrightarrow & \mathbf{R} \\ x & \longmapsto & x^2 \end{array}$$

n'est ni injective ni surjective car $f(-1) = 1 = f(1)$ et -1 n'est pas le carré d'un nombre réel. L'application

$$\begin{array}{rcl} f_1 : \mathbf{R}_+ & \longrightarrow & \mathbf{R} \\ x & \longmapsto & x^2 \end{array}$$

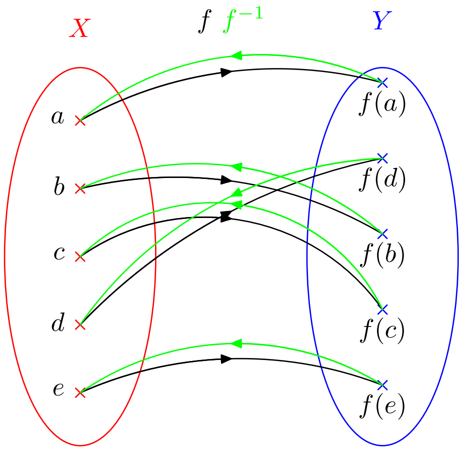


Figure 12. Application réciproque

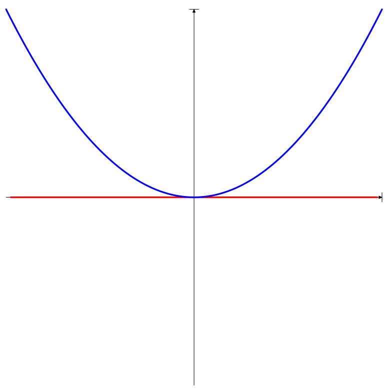


Figure 13. L'élévation au carré

et injective mais pas surjective. L'application

$$\begin{aligned} f_2 : \mathbf{R} &\longrightarrow \mathbf{R}_+ \\ x &\longmapsto x^2 \end{aligned}$$

est surjective mais pas injective. L'application

$$\begin{aligned} f_3 : \mathbf{R}_+ &\longrightarrow \mathbf{R}_+ \\ x &\longmapsto x^2 \end{aligned}$$

est bijective de réciproque l'application $f_3^{-1} : x \mapsto \sqrt{x}$. Cela souligne que lorsqu'on considère les questions d'injectivité ou de surjectivité, il faut faire attention à l'ensemble de départ et à l'ensemble d'arrivée.

Revenons maintenant au cas des applications linéaires :

Proposition 2.8.9

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire. Alors

a) L'application f est injective si et seulement si $\mathrm{Ker}(f) = \{\overrightarrow{0}\}$.

b) L'application f est surjective si et seulement si $\mathrm{Im}(f) = \mathbf{R}^m$.

L'assertion b) résulte des définitions. Nous allons donc seulement prouver la première assertion.

Preuve de l'assertion a). — Démontrons l'implication $\Rightarrow$. Supposons donc f injective. Comme l'application f est linéaire, $f\left(\overrightarrow{0}_{\mathbf{R}^n}\right) = \overrightarrow{0}_{\mathbf{R}^m}$. Si $\vec{u} \in \mathrm{Ker}(f)$, alors $f(\vec{u}) = \overrightarrow{0}_{\mathbf{R}^m}$ et comme f est supposée injective, $\vec{u} = \overrightarrow{0}_{\mathbf{R}^n}$. Par conséquent l'unique élément de $\mathrm{Ker}(f)$ est le vecteur nul de $\mathbf{R}^n$.

Démontrons maintenant la réciproque. Supposons $\mathrm{Ker}(f) = \{\overrightarrow{0}\}$. Soient $\vec{u}_1, \vec{u}_2 \in \mathbf{R}^n$ tels que $f(\vec{u}_1) = f(\vec{u}_2)$. alors, comme l'application f est linéaire,

$$f(\vec{u}_1 - \vec{u}_2) = f(\vec{u}_1) - f(\vec{u}_2) = \overrightarrow{0}_{\mathbf{R}^m}.$$

Par définition du noyau cela signifie que $\vec{u}_1 - \vec{u}_2 \in \mathrm{Ker}(f)$. Mais par hypothèse $\overrightarrow{0}_{\mathbf{R}^n}$ est l'unique élément du noyau. Donc $\vec{u}_1 - \vec{u}_2 = \overrightarrow{0}_{\mathbf{R}^n}$ et donc $\vec{u}_1 = \vec{u}_2$ ce qui conclut la preuve de l'injectivité de f. □

Exemple 2.8.10. — Vérifions que l'application linéaire f de matrice $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ est injective.

Écrivons explicitement f. L'image de $(x,y) \in \mathbf{R}^2$ est donnée par le produit matriciel

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix}$$

Autrement dit f est donnée par $(x,y) \mapsto (x + 2y, 3x + 4y)$. Par conséquent un couple (x,y) appartient au noyau si et seulement si il est solution du système

$$\begin{cases} x + 2y = 0 \\ 3x + 4y = 0 \end{cases}$$

En notant L_i la i-ème ligne du système la transformation $L_2 \leftarrow L_2 - 3L_1$ donne le système équivalent

$$\left\{\begin{array}{ccccc} x & + & 2y & = & 0 \\ & - & 2y & = & 0 \end{array}\right.$$

qui donne $y = 0$ puis $x = 0$. Donc $\mathrm{Ker}(f) = \{(0,0)\}$ et f est injective.

2.9. Base d'un sous-espace vectoriel, dimension. — Nous allons utiliser la notion de bijectivité pour définir une base.

Définition 2.9.1

Soit F un sous-espace vectoriel de $\mathbf{R}^m$ et soient $\vec{f}_1, \ldots, \vec{f}_n$ des vecteurs de F. On dit que la famille $(\vec{f}_1, \ldots, \vec{f}_n)$ est une *base de F* si et seulement si l'application

$$\begin{array}{ccc} \mathbf{R}^n & \longrightarrow & F \\ (x_1, \ldots, x_n) & \longmapsto & x_1\vec{f}_1 + \cdots + x_n\vec{f}_n \end{array}$$

est *bijective*. Si $(\vec{f}_1, \ldots, \vec{f}_n)$ est une base de F, pour tout $\vec{u} \in F$, il existe alors un unique n-uplet $(x_1, \ldots, x_n) \in \mathbf{R}^n$ tel que

$$\vec{u} = x_1\vec{f}_1 + \ldots x_n\vec{f}_n.$$

Ce n-uplet $(x_1, \ldots, x_n)$ s'appelle *les coordonnées de $\vec{u}$ dans la base $(\vec{f}_1, \ldots, \vec{f}_n)$*. L'entier n ne dépend que du sous-espace vectoriel F, on l'appelle la *dimension de F* et on le note $\dim(F)$.

Un sous-espace vectoriel de dimension 1 est appelé *droite vectorielle* et un sous espace vectoriel de dimension 2 est appelé *plan vectoriel*.

***Exemples 2.9.2.* —** a) L'unique sous-espace de dimension 0 est le sous-espace vectoriel $\left\{\overrightarrow{0}_{\mathbf{R}^m}\right\}$. En effet $\mathbf{R}^0$ possède un unique élément $\overrightarrow{0}_{\mathbf{R}^0} = ()$ où $()$ est l'unique 0-uplet.

b) Soit $\vec{e}_j = (\delta_{1,j}, \ldots, \delta_{n,j})$ pour $j \in \{1, \ldots, n\}$, où δ désigne le symbole de Kronecker. D'après la formule (8),

$$x_1\vec{e}_1 + \cdots + x_n\vec{e}_n = (x_1, \ldots, x_n).$$

Donc pour un n-uplet $\vec{u} = (u_1, \ldots, u_n) \in \mathbf{R}^n$, l'égalité

$$\vec{u} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$$

équivaut à $(x_1, \ldots, x_n) = (u_1, \ldots, u_n)$. Il y a donc bien une unique possibilité pour $(x_1, \ldots, x_n)$. Donc $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base de $\mathbf{R}^n$ et les coordonnées d'un n-uplet $\vec{u} = (u_1, \ldots, u_n)$ dans cette base sont $(u_1, \dots, u_n)$ lui-même. La base $(\vec{e}_1, \dots, \vec{e}_n)$ est appelée la *base usuelle* (ou, par certains auteurs, base canonique) *de* $\mathbf{R}^n$.

c) L'exercice corrigé 2.1 donne un exemple de recherche d'une base dans un plan vectoriel.

Proposition 2.9.3

Soient F_1 et F_2 des sous-espaces vectoriels de $\mathbf{R}^m$ tels que $F_1 \subset F_2$ (F_1 est contenu dans F_2) alors

a) On a l'inégalité $\dim(F_1) \leqslant \dim(F_2)$;

b) L'égalité de parties $F_1 = F_2$ équivaut à l'égalité $\dim(F_1) = \dim(F_2)$.

Exemple 2.9.4. — Si F est une droite vectorielle de $\mathbf{R}^m$, les seuls sous-espaces vectoriels contenus dans F sont $\{\overrightarrow{0}\}$ et F. En effet, si F' est un sous-espace vectoriel de $\mathbf{R}^m$ contenu dans F, alors $\dim(F') \leqslant \dim(F) = 1$. Donc $\dim(F') \in \{0, 1\}$. Si $\dim(F') = 0$, alors $F' = \{\overrightarrow{0}\}$. Sinon $\dim(F') = 1 = \dim(F)$ et donc $F' = F$.

Théorème du rang 2.9.5

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire alors

$$n = \dim(\operatorname{Ker}(f)) + \dim(\operatorname{Im}(f)).$$

2.10. Bases de $\mathbf{R}^n$, isomorphisme, matrice inversible. — Les bases de $\mathbf{R}^n$ sont particulièrement utiles, l'énoncé suivant permet de les caractériser.

Proposition 2.10.1

Soient $\vec{f}_1, \dots, \vec{f}_n$ des vecteurs de $\mathbf{R}^n$ et soit $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ l'endomorphisme correspondant donné par

$$(x_1, \dots, x_n) \longmapsto x_1\vec{f}_1 + \dots + x_n\vec{f}_n.$$

On note $A = \operatorname{Mat}(\varphi)$. Les sept assertions suivantes sont équivalentes :

(i) $(\vec{f}_1, \ldots, \vec{f}_n)$ est une base de $\mathbf{R}^n$;

(ii) L'endomorphisme φ est bijectif ;

(iii) L'endomorphisme φ est injectif ;

(iv) L'endomorphisme φ est surjectif ;

(v) $\mathrm{Ker}(\varphi) = \left\{\overrightarrow{0}\right\}$;

(vi) $\mathrm{Im}(\varphi) = \mathbf{R}^n$;

(vii) Il existe une matrice $B \in \mathscr{M}_n(\mathbf{R})$ telle que $BA = I_n$;

(viii) Il existe une matrice $B \in \mathscr{M}_n(\mathbf{R})$ telle que $AB = I_n$.

Si ces conditions sont vérifiées, alors

a) La matrice B des assertions (vii) et (viii) est unique, c'est la matrice de la réciproque φ^{-1} de φ qui est une application linéaire.

b) Si $\vec{u} = (x_1, \ldots, x_n) \in \mathbf{R}^n$, les coordonnées de $\vec{u}$ dans la base $(\vec{f}_1, \ldots, \vec{f}_n)$ sont données par le vecteur colonne $B\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

On prendra garde au fait que la propriété ne fonctionne que lorsque le nombre de vecteurs $\vec{f}_1, \ldots, \vec{f}_n$ est précisément n, c'est-à-dire la dimension de $\mathbf{R}^n$.

Idée de la preuve. — L'équivalence (i)$\Leftrightarrow$(ii) résulte de la définition d'une base.

Les implications (ii)$\Rightarrow$(iii) et (ii)$\Rightarrow$(iv) résulte de la définition de la bijectivité (définition 2.8.1).

Les équivalences (iii)$\Leftrightarrow$(v) et (iv)$\Leftrightarrow$(vi) sont l'objet de la proposition 2.8.9.

Démontrons l'implication (v)$\Rightarrow$(ii). Si $\mathrm{Ker}(\varphi) = \left\{\overrightarrow{0}\right\}$, alors $\dim(\mathrm{Ker}(\varphi)) = 0$. Par le théorème du rang,

$$\dim(\mathrm{Im}(\varphi)) = n - \dim(\mathrm{Ker}(\varphi)) = n = \dim(\mathbf{R}^n).$$

L'assertion b) de la proposition 2.9.3 donne que $\mathrm{Im}(\varphi) = \mathbf{R}^n$. Donc φ est injective et surjective et donc bijective.

L'implication (vi)$\Rightarrow$(ii) se démontre de manière similaire : si $\mathrm{Im}(\varphi) = \mathbf{R}^n$, alors $\dim(\mathrm{Im}(\varphi)) = n$. Par le théorème du rang,

$$\dim(\mathrm{Ker}(\varphi)) = n - \dim(\mathrm{Im}(\varphi)) = n - n = 0.$$

Par l'exemple 2.9.2 a), $\mathrm{Ker}(\varphi) = \{\overrightarrow{0}\}$. Donc φ est injective et surjective et donc bijective.

Vérifions que si φ est bijective, sa réciproque est également linéaire : Soient $\vec{u}, \vec{v} \in \mathbf{R}^n$, alors

$$\varphi^{-1}(\vec{u}+\vec{v}) = \varphi^{-1}\big(\varphi(\varphi^{-1}(\vec{u})) + \varphi(\varphi^{-1}(\vec{v}))\big) = \varphi^{-1}\big(\varphi(\varphi^{-1}(\vec{u}) + \varphi^{-1}(\vec{v}))\big) = \varphi^{-1}(\vec{u}) + \varphi^{-1}(\vec{v})$$

où l'avant-dernière égalité résulte de la linéarité de φ, les autres résultant de ce que $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = \mathrm{Id}_{\mathbf{R}^n}$. De même soit $\lambda \in \mathbf{R}$ et soit $\vec{u} \in \mathbf{R}^n$

$$\varphi^{-1}(\lambda\vec{u}) = \varphi^{-1}\big(\lambda\varphi(\varphi^{-1}(\vec{u}))\big) = \varphi^{-1}\big(\varphi(\lambda\varphi^{-1}(\vec{u}))\big) = \lambda\varphi^{-1}(\vec{u}).$$

Donc φ^{-1} est linéaire.

Démontrons maintenant que (ii)⇒((vii) et (viii)) Soit $B = \mathrm{Mat}(\varphi^{-1})$. Les relations $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = \mathrm{Id}_{\mathbf{R}^n}$ impliquent les relations matricielles

$$AB = \mathrm{Mat}(\varphi \circ \varphi^{-1}) = BA = \mathrm{Mat}(\varphi^{-1} \circ \varphi) = \mathrm{Mat}(\mathrm{Id}_{\mathbf{R}^n}) = I_n.$$

Passons à l'implication (vii)⇒(v) Si $BA = I_n$, soit ψ l'application de matrice B. Alors

$$\mathrm{Mat}(\psi \circ \varphi) = BA = I_n = \mathrm{Mat}(\mathrm{Id}_{\mathbf{R}^n})$$

et donc $\psi \circ \varphi = \mathrm{Id}_{\mathbf{R}^n}$ Donc si $\vec{u} \in \mathrm{Ker}(\varphi)$, alors $\vec{u} = \psi \circ \varphi(\vec{u}) = \psi(\overrightarrow{0}) = \overrightarrow{0}$. Donc $\mathrm{Ker}(\varphi) = \{\overrightarrow{0}\}$.

Il nous reste à démontrer l'implication (viii)⇒(iii) Si $AB = I_n$, soit ψ l'application de matrice B. Alors $\mathrm{Mat}(\varphi \circ \psi) = AB = I_n = \mathrm{Mat}(\mathrm{Id}_{\mathbf{R}^n})$ et donc $\varphi \circ \psi = \mathrm{Id}_{\mathbf{R}^n}$. Donc si $\vec{v} \in \mathbf{R}^n$, alors $\vec{v} = \varphi \circ (\psi(\vec{v}))$. Donc φ est surjective.

L'assertion a) découle de ce qui précède et l'assertion b) de la définition des coordonnées. □

Définition 2.10.2

Soit $A \in \mathscr{M}_n(\mathbf{R})$. On dit que la matrice A est inversible s'il existe une matrice $B \in \mathbf{R}^n$ telle que

$$AB = BA = I_n.$$

La matrice B est uniquement déterminée par A, on l'appelle l'inverse de A et on la note A^{-1}.

Remarque 2.10.3. — On notera que si φ est un endomorphisme *bijectif*, alors

$$\mathrm{Mat}(\varphi^{-1}) = \mathrm{Mat}(\varphi)^{-1}.$$

L'exercice corrigé 2.2 donne un exemple numerique de calcul de l'inverse d'une matrice. La remarque 2.0.1 explique comment faire ce calcul avec une matrice augmentée.

Définition 2.10.4

Le déterminant d'une matrice de $\mathscr{M}_2(\mathbf{R})$ est donnée par

$$\text{dét}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

et le déterminant d'une matrice de $\mathscr{M}_3(\mathbf{R})$ par

$$\text{dét}\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2}$$
$$- a_{3,1}a_{2,2}a_{1,3} - a_{3,2}a_{2,3}a_{1,1} - a_{3,3}a_{2,1}a_{1,2}$$

calculé avec la règle de Sarrus

$$\begin{array}{|ccc|cc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \end{array}$$

***Remarques 2.10.5.* —** a) On peut aussi définir le déterminant d'une matrice de $\mathscr{M}_1(\mathbf{R})$ par

$$\text{dét}(a) = a.$$

b) Plus généralement, il est possible de définir le déterminant d'une matrice carrée de taille arbitraire en utilisant la formule de développement suivant la première ligne :

$$\text{dét}\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} = \sum_{i=1}^{n}(-1)^i a_{1,i}\,\text{dét}\begin{pmatrix} a_{2,1} & \cdots & a_{2,i-1} & a_{2,i+1} & \cdots & a_{2,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{m,i-1} & a_{m,i+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Le lecteur pourra vérifier que cette formule vaut aussi pour $n = 2$ ou $n = 3$.

Proposition 2.10.6

Si $A \in \mathscr{M}_n(\mathbf{R})$, alors la matrice A est inversible si et seulement si $\text{dét}(A) \neq 0$.

Exemple 2.10.7. — En prenant la matrice de l'exercice corrigé 2.2

$$\text{dét}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} = 14 + 6 + 4 - 4 - 12 - 7 = 1.$$

Formules

Soient $A, B \in \mathscr{M}_n(\mathbf{R})$,

$$\text{dét}(AB) = \text{dét}(A)\text{dét}(B)$$

Interprétation géométrique 2.10.8. — Si $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ sont des vecteurs de $\mathbf{R}^2$, alors on peut définir

$$\text{dét}(\vec{u}, \vec{v}) = \text{dét}\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$$

Alors $|\text{dét}(\vec{u}, \vec{v})|$ est l'aire du parallélogramme $(OACB)$ où $O = (0,0)$, $A = O + \vec{u}$, $B = O + \vec{v}$ et $C = O + \vec{u} + \vec{v}$ (figure 14).

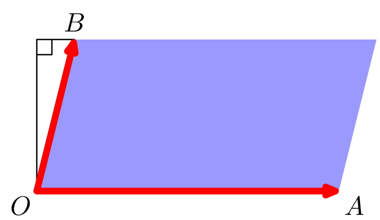


FIGURE 14. Aire du parallélogramme

De même, si $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ et $\vec{w} = (w_1, w_2, w_3)$ sont des vecteurs de $\mathbf{R}^3$, on peut définir

$$\text{dét}(\vec{u}, \vec{v}, \vec{w}) = \text{dét}\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

Alors $|\text{dét}(\vec{u}, \vec{v}, \vec{w})|$ est le volume du parallélépipède défini par les trois vecteurs $\vec{u}$, $\vec{v}$ et $\vec{w}$ (figure 15).

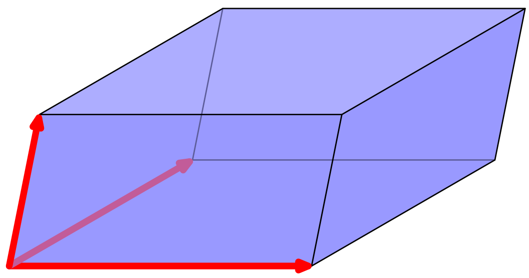


FIGURE 15. Volume du parallélépipède

Cette interprétation géométrique du déterminant comme volume est à la base du résultat suivant qui est une application très utile du déterminant :

Théorème 2.10.9

Soient U et V des ouverts de $\mathbf{R}^n$. Soit $\varphi : U \to V$ une bijection différentiable. On suppose que $f : V \to \mathbf{R}$ est une fonction intégrable. Alors la fonction $x \mapsto f \circ \varphi(x)|\text{dét}(\text{Jac}_x(\varphi))|$ est intégrable sur U et

$$\int_U f \circ \varphi(x)|\text{dét}(\text{Jac}_x(\varphi))|\mathrm{d}x_1 \dots \mathrm{d}x_n = \int_V f(y)\mathrm{d}y_1 \dots \mathrm{d}y_n$$

Remarques 2.10.10. — a) Pour $n = 1$ cela redonne une formule de changement de variable classique que vous avez pu voir sous la forme : si $y = \varphi(x)$, $\mathrm{d}y = |\varphi'(x)|\mathrm{d}x$

b) Le symbole $\int$ désigne ici en fait une intégrale multiple en n variables (*cf.* l'exemple 1.7.3 b)).

c) Les subtilités de l'intégration en plusieurs variables sortent du cadre de ce cours, nous ne définissons donc pas formellement « intégrable ». Nous nous contentons donc de donner le sens étymologique « qu'on peut intégrer ». Par exemple, une fonction continue sur un produit d'intervalles fermés et bornés est intégrable, de même que sa restriction à un ouvert contenu dans ce produit.

Exemples 2.10.11. — a) On considère à nouveau les coordonnées circulaires :

$$\varphi : (r, \theta) \longmapsto (r\cos(\theta), r\sin(\theta))$$

qui définit une bijection de $\mathbf{R}_+ \times [0, 2\pi[$ sur $\mathbf{R}^2 - \{0\}$. Alors

$$\begin{aligned}|\text{dét}(\text{Jac}_{(r,\theta)}(\varphi))| &= \left|\text{dét}\begin{pmatrix}\cos(\theta) & -r\sin(\theta)\\ \sin(\theta) & r\cos(\theta)\end{pmatrix}\right|\\ &= r(\cos(\theta)^2 + \sin(\theta)^2)\\ &= r.\end{aligned}$$

Cela donne la formule

$$\int_U f(r\cos(\theta), r\sin(\theta)) r \mathrm{d}r\,\mathrm{d}\theta = \int_{\varphi(U)} f(x,y)\mathrm{d}x\,\mathrm{d}y$$

b) Comme cas particulier de la formule précédente, calculons $\int_{-\infty}^{+\infty} e^{-x^2}\mathrm{d}x$.

$$\begin{aligned}\left(\int_{-\infty}^{+\infty} e^{-x^2}\mathrm{d}x\right)^2 &= \int_{-\infty}^{+\infty} e^{-y^2}\int_{-\infty}^{+\infty} e^{-x^2}\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\mathbf{R}^2} e^{-(x^2+y^2)}\mathrm{d}x\,\mathrm{d}y\\ &= \int_0^{2\pi}\int_0^{+\infty} e^{-r^2} r\mathrm{d}r\,\mathrm{d}\theta\\ &= 2\pi\left[-\frac{1}{2}e^{-t^2}\right]_0^{+\infty}\\ &= \pi\end{aligned}$$

Ce qui donne $\int_{-\infty}^{+\infty} e^{-x^2}\mathrm{d}x = \sqrt{\pi}$. Plus généralement, on peut en déduire, si $\mu \in \mathbf{R}$ et $\sigma > 0$ la formule

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\mathrm{d}x = 1.$$

Donc la fonction $x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ est une fonction densité d'une loi de probabilité sur $\mathbf{R}$ appelée *loi gaussienne* de moyenne μ et d'écart type σ.

Changer de base permet de choisir un système de coordonnées mieux adapté au problème considéré.

Terminologie 2.10.12

Un *endomorphisme linéaire* (ou simplement *endomorphisme*) est une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^n$.

Définition 2.10.13

Soit $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire. Soient $\boldsymbol{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ une base de $\mathbf{R}^n$ et $\boldsymbol{f} = (\vec{f}_1, \ldots, \vec{f}_m)$ une base de $\mathbf{R}^m$. La *matrice de φ dans les bases $\boldsymbol{e}$ et $\boldsymbol{f}$* est la matrice dont la j-ème colonne sont les coordonnées de $\varphi(\vec{e}_j)$ dans la base $(\vec{f}_1, \ldots, \vec{f}_m)$. On la note

$$\mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi)$$

Dans le cas d'un endomorphisme $\psi : \mathbf{R}^n \to \mathbf{R}^n$, on note également

$$\mathrm{Mat}_{\boldsymbol{e}}(\psi) = \mathrm{Mat}_{\boldsymbol{e},\boldsymbol{e}}(\psi).$$

Soient $\boldsymbol{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ et $\boldsymbol{e}' = (\vec{e}_1{}', \ldots, \vec{e}_n{}')$ des bases de $\mathbf{R}^n$. La *matrice de passage de $\boldsymbol{e}$ à $\boldsymbol{e}'$* est la matrice dont la j-ème sont les coordonnées de $\vec{e}_j{}'$ dans la base $(\vec{e}_1, \ldots, \vec{e}_n)$. Dans ce cours, on la note $P_{\boldsymbol{e}}^{\boldsymbol{e}'}$.

Proposition 2.10.14

Soient $\boldsymbol{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ une base de $\mathbf{R}^n$, $\boldsymbol{f} = (\vec{f}_1, \ldots, \vec{f}_m)$ une base de $\mathbf{R}^m$ et soit $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire. Soit $\vec{u} \in \mathbf{R}^n$. On note X le vecteur colonne de ses coordonnées dans la base $\boldsymbol{e}$ et Y le vecteur colonne des coordonnées de $\varphi(\vec{u}) \in \mathbf{R}^m$ dans la base $\boldsymbol{f}$. Alors

$$Y = \mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi) X.$$

Proposition 2.10.15

Soient $\boldsymbol{e} = (\vec{e}_1, \dots, \vec{e}_n)$ une base de $\mathbf{R}^n$, $\boldsymbol{f} = (\vec{f}_1, \dots, \vec{f}_m)$ une base de $\mathbf{R}^m$ et $\boldsymbol{g} = (\vec{g}_1, \dots, \vec{g}_p)$ une base de $\mathbf{R}^p$. et soient $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ et $\psi : \mathbf{R}^m \to \mathbf{R}^p$ des applications linéaires, alors

$$\mathrm{Mat}_{\boldsymbol{e},\boldsymbol{g}}(\psi \circ \varphi) = \mathrm{Mat}_{\boldsymbol{f},\boldsymbol{g}}(\psi)\,\mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi).$$

Formules (Changement de base)

Soient $\boldsymbol{e} = (\vec{e}_1, \dots, \vec{e}_n)$, $\boldsymbol{e}' = (\vec{e}'_1, \dots, \vec{e}'_n)$ et $\boldsymbol{e}'' = (\vec{e}''_1, \dots, \vec{e}''_n)$ des bases de $\mathbf{R}^n$ et soient $\boldsymbol{f} = (\vec{f}_1, \dots, \vec{f}_m)$ et $\boldsymbol{f}' = (\vec{f}'_1, \dots, \vec{f}'_m)$ des bases de $\mathbf{R}^m$. Soit $\vec{u}$ un vecteur de $\mathbf{R}^n$, soit U le vecteur colonne de ses coordonnées dans la base $\boldsymbol{e}$ et U' celui de ses coordonnées dans la base $\boldsymbol{e}'$. Soit $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire. Alors

$$P_{\boldsymbol{e}}^{\boldsymbol{e}''} = P_{\boldsymbol{e}}^{\boldsymbol{e}'} P_{\boldsymbol{e}'}^{\boldsymbol{e}''} \qquad P_{\boldsymbol{e}'}^{\boldsymbol{e}} = \left(P_{\boldsymbol{e}}^{\boldsymbol{e}'}\right)^{-1}$$

$$U' = (P_{\boldsymbol{e}}^{\boldsymbol{e}'})^{-1} U \qquad \mathrm{Mat}_{\boldsymbol{e}',\boldsymbol{f}'}(\varphi) = \left(P_{\boldsymbol{f}}^{\boldsymbol{f}'}\right)^{-1} \mathrm{Mat}_{\boldsymbol{e}',\boldsymbol{f}'}(\varphi) P_{\boldsymbol{e}}^{\boldsymbol{e}'}$$

Explications. — Introduisons des notations pour les coefficients des matrices de passage de $\boldsymbol{e}$ à $\boldsymbol{e}'$ et de $\boldsymbol{e}'$ à $\boldsymbol{e}''$:

$$P_{\boldsymbol{e}}^{\boldsymbol{e}'} = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \quad \text{et} \quad P_{\boldsymbol{e}'}^{\boldsymbol{e}''} = (b_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

Soit $i \in \{1, \dots, n\}$.

$$\vec{e}''_k = \sum_{j=1}^{n} b_{j,k} \vec{e}'_k = \sum_{j=1}^{n} b_{j,k} \sum_{i=1}^{n} a_{i,j} \vec{e}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} b_{j,k} \right) \vec{e}_i$$

ce qui prouve l'égalité $P_{\boldsymbol{e}}^{\boldsymbol{e}''} = P_{\boldsymbol{e}}^{\boldsymbol{e}'} P_{\boldsymbol{e}'}^{\boldsymbol{e}''}$. En appliquant cette formule avec $\boldsymbol{e}'' = \boldsymbol{e}$, on obtient que $P_{\boldsymbol{e}}^{\boldsymbol{e}'} P_{\boldsymbol{e}'}^{\boldsymbol{e}} = I_n$ ce qui fournit la seconde formule. Notons $(u_1, \dots, u_n)$ (resp. $(u'_1, \dots, u'_n)$) les coordonnées de $\vec{u}$ dans la base $\boldsymbol{e}$ (resp. $\boldsymbol{e}'$). Alors

$$\vec{u} = u'_1 \vec{e}'_1 + \cdots + u'_n \vec{e}'_n = \sum_{j=1}^{n} u'_j \left(\sum_{i=1}^{n} a_{i,j} \vec{e}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} u'_j \right) \vec{e}_i$$

ce qui prouve l'égalité $U = P_{\boldsymbol{e}}^{\boldsymbol{e}'} U'$ et, par conséquent, la troisième formule.

Notons $\vec{v} = \varphi(\vec{u})$, V le vecteur colonne des coordonnées de $\vec{v}$ dans la base $\boldsymbol{f}$ et V' celui de ses coordonnées dans la base $\boldsymbol{f}'$. Alors on a les relations $V' = (P_{\boldsymbol{f}}^{\boldsymbol{f}'})^{-1} V$, $V = \mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi) U$ et $U = P_{\boldsymbol{e}}^{\boldsymbol{e}'} U'$. En combinant ces formules, on obtient la seconde égalité dans

$$\mathrm{Mat}_{\boldsymbol{e}',\boldsymbol{f}'}(\varphi) U' = V' = (P_{\boldsymbol{f}}^{\boldsymbol{f}'})^{-1} \mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi) P_{\boldsymbol{e}}^{\boldsymbol{e}'} U'$$

Comme cette égalité est vraie pour n'importe quel vecteur vecteur $\vec{u}$, elle est en particulier vraie pour tout les vecteurs $\vec{e}_i'$ pour $i \in \{1, \ldots, n\}$. On en déduit que les colonnes des deux matrices $\mathrm{Mat}_{\boldsymbol{e}',\boldsymbol{f}'}(\varphi)$ et $(P_{\boldsymbol{f}}^{\boldsymbol{f}'})^{-1} \mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi) P_{\boldsymbol{e}}^{\boldsymbol{e}'}$ sont égales ce qui prouve la dernière égalité du formulaire. $\square$

Corollaire 2.10.16

Soient $\boldsymbol{e} = (\vec{e}_1, \ldots, \vec{e}_n)$, $\boldsymbol{e}' = (\vec{e}_1', \ldots, \vec{e}_n')$ des bases de $\mathbf{R}^n$ et soit $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ un endomorphisme. Alors

$$\mathrm{Mat}_{\boldsymbol{e}'}(\varphi) = (P_{\boldsymbol{e}}^{\boldsymbol{e}'})^{-1} \mathrm{Mat}_{\boldsymbol{e}}(\varphi) P_{\boldsymbol{e}}^{\boldsymbol{e}'}.$$

Terminologie 2.10.17

On dit que des matrices M et $N \in \mathscr{M}_n(\mathbf{R})$ sont *semblables* s'il existe une matrice inversible $P \in \mathscr{M}_n(\mathbf{R})$ telle que

$$N = P^{-1} M P.$$

Remarque 2.10.18. — On notera que des matrices sont semblables si et seulement si ce sont les matrices d'un même endomorphisme dans des bases différentes.

2.11. Valeurs propres, vecteurs propres et diagonalisation. — Un exemple particulièrement utile de changement de base concerne la diagonalisation des applications linéaires de $\mathbf{R}^n$ dans $\mathbf{R}^n$.

Terminologie 2.11.1

Une matrice *diagonale* est une matrice de la forme

$$\begin{pmatrix} \lambda_1 & 0 & & & 0 \\ 0 & \lambda_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \lambda_{n-1} & 0 \\ 0 & & & 0 & \lambda_n \end{pmatrix} \in \mathscr{M}_n(\mathbf{R}).$$

où $\lambda_1, \ldots, \lambda_n$ sont des nombres réels.

Définition 2.11.2

Soit $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ un endomorphisme et soit $\lambda \in \mathbf{R}$. On dit que λ est une *valeur propre de* φ s'il existe un vecteur *non nul* $\vec{u} \in \mathbf{R}^n$ tel que

$$\varphi(\vec{u}) = \lambda\vec{u}.$$

Un tel vecteur $\vec{u}$ est appelé *vecteur propre de* φ *pour* λ

On dit que φ est *diagonalisable* s'il vérifie une des conditions équivalentes suivantes :

(i) Il existe une base $(\vec{f}_1, \ldots, \vec{f}_n)$ de $\mathbf{R}^n$ formée de vecteurs propres de φ ;

(ii) Il existe une base $\boldsymbol{f} = (\vec{f}_1, \ldots, \vec{f}_n)$ de $\mathbf{R}^n$ telle que $\mathrm{Mat}_{\boldsymbol{f}}(\varphi)$ soit diagonale ;

(iii) La matrice $\mathrm{Mat}(\varphi)$ est semblable à une matrice diagonale.

On dit aussi que φ se *diagonalise* dans la base $(\vec{f}_1, \ldots, \vec{f}_n)$

Remarque 2.11.3. — Diagonaliser un endomorphisme est notamment pratique pour calculer les puissances itérées d'un endomorphisme : soit $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ un endomorphisme. On définit par récurrence φ^k par $\varphi^0 = \mathrm{Id}_{\mathbf{R}^n}$ et $\varphi^{k+1} = \varphi \circ \varphi^k$ pour tout $k \geqslant 0$. (Autrement dit, $\varphi^k = \varphi \circ \cdots \circ \varphi$

où φ est répété k fois.) En effet si $\boldsymbol{f} = (\vec{f}_1, \ldots, \vec{f}_n)$ diagonalise φ alors

$$\mathrm{Mat}_{\boldsymbol{f}}(\varphi^k) = \mathrm{Mat}_{\boldsymbol{f}}(\varphi)^k = \begin{pmatrix} \lambda_1^k & 0 & & & 0 \\ 0 & \lambda_2^k & & & \\ & & \ddots & & \\ & & & \lambda_{n-1}^k & 0 \\ 0 & & & 0 & \lambda_n^k \end{pmatrix}$$

Terminologie 2.11.4

On dit également qu'une matrice est *diagonalisable* si elle est semblable à une matrice diagonale.

Remarque 2.11.5. — Soit φ un endomorphisme de $\mathbf{R}^n$ et soit $A = \mathrm{Mat}(\varphi)$ sa matrice dans la base usuelle. Les assertions suivantes sont équivalentes

(i) λ est valeur propre de φ ;

(ii) $\exists \vec{u} \in \mathbf{R}^n - \{0\}, \quad \varphi(\vec{u}) = \lambda\vec{u}$;

(iii) $\exists \vec{u} \in \mathbf{R}^n - \{0\}, \quad \varphi(\vec{u}) - \lambda\vec{u} = \overrightarrow{0}$;

(iv) $\exists \vec{u} \in \mathbf{R}^n - \{0\}, \quad (\varphi - \lambda \mathrm{Id}_{\mathbf{R}^n})(\vec{u}) = \overrightarrow{0}$;

(v) $\mathrm{Ker}(\varphi - \lambda \mathrm{Id}_{\mathbf{R}^n}) \neq \{\overrightarrow{0}\}$;

(vi) La matrice $A - \lambda I_n$ n'est pas inversible ;

Si $n \in \{1, 2, 3\}$, cela équivaut également à

(vii) $\mathrm{dét}(\lambda I_n - A) = 0$.

Définition 2.11.6

Si $n \in \{1, 2, 3\}$, le polynôme caractéristique de φ est le polynôme de degré n

$$\chi_\varphi(X) = \mathrm{dét}(XI_n - A).$$

Proposition 2.11.7

L'ensemble
$$\{\lambda \in \mathbf{R} \mid \chi_\varphi(\lambda) = 0\}$$
est l'ensemble des valeurs propres réelles de φ (aussi appelé *spectre* réel de φ).

Proposition 2.11.8

Si l'ensemble $\{\lambda \in \mathbf{R} \mid \chi_\varphi(\lambda) = 0\}$ contient n éléments distincts, alors φ est diagonalisable.

Définition 2.11.9

Une base $(\vec{u}_1, \dots, \vec{u}_n)$ d'un sous-espace vectoriel F de $\mathbf{R}^m$ est dite *orthonormée* si elle vérifie
$$\vec{u}_i \cdot \vec{u}_j = \delta_{i,j}$$
pour tous $i, j \in \{1, \dots, n\}$, où $\delta_{i,j}$ est le symbole de Kronecker.

Remarques 2.11.10. — a) Soit $(\vec{e}_1, \dots, \vec{e}_n)$ une base orthonormée d'un sous-espace vectoriel F de $\mathbf{R}^m$. Soient $\vec{u}$ et $\vec{v}$ des vecteurs de F de coordonnées respectives $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ dans cette base. Alors
$$\begin{aligned}\vec{u} \cdot \vec{v} &= (x_1\vec{e}_1 + \dots + x_n\vec{v}_n) \cdot (y_1\vec{e}_1 + \dots + y_n\vec{e}_n) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \vec{e}_i \cdot \vec{e}_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \delta_{i,j} \\ &= x_1 y_1 + \dots + x_n y_n.\end{aligned}$$

b) Une famille de n-vecteurs $\boldsymbol{u} = (\vec{u}_1, \dots, \vec{u}_n)$ de $\mathbf{R}^n$ est une base orthonormée si et seulement si la matrice $P_{\boldsymbol{e}}^{\boldsymbol{u}}$ dont la j-ème colonne sont les coordonnées de $\vec{u}_j$ dans la base usuelle $\boldsymbol{e} = (\vec{e}_1, \dots, \vec{e}_n)$ est orthogonale.

Théorème 2.11.11

Si $\mathrm{Mat}(\varphi)$ est *symétrique*, alors il existe une base orthonormée de $\mathbf{R}^n$ qui diagonalise φ.

Remarque 2.11.12. — Notons qu'inversement, si φ se diagonalise dans une base orthonormée, alors il existe une matrice orthogonale P et une matrice diagonale D telles que

$$\mathrm{Mat}(\varphi) = PDP^{-1} = PD{}^tP$$

Donc

$${}^t\,\mathrm{Mat}(\varphi) = {}^t(PD{}^tP) = {}^{tt}P{}^tD{}^tP = \mathrm{Mat}(\varphi).$$

Donc $\mathrm{Mat}(\varphi)$ est symétrique.

2.12. Application à l'étude des fonctions en plusieurs variables. — Appliquons ce qui précède à l'étude des fonctions en plusieurs variables.

Définition 2.12.1

Si f est une fonction $\mathbf{R}^n$ dans $\mathbf{R}$ définie sur un ouvert sur lequel elle est de classe $\mathscr{C}^2$, c'est-à-dire que toutes ses dérivées partielles secondes sont bien définies et continues, alors pour tout $A \in \mathscr{D}_f$, on définit une application

$$\mathrm{d}^2 f_A : \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R}$$

de la manière suivante : si $\vec{u} = (u_1, \ldots, u_n)$ et $\vec{v} = (v_1, \ldots, v_n)$, on considère les vecteurs colonnes correspondants

$$U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad \text{et} \quad V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

et on pose

$$\mathrm{d}^2 f_A(\vec{u}, \vec{v}) = {}^t U \,\mathrm{Hess}_A(f) V$$

où $\mathrm{Hess}_A(f)$ est la matrice hessienne de f en A :

$$\mathrm{Hess}_A(f) = \left(\frac{\partial^2 f}{\partial X_i \partial X_j}\right)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}.$$

Théorème 2.12.2

Sous les hypothèses de la définition précédente,

$$f(M) = f(A) + \mathrm{d}f_A\left(\overrightarrow{AM}\right) + \frac{1}{2}\mathrm{d}^2 f_A\left(\overrightarrow{AM}, \overrightarrow{AM}\right) + o(AM^2).$$

Cette formule s'appelle le *développement de* Taylor *à l'ordre 2 de f en A*

Nous allons maintenant utiliser deux résultats vus auparavant :

Rappels 2.12.3. — a) Par le théorème de Schwarz (théorème 1.9.5), la matrice hessienne $\mathrm{Hess}_A(f)$ est symétrique.

b) Une matrice symétrique se diagonalise dans une base orthonormée (théorème 2.11.11).

Remarque 2.12.4. — On se place sous les hypothèses de la définition, les rappels fournissent une base orthonormée $(\vec{e}_1, \dots, \vec{e}_n)$ qui diagonalise $\mathrm{Hess}_A(f)$. Cela signifie que si E_i est le vecteur colonne correspondant à $\vec{e}_i$ alors $\mathrm{Hess}_A(f)E_i = \lambda_i E_i$ pour un $\lambda_i \in \mathbf{R}$. Avec ces notations, si un vecteur $\vec{u} \in \mathbf{R}^n$ a pour coordonnées $(x_1, \dots, x_n)$ dans cette base, on aura la formule

$$\mathrm{d}^2 f_A(\vec{u}, \vec{u}) = \sum_{i=1}^{n} \lambda_i x_i^2.$$

Définition 2.12.5

Soit f une fonction différentiable en un point $A \in \mathscr{D}_f$. On dit A est un *point critique de f* si

$$\mathrm{d}f_A = 0,$$

(ce qui équivaut à $\mathrm{grad}_A(f) = 0$).

Rappel 2.12.6. — Si l'application f admet un minimum ou un maximum local en A, alors A est un point critique de f (proposition 1.10.20).

Mais comme on l'a déjà vu, la réciproque est fausse. Cependant on a le critère suivant :

Critère 2.12.7

Soit f une application en n variables réelles à valeur réelle, définie sur un ouvert et de classe $\mathscr{C}^2$. Soit A un point *critique* de f. Alors

a) Si toutes les valeurs propres de $\text{Hess}_A(f)$ sont toutes *strictement* positives, alors f admet un minimum local en A ;

b) Si toutes les valeurs propres de $\text{Hess}_A(f)$ sont toutes *strictement* négatives, alors f admet un maximum local en A ;

c) Si $\text{Hess}_A(f)$ admet à la fois des valeurs propres strictement positives et des valeurs propres strictement négatives, alors f n'admet ni maximum ni minimum local en A. On dit que A est un *point selle* pour f.

Idée de la preuve. — Démontrons le point a), on suppose donc que toute les valeurs propres λ_i sont strictement positives. Choisissons une base orthonormée $(\vec{e}_1, \dots, \vec{e}_n)$ qui diagonalise $\text{Hess}_A(f)$. Alors, comme A est un point critique de f, le développement de Taylor à l'ordre deux de f s'écrit :

$$f(M) = f(A) + 0 + \frac{1}{2}\sum_{i=1}^{n}\lambda_i x_i^2 + o(AM^2)$$

si $\overrightarrow{AM} = \sum_{i=1}^{n} x_i \vec{e}_i$. Notons alors $c = \frac{1}{2}\min_{1\leqslant i\leqslant n}(\lambda_i)$. Alors

$$\frac{1}{2}\sum_{i=1}^{n}\lambda_i x_i^2 \geqslant cAM^2$$

Et par définition de la notation $o(g)$ (définition 1.10.2), il existe un nombre réel $\eta > 0$ tel que pour tout point $M \in \mathscr{D}_f$ vérifiant $AM < \eta$, on ait que le terme $o(AM^2)$ est borné, en valeur absolue, par $\frac{c}{2}AM^2$. En définitive, si $AM < \eta$, on obtient

$$f(M) \geqslant f(A) + cAM^2 - \frac{c}{2}AM^2 \geqslant f(A).$$

Cela prouve que f admet un minimum local en A. □

Remarque 2.12.8. — Soit $n \in \{1, 2, 3\}$ Notons que si $M \in \mathscr{M}_n(\mathbf{R})$ est une matrice diagonalisable c'est à dire que

$$M = P \begin{pmatrix} \lambda_1 & 0 & & & 0 \\ 0 & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & 0 \\ 0 & & & 0 & \lambda_n \end{pmatrix} P^{-1}$$

alors

$$\text{dét}(M) = \lambda_1 \ldots \lambda_n$$

Pour $n = 2$, on en déduit le critère suivant en un point critique A :

a) Si $\text{dét}(\text{Hess}_A(f)) > 0$ alors les valeurs propres de $\text{Hess}_A(f)$ sont toutes les deux strictement positives ou strictement négatives, et la fonction possède un minimum local ou un maximum local en A ;

b) Si $\text{dét}(\text{Hess}_A(f)) < 0$ alors les valeurs propres sont de signes opposés et donc la fonction f a un point selle en A.

On a donc obtenu une méthode pour rechercher les extrema locaux :

1. Chercher les points critiques ;
2. En un point critique A, cherche le signe des valeurs propres de $\text{Hess}_A(f)$.

Exercices corrigés

Dans ce qui suit, le texte en bleu est un commentaire explicatif qui ne fait pas partie de la solution elle-même. Commençons par la recherche d'une base pour un sous-espace vectoriel.

Énoncé 2.1

Soit
$$F = \{ (x,y,z) \in \mathbf{R}^3 \mid 5x + 3y + z = 0 \}.$$

1. Vérifier que F est un sous-espace vectoriel de $\mathbf{R}^3$.
2. Trouver une base de F et donner sa dimension.

Solution. — 1. L'application $f(x,y,z) \mapsto 5x + 3y + z$ est l'application linéaire de matrice $\begin{pmatrix} 5 & 3 & 1 \end{pmatrix}$ et $F = \mathrm{Ker}(f)$. donc F est bien un sous-espace vectoriel de $\mathbf{R}^3$.

2. Il faut comprendre qu'une base n'est pas unique. L'énoncé nous demande simplement d'en exhiber une. Soit $(x,y,z) \in F$, alors $5x + 3y + z = 0$ et donc $z = -5x - 3y$. Cela donne les égalités
$$\begin{aligned}(x,y,z) &= (x,y,-5x-3y)\\ &= x(1,0,-5) + y(0,1,-3).\end{aligned}$$
Posons $\vec{u}_1 = (1,0,-5)$ et $\vec{u}_2 = (0,1,-3)$. D'après ce qui précède, l'application
$$\begin{aligned} g : \mathbf{R}^2 &\longrightarrow F \\ (x,y) &\longmapsto x\vec{u}_1 + y\vec{u}_2 \end{aligned}$$
est surjective. Vérifions qu'elle est injective en considérant son noyau. Soit $(y,z) \in \mathbf{R}^2$, l'égalité $g(y,z) = (0,0,0)$ équivaut au système
$$\begin{cases} -3y & -5z & = & 0 \\ y & & = & 0 \\ & z & = & 0 \end{cases}$$
qui implique $(y,z) = (0,0)$. Donc l'application g est également injective.

Conclusion. $((1,0,-5),(0,1,-3))$ est une base de F qui est un plan vectoriel. □

Donnons maintenant un exemple de calcul de l'inverse d'une matrice.

Énoncé 2.2

On se donne la matrice
$$A=\begin{pmatrix}1&1&1\\1&2&3\\2&4&7\end{pmatrix}.$$

1. Décrire explicitement l'application linéaire φ associée à A.
2. Vérifier que φ est bijective et déterminer sa réciproque.
3. En déduire A^{-1}.
4. Vérifier le résultat à l'aide d'un produit matriciel.

Solution. — 1. On fait le produit matriciel
$$\begin{pmatrix}1&1&1\\1&2&3\\2&4&7\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+y+z\\x+2y+3z\\2x+4y+7z\end{pmatrix}.$$

Donc φ est donnée par
$$(x,y,z)\longmapsto(x+y+z,x+2y+3z,2x+4y+7z).$$

2. Soit (a,b,c) un triplet arbitraire de nombre réels. On veut montrer qu'il existe un unique $(x,y,z)\in\mathbf{R}^3$ tel que $\varphi(x,y,z)=(a,b,c)$. Mais l'égalité $\varphi(x,y,z)=(a,b,c)$ est vérifiée si et seulement si (x,y,z) est solution du système où x, y et z sont des inconnues et a, b et c des constantes
$$\left\{\begin{array}{rcrcrcrcrcr} x&+&y&+&z&=&a&&&&\\ x&+&2y&+&3z&=&&&b&&\\ 2x&+&4y&+&7z&=&&&&&c.\end{array}\right.$$

On va résoudre ce système à l'aide de la méthode du pivot de Gauss. Les opérations $L_2\leftarrow L_2-L_1$ et $L_3\leftarrow L_3-2L_1$ permettent d'éliminer la variable x des deux dernières équations, donnant le système équivalent
$$\left\{\begin{array}{rcrcrcrcrcr} x&+&y&+&z&=&a&&&&\\ &&y&+&2z&=&-a&+&b&&\\ &&2y&+&5z&=&-2a&&&+&c.\end{array}\right.$$

L'opération $L_3 \leftarrow L_3 - 2L_2$ permet alors déliminer la variable y de la dernière équation donnant le système

$$\left\{\begin{array}{rcrcrcrcrcr} x & + & y & + & z & = & a & & & & \\ & & y & + & 2z & = & -a & + & b & & \\ & & & & z & = & & & -2b & + & c. \end{array}\right.$$

On a maintenant un système triangulaire Les opérations $L_1 \leftarrow L_1 - L_3$ et $L_2 \leftarrow L_2 - 2L_3$ permettent alors d'éliminer la variable z des deux premières équations

$$\left\{\begin{array}{rcrcrcrcrcr} x & + & y & & & = & a & + & 2b & - & c \\ & & y & & & = & -a & + & 5b & - & 2c \\ & & & & z & = & & & -2b & + & c. \end{array}\right.$$

Enfin l'opération $L_1 \leftarrow L_1 - L_2$ permet d'éliminer y de la première équation

$$\left\{\begin{array}{rcrcrcrcrcr} x & & & & & = & 2a & - & 3b & + & c \\ & & y & & & = & -a & + & 5b & - & 2c \\ & & & & z & = & & & -2b & + & c. \end{array}\right.$$

Donc l'application φ est bijective, de réciproque

$$\varphi^{-1} : (a,b,c) \longmapsto (2a - 3b + c, -a + 5b - 2c, -2b + c).$$

3. En considérant les images par φ^{-1} de $(1,0,0)$, $(0,1,0)$ et $(0,0,1)$, on obtient que

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

4. On vérifie avec le calcul

$$A^{-1}A = \begin{pmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Donc $\left(\begin{smallmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{smallmatrix}\right)$ est bien l'inverse de A.

□

***Remarque 2.0.1.* —** Pour calculer l'inverse on a juste besoin de faire les opérations sur les coefficients du système. Cela revient à manipuler ce qu'on appelle parfois des *matrices augmentées*. On démarre le processus en écrivant côte à côte les matrices A et I_n. La méthode de Gauss permet, à l'aide d'opérations sur les lignes de la matrices à n lignes et $2n$ colonnes de se ramener à une matrice triangulaire à gauche, puis, si la matrice A est inversible, à la matrice I_n à gauche.

La matrice obtenue à droite est alors l'inverse de A. Appliquons cette méthode sur l'exemple de l'exercice :

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 4 & 7 & 0 & 0 & 1 \end{array}\right).$$

Les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 2L_1$ donnent

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 2 & 5 & -2 & 0 & 1 \end{array}\right).$$

L'opération $L_3 \leftarrow L_3 - 2L_2$ donne

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right).$$

On effectue alors $L_1 \leftarrow L_1 - L_3$ et $L_2 \leftarrow L_2 - 2L_3$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & 0 & -1 & 5 & -2 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right)$$

et on termine avec $L_1 \leftarrow L_1 - L_2$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -3 & 1 \\ 0 & 1 & 0 & -1 & 5 & -2 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right)$$

ce qui donne $A^{-1} = \begin{pmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix}$. Le lecteur pourra noter que les opérations faites sont exactement celles de l'exercice et que la matrice augmentée à chaque étape correspondent aux coefficients à gauche et à droite des systèmes d'équations.

Passons maintenant à un exercice de diagonalisation.

Énoncé 2.3

Soit $\varphi : \mathbf{R}^3 \to \mathbf{R}^3$ l'endomorphisme de matrice

$$A = \begin{pmatrix} -1 & 1 & 0 \\ -2 & -2 & 2 \\ -4 & -8 & 6 \end{pmatrix}$$

1. Trouver les valeurs propres de φ et en déduire que φ est diagonalisable.
2. Trouver une base $\boldsymbol{f} = (\vec{f}_1, \vec{f}_2, \vec{f}_3)$ formée de vecteurs propres de φ.
3. Déterminer la matrice de passage de la base usuelle $\boldsymbol{e}$ à la base $\boldsymbol{f}$ et calculer son inverse.
4. Calculer les coefficients de la matrice A^n pour tout $n \in \mathbf{N}$.

Solution. — 1. On utilise la règle de Sarrus pour calculer le polynôme caractéristique :

$$\begin{aligned} \text{dét}(XI_3 - A) &= \text{dét}\begin{pmatrix} X+1 & -1 & 0 \\ 2 & X+2 & -2 \\ 4 & 8 & X-6 \end{pmatrix} \\ &= (X+1)(X+2)(X-6) + 8 + 16(X+1) + 2(X-6) \\ &= X^3 - 3X^2 + 2X \\ &= X(X-1)(X-2) \end{aligned}$$

Donc φ admet 3 valeurs propres distinctes 0, 1 et 2. Il en résulte que φ est diagonalisable.

2. L'application φ est donnée par :

$$(x, y, z) \longmapsto (-x + y, -2x - 2y + 2z, -4x - 8y + 6z)$$

Valeur propre 0 : on veut trouver un vecteur non nul dans $\text{Ker}(\varphi)$. Mais un triplet (x, y, z) appartient à ce noyau si et seulement s'il est solution du système

$$\begin{cases} -X + Y = 0 \\ -2X - 2Y + 2Z = 0 \\ -4X - 8Y + 6Z = 0 \end{cases}$$

qu'on résoud à l'aide de la méthode du pivot de Gauss. Les opérations $L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 - 4L_1$ qui permettent d'éliminer la variable X des deux dernières équations donnent le système équivalent

$$\left\{\begin{array}{rcrcrcl} -X & + & Y & & & = & 0 \\ & - & 4Y & + & 2Z & = & 0 \\ & - & 12Y & + & 6Z & = & 0 \end{array}\right.$$

L'opération $L_3 \leftarrow L_3 - 3L_2$ permet de retirer la troisième ligne, puis $L_2 \leftarrow \frac{1}{2}L_2$ donne le système équivalent :

$$\left\{\begin{array}{rcrcrcl} -X & + & Y & & & = & 0 \\ & - & 2Y & + & Z & = & 0 \end{array}\right.$$

C'est à dire

$$\left\{\begin{array}{rcl} Y & = & X \\ Z & = & 2X \end{array}\right.$$

En prenant $X = 1$ on obtient une solution non nulle $(1, 1, 2)$. Par conséquent on peut poser $\vec{f}_1 = (1, 1, 2) \in \mathrm{Ker}(\varphi)$. On notera qu'ici il suffit de donner un vecteur non nul dans $\mathrm{Ker}(\varphi)$. Tout vecteur $\lambda\vec{f}_1$ avec $\lambda \in \mathbf{R}^*$ conviendrait donc également. Il n'y a pas unicité de la réponse. Néanmoins, nous avons choisi une réponse de longueur minimale parmi celles avec des coordonnées entières.

Valeur propre 1 : un vecteur $\vec{u} = (x, y, z)$ vérifie $\varphi(\vec{u}) = \vec{u}$ si et seulement s'il est solution du système

$$\left\{\begin{array}{rcrcrcl} -X & + & Y & & & = & X \\ -2X & - & 2Y & + & 2Z & = & Y \\ -4X & - & 8Y & + & 6Z & = & Z \end{array}\right.$$

c'est-à-dire du système

$$\left\{\begin{array}{rcrcrcl} -2X & + & Y & & & = & 0 \\ -2X & - & 3Y & + & 2Z & = & 0 \\ -4X & - & 8Y & + & 5Z & = & 0 \end{array}\right.$$

Les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 2L_1$ donnent le système équivalent

$$\left\{\begin{array}{rcrcrcl} -2X & + & Y & & & = & 0 \\ & - & 4Y & + & 2Z & = & 0 \\ & - & 10Y & + & 5Z & = & 0 \end{array}\right.$$

L'opération $L_3 \leftarrow L_3 - \frac{5}{2}L_2$ permet d'éliminer la troisième ligne. Puis l'opération $L_2 \leftarrow \frac{1}{2}L_2$ donne

$$\left\{\begin{array}{rcrcrcl} -2X & + & Y & & & = & 0 \\ & - & 2Y & + & Z & = & 0 \end{array}\right.$$

et $\vec{f_2} = (1, 2, 4)$ est une solution non nulle de ce système et donc un vecteur propre pour la valeur propre 1.

Valeur propre 2 : un vecteur $\vec{u} = (x, y, z)$ vérifie $\varphi(\vec{u}) = 2\vec{u}$ si et seulement s'il est solution du système

$$\begin{cases} -X + Y = 2X \\ -2X - 2Y + 2Z = 2Y \\ -4X - 8Y + 6Z = 2Z \end{cases}$$

c'est-à-dire du système

$$\begin{cases} -3X + Y = 0 \\ -2X - 4Y + 2Z = 0 \\ -4X - 8Y + 4Z = 0 \end{cases}$$

L'opérations $L_3 \leftarrow L_3 - 2L_2$ permet d'éliminer la troisième ligne, si bien que le système équivaut à

$$\begin{cases} -3X + Y = 0 \\ -X - 2Y + Z = 0 \end{cases}$$

dont $\vec{f_3} = (1, 3, 7)$ est une solution non nulle.

3. On obtient la matrice de passage

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix}$$

qui est exactement celle de l'exercice précédent, ce qui nous permet de reprendre les calculs faits. Avec la méthode de la matrice augmentée :

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 4 & 7 & 0 & 0 & 1 \end{array}\right).$$

Les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - 2L_1$ donnent

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 2 & 5 & -2 & 0 & 1 \end{array}\right).$$

L'opération $L_3 \leftarrow L_3 - 2L_2$ donne

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right).$$

On effectue alors $L_1 \leftarrow L_1 - L_3$ et $L_2 \leftarrow L_2 - 2L_3$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & 0 & -1 & 5 & -2 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right)$$

et on termine avec $L_1 \leftarrow L_1 - L_2$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -3 & 1 \\ 0 & 1 & 0 & -1 & 5 & -2 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right)$$

ce qui donne $P^{-1} = \begin{pmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix}$

4. Pour $n \geqslant 1$, on obtient les égalités

$$\begin{aligned} A^n &= P \begin{pmatrix} 0^n & 0 & 0 \\ 0 & 1^n & 0 \\ 0 & 0 & 2^n \end{pmatrix} P^{-1} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2^n \\ 0 & 2 & 3 \times 2^n \\ 0 & 4 & 7 \times 2^n \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ -1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 5 - 2 \times 2^n & -2 + 2^n \\ -2 & 10 - 6 \times 2^n & -4 + 3 \times 2^n \\ -4 & 20 - 14 \times 2^n & -8 + 7 \times 2^n \end{pmatrix} \end{aligned}$$

À titre de vérification pour $n = 1$, on obtient $\begin{pmatrix} -1 & 1 & 0 \\ -2 & -2 & 2 \\ -4 & -8 & 6 \end{pmatrix} = A$.

$\square$

Fiche de révision

2.1. Produit

Définition 2.1.1

Soient m, n, p trois entiers strictement positifs. Soit $A = (a_{i,j})$ une matrice de $\mathscr{M}_{m,n}(\mathbf{R})$ et soit $B = (b_{j,k})$ une matrice de $\mathscr{M}_{n,p}(\mathbf{R})$. On appelle produit matriciel de A par B la matrice $C \in \mathscr{M}_{m,p}(\mathbf{R})$ dont le terme général $c_{i,k}$ est défini, pour tout $i = 1,\ldots,m$ et pour tout $k \in 1,\ldots,p$ par :

$$c_{i,k} = \sum_{j=1}^{n} a_{i,j}\, b_{j,k}\,.$$

2.2. Applications linéaires

Définition 2.2.1

Une application $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ est dite *linéaire* si elle vérifie les deux conditions suivantes :

(i) *Compatibilité avec l'addition :* Pour tous vecteurs $\vec{u}, \vec{v} \in \mathbf{R}^n$,

$$\varphi(\vec{u}+\vec{v}) = \varphi(\vec{u}) + \varphi(\vec{v});$$

(ii) *Compatibilité avec la multiplication par un nombre réel :* Pour tout $\lambda \in \mathbf{R}$, pour vecteur $\vec{u} \in \mathbf{R}^n$,

$$\varphi(\lambda\vec{u}) = \lambda\varphi(\vec{u}).$$

Proposition 2.2.2

Soient $\boldsymbol{e} = (\vec{e}_1,\ldots,\vec{e}_n)$ une base de $\mathbf{R}^n$, $\boldsymbol{f} = (\vec{f}_1,\ldots,\vec{f}_m)$ une base de $\mathbf{R}^m$ et $\boldsymbol{g} = (\vec{g}_1,\ldots,\vec{g}_p)$ une base de $\mathbf{R}^p$ et soient $\varphi : \mathbf{R}^n \to \mathbf{R}^m$ et $g : \mathbf{R}^m \to \mathbf{R}^p$ des applications linéaires, alors

$$\mathrm{Mat}_{\boldsymbol{e},\boldsymbol{g}}(\psi\circ\varphi) = \mathrm{Mat}_{\boldsymbol{f},\boldsymbol{g}}(\psi)\,\mathrm{Mat}_{\boldsymbol{e},\boldsymbol{f}}(\varphi).$$

2.3. Sous-espaces

Définition 2.3.1

Un *sous-espace vectoriel* de $\mathbf{R}^n$ est une partie E de $\mathbf{R}^n$ qui satisfait les trois conditions suivantes :

(i) La partie E contient le vecteur nul : $\overrightarrow{0} \in E$;

(ii) La partie E est *stable par addition* :

$$\forall \vec{u}, \vec{v} \in E, \quad \vec{u} + \vec{v} \in E;$$

(iii) La partie E est *stable par multiplication par un nombre réel* :

$$\forall \lambda \in \mathbf{R}, \forall \vec{u} \in E, \quad \lambda \vec{u} \in E.$$

Proposition 2.3.2

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application *linéaire*. Alors

a) L'ensemble

$$f^{-1}(\{0\}) = \{\vec{u} \in \mathbf{R}^n \mid f(\vec{u}) = \overrightarrow{0}\}$$

est un sous-espace vectoriel de $\mathbf{R}^n$ qu'on appelle le *noyau de* f et qu'on note $\mathrm{Ker}(f)$.

b) L'ensemble des valeurs de f c'est-à-dire des vecteurs de $\mathbf{R}^m$ de la forme $f(\vec{u})$ pour un $\vec{u} \in \mathbf{R}^n$ (avec des notations ensemblistes

$$f(\mathbf{R}^m) = \{\vec{v} \in \mathbf{R}^m \mid \exists \vec{u} \in \mathbf{R}^n, \vec{v} = f(\vec{u})\}$$

est un sous-espace vectoriel de $\mathbf{R}^m$ qu'on appelle l'*image de* f et qu'on note $\mathrm{Im}(f)$.

2.4. Dimension

Proposition 2.4.1

Soient F_1 et F_2 des sous-espaces vectoriels de $\mathbf{R}^m$ tels que $F_1 \subset F_2$ (F_1 est contenu dans F_2) alors

a) On a l'inégalité $\dim(F_1) \leqslant \dim(F_2)$;

b) L'égalité de parties $F_1 = F_2$ équivaut à l'égalité $\dim(F_1) = \dim(F_2)$.

Théorème du rang 2.4.2

Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application linéaire alors
$$n = \dim(\operatorname{Ker}(f)) + \dim(\operatorname{Im}(f)).$$

2.5. Application

Théorème 2.5.1

Soit f une application en n variables à valeurs réelles de classe $\mathscr{C}^2$, alors
$$f(M) = f(A) + \mathrm{d}f_A\left(\overrightarrow{AM}\right) + \frac{1}{2}\mathrm{d}^2 f_A\left(\overrightarrow{AM}, \overrightarrow{AM}\right) + o(AM^2).$$
Cette formule s'appelle le *développement de* Taylor *à l'ordre* 2 *de f en A*.

Entraînement

2.1. Exercices

2.1.1. *Rappels sur le pivot de Gauss*

Exercice 2.1. Résoudre les systèmes d'équations linéaires suivants :

$$\begin{cases} X+Y+Z=3 \\ \phantom{X+{}}Y+Z=2 \\ \phantom{X+Y+{}}Z=1 \end{cases} \qquad \begin{cases} X+Y+Z=3 \\ \phantom{X+{}}Y+Z=2 \\ X+Y\phantom{{}+Z}=2 \end{cases} \qquad \begin{cases} 2X+3Y=4 \\ X-2Z=5 \end{cases}$$

Exercice 2.2. Pour chacun des systèmes suivants, tracer la droite de $\mathbf{R}^2$ correspondant à chaque équation $ax+by=c$ du système et trouver graphiquement l'ensemble des solutions. Vérifier le résultat en utilisant le pivot de Gauss.

$$\begin{cases} X+Y=1 \\ X-Y=0 \end{cases} \qquad \begin{cases} X-Y=2 \\ Y-X=1 \end{cases} \qquad \begin{cases} X-3Y=-3 \\ X-Y=1 \\ X+Y=5 \end{cases}$$

Exercice 2.3. Résoudre par la méthode du pivot de Gauss les systèmes linéaires suivants :

$$\begin{cases} 3X+Y=2 \\ X+2Y=1 \end{cases} \qquad \begin{cases} 2X+3Y=1 \\ 5X+7Y=3 \end{cases}$$

$$\begin{cases} 3X+Y=2 \\ 6X+2Y=1 \end{cases} \qquad \begin{cases} 2X+4Y=10 \\ 3X+6Y=15 \end{cases}$$

Exercice 2.4. Résoudre par la méthode du pivot de Gauss les systèmes linéaires suivants :

$$\begin{cases} X+Y+Z=1 \\ X+2Y+2Z=0 \\ Y+4Z=-4 \end{cases} \qquad \begin{cases} X+Y+Z=2 \\ X+Y+2Z=0 \\ X+2Y-Z=0 \end{cases}$$

$$\begin{cases} 2X+Y-3Z=5 \\ 3X-2Y+2Z=5 \\ 5X-3Y-Z=16 \end{cases} \qquad \begin{cases} X+Y+Z+T=2 \\ X+Y+2Z+2T=0 \\ X+2Y-Z-T=1 \\ Z-T=0 \end{cases}$$

Exercice 2.5. 1. Déterminer suivant les valeurs du paramètre a les solutions des systèmes suivants :

$$\begin{cases} X+aY=2 \\ aX+Y=2 \end{cases} \qquad \begin{cases} X-2Y=2 \\ X-aY=a \end{cases} \qquad \begin{cases} X+Y=3 \\ aX+Y=a. \end{cases}$$

2. Déterminer suivant les valeurs des paramètres réels a et b les solutions des systèmes linéaires suivants :
$$\begin{cases} X - aY = 2 \\ aX + Y = b \end{cases} \qquad \begin{cases} aX + 8Y = b \\ X - bY = a \end{cases} \qquad \begin{cases} aX + bY = 1 \\ bX + aY = 1 \end{cases}$$
3. Interprétez les résultats des questions précédentes en termes d'intersections de droites dans le plan.

Exercice 2.6. 1. Résoudre, en appliquant soigneusement la méthode du pivot de Gauss et en indiquant chaque opération effectuée, le système d'équations linéaires
$$\begin{cases} X + Y + Z + T = 5 \\ 2X + 3Y - 2Z - 2T = 4 \\ X + 3Y - 6Z + 3T = 4 \\ X - Y + Z - T = -1. \end{cases}$$
2. Vérifier que la ou les solutions obtenues vérifient le système initial.

Exercice 2.7. Equilibrer les reactions chimiques suivantes en interpretant la conservation des divers éléments comme une condition linéaire sur les quantités de réactif :

1. $NO_2 + H_2O \longrightarrow HNO_3 + NO$;
2. $Fe_7S_8 + O_2 \longrightarrow Fe_3O_4 + SO_2$;
3. $C_2H_6 \longrightarrow C_2H_4 + H_2$.

2.1.2. *Calcul matriciel*

Exercice 2.8. On considère la matrice suivante.
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

1. Calculer A et A^2. Exprimer simplement A^2 en termes de A.
2. Pour un entier $k \geqslant 1$, démontrer par récurrence que $A^k = 3^{k-1}A$.
3. En déduire l'expression de A^k en fonction de k.

Exercice 2.9. Pour chacune des application linéaires suivantes de $\mathbf{R}^n$ dans $\mathbf{R}^m$, écrire la matrice qui lui correspond.
$$f(x,y,z) = x + 2y + 3z \qquad f(x) = (x, -x, 2x)$$
$$f(x,y) = (x+y, x-y) \qquad f(x,y,z) = (x+2y+3z, x+y+z, x-y-z)$$
$$f(x,y,z) = (x-2y, 3y) \qquad f(x,y,z,t) = (x+y-2z+t, x+y+t)$$

Exercice 2.10. Dans cet exercice, on note A la matrice

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

On note B la matrice $A - 2I_3$.

1. Calculer les matrices A^2 et A^3.
2. (a) Calculer la matrice B.
   (b) Calculer les matrices B^2 et B^3.
   (c) Calculer les produits de matrices $(2I_3)B$ et $B(2I_3)$, comparer les résultats obtenus.
3. Dans cette question, on pourra utiliser, sans le prouver, que si M et $N \in \mathscr{M}_3(\mathbf{R})$ sont des matrices telles que $MN = NM$, alors pour tout entier n, la formule du binôme s'applique :

$$(M+N)^n = \sum_{k=0}^{n} \binom{n}{k} M^k N^{n-k}$$

   où $\binom{n}{k}$ désigne ici le coefficient binomial $\frac{n!}{k!(n-k)!}$.

   En utilisant cette formule et la question précédente calculer les neuf coefficients de la matrice A^n pour tout entier $n \in \mathbf{N}$.
4. Vérifier pour les trois entiers $n = 1$, 2 et 3 que l'expression obtenue dans la question précédente pour A^n redonne bien les matrices A, A^2 et A^3 obtenues auparavant.

Exercice 2.11. On considère la matrice suivante.

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

On note φ l'endomorphisme de $\mathbf{R}^3$ défini par la matrice A. On pose $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ et $\vec{e}_3 = (0,0,1)$.

1. Démontrer que $\varphi \circ \varphi(\vec{e}_1) = \varphi(\vec{e}_2) = 0$. Démontrer que $\varphi \circ \varphi(\vec{e}_3) = \varphi(\vec{e}_3)$.
2. En déduire A^2. Vérifier en effectuant le produit matriciel.
3. Démontrer que $A^3 = A^2$ sans effectuer le produit matriciel, puis vérifier en l'effectuant.
4. Donner une base de $\mathrm{Ker}(\varphi)$ et une base de $\mathrm{Im}(\varphi)$

Exercice 2.12. On considère la matrice suivante.

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

On note φ l'endomorphisme de $\mathbf{R}^3$ défini par la matrice A. On pose $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ et $\vec{e}_3 = (0,0,1)$.

1. Pour $i = 1,2,3$, déterminer $\varphi \circ \varphi(\vec{e}_i)$, puis $\varphi \circ \varphi \circ \varphi(\vec{e}_i)$.
2. En déduire que $A^2 = A^{-1}$. Vérifier en calculant le produit matriciel.

Exercice 2.13. On considère la matrice suivante.

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

On note φ l'endomorphisme de $\mathbf{R}^3$ défini par la matrice A. On pose $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ et $\vec{e}_3 = (0,0,1)$.

1. Pour $i = 1,2,3$, déterminer $\varphi \circ \varphi(\vec{e}_i)$, puis $\varphi \circ \varphi \circ \varphi(\vec{e}_i)$.
2. En déduire A^2 et A^3.
3. Pour $k \in \mathbf{N}^*$, donner une expression de $(I_3 + A)^k$ en fonction de k. Vérifier votre expression pour $k = 3$ en effectuant le produit matriciel.
4. Reprendre la question précédente pour $(I_3 - A)^k$, puis pour $(3I_3 - 2A)^k$.

2.1.3. *Algèbre linéaire*

Exercice 2.14. Les parties suivantes sont-elles des sous-espaces vectoriels de $\mathbf{R}^2$? (justifier la réponse)

1. $A = \{(x,y) \in \mathbf{R}^2 \mid x \leqslant y\}$
2. $B = \{(x,y) \in \mathbf{R}^2 \mid xy = 0\}$
3. $C = \{(x,y) \in \mathbf{R}^2 \mid x = y\}$
4. $D = \{(x,y) \in \mathbf{R}^2 \mid x + y = 1\}$
5. $E = \{(x,y) \in \mathbf{R}^2 \mid x^2 - y^2 = 0\}$
6. $F = \{(x,y) \in \mathbf{R}^2 \mid x^2 + y^2 = 0\}$

Exercice 2.15. On considère les parties suivantes de $\mathbf{R}^2$:

$$A = \{\,(x,y) \in \mathbf{R}^2 \mid y - 2x = 0\,\} \quad \text{et} \quad B = \{\,(x,y) \in \mathbf{R}^2 \mid y - x^2 = 0\,\}.$$

1. Dessiner, sur deux dessins différents, les parties A et B.
2. Déterminer, en justifiant vos réponses, si A ou B sont des sous-espaces vectoriels de $\mathbf{R}^2$.

Exercice 2.16. On considère la partie F de $\mathbf{R}^3$ définie par

$$F = \{ (x, y, z) \in \mathbf{R}^3 \mid x + 2y - z = 0 \}$$

1. Démontrer, en utilisant la définition, que l'ensemble F est un sous-espace vectoriel de $\mathbf{R}^3$.
2. Utiliser la notion de noyau pour donner une autre justification du fait que F est un sous-espace vectoriel de $\mathbf{R}^3$.

Exercice 2.17. On considère l'application φ de $\mathbf{R}^3$ dans $\mathbf{R}$ définie par

$$(x, y, z) \mapsto x - y + z.$$

1. Démontrer que φ est linéaire.
2. Démontrer que φ est surjective.
3. Soit $P = \mathrm{Ker}(\varphi)$. Quelle est la dimension de P ?
4. Soit $\vec{u} = (1, 0, 0)$ et soit $D = \mathrm{Vect}(\vec{u})$. Que vaut $\varphi(\vec{u})$?
5. Soit $\vec{v} \in \mathbf{R}^3$.
   (a) Si $\vec{v} = \vec{v}_1 + \vec{v}_2$ avec $\vec{v}_1 \in D$ et $\vec{v}_2 \in P$, justifier que $\varphi(v) = \varphi(v_1)$ puis que $\vec{v}_1 = \frac{\varphi(v)}{\varphi(u)}\vec{u}$.
   (b) Prouver qu'il existe un unique couple $(\vec{v}_1, \vec{v}_2)$ avec $\vec{v}_1 \in D$ et $\vec{v}_2 \in P$ tel que $\vec{v} = \vec{v}_1 + \vec{v}_2$.

Exercice 2.18. Certaines des transformations qui servent dans les logiciels de graphisme ou de gestion d'images sont des endomorphimes linéaires du plan $\mathbf{R}^2$. Par exemple, la rotation d'un quart de tour à droite est donnée par l'application linéaire $(x, y) \longmapsto (y, -x)$. Pour chaque application linéaire ci-dessous, décrire en termes simples la transformation géométrique correspondante.

$$(x, y) \longmapsto (-x, y) \qquad (x, y) \longmapsto (x/2, y/2)$$
$$(x, y) \longmapsto (2x, y) \qquad (x, y) \longmapsto (x + y, y).$$

Exercice 2.19. Parmi les applications de $\mathbf{R}^2$ dans $\mathbf{R}^2$ définies par les relations qui suivent, déterminer lesquelles sont linéaires.

$$f_1(x, y) = (x + y, x - y) \quad f_2(x, y) = (|x| + |y|, 2) \quad f_3(x, y) = (x, -y)$$
$$f_4(x, y) = (xy, y) \quad f_5(x, y) = (x - y + 1, x) \quad f_6(x, y) = \left(\frac{1}{1 + x^2}, \frac{1}{1 + y^2}\right)$$

pour tous $x, y \in \mathbf{R}$.

Exercice 2.20. Pour chacune des applications ci-dessous, vérifier qu'elle est linéaire et determiner son noyau, son image ainsi qu'une base de chacun de ces sous-espaces.

$$f_1 : \begin{array}{ccc} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ (x,y) & \longmapsto & (2x+3y, 3x-y) \end{array} \qquad f_2 : \begin{array}{ccc} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ (x,y) & \longmapsto & (2x+3y, -4x-6y) \end{array}$$

$$f_3 : \begin{array}{ccc} \mathbf{R}^3 & \longrightarrow & \mathbf{R} \\ (x,y,z) & \longmapsto & (2x+y+z) \end{array} \qquad f_4 : \begin{array}{ccc} \mathbf{R}^3 & \longrightarrow & \mathbf{R}^2 \\ (x,y,z) & \longmapsto & (x+y+z, 2x+y-z) \end{array}$$

Exercice 2.21. Pour chacune des familles suivantes de vecteurs de $\mathbf{R}^n$ ($n = 2, 3, 4$ suivant le cas) déterminer si elle forme une base de cet espace. Si c'est le cas, trouver les coordonnées d'un n-uplet arbitraire $(x_1, x_2, \dots, x_n)$ dans cette base.

1. $((2,3),(1,1))$;
2. $((2,0),(1,1),(1,2))$;
3. $((1,1,0),(1,0,1),(0,1,1))$;
4. $((-2,4,0),(0,1,1),(1,1,1))$;
5. $((-1,5,4),(0,1,1),(1,-4,-3))$;
6. $((-1,5,4,0),(0,1,1,2),(1,-4,-3,-1))$;
7. $((-1,5,4,0),(0,1,1,2),(1,-4,-3,-1),(0,2,2,1))$;

Exercice 2.22. Donner une condition nécessaire et suffisante portant sur les paramètres réels a et b pour que chacune des familles suivantes soient des bases de $\mathbf{R}^3$.

1. $((1,1,1),(0,a,1),(0,0,b))$,
2. $((1,0,1),(a,b,1),(b,a,1))$,
3. $((1,a,b),(a,1,a),(b,b,1))$,
4. $((a,a,b),(a,b,a),(b,a,a))$,
5. $((0,a,b),(a,0,b),(a,b,0))$.

Exercice 2.23. Pour les applications linéaires suivantes de $\mathbf{R}^n$ dans $\mathbf{R}^m$, préciser les entiers n et m, donner la matrice de f puis donner une base de $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$.

1. $f_1 : (x,y) \mapsto (x+3y, 2x-y, -x+5y)$;
2. $f_2 : (x,y,z) \mapsto (x+y+2z, x-y)$;
3. $f_3 : (x,y,z) \mapsto (-x-2y+z, 2x-y, x-3y+z)$;
4. $f_4 : (x,y,z) \mapsto (x+y+z, 2x-y, z)$.

Exercice 2.24. Calculer le déterminant des matrices suivantes. Sont-elles inversibles ? Inverser celles qui le sont.

$$A=\begin{pmatrix}2 & -2\\ 3 & -5\end{pmatrix},\quad B=\begin{pmatrix}4 & 1\\ 12 & 3\end{pmatrix},\quad C=\begin{pmatrix}2 & 1\\ -5 & -1\end{pmatrix}$$

Exercice 2.25. Soit $\vec{u}\in\mathbf{R}^2$ un vecteur de norme 1 et soit D la droite vectorielle engendrée par $\vec{u}$.

1. Écrire la matrice de la projection orthogonale sur D.
2. Écrire la matrice de la symétrie orthogonale par rapport à D. Calculer le déterminant de cette matrice.
3. Soit s_1 et s_2 des symétries orthogonales par rapport à des droites vectorielles dans $\mathbf{R}^2$. En utilisant le calcul matriciel justifier que la composée $s_1\circ s_2$ est une rotation.

Exercice 2.26. On pose

$$\vec{u}=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right);$$

on note D la droite vectorielle engendrée par $\vec{u}$ et P le plan vectoriel d'équation $X+Y+Z=0$.

1. Trouver deux vecteurs $\vec{f_1}$ et $\vec{f_2}$ orthogonaux et de norme 1 dans P.
2. (a) Trouver une base orthonormée $(\vec{e}_1,\vec{e}_2,\vec{e}_3)$ de $\mathbf{R}^3$, avec $\vec{e}_1\in D$.
   (b) Donner la matrice de changement de base Q de la base usuelle de $\mathbf{R}^3$ vers la base $(\vec{e}_1,\vec{e}_2,\vec{e}_3)$.
   (c) Calculer son déterminant.
   (d) Quitte à échanger $\vec{e}_2$ et $\vec{e}_3$, expliquer comment on peut se ramener au cas où $\text{dét}(Q)=1$.
   (e) Que vaut le produit ${}^tQQ$?
   (f) La matrice Q est-elle inversible ? Trouver son inverse.
3. Écrire dans la base $(\vec{e}_1,\vec{e}_2,\vec{e}_3)$ la matrice de la rotation r_θ d'axe la droite D orientée par $\vec{u}$ et d'angle $2\pi/3$.
4. Donner la matrice M de r_θ dans la base usuelle de $\mathbf{R}^3$.
5. Que vaut M^3 ?
6. Calculer M^k pour tout entier $k\in\mathbf{Z}$.

Exercice 2.27. Soit

$$M=\begin{pmatrix}\sqrt{3}-1 & 0 & -2\sqrt{3}\\ 0 & 2 & 0\\ \sqrt{3} & 0 & -1-\sqrt{3}\end{pmatrix}.$$

1. Calculer dét(M).
2. Calculer M^2 et M^3.
3. En déduire l'inverse de M.
4. Calculer M^k pour tout $k \in \mathbf{N}$. On pourra écrire $k = 3m + r$ avec $m \in \mathbf{N}$ et $r = 0, 1$ ou 2.
5. Pour tout $k \in \mathbf{N}$, calculer dét(M^k) de deux manières différentes.

Exercice 2.28. Calculer le déterminant des matrices suivantes. Sont-elles inversibles ? Inverser celles qui le sont. Calculer noyau et image de l'application linéaire associée pour celles qui ne sont pas inversibles.

$$\begin{pmatrix} 2 & 3 & 6 \\ 4 & -1 & 1 \\ 6 & 2 & 7 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 3 & 6 \\ 4 & -1 & -1 \\ 6 & 2 & 1 \end{pmatrix}$$

Exercice 2.29. Soit $f : \mathbf{R}^3 \to \mathbf{R}^3$ l'application linéaire dont la matrice dans la base usuelle de $\mathbf{R}^3$ est

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

1. Montrer que l'application f possède exactement deux valeurs propres et que l'une d'entre elle est 3.
2. Donner une base $\boldsymbol{e}$ de $\mathbf{R}^3$ formée de vecteurs propres pour f.
3. Quelle est la matrice de f dans la base $\boldsymbol{e}$?

Exercice 2.30. Pour chaque matrice suivante, Donner l'application linéaire f correspondante. Déterminer si l'application linéaire est diagonalisable ? Si c'est le cas, donner une base dans laquelle la matrice de l'application linéaire est diagonale.

$$\begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -5 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{3}{2} & \frac{3}{2} & 1 \\ -\frac{3}{2} & -\frac{1}{2} & 0 \end{pmatrix}.$$

2.1.4. *Applications aux fonctions en plusieurs variables*

Exercice 2.31. On considère la fonction de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par

$$f : (x, y) \mapsto (x^2 + y^2)^2 - 2x^2 + 2y^2.$$

1. Quel est le domaine de définition de f ?

2. Pour tout $A = (x, y) \in \mathscr{D}_f$ calculer $\mathrm{grad}_A(f)$.
3. Prouver que la fonction f possède exactement trois points critiques que l'on déterminera.
4. Calculer la matrice hessienne de A en un point $A = (x, y)$.
5. Pour chacun des trois points trouvés à la question 3 :
   (a) Déterminer les valeurs propres de la matrice hessienne ;
   (b) En déduire si la fonction f a un minimum, un maximum ou un point selle en ce point.

Exercice 2.32. On considère la fonction de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par

$$f : (x, y) \mapsto \exp(-x^2 - y^2).$$

1. Quel est le domaine de définition de f ?
2. Écrire la fonction comme la composée de deux applications et en déduire directement que f admet un maximum en un unique point que l'on précisera. Donner la valeur de ce maximum.
3. Pour tout $A = (x, y) \in \mathscr{D}_f$, calculer $\mathrm{grad}_A(f)$.
4. Trouver l'ensemble des points critiques de f.
5. Calculer la matrice hessienne de A en un point $A = (x, y)$.
6. Déterminer les valeurs propres de la matrice hessienne de f en chacun de ses points critiques. Quel est le lien avec la question 2 ?

Exercice 2.33. On considère la fonction de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par

$$f : (x, y) \mapsto xy - x + y.$$

1. Quel est le domaine de définition de f ?
2. Pour tout $A = (x, y) \in \mathscr{D}_f$, calculer $\mathrm{grad}_A(f)$.
3. Prouver que la fonction f possède exactement un point critique que l'on déterminera.
4. Calculer la matrice hessienne de A en un point $A = (x, y)$.
5. Déterminer les valeurs propres de la matrice hessienne en le point critique.
6. En déduire si la fonction f a un minimum, un maximum ou un point selle en ce point.

Annales

Énoncé partiel 2023

Les calculatrices, documents et objets connectables au réseau sont interdits.

La clarté des justifications des réponses données est un facteur important d'appréciation des copies. Prenez le temps de vérifier vos réponses.

Question de cours 1. On se donne des fonctions f et g en n variables réelles et à valeurs réelles que l'on suppose différentiables. On note A un point appartenant à l'intersection des domaines de définitions $\mathscr{D}_f$ et $\mathscr{D}_g$.

Donner les formules pour les gradients $\operatorname{grad}_A(f+g)$, $\operatorname{grad}_A(fg)$ et, dans le cas où $g(A) \neq 0$, pour $\operatorname{grad}_A(f/g)$.

Soit h une fonction de $\mathbf{R}$ dans $\mathbf{R}$ dérivable en $f(A)$; donner la formule pour le gradient $\operatorname{grad}_A(h \circ f)$.

Exercice. On considère la fonction g de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par $g : (x, y) \mapsto \frac{1}{1+x^2+y^2}$.

1. Déterminer le domaine de définition de g.
2. En écrivant g comme une application composée, justifier que g admet un maximum en $(0, 0)$ et calculer sa valeur. Ce maximum est-il global ?
3. Justifier que g est différentiable et calculer le gradient $\operatorname{grad}_A(g)$ pour tout $A = (x, y) \in \mathbf{R}^2$.
4. Déterminer pour quels points $A \in \mathbf{R}^2$ on a l'égalité $\operatorname{grad}_A(g) = (0, 0)$.
5. Quel lien y a-t-il entre la question 2 et la question 4 ?

Problème. On considère l'application f de $\mathbf{R}^2$ dans $\mathbf{R}$ donnée par $f : (x, y) \mapsto x^2 - y^2$.

1. Calculer les nombres $f(1, 1)$, $f(1, -1)$, $f(1, 0)$, $f(0, 1)$, $f(-1, 0)$ et $f(0, -1)$.
2. Représenter sur un unique dessin :
   (a) L'ensemble X_0 des couples $(x, y) \in \mathbf{R}^2$ tels que $f(x, y) = 0$;
   (b) L'ensemble X_+ des couples $(x, y) \in \mathbf{R}^2$ tels que $f(x, y) > 0$;

(c) L'ensemble X_- des couples $(x, y) \in \mathbf{R}^2$ tels que $f(x, y) < 0$;

3. Soit r un nombre réel strictement positif quelconque. On pose $A = (\frac{r}{2}, 0)$ et $B = (0, \frac{r}{2})$ et on note $O = (0, 0)$ l'origine de $\mathbf{R}^2$. On rappelle que $\mathrm{B}(O, r)$ désigne la boule ouverte de centre O et de rayon r.
   (a) Vérifier que $A \in \mathrm{B}(O, r)$ et que $f(A) > 0$.
   (b) Que peut-on dire d'analogue pour le point B ?
   (c) Prouver que la fonction f n'admet ni minimum local ni maximum local en O.
4. (a) Justifier que l'application f est différentiable et calculer le gradient $\mathrm{grad}_A(f)$ pour tout $A = (x, y) \in \mathbf{R}^2$.
   (b) Déterminer pour quels points $A \in \mathbf{R}^2$ on a l'égalité $\mathrm{grad}_A(f) = (0, 0)$. Dire pour chacun d'entre eux si f admet un minimum local ou un maximum local en ce point.
5. Soit c un nombre réel non nul. On considère la courbe paramétrée
$$g_c : t \longmapsto \left(\frac{1}{2}\left(t + \frac{c}{t}\right), \frac{1}{2}\left(t - \frac{c}{t}\right)\right)$$
   (a) Quel est le domaine de définition $\mathscr{D}_{g_c}$ de g_c (c'est-à-dire pour quels nombres $t \in \mathbf{R}$ le couple $g_c(t)$ est-il bien défini) ?
   (b) Donner une expression simple pour la fonction composée $f \circ g_c$.
   (c) Calculer la vitesse $g_c'(t)$ pour tout $t \in \mathscr{D}_{g_c}$.
   (d) Soit $t \in \mathscr{D}_{g_c}$. Calculer le produit scalaire $g_c'(t).\, \mathrm{grad}_{g_c(t)}(f)$.
   (e) Peut-on retrouver le résultat de la question précédente sans calculer $g_c'(t)$?
   (f) Soit $(x, y) \in \mathbf{R}^2$ tel que $x \neq y$ et $x \neq -y$. Trouver des nombres réels $c \in \mathbf{R}^*$ et $t \in \mathbf{R}^*$ tels que $(x, y) = g_c(t)$. (Indication : on pourra commencer par noter que si $(x, y) = g_c(t)$ alors $f(x, y) = f \circ g_c(t)$)
   (g) Si $c \neq 0$, que peut-on dire de la ligne de niveau $f^{-1}(\{c\})$ et de la fonction g_c ?
6. Dessiner l'allure des lignes de niveau $f^{-1}(\{c\})$ pour les nombres réels $c \in \{0, -1, 1, -4, 4\}$.

Corrigé partiel 2023

Question de cours 2. Sous les hypothèses de l'énoncé, on a les formules

$$\begin{aligned}
\operatorname{grad}_A(f+g) &= \operatorname{grad}_A(f) + \operatorname{grad}_A(g)\\
\operatorname{grad}_A(fg) &= f(A)\operatorname{grad}_A(g) + g(A)\operatorname{grad}_A(f)\\
\operatorname{grad}_A(f/g) &= \frac{1}{g(A)^2}\big(g(A)\operatorname{grad}_A(f) - f(A)\operatorname{grad}_A(g)\big)\\
\operatorname{grad}_A(h\circ f) &= h'(f(A))\operatorname{grad}_A(f).
\end{aligned}$$

Exercice.

1. Pour tout $(x,y)\in\mathbf{R}^2$, on a la minoration $1+x^2+y^2 \geqslant 1$ donc $1+x^2+y^2\neq 0$ et donc $\mathscr{D}_g = \mathbf{R}^2$.

2. Soit h la fonction d'une variable réelle donnée par $h : t\mapsto \frac{1}{1+t}$. Elle est définie sur $\mathscr{D}_h = \mathbf{R} - \{-1\}$. On note également f l'application donnée par $f : (x,y)\mapsto x^2+y^2$. Alors

$$h\circ f(x,y) = \frac{1}{1+(x^2+y^2)} = g(x,y)$$

pour tout $(x,y)\in\mathbf{R}^2$ et donc $g = h\circ f$. La fonction h est dérivable sur son domaine de définition et $h'(t) = \frac{-1}{(1+t)^2} < 0$ pour tout $t\in\mathbf{R}-\{-1\}$. L'application h est donc strictement décroissante sur l'intervalle ouvert $]-1,+\infty[$. Comme $f(x,y)\geqslant 0$ pour tout $(x,y)\in\mathbf{R}^2$, on obtient que

$$g(x,y) = h(f(x,y)) \leqslant h(0) = 1 = g(0,0).$$

Donc g admet un maximum global en $(0,0)$ qui vaut 1.

3. Notons $X : (x,y)\mapsto x$ et $Y : (x,y)\mapsto y$ les applications correspondant aux deux composantes du couple. Alors la fonction $u = 1+X^2+Y^2$ est différentiable et ne prend pas la

valeur 0 sur $\mathbf{R}^2$. Par conséquent la fonction g est différentiable et

$$\begin{aligned}\operatorname{grad}_A(g) &= \frac{-1}{u(A)^2}\operatorname{grad}_A(u)\\ &= \frac{-1}{(1+x^2+y^2)^2}(2X(A)\operatorname{grad}_A(X)+2Y(A)\operatorname{grad}_A(Y))\\ &= \frac{-1}{(1+x^2+y^2)^2}(2x(1,0)+2y(0,1))\\ &= \left(\frac{-2x}{(1+x^2+y^2)^2}, \frac{-2y}{(1+x^2+y^2)^2}\right)\end{aligned}$$

4. La relation $\operatorname{grad}_A(g) = (0,0)$ équivaut à $x = 0$ et $y = 0$ donc $\operatorname{grad}_A(g)$ ne s'annule que lorsque $A = (0,0)$.
5. Par la question 2, la fonction g admet un maximum global en $(0,0)$ et elle est différentiable en ce point. Donc $\operatorname{grad}_{(0,0)}(g) = (0,0)$.

Problème. 1. La formule $f(x,y) = x^2 - y^2$ pour tout $(x,y) \in \mathbf{R}^2$ donne en spécialisant :

$$\begin{aligned} f(1,1) &= 0 & f(1,-1) &= 0\\ f(1,0) &= 1 & f(0,1) &= -1\\ f(-1,0) &= 1 & f(0,-1) &= -1\end{aligned}$$

2. Soit $(x,y) \in \mathbf{R}^2$.
   (a) On a les équivalences

$$\begin{aligned} f(x,y) = 0 &\iff x^2 - y^2 = 0\\ &\iff (x-y)(x+y) = 0\\ &\iff y = x \text{ ou } y = -x;\end{aligned}$$

   (b) Ainsi que

$$\begin{aligned} f(x,y) > 0 &\iff x^2 - y^2 > 0\\ &\iff (x-y)(x+y) > 0\\ &\iff (y < x \text{ et } y > -x) \text{ ou } (y > x \text{ et } y < -x);\end{aligned}$$

(c) Et

$$\begin{aligned} f(x,y) < 0 &\iff x^2 - y^2 < 0 \\ &\iff (x-y)(x+y) < 0 \\ &\iff (y < x \text{ et } y < -x) \text{ ou } (y > x \text{ et } y > -x). \end{aligned}$$

ce qui donne la figure 1

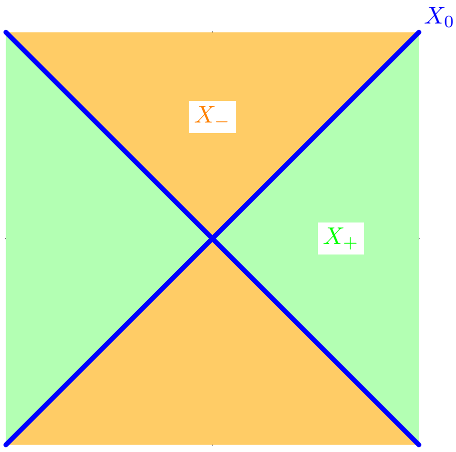


FIGURE 1. Les ensembles X_0, X_+ et X_-

.

3. (a) La distance OA est égale à $\frac{r}{2} < r$ donc $A \in \mathrm{B}(O,r)$ et $f(A) = \frac{r^2}{4} > 0$.
   (b) De même, $OB = \frac{r}{2}$ donc $B \in \mathrm{B}(0,r)$ et $f(B) = -\frac{r^2}{4} < 0$.
   (c) Par les deux questions précédentes, pour tout nombre $r \in \mathbf{R}_+^*$, il existe des points A et B dans la boule $\mathrm{B}(O,r)$ tels que
   $$f(B) < 0 = f(O) < f(A).$$
   Donc la fonction f d'admet ni minimum local ni maximum local en O.
4. (a) Avec les notation usuelles pour les fonctions définies par les composantes, $f = X^2 - Y^2$ est différentiable et
   $$\begin{aligned} \mathrm{grad}_A(f) &= 2X(A)\,\mathrm{grad}_A(X) - 2Y(A)\,\mathrm{grad}_A(Y) \\ &= 2x(1,0) - 2y(0,1) = (2x, -2y). \end{aligned}$$

(b) Soit $A = (x, y) \in \mathbf{R}^2$. La relation $\mathrm{grad}_A(f) = (0,0)$ équivaut à $(2x, -2y) = (0,0)$ et donc à $x = y = 0$. Donc le gradient ne s'annule qu'en O. Par la question 3.(c), la fonction n'admet ni minimum local ni maximum local en O.

5. (a) Les composantes de $g_c(t)$ sont bien définies dès que $t \neq 0$. Donc $\mathscr{D}_{g_c} = \mathbf{R} - \{0\}$.

(b) Soit $t \in \mathbf{R} - \{0\}$.

$$f \circ g_c(t) = f\left(\frac{1}{2}\left(t + \frac{c}{t}\right), \frac{1}{2}\left(t - \frac{c}{t}\right)\right) =$$
$$\frac{1}{4}\left(\left(t + \frac{c}{t}\right)^2 - \left(t - \frac{c}{t}\right)^2\right) = \frac{1}{4}\left(t + \frac{c}{t} + t - \frac{c}{t}\right)\left(t + \frac{c}{t} - t + \frac{c}{t}\right) = c.$$

Donc $f \circ g_c$ est l'application constante de valeur c.

(c) Les composantes de g_c sont dérivables sur $\mathbf{R}^*$ et

$$g_c'(t) = \left(\frac{1}{2}\left(1 - \frac{c}{t^2}\right), \frac{1}{2}\left(1 + \frac{c}{t^2}\right)\right)$$

pour tout $t \neq 0$.

(d) En appliquant la question précédente, la question 4.(a) et la définition du produit scalaire, on obtient pour tout t in $\mathbf{R}^*$ les égalités

$$\begin{aligned}
&g_c'(t) \cdot \mathrm{grad}_{g_c(t)}(f) \\
&= \left(\frac{1}{2}\left(1 - \frac{c}{t^2}\right), \frac{1}{2}\left(1 + \frac{c}{t^2}\right)\right) \cdot \left(2\left(t + \frac{c}{t}\right), -2\left(t - \frac{c}{t}\right)\right) \\
&= \frac{1}{t}\left(\left(t - \frac{c}{t}\right)\left(t + \frac{c}{t}\right) - \left(t + \frac{c}{t}\right)\left(t - \frac{c}{t}\right)\right) \\
&= 0.
\end{aligned}$$

(e) Par la question (b), la composée $f \circ g_c$ est constante donc

$$g_c'(t) \cdot \mathrm{grad}_{g_c(t)}(f) = 0$$

pour tout $t \in \mathscr{D}_{g_c}$.

(f) Soit $(x, y) \in \mathbf{R}^2$ avec $y \neq x$ et $y \neq -x$.

On cherche des nombres c et t de sorte que $(x, y) = g_c(t)$. Mais cette relation implique d'une part les égalités

$$f(x, y) = f(g_c(t)) = f \circ g_c(t)) = c$$

la dernière égalité résultant de la question (b). La relation $(x, y) = g_c(t)$ implique d'autre part que $x+y=t$. Les seules possibilités pour c et t sont donc $c = f(x, y) = x^2 - y^2$ et $t = x + y$.

Inversement vérifions que les nombres $c = x^2 - y^2$ et $t = x + y$ conviennent. Alors comme $y \neq x$ et $y \neq -x$, les nombres c et t sont tous les deux non nuls et $\frac{c}{t} = x - y$, ce qui donne

$$g_c(t) = \left(\frac{1}{2}(x+y+x-y), \frac{1}{2}(x+y-x+y)\right) = (x, y).$$

Donc les nombres $c = x^2 - y^2$ et $t = x + y$ vérifient bien $g_c(t) = (x, y)$.

(g) Par les questions (b) et (f) $f^{-1}(\{c\})$ est l'ensemble des valeurs de la fonction g_c.

6. La courbe de niveau $f^{-1}(\{0\})$ a été décrite dans la question 2.(a), c'est la réunion des droites $X = Y$ et $X = -Y$. On notera que $f(-x, y) = f(x, y) = f(x, -y)$, ce qui implique que les courbes de niveau sont symétriques par rapport aux axes de coordonnées. De plus la distance de $g_c(t)$ au point $\frac{1}{2}(t, t)$ tend vers 0 lorsque t tend vers $+\infty$ et la distance de $g_c(t)$ au point $\frac{c}{2}\left(\frac{1}{t}, \frac{-1}{t}\right)$ tend vers 0 lorsque t tend vers 0, ce qui signifie que les courbes de niveau admettent toutes les deux droites $X = Y$ et $X = -Y$ comme asymptotes.

L'allure des lignes de niveau demandées est représentée dans la figure 2

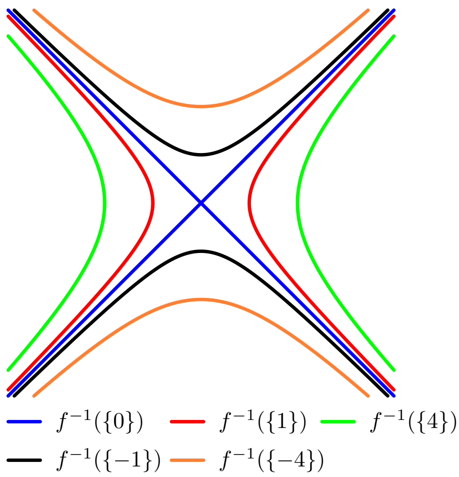


FIGURE 2. Courbes de niveau

LISTE DES FIGURES

GLOSSAIRE

INDEX