Cours - TD: Mathématiques appliquées 2 (semestre 2)

U.F.R. Sciences Économiques et de Gestion

L1 Économie (2016-17)

Cours: J.-M. Barbaroux.
Travaux dirigés: J.-M. Barbaroux & E. Leopold
email: barbarou@univ-tln.fr
web: http://barbarou.univ-tln.fr

Introduction

Que ce soit en Économie, en Physique (mécanique, optique, electromagnétisme, mécanique quantique, astrophysique, etc.), en Biologie (dynamique des populations, neurosciences), ou en Sciences Humaines (sociologie, étude de réseaux complexes), lorsqu'un problème est formalisé par les mathématiques, par exemple à l'aide d'équations, il est rarement sous une forme "simple".

Certains problèmes se formulent à l'aide d'une branche des mathématiques, que l'on nomme **algèbre linéaire**, soit comme une modélisation directe du problème lui-même, soit comme approximation dans certaines conditions.

L'avantage de l'algèbre linéaire, sans être une branche triviale des mathématiques, et qu'il est possibe, grâce à une formalisation que nous allons découvrir, de résoudre explicitement certains problèmes dits "linéaires', et/ou de les programmer pour que les machines les résolvent à notre place.

En Économie, le paradigme de l'algèbre linéaire est le modèle de Leontieff (Prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel en 1973).

Nous allons introduire les bases de l'algèbre linéaire, dans un cas simple qui couvre déjà de nombreux problèmes.

Ce cours d'algèbre linéaire sera partiellement utilisé dans la seconde partie. La deuxième partie du cours est dédiée à la recherche d'extremums de fonctions de plusieurs variables (recherche de maximum ou minimu, local ou global). Dans cette partie **Analyse** du cours, nous donnerons la définition de fonctions de plusieurs variables, et étendrons les notions de continuité et dérivabilité vues en cours de premier semestre pour les fonctions d'une variable réelle, au cas de fonctions de plusieurs variables.

Le plan du cours d'**algèbre linéaire** sera le suivant:

(1) Systèmes linéaires - Résolution par pivot de Gauss
(2) Espaces vectoriels $\mathbb{R}^n$ - Familles libres, familles génératrices, bases.
(3) Matrices - Calcul matriciel - Déterminants - Méthodes de calcul de déterminants
(4) Applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Représentation d'une application linéaire par une matrice. Rang d'une application linéaire.

Le plan du cours d'**analyse** sera le suivant:

(0) Rappels sur: dérivée d'une fonction d'une variable; composition de fonctions et dérivé e de fonctions composées; Théorème de Rolle et Théorème des accroissements finis; équation de la tangente en un point à une fonction.
(1) Fonctions de plusieurs variables - Dérivées - Différentiabilité.
(2) Méthodes de recherche d'extrema.

L'objectif de ce cours est double:

▶ Savoir repérer un problème d'algèbre linéaire ou d'analyse déjà rencontré, et le résoudre en appliquant les méthodes mathématiques vues en cours.

▶ Se familiariser avec les concepts mathématiques vus en cours pour être capable de modéliser les problèmes, d'interpréter les résultats donnés par l'application des méthodes de calculs, et, le cas échéant, créer un algorithme pour mener les calculs.

Le premier objectif n'est pas difficile. Il suffit d'être attentif à tout ce qui est dit en cours, et de bien suivre les exemples traités en cours. Il suffit ensuite d'appliquer directement les méthodes vues, en préparant ses feuilles de TD avant la séance de TD.

Le deuxième objectif demande plus de temps et aussi de la régularité dans le travail. Notamment, il faudra relire son cours avant le TD suivant. Il est normal de ne pas comprendre tout de suite les subtilités, ou plus simplement l'utilité de certaines notions. N'hésitez pas, en début de cours, à poser des questions sur le cours précédent.

Chaque exercice du fascicule de TD est marqué d'un symbole:

- Le symbole ♡ mentionne les exercices fondamentaux; ces exercices sont pour la plupart élémentaires et on trouve leurs solutions dans de nombreux ouvrages de référence ou en ligne. Ces exercices doivent être tous systématiquement préparés. Ils sont indispensables à la bonne *utilisation des méthodes* vues en cours, et sont déjà une première base pour acquérir les notions mathématiques.

- Le symbole ♣ est utilisé pour les exercices de difficulté moyenne. Ce sont en général de bons exercices pour s'entraîner et vérifier que l'on a bien acquis les notions du cours. Ces exercices ne se résolvent pas nécessairement par application directe d'une méthode vue en cours. Ils requièrent de passer un peu plus de temps dessus. C'est avec ces exercices que vous *comprendrez* les notions du cours.

- Le symbole ♠ accompagne les exercices qui présentent souvent plusieurs difficultés. Ils sont utiles pour aller un peu plus loin.

Seuls des exercices de type ♡ et ♣ seront au programme des examens.

ALGÈBRE LINÉAIRE

1. Fiche de cours: Systèmes linéaires

Définition 1.1. *On appelle équation linéaire réelle (respectivement complexe) d'inconnues $x_1, x_2, \cdots, x_n$ toute relation du type*

$$a_1x_1 + a_2x_2 + \cdots + a_px_p = b, \tag{1}$$

où $a_1, a_2, \cdots, a_p, b$ sont des nombres réels (respectivement complexes).
Cette expression s'écrit aussi

$$a_1x_1 + a_2x_2 + \cdots + a_px_p - b = 0, \tag{2}$$

Exemple: L'équation de droite

$$y - 3x = 1 \quad (\text{ou encore } y = 3x + 1).$$

est une équation linéaire réelle. Dans cet exemple, les variables x_1 et x_2 sont respectivement y et x, et $a_1 = 1$, $a_2 = -3$, $b = 1$.

Remarque. *Les inconnues $x_1, x_2, \cdots, x_n$ sont aussi appelées les variables de l'équation.*

Définition 1.2 (Systèmes linéaires). *On appelle système linéaire de n équations à p inconnues $x_1, x_2, \cdots, x_p$, une liste de n équations de la forme*

$$\left\{\begin{array}{llllll} a_{11}x_1 & +a_{12}x_2 & +a_{13}x_3 & +\cdots & +a_{1p}x_p & = b_1 \quad (\textit{ligne 1, notée } L_1) \\ a_{21}x_1 & +a_{22}x_2 & +a_{23}x_3 & +\cdots & +a_{2p}x_p & = b_2 \quad (\textit{ligne 2, notée } L_2) \\ \vdots & \vdots & \vdots & & \vdots & = \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +a_{n3}x_3 & +\cdots & +a_{np}x_p & = b_n \quad (\textit{ligne n, notée } L_n) \end{array}\right. \tag{3}$$

Remarque. *Pour la résolution des systèmes par la méthode du pivot de Gauss (voir plus bas), il est commun de numéroter chaque ligne par $L_1, L_2, \cdots, L_i, \cdots, L_n$.*

L'indice i dans a_{ij} correspond à l'indice de "ligne", c'est à dire que a_{ij} intervient dans la i-ème ligne L_i du système

L'indice j dans a_{ij} correspond à l'indice de "colonne", c'est à dire que a_{ij} est le coefficient devant la j-ème inconnue x_j.

Définition 1.3. *Une solution du système linéaire* (3) *est une liste de p nombres réels (complexes) $(s_1, s_2, \cdots, s_p)$ qui mis à la place de $(x_1, x_2, \cdots, x_n)$ vérifient l'équation* (3).

Résoudre une équation linéaire réelle (resp. complexe) consiste à décrire l'ensemble des listes de p nombres réels (resp. complexes) qui vérifient la relation (1)

Remarque. *Une liste de p nombres réels $(s_1, s_2, \cdots, s_p)$ est appelée un p-uplet. Un p-uplet se note avec des parenthèses et chaque réel est séparé des autres par une virgule ou un point-virgule. Cette notation est importante, et l'utilisation d'accolades $\{\ \}$ pour un p-uplet est incorrecte, car cela est réservé à la notation des ensembles.*
⚠ *L'ordre des réels dans le p-uplet est important, car il correspond à l'ordre dans lequel on remplace les valeurs dans les variables. Ainsi, $(1,2) \neq (2,1)$.*

Définition 1.4.

- *On appelle coefficients d'un système linéaire les valeurs a_{ij}, où $i \in \{1, \cdots, n\}$, $j \in \{1, \cdots, p\}$.*
- *On appelle second membre d'un système linéaire le n-uplet $(b_1, b_2, \cdots, b_n)$.*
- *On dira qu'un système linéaire est homogène si le second membre est nul, i.e., si $(b_1, b_2, \cdots, b_n) = (0, 0, \cdots, 0)$.*

Exemple: Le système

$$\left\{\begin{array}{cccc} x_1 & -x_2 & +2x_3 & = & 1 \\ 3x_1 & & -4x_3 & = & -1 \end{array}\right. \tag{4}$$

admet comme solution $(1,2,1)$. Ce n'est pas la seule solution de ce système. Ce n'est pas un système homogène. Il admet pour second membre $(1,-1)$.

Définition 1.5 (Systèmes équivalents). *On dira que deux systèmes linéaires sont équivalents s'ils ont le même ensemble de solutions.*

Exemple: On peut montrer que le système (4) est équivalent au système:

$$\left\{\begin{array}{ccccc} x_1 & -x_2 & +2x_3 & = & 1 \\ & 3x_2 & -10x_3 & = & -4 \end{array}\right.$$

La fin de ce chapitre est consacrée aux méthodes permettant d'obtenir, à partir d'un système linéaire donné, des systèmes linéaires équivalents. Nous verrons aussi une méthode standard de résolution des systèmes linéaires: *Le pivot de Gauss.* Dans les chapitres ultérieurs, nous verrons d'autres méthodes de résolution.

Théorème 1.6. *Pour tout système linéaire de la forme (3), on a l'un des 3 cas suivants:*

(1) Le système n'admet aucune solution
(2) Le système admet un p-uplet de solution et un seul
(3) Le système admet une infinité de p-uplets solutions.

Remarque. *i) Les systèmes homogènes admettent toujours comme solution (au moins) le p-uplet* $(0, 0 \cdots, 0)$.
ii) D'après ce théoreme, il n'est pas possible de trouver seulement deux (ou un nombre fini strictement supérieur à un) solutions à un système linéaire. Dès qu'on a plus d'une solution, il y en a forcément une infinité.

Résolution de systèmes linéaires par la méthode du pivot de Gauss

Pour résoudre un système linéaire par la méthode du pivot de Gauss, on va remplacer le système initial par des systèmes équivalents (au sens de la définition 1.5) grâce à certaines opérations sur les lignes. L'objectif est d'aboutir à un système que l'on peut résoudre simplement, qu'on nomme "système échelonné".

Définition 1.7. *Un système est dit échelonné si le nombre de coefficients nuls commençant une ligne est strictement croissant d'une ligne à l'autre.*

Exemple. Les systèmes suivants sont échelonnés.

$$\left\{\begin{array}{rrrcr} x_1 & -x_2 & +2x_3 & = & 1 \\ & 3x_2 & -10x_3 & = & -4 \end{array}\right.$$

$$\left\{\begin{array}{rrrrcr} x_1 & -x_2 & +x_3 & +5x_4 & = & 1 \\ & x_2 & -10x_3 & -4x_4 & = & -4 \\ & & 2x_3 & +x_4 & = & 0 \\ & & & -6x_4 & = & -1 \end{array}\right.$$

$$\left\{\begin{array}{rrrrcr} x_1 & -3x_2 & +x_3 & +\sqrt{5}x_4 & = & 1 \\ & & 3x_3 & -2x_4 & = & -4 \\ & & & +7x_4 & = & 0 \end{array}\right.$$

Définition 1.8. *Etant donné un système linéaire, de n équations à p inconnues $x_1, \cdots, x_p$, on transforme le système en un système équivalent si on procède à l'une des opérations suivantes sur les lignes.*

(1) *On remplace une ligne par un multiple d'elle-même:*
$L_i \longrightarrow \lambda L_i$, *où λ est un nombre non nul quelconque.*

(2) *On remplace une ligne par elle-même plus un multiple d'une autre ligne:*
$L_i \longrightarrow L_i + \lambda L_j \quad (i \neq j)$.

(3) *On intervertit deux lignes quelconques:*
$L_i \longleftrightarrow L_j$.

Exemple. Voici un exemple d'applications des règles ci-dessus qui permettent de transformer le système initial en des systèmes équivalents. On écrira le symbole $\Leftrightarrow$ entre deux systèmes pour mettre en évidence que les systèmes sont équivalents. On écrira aussi dans chaque système l'opération sur les lignes que l'on va effectuer et qui permettra d'obtenir le système suivant.

$$\left\{\begin{array}{rrrcrl} 2x_1 & +x_2 & +x_3 & = & 5 & \quad L_1 \longleftrightarrow L_2 \\ x_1 & -x_2 & +2x_3 & = & 1 & \\ 3x_1 & & -4x_3 & = & -1 & \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{rrrcrl} x_1 & -x_2 & +2x_3 & = & 1 & \\ 2x_1 & +x_2 & +x_3 & = & 5 & \quad L_2 \longrightarrow L_2 - 2L_1 \\ 3x_1 & & -4x_3 & = & -1 & \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{rrrcrl} x_1 & -x_2 & +2x_3 & = & 1 & \\ & +3x_2 & -3x_3 & = & 3 & \\ 3x_1 & & -4x_3 & = & -1 & \quad L_3 \longrightarrow L_3 - 3L_1 \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{rrrcrl} x_1 & -x_2 & +2x_3 & = & 1 & \\ & 3x_2 & -3x_3 & = & 3 & \quad L_2 \longrightarrow \frac{1}{3}L_2 \\ & 3x_2 & -10x_3 & = & -4 & \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{rrrcrl} x_1 & -x_2 & +2x_3 & = & 1 & \\ & x_2 & -x_3 & = & 1 & \quad L_3 \longrightarrow L_3 - 3L_2 \\ & 3x_2 & -10x_3 & = & -4 & \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{rrrcrl} x_1 & -x_2 & +2x_3 & = & 1 & \\ & x_2 & -x_3 & = & 1 & \quad L_3 \longrightarrow -\frac{1}{6}L_3 \\ & & -7x_3 & = & -7 & \end{array}\right. \quad \Leftrightarrow \left\{\begin{array}{rrrcr} x_1 & -x_2 & +2x_3 & = & 1 \\ & x_2 & -x_3 & = & 1 \\ & & x_3 & = & 1 \end{array}\right.$$

Le système de départ a donc les mêmes solutions que le dernier système. L'avantage est que le dernier système se résout simplement. On trouve $x_3 = 1$, qui donne alors avec la ligne 2 (du dernier système) $x_2 = 1 + x_3 = 1 + 1 = 2$, puis avec la ligne 1, $x_1 = x_2 - 2x_3 + 1 = 2 - 2 + 1 = 1$. Le système admet donc comme unique solution le triplet: $(1, 2, 1)$.

C'est un exemple d'application de la méthode de résolution par le pivot de Gauss.

Pour la résolution d'un système linéaire par la méthde du pivot de Gauss, les étapes sont les suivantes:

(1) On réécrit le système en plaçant en ligne 1 une ligne où le premier coefficient est non nul. Si possible, on choisit une ligne où le premier coefficient vaut 1.
(2) Si le premier coefficient de la ligne 1 est différent de 1, on remplace la première ligne par elle-même divisée par la valeur du premier coefficient de cette ligne. On obtient alors une ligne où le premier coefficient a pour valeur 1.

(3) On laisse inchangée la ligne 1, et on remplace chaque ligne L_i (successivement pour $i = 2, \cdots, n$) par elle-même moins la première ligne fois le premier coefficient de la ligne i: $L_i \longrightarrow L_i - a_{i1}L_1$. On obtient alors un système où les lignes 2 à n ont leur premier coefficient nul.

(4) On conserve la ligne 1 inchangée (et cela jusqu'à la fin), et on procède avec les lignes 2 à n en recommençant au point 1.

Exercice. Vérifiez que l'on a bien procédé ainsi dans l'exemple précédent.

Exercice. Résolvez les systèmes suivants par la méthode du pivot de Gauss.

$$(S1) : \left\{ \begin{array}{rrcr} 2x_1 & -x_2 & = & -6 \\ x_1 & +2x_2 & = & 7 \end{array} \right. \qquad (S2) : \left\{ \begin{array}{rrrcr} x_1 & +x_2 & +2x_3 & = & 1 \\ 2x_1 & -x_2 & -x_3 & = & 1 \end{array} \right.$$

$$(S3) : \left\{ \begin{array}{rrrrcr} x_1 & +x_2 & +4x_3 & +x_4 & = & 0 \\ 2x_1 & -x_2 & +6x_3 & +x_4 & = & 1 \\ -x_1 & +3x_2 & +x_3 & +4x_4 & = & 0 \\ 2x_1 & +x_2 & -x_3 & -3x_4 & = & 4 \end{array} \right. \qquad (S4) : \left\{ \begin{array}{rrrcr} 2x_1 & +x_2 & +4x_3 & = & 1 \\ 2x_1 & -x_2 & +x_3 & = & -2 \\ -x_1 & +3x_2 & +4x_3 & = & 4 \\ x_1 & +x_2 & -x_3 & = & 4 \end{array} \right.$$

2. Fiche de cours: Espaces vectoriels $\mathbb{R}^n$ - Familles libres, familles génératrices, bases.

2.1. **Espace vectoriel $\mathbb{R}^n$.** L'algèbre linéaire requiert de travailler sur des ensembles qui ont une certaine structure. Pour faire simple, on aura besoin de travailler avec des ensembles dans lesquels on peut "sommer" les éléments et le résultat obtenu reste dans l'ensemble, et on peut les "multiplier par un nombre" (réel ou complexe), que l'on appellera un scalaire, et le résultat obtenu reste dans l'ensemble. On appelle ces propriétés *stabilité par addition* et *stabilité par multiplication par un scalaire.* Notez qu'il n'est pas nécessaire de pouvoir multiplier ces éléments entre eux, même si on verra des cas où c'est possible. En fin de chapitre, nous donnerons une définition de ces espaces, que l'on appellera espaces vectoriels.

Exercice. Donnez des exemples d'ensembles que vous connaissez et qui respectent ces deux conditions (stabilité par addition et stabilité par multiplication par un réel). Donnez aussi des exemples d'ensembles qui ne respectent pas l'une au moins ces deux conditions.

Pour l'essentiel de ce chapitre et de cette année, nous allons nous intéresser aux cas particuliers des espaces vectoriels $\mathbb{R}^n$ ($n \in \mathbb{N}^*$).

Définition 2.1. *Soit n un entier non nul. On appelle $\mathbb{R}^n$, noté aussi $\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ fois}}$, l'ensemble des n-uplets définis par*

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

Les éléments de $\mathbb{R}^n$ sont appelés vecteurs.

Remarque. *On définit de la même manière $\mathbb{C}^n$. Dans la suite, on notera les éléments de $\mathbb{R}^n$ par des vecteurs colonnes $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Le choix de notation en ligne ou en colonne n'est pas identique, même si pour l'instant, la distinction entre les deux notations semble être sans importance. C'est pourquoi on adoptera plutôt la notation en colonne pour les vecteurs de $\mathbb{R}^n$ dans la suite.*

Exemple. L'élément $\begin{pmatrix} 1 \\ -1 \\ \sqrt{2} \\ \frac{3}{2} \end{pmatrix}$ est un élément de $\mathbb{R}^4$. Le 4-uplet $\begin{pmatrix} -1 \\ \frac{3}{2} \\ \sqrt{2} \\ 1 \end{pmatrix}$ en est un autre.

Définition 2.2 (Somme - Produit extérieur). *Soient* $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ *et* $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ *deux éléments de* $\mathbb{R}^n$*. Alors:*

$$u + v = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix} \quad \textit{(somme)}$$

$$\lambda u = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix} \quad \textit{(produit extérieur par un scalaire } \lambda \in \mathbb{R}\textit{)}$$

Le vecteur nul (qui sera appelé élément neutre) est $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$*. L'opposé d'un vecteur* u *est le vecteur noté* $-u$ *et qui vérifie* $u + (-u) = 0$*, c'est donc le vecteur* $-u = \begin{pmatrix} -u_1 \\ \vdots \\ -u_n \end{pmatrix}$.

Remarque. *i) Plutôt que la notation* u*, on utilise parfois* $\vec{u}$ *(notation du lycée).*
ii) Dans $\mathbb{R}^2$*, on peut aussi représenter la somme de vecteurs, et le produit extérieur, par un schéma.*

La structure d'espace vectoriel de $\mathbb{R}^n$ permet d'avoir les propriétés suivantes qui sont bien connues dans $\mathbb{R}$.

Proposition 2.3. *Pour* u*,* v *et* w *vecteurs quelconques de* $\mathbb{R}^n$ *et* λ *et* μ *deux réels quelconques, on a*

1.	$u + v = v + u$;	2.	$u + (v + w) = (u + v) + w$
3.	$u + 0 = 0 + u = u$	4.	$u + (-u) = 0$
5.	$1u = u$	6.	$0u = 0$
7.	$\lambda(u + v) = \lambda u + \lambda v$	8.	$\lambda(\mu u) = \lambda(\mu u)$

L'espace vectoriel $\mathbb{R}$ est usuellement représenté par l'axe des abscisses.
L'espace vectoriel $\mathbb{R}^2$ est représenté par le plan.

Exercice. Démontrez les propriétés de la proposition 2.3 dans le cas $\mathbb{R}^2$.

2.2. Familles génératrices, bases de $\mathbb{R}^n$.

DÉFINITION 2.4 (Combinaisons linéaires). *Soient $v^{(1)}, \cdots, v^{(p)}$, p vecteurs de $\mathbb{R}^n$. On appelle combinaison linéaire des vecteurs $v^{(1)}, \cdots, v^{(p)}$, tout vecteur de la forme*

$$\lambda_1 v^{(1)} + \lambda_2 v^{(2)} + \cdots + \lambda_p v^{(p)},$$

où $\lambda_1, \lambda_2 \cdots, \lambda_p$ sont p nombres réels. Les scalaires $\lambda_1, \lambda_2 \cdots, \lambda_p$ sont les coefficients de la combinaison linéaire.

DÉFINITION 2.5 (Famille génératrice). *Soit $\{v^{(1)}, v^{(2)}, \cdots, v^{(p)}\}$ une famille de p vecteurs de $\mathbb{R}^n$. On dit que cette famille est génératrice de $\mathbb{R}^n$ si tout vecteur u de $\mathbb{R}^n$ peut s'écrire comme combinaison linéaire de vecteurs de $\{v^{(1)}, v^{(2)}, \cdots, v^{(p)}\}$:*

$$\forall u \in \mathbb{R}^n,\ \exists(\lambda_1, \lambda_2, \cdots, \lambda_p) \text{ tels que } u = \lambda_1 v^{(1)} + \lambda_2 v^{(2)} + \cdots + \lambda_p v^{(p)}.$$

Exemple. La famille $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ est génératrice de $\mathbb{R}^3$.

La famille $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} ; \begin{pmatrix} 2 \\ -1 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ est génératrice de $\mathbb{R}^2$.

Exercice. Pouvez-vous donnez une famille (la plus “simple possible”) qui soit génératrice de $\mathbb{R}^4$?

Est-il possible de contruire une famille de deux vecteurs dans $\mathbb{R}^3$ qui soit une famille génératrice? (pensez à la représentation géométrique).

DÉFINITION 2.6 (Base). *On appelle base de l'espace vectoriel $\mathbb{R}^n$ toute famille génératrice de $\mathbb{R}^n$ qui contient exactement n vexteurs.*

REMARQUE. *Nous verrons plus loin (voir Proposition 3.18), à l'aide des déterminants, un critère simple qui permet de dire si une famille est une base de $\mathbb{R}^n$.*

Exemple. La famille $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ est une base de $\mathbb{R}^3$. On peut montrer que la famille $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} ; \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$ est aussi une base de $\mathbb{R}^3$.

3. Fiche de cours: Matrices - Calcul matriciel - Déterminants - Méthodes de calcul de déterminants

Dans ce chapitre, nous définissons de nouveaux objets mathématiques, les matrices, et une quantité associée à chaque matrice, le déterminant. Ces objets seront très utiles dans les deux paragraphes suivants, mais aussi pour la résolution de systèmes linéaires de n équations à n inconnues.

3.1. **Définitions.**

Définition 3.1 (Matrice).

- *Une matrice est un tableau rectangulaire d'éléments de $\mathbb{R}$ (ou $\mathbb{C}$). Les matrices sont notées usuellement avec des lettres majuscules.*
- *Une matrice est de taille $n \times p$ si c'est un tableau qui possède n lignes et p colonnes. L'ensemble des matrices à n lignes et p colonnes est noté $\mathfrak{M}_{n,p}$.*
- *Les valeurs dans le tableau sont appelées les coefficients de la matrice.*
- *Le coefficient situé à la i-ème ligne et à la j-ème colonne de la matrice A est noté $a_{i,j}$ ou a_{ij}. (si la matrice est notée B, on notera les coefficients par b_{ij}, etc.). On utilise alors la notation $A = (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ ou encore $A = (a_{ij})$.*

Exemple. La matrice $A = \begin{pmatrix} 1 & 0 & -1 & 3 & 4 \\ 0 & 2 & \sqrt{2} & \pi & 0 \\ -1 & \sqrt{3} & 12 & 5 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ est une matrice 4×5, puisqu'elle possède 4 lignes et 5 colonnes. L'élément $a_{2,4}$ est π et l'élément $a_{4,2}$ est 1.

Définition 3.2 (Quelques matrices remarquables).

- *Une matrice $n \times n$, i.e., ayant le même nombre de lignes et de colonnes, est une matrice carrée. L'ensemble des matrices carrées $n \times n$ est noté $\mathfrak{M}_{n,n}$ ou $\mathfrak{M}_n$. Une matrice $n \times n$ est appelée matrice d'ordre n.*
- *Dans une matrice carrée, on appelle diagonale les éléments $a_{i,i}$ ayant même indice de ligne et de colonne.*
- *Une matrice carrée est triangulaire supérieure si tous ses coefficients au dessous de la diagonale sont nuls. Elle est triangulaire inférieure si tous ses coefficients au dessus de la diagonale sont nuls.*
- *Une matrice carrée dont tous les éléments sont nuls sauf les éléments diagonaux, est appelée matrice diagonale.*
- *Une matrice (carrée) diagonale dont tous les éléments de la diagonale valent 1 est appelée matrice unité ou matrice identité. On la note I_n ou seulement I s'il n'y a pas d'ambiguïté.*
- *Une matrice dont tous les éléments sont nuls est appelée matrice nulle.*
- *Une matrice $n \times 1$ est appelée matrice colonne (ou vecteur colonne)*

- *Une matrice $1 \times p$ est appelée matrice ligne (ou vecteur ligne)*

Exemples. Dans la matrice carrée suivante, la diagonale est en rouge: $\begin{pmatrix} 2 & 5 & -1 & 7 \\ 3 & 4 & 0 & 4 \\ 2 & 0 & 0 & 6 \\ 8 & -2 & 0 & 9 \end{pmatrix}$.

La matrice $\begin{pmatrix} 2 & 1 & -7 & 8 \\ 0 & 6 & 9 & 4 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ est triangulaire supérieure. La matrice $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est une matrice diagonale. La matrice $B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ est la matrice identité I_4. La matrice $X = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}$ est une matrice colonne, et $Y = \begin{pmatrix} 1 & 2 & 0 & -1 \end{pmatrix}$ est une matrice ligne.

3.2. **Opérations sur les matrices.**

DÉFINITION 3.3 (Somme). *Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices $n \times p$. Alors la matrice $C = A + B$ est une matrice $n \times p$ obtenue en sommant terme à terme chaque coefficient des matrices A et B:*

$$A + B = C = (c_{ij}) = (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1p} + b_{1p} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2p} + b_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{np} + b_{np} \end{pmatrix}$$

DÉFINITION 3.4 (Produit par un réel (ou un complexe)). *Soit $A = (a_{i,j})$ une matrice et soit $\lambda \in \mathbb{R}$. Alors, la matrice λA est la matrice obtenue en multipliant chaque coefficient de la matrice A par λ. La matrice λA a donc pour coefficients $(\lambda a_{i,j})$.*

REMARQUE. *On ne peut pas additionner deux matrices qui n'auraient pas le même nombre de lignes ou de colonnes.*

Si A est une matrice $n \times p$ et si 0 est la matrice nulle $n \times p$, alors $A+0 = 0+A = A$.

Grâce au produit d'une matrice par un scalaire (réel ou complexe), l'ensemble des matrices $\mathfrak{M}_{n,p}$ est un espace vectoriel. Il vérifie donc toutes les propriétés de la proposition 2.3. En particulier, l'opposé $-A$ d'une matrice $A = (a_{i,j})$ de $\mathfrak{M}_{n,p}$ est dans $\mathfrak{M}_{n,p}$, et a pour coefficients $(-a_{i,j})$.

On définit ensuite le produit de matrices. La définition ci-dessous est utile pour des raisonnements abstraits, que nous éviterons pour cette année. C'est par un exemple que vous comprendrez mieux comment se fait le produit de matrices.

Définition 3.5 (Produit de matrices). *Soient $A \in \mathfrak{M}_{n,p}$ et $B \in \mathfrak{M}_{p,\ell}$. Alors la matrice $C = AB$ a ses coefficients $c_{i,j}$ qui sont obtenus, pour tout $i \in \{1, \cdots n\}$ pour tout $j \in \{1, \cdots, \ell\}$, par*

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^{p} a_{i,k} b_{k,j} \\ &= a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,p}b_{p,j} \end{aligned}$$

Remarque. *i)* ⚠ *Notez bien que les deux matrices A et B ne sont pas nécessairement dans les mêmes ensembles de matrices, mais que le nombre de colonnes de A doit être égal au nombre de lignes de B.*
ii) ⚠ *Les matrices ne vérifient en général pas la propriété de commutation. Cela signifie qu'en général $AB \neq BA$. Vous remarquerez d'ailleurs que l'un des deux produits peut être bien défini, alors que l'autre ne l'est pas forcément, à cause du nombre de lignes et de colonnes de A et B.*

Exemple. Soient $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 1 & -1 \end{pmatrix} \in \mathfrak{M}_{2,4}$ et $B = \begin{pmatrix} 3 & 1 & 9 \\ 1 & 0 & 4 \\ 9 & 2 & 5 \\ 2 & -1 & 2 \end{pmatrix} \in \mathfrak{M}_{4,3}$; Alors, la matrice $C = AB$ est dans $\mathfrak{M}_{2,3}$ et:

$$\begin{array}{cc} & \begin{pmatrix} 3 & 1 & 9 \\ 1 & 0 & 4 \\ 9 & 2 & 5 \\ 2 & -1 & 2 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 1 & -1 \end{pmatrix} & \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{2,3} \end{pmatrix} \end{array}$$

où $c_{2,3} = 2 \times 9 \ + \ 5 \times 4 \ + \ 1 \times 5 \ + \ (-1) \times 2 = 41$. Le calcul des autres coefficients donne $C = \begin{pmatrix} 40 & 3 & 40 \\ 18 & 5 & 41 \end{pmatrix}$.

On ne peut pas calculer BA.

Exercice. Soit $A = \begin{pmatrix} 1 & 3 & -4 \\ 2 & 6 & -8 \\ 1 & 3 & -4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 1 \\ 2 & -3 & 4 \end{pmatrix}$ et $I = I_3$ la matrice identité. Calculer AB et BA, puis IA, IB, AI et BI.

Définition 3.6 (Transposition). *Si $A \in \mathfrak{M}_{n,p}$, alors sa matrice transposée, notée ${}^tA$ est dans $\mathfrak{M}_{p,n}$ et ses coefficients sont obtenus en échangeant le rôle du numéro de ligne et du numéro de colonne. La matrice ${}^tA = (\alpha_{i,j})$ vérifie donc $\alpha_{i,j} = a_{j,i}$*

Exemple. La matrice transposée de la matrice A de l'exemple ci-dessus est: $\begin{pmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 1 \\ 4 & -1 \end{pmatrix}$.

Proposition 3.7 (Propriétés du produit de matrices).

(1) $A(BC) = (AB)C$. C'est l'associativité du produit de matrices.
(2) $A(B+C) = AB + AC$. C'est la distributivité de la multiplication par rapport à l'addition.

3.3. **Inverse d'une matrice et méthode de calcul.**

Définition 3.8 (Inverse d'une matrice). *Soit $A \in \mathfrak{M}_n$ une matrice carrée d'ordre n. S'il existe une matrice $B \in \mathfrak{M}_n$ telle que $AB = BA = I_n$, alors on dit que A est inversible et son inverse est B. On note alors A^{-1}l'inverse de la matrice A.*

Remarque. *i) On ne parlera pas d'inverse pour une matrice qui n'est pas carrée.*
ii) Si A est inversible, alors $(A^{-1})^{-1} = A$.
iii) Les matrices carrées ne sont pas toutes inversibles. On peut par exemple montrer que $\begin{pmatrix} 3 & 2 \\ 6 & 4 \end{pmatrix}$ n'est pas inversible. On verra dans le paragraphe suivant un critère permettant de montrer qu'une matrice est inversible ou pas.
iv) L'inverse d'une matrice A, s'il existe, est unique.

Proposition 3.9. *Soient A et B deux matrices carrées suposées inversibles. Alors AB est inversible et on a $(AB)^{-1} = B^{-1}A^{-1}$.*

Méthode de Gauss-Jordan d'inversion des matrices. Dans ce qui suit, nous allons voir une méthode pour inverser une matrice, basée sur la méthode du pivot de Gauss. Il existe d'autres méthodes (en particulier à l'aide du déterminant). Nous n'aurons pas le temps de les voir.

La méthode de calcul de l'inverse d'une matrice A consiste à écrire côte à côte la matrice A que l'on doit inverser et la matrice unité. On écrit $(A \mid I)$. C'est ce que l'on appelle la matrice augmentée. On effectue ensuite sur les lignes de A les opérations vues dans la définition 1.8 (et seulement ce type d'opérations) pour transformer A en la matrice identité. Pour chaque opération faite sur les lignes de A, on effectue la même opération sur les lignes de la matrice à droite de A. À la fin, quand A est devenue I, alors I est devenue A^{-1}. La stratégie consiste d'abord à faire les mêmes opérations que pour le pivot de Gauss, i.e., on cherche à obtenir une matrice triangulaire.

Voici un exemple de calcul d'inverse de matrice par cette méthode. À chaque étape, on note l'opération sur les lignes qui sera effectuée à l'étape suivante.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \qquad L_3 \longrightarrow L_3 + L_1$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 3 & 1 & 1 & 0 & 1 \end{array}\right) \qquad L_2 \longrightarrow \frac{1}{2}L_2 \quad \text{puis } L_3 \longrightarrow L_3 - 3L_2$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1/2 & 0 \\ 0 & 0 & -2 & 1 & -3/2 & 1 \end{array}\right) \qquad L_3 \longrightarrow -\frac{1}{2}L_3 \quad \text{puis } L_2 \longrightarrow L_2 - L_3$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & -1/4 & 1/2 \\ 0 & 0 & 1 & -1/2 & 3/4 & -1/2 \end{array}\right) \qquad L_1 \longrightarrow L_1 - 2L_2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1/2 & -1 \\ 0 & 1 & 0 & 1/2 & -1/4 & 1/2 \\ 0 & 0 & 1 & -1/2 & 3/4 & -1/2 \end{array}\right)$$

L'inverse de $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 2 \\ -1 & 1 & 1 \end{pmatrix}$ est $A^{-1} = \begin{pmatrix} 0 & 1/2 & -1 \\ 1/2 & -1/4 & 1/2 \\ -1/2 & 3/4 & -1/2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 0 & 2 & -4 \\ 2 & -1 & 2 \\ -2 & 3 & -2 \end{pmatrix}$.

Applicaton du calcul d'inverse de matrice. Une application simple du calcul de l'inverse d'une matrice est pour la résolution de systèmes linéaires de n équations à n inconnues. En effet, tout système linéaire de n équations à n inconnues $x_1, x_2, \cdots, x_n$ avec second membre $b_1, b_2, \cdots, b_n$ peut se réécrire sous la forme $AX = B$ où X et B sont les matrices colonnes $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. Si la matrice A est inversible, on a alors un seul n-uplet solution donné par $X = A^{-1}B$. Si la matrice A n'est pas inversible, on a alors soit aucune solution, soit une infinité de solutions.

Exemple. Trouver l'ensemble des solutions de

$$\left\{\begin{array}{cccc} x_1 & +2x_2 & & = -2 \\ & 2x_2 & +2x_3 & = 4 \\ -x_1 & +x_2 & +x_3 & = 12 \end{array}\right.$$

Ce système se réécrit $AX = B$, avec $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 2 \\ -1 & 1 & 1 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $B = \begin{pmatrix} -2 \\ 4 \\ 12 \end{pmatrix}$.

Le calcul de A^{-1} donne $A^{-1} = \begin{pmatrix} 0 & 1/2 & -1 \\ 1/2 & -1/4 & 1/2 \\ -1/2 & 3/4 & -1/2 \end{pmatrix}$ donc, il existe un unique triplet solution du système:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A^{-1}B = \begin{pmatrix} 0 & 1/2 & -1 \\ 1/2 & -1/4 & 1/2 \\ -1/2 & 3/4 & -1/2 \end{pmatrix} \begin{pmatrix} -2 \\ 4 \\ 12 \end{pmatrix} = \begin{pmatrix} -10 \\ 4 \\ -2 \end{pmatrix}.$$

3.4. **Déterminant de matrice carrée - Méthode de calcul de déterminants.**

À toute matrice carrée $A \in \mathfrak{M}_n$, on peut associer un unique nombre, appelé déterminant, que l'on notera $\det(A)$ ou $|A|$.

La définition mathématique permet, comme il se doit, de caractériser de façon unique le déterminant. Cependant, cette définition ne donne pas directement une méthode de calcul. C'est pourquoi nous allons ci-dessous donner directement les méthodes permettant de calculer un déterminant de matrices carrées.

Définition 3.10 (Déterminants 1×1 et 2×2).
i) Soit $A \in \mathfrak{M}_1$ une matrice carrée d'ordre 1 *(ne contenant donc qu'un seul coefficient. $A = (a_{11})$. Alors,*

$$\det(A) = |A| = a_{11}.$$

ii) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}_2$ une matrice d'ordre 2*. Alors*

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb.$$

Proposition 3.11. *Le déterminant d'une matrice triangulaire inférieure ou supérieure d'ordre n est égal au produit de ses éléments diagonaux. C'est vrai aussi pour une matrice diagonale.*

Exemple. On a $\begin{vmatrix} 2 & 1 & -7 & 8 \\ 0 & 6 & 9 & 4 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2 \times 6 \times 3 \times 1 = 36.$

Définition 3.12 (sous-déterminants). *Soit $A \in \mathfrak{M}_n$ une matrice carrée d'ordre n. On dira que B est une sous-matrice de A si elle est obtenue en enlevant des lignes et/ou des colonnes à A.*

Si $A \in \mathfrak{M}_n$, on notera $A^*_{ij}* \in \mathfrak{M}_{n-1}$ la sous-matrice de A à qui on a ôté la ligne i et la colonne j.

Exemple. Si $A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ -1 & 0 & 2 & 2 \\ 9 & -1 & 1 & 1 \\ 2 & -5 & -2 & 4 \end{pmatrix}$ alors $A^*_{23} = \begin{pmatrix} 1 & 2 & 4 \\ 9 & -1 & 1 \\ 2 & -5 & 4 \end{pmatrix}$

Définition 3.13 (Déterminants $n \times n$). *Soit $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \in \mathfrak{M}_n$ une matrice carrée d'ordre n. Alors*

$$\det(A) = a_{1,1}|A^*_{11}| - a_{1,2}|A^*_{12}| + a_{1,3}|A^*_{13}| + \cdots + (-1)^{n+1}|A^*_{1n}|.$$

⚠ *Notez bien l'alternance des signes $+$ et $-$ dans cette formule.*

On a effectué ici ce que l'on appelle un développement par rapport à la première ligne de A.

Cette définition permet de ramener le calcul d'un déterminant d'ordre n au calcul de n déterminants d'ordre $n-1$. Par exemple, on calcule un déterminant d'ordre 3 à l'aide de 3 déterminants d'ordre 2:

$$\begin{vmatrix} 3 & -1 & 6 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} = 3 \times \begin{vmatrix} 2 & 0 \\ 0 & 5 \end{vmatrix} - (-1) \times \begin{vmatrix} 4 & 0 \\ 1 & 5 \end{vmatrix} + 6 \times \begin{vmatrix} 4 & 2 \\ 1 & 0 \end{vmatrix} = 3 \times 10 + 1 \times 20 + 6 \times (-2) = 38$$

Proposition 3.14. *On peut effectuer le calcul d'un déterminant $n \times n$ en développant par rapport à une autre ligne que la première ligne. La formule est similaire. Il faut toutefois faire attention à l'alternance des signes $+$ et $-$. Si on développe par rapport à une ligne impaire (comme la première, la troisième, etc.), on commence par un "$+$", et on alterne. Si on développe par rapport à une ligne paire (la deuxième, la quatrième, etc.), on commence alors par le signe "$-$", et on alterne.*

Ainsi, soit $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \in \mathfrak{M}_n$ *une matrice carrée d'ordre* n.
Alors en développant par rapport à la i*-ème ligne, on obtient:*

$$\det(A) = (-1)^{i+1}a_{i,1}|A^*_{i1}| + (-1)^{i+2}a_{i,2}|A^*_{i2}| + (-1)^{i+3}a_{i,3}|A^*_{i3}| + \cdots + (-1)^{i+n}a_{i,n}|A^*_{i,n}|.$$

Dans le développement du déterminant, on a une alternance de + et de −. Le premier signe est un + si on développe par rapport à un numéro de ligne ou de colonne impair et c'est un − sinon.

Voyons tout de suite ce que cela donne sur un exemple. On reprend la matrice précédente, mais on développe cette fois par rapport à la deuxième ligne. Faites bien attention. Les coefficients choisis changent, les signes changent, mais aussi les sous-matrices associées changent. Par contre, le résultat final pour le déterminant est le même.

$$\begin{vmatrix} 3 & -1 & 6 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} = -4 \times \begin{vmatrix} -1 & 6 \\ 0 & 5 \end{vmatrix} + 2 \times \begin{vmatrix} 3 & 6 \\ 1 & 5 \end{vmatrix} - 0 \times \begin{vmatrix} 3 & -1 \\ 1 & 0 \end{vmatrix} = (-4)\times(-5) + 2\times(15-6) = 38$$

On retrouve bien entendu la même valeur du déterminant.

Proposition 3.15. *On peut effectuer le calcul d'un déterminant* $n \times n$ *en développant par rapport à une colonne. Là encore, la formule est similaire au cas où on développe par rapport à une ligne. Si on développe par rapport à une colonne impaire (la première, la troisième, etc.), on commence par un "+", et on alterne. Si on développe par rapport à une colonne paire (la deuxième, la quatrième, etc.), on commence alors par le signe "−", et on alterne.*

Ainsi, soit $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \in \mathfrak{M}_n$ *une matrice carrée d'ordre* n.
Alors en développant par rapport à la j*-ème colonne, on obtient:*

$$\det(A) = (-1)^{j+1}a_{1,j}|A^*_{1j}| - (-1)^{j+1}a_{2,j}|A^*_{2j}| + (-1)^{j+1}a_{3,j}|A^*_{3j}| + \cdots + (-1)^n(-1)^{j+1}a_{n,j}|A^*_{nj}|.$$

Reprenons la matrice précédente, et développons par rapport à la 3ème colonne.

$$\begin{vmatrix} 3 & -1 & 6 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} = 6 \times \begin{vmatrix} 4 & 2 \\ 1 & 0 \end{vmatrix} - 0 \times \begin{vmatrix} 3 & -1 \\ 1 & 0 \end{vmatrix} + 5 \times \begin{vmatrix} 3 & -1 \\ 4 & 2 \end{vmatrix} = 6 \times (-2) + 5 \times 10 = 38$$

Une dernière série de propriétés utiles pour le calcul de déterminant.

Proposition 3.16. *Soit A et B deux matrices carrées de $\mathfrak{M}_n$.*

(1) *Si on intervertit deux lignes de A, on change son déterminant en son opposé.*
(2) *Si on intervertit deux colonnes de A, on change son déterminant en son opposé.*
(3) *Si on remplace une ligne de A par elle-même plus une combinaison linéaire des autres lignes, on ne change pas son déterminant.*
(4) *Si on remplace une colonne de A par elle-même plus une combinaison linéaire des autres colonnes, on ne change pas son déterminant.*
(5) *Soit $\lambda \in \mathbb{R}$ et $A \in \mathfrak{M}_n$, alors $\det(\lambda A) = \lambda^n A$.*
(6) *Si on multiplie une ligne (ou une colonne) de A par λ, alors on multiplie son déterminant par λ.*
(7) $\det(AB) = \det(A)\det(B)$
(8) *Si A a deux lignes identiques, alors $\det(A) = 0$.*
(9) *Si une ligne de A est combinaison linéaire des autres lignes, alors $\det(A) = 0$.*
(10) *Si A a deux colonnes identiques, alors $\det(A) = 0$.*
(11) *Si une colonne de A est combinaison linéaire des autres colonnes, alors $\det(A) = 0$.*
(12) *Le déterminant d'une matrice et de sa transposée son égaux:* $\det(A) = \det({}^tA)$.

Remarque. ⚠ *En général on a $\det(A+B) \neq \det(A) + \det(B)$.*

Exemples. Appliquons quelques-unes des propriétés ci-dessus.

$$\begin{vmatrix} 3 & -1 & 6 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} = -\begin{vmatrix} 4 & 2 & 0 \\ 3 & -1 & 6 \\ 1 & 0 & 5 \end{vmatrix} \quad \text{propriété (1).}$$

$$\begin{aligned}
\begin{vmatrix} 3 & -1 & 6 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} &= \begin{vmatrix} 0 & -1 & -9 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} \quad \text{(on a fait } L_1 \longrightarrow L_1 - 3L_3 \text{ et propriété (3))} \\
&= \begin{vmatrix} 0 & -1 & -9 \\ 0 & 2 & -20 \\ 1 & 0 & 5 \end{vmatrix} \quad \text{(on a fait } L_2 \longrightarrow L_2 - 4L_3 \text{ et propriété (3))} \\
&= 1 \times \begin{vmatrix} -1 & -9 \\ 2 & -20 \end{vmatrix} \quad \text{(développement par rapport à la 1ère colonne)} \\
&= -1 \times (-20) - 2 \times (-9) = 38.
\end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} 30 & -1 & 6 \\ 40 & 2 & 0 \\ 10 & 0 & 5 \end{vmatrix} &= \begin{vmatrix} 10\times 3 & -1 & -9 \\ 10\times 4 & 2 & 0 \\ 10\times 1 & 0 & 5 \end{vmatrix} \qquad (\text{ propriété } (6)) \\
&= 10 \begin{vmatrix} 3 & -1 & -9 \\ 4 & 2 & 0 \\ 1 & 0 & 5 \end{vmatrix} \qquad (\text{ propriété } (6)) \\
&= 10 \times 38 = 380.
\end{aligned}$$

$$\begin{vmatrix} 3 & 7 & 6 \\ 4 & 2 & 0 \\ 34 & 72 & 60 \end{vmatrix} = 0 \quad (\text{on remarque que } L_3 = 10L_1 + L_2 \text{ et propriété } (10))$$

Voici un théorème important sur l'inversibilité d'une matrice.

Théorème 3.17. *Soit $A \in \mathfrak{M}_n$. Alors:*

$$A \textit{ est inversible} \quad \Leftrightarrow \quad \det(A) \neq 0$$

La proposition suivante donne une propriété simple pour montrer qu'une famille de n vecteurs de $\mathbb{R}^n$ est une base.

Proposition 3.18. *Soit $(f_1, f_2, \cdots f_n)$, n vecteurs dans $\mathbb{R}^n$. Alors, $(f_1, f_2, \cdots f_n)$ est une base de $\mathbb{R}^n$ si et seulement si $\det(f_1, f_2, \cdots f_n) \neq 0$.*

Remarque. *Les propriétés précédentes donnent la règle suivante, appelée règle de Sarrus, pour calculer le déterminant de matrices* 3×3.

Le déterminant de la matrice

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

est égal à

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

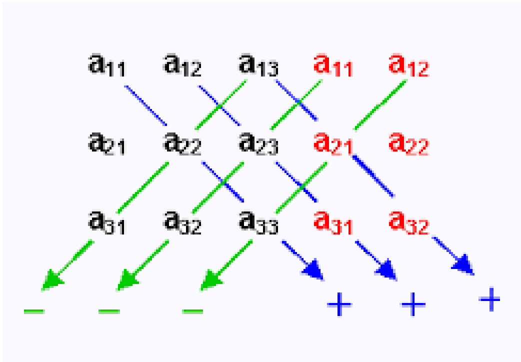

Figure 1. Règle de Sarrus

4. Fiche de cours: Application linéaires

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

Définition 4.1 (Application). *Soient E et F deux ensembles quelconques. On appelle application de E vers F, une fonction qui a tout élément de E associe un élément de F.*

Exemples. i) $E = \mathbb{R}$, $F = \mathbb{R}$ et $f(x) = \sin(x)$.
ii) $E = \mathbb{R}$, $F = \mathbb{R}$ et $g_1(x) = 3x$. Si on prend $g_2(x) = \sqrt{x}$, ce n'est pas une application sur E (seulement une fonction), car elle n'est pas définie sur tout E.
iii) $E = \mathbb{R}^2$, $F = \mathbb{R}$ et $h((x_1, x_2)) = 3x_1x_2 + {x_1}^2 \exp x_2$.
iv) $E = \mathbb{R}^2$, $F = \mathbb{R}$ et $h((x, y)) = 3y - 4x$.
v) $E = \mathbb{R}^3$, $F = \mathbb{R}^2$ et $j(\ (x_1,\ x_2,\ x_3)\ ) = (x_1 - 5x_2,\ x_3 + 9x_2 - x_1)$.

Définition 4.2 (Application linéaire). *Soit f une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$. On dit que f est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ si*

(a) *Pour tout u et v dans $\mathbb{R}^n$, on a $f(u + v) = f(u) + f(v)$.*
(b) *Pour tout $u \in \mathbb{R}^n$ et tout $\lambda \in \mathbb{R}$, $f(\lambda u) = \lambda f(u)$.*

L'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$ est noté $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$.

Proposition 4.3 (Application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$). *Avec la définition 4.2 ci-dessus, on peut montrer que toute application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$ est de la forme:*

$$F : \begin{array}{lcl} \mathbb{R}^n & \to & \mathbb{R} \\ (x_1, x_2, \cdots, x_n) & \mapsto & \alpha_1 x_1 + \alpha_2 x_2 + \cdots \alpha_n x_n \end{array} \tag{5}$$

où $(x_1, x_2, \cdots, x_n)$ est la variable de l'application, et les $\alpha_1, \alpha_2, \cdots, \alpha_n$ sont des constantes réelles.

Exemple. Dans les exemples d'applications i) à v) ci-dessus, dire celles qui sont des applications linéaires et celles qui ne le sont pas.

Proposition 4.4. *Si f est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$, alors on a*

(1) $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^p}$.
(2) Pour tout $u \in \mathbb{R}^n$, $f(-u) = -f(u)$.

où $0_{\mathbb{R}^n}$ est le vecteur nul $\underbrace{(0, 0, \cdots, 0)}_{n \text{ fois}}$ de $\mathbb{R}^n$ et $0_{\mathbb{R}^p}$ est le vecteur nul de $\mathbb{R}^p$.

Proposition 4.5. *Si f et g sont deux applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$, alors*

(1) $f + g$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$
(2) Pour tout $\lambda \in \mathbb{R}$, λf est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$.
(3) Si f est une application linéaire bijective de $\mathbb{R}^n$ dans $\mathbb{R}^n$, alors son application réciproque est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$.

A savoir et savoir faire pour les contrôles de connaissances: Algèbre linéaire

Au moins 80% de la partie de l'examen qui portera sur l'algèbre linéaire (chapitres 1,2,3 et 4) concernera les points suivants:

- Pour le Chapitre 1
  - Reconnaître une équation linéaire et un système d'équations linéaires, et savoir les réécrire, si nécessaire, sous la forme donnée dans les définitions (i.e., avec les termes constants au second membre).
  - Connaître le Théorème 1.6, et savoir qu'un système homogène admet toujours au moins la solution "nulle" comme solution.
  - Résoudre un système linéaire par la méthode du Pivot de Gauss.
- Pour le Chapitre 2
  - Connaître par coeur la Definition 2.1 de l'espace $\mathbb{R}^n$.
  - Connaître la Définition 2.6 d'une base, et, à l'aide de la Proposition 3.18, savoir si une famille de vecteurs de $\mathbb{R}^n$ forme une base.
- Pour le Chapitre 3
  - Savoir reconnaître une matrice carrée, triangulaire (supérieure ou inférieure), diagonale, identité, colonne, ligne.
  - Savoir additionner deux matrices et multiplier une matrice par un réel.
  - Pour deux matrices A et B, savoir si on peut calculer le produite AB, et savoir le calculer s'il est défini.
  - Savoir calculer l'inverse d'une matrice à l'aide de la méthode de Gauss-Jordan; savoir utiliser le calcul de l'inverse d'une matrice pour résoudre des systèmes du type $AX = B$ (A et B matrices): voir "Application du calcul d'inverse de matrice" en fin de paragraphe 3.3.
  - Savoir calculer un déterminant 2×2 et un déterminant 3×3.
  - À l'aide de la Proposition 3.18, savoir si une famille de vecteurs de $\mathbb{R}^n$ forme une base (voir chapitre 2).
- Pour le Chapitre 4
  - Aucune notion exigible cette année sur le chapitre 4.

ANALYSE

5. Fiche de cours: Fonctions de plusieurs variables

5.1. **Définition.**

Définition 5.1 (Fonctions de plusieurs variables). *Soit X un sous-ensemble de $\mathbb{R}^n$. On appelle application numérique, de n variables réelles, définie sur X, toute fonction f telle que pour tout $x = (x_1, x_2, x_3, \cdots, x_n) \in X$, $f(x)$ est bien défini, avec $f(x) \in \mathbb{R}$.*

Remarque 5.2. *i) Voir aussi la définition 4.1.*
ii) L'ensemble $X \subset \mathbb{R}^n$ sur lequel la fonction f est définie est appelé domaine de définition de f, et on le note $\mathcal{D}_f$.

Exemples: i) La fonction $f : (x, y, z) \to \sqrt{x}(y^2 - \frac{1}{z}) + \mathrm{e}^z$ est une application numérique sur $\mathbb{R}^3$ dont le domaine de définition est donné par les triplets (x, y, z) de $\mathbb{R}^3$ tels que $x \geq 0$ et $z \neq 0$. On a donc $\mathcal{D}_f = \{(x, y, z) \in \mathbb{R}^3 \mid x \geq 0 \text{ et } z \neq 0\}$.
ii) $f(x) = f((x_1, x_2)) = x_1^2 + \cos(x_1 x_2)$ est une application numérique sur $X = \mathbb{R}^2$.
iii) $f(x, y) = \frac{x}{y} - x^2 \mathrm{e}^y$ est une application numérique sur $X = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0\}$.
iv) Pour $f(x_1, x_2, x_3) = x_3 \ln(x_1) - x_2^3 + 3x_3 x_1^3$, $\mathcal{D}_f = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 > 0\}$.

Remarque 5.3. *i) Dans le cas d'une fonction de deux variables $f(x, y)$ de $X \subset \mathbb{R}^2$ dans $\mathbb{R}$, on peut visualiser dans $\mathbb{R}^3$ la fonction en représentant horizontalement la plan (Ox, Oy), et en traçant l'ensemble des points $M(x, y, z)$ dont la côte ou altitude est $z = f(x, y)$ (voir Figure 2 ci-dessous).*

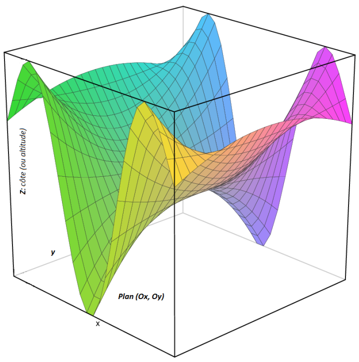

Figure 2. Représentation graphique de $f(x, y) = x^2 - \cos(x)y^2$

ii) On a déjà étudié des fonctions de plusieurs variables particulières: les applications linéaires. Le graphe d'une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ est un plan (voir Figure 3).

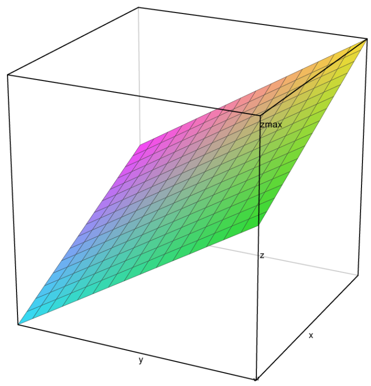

FIGURE 3. Représentation graphique de l'application linéaire $f(x_1, x_2) = 2x_1 - 3x_2$

5.2. **Continuité.**

DÉFINITION 5.4 (Continuité en un point). *Soit $f : X \subset \mathbb{R}^n \to \mathbb{R}$ une application numérique de n variables réelles. On dit que f est continue en $a \in X$ si*

$$\lim_{x \to a} f(x) = f(a).$$

DÉFINITION 5.5 (Continuité sur un ensemble). *Soit $f : X \subset \mathbb{R}^n \to \mathbb{R}$. On dit que f est continue sur $J \subset X$ si f est continue en tout point de J.*

REMARQUE 5.6. *i) Ces définitions sont identiques à celles pour les fonctions numériques d'une seule variable. ⚠ La différence est que, dans la définition 5.4, x et a sont des éléments de $\mathbb{R}^n$, i.e., $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ et $a = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n$. C'est cette définition que nous utiliserons pour montrer la continuité d'une fonction.*
ii) Toute la difficulté de cette définition est contenue dans la notion de limite. On donne ci-dessous une définition rigoureuse de la notion de limite, qui est valable aussi pour les fonctions d'une seule variable. On ne travaillera pas avec cette définition et elle ne sera pas exigible à l'examen. On utilisera plutôt des théorèmes généraux sur la continuité (voir les théorèmes 5.12 et 5.13).

DÉFINITION 5.7 (Distance euclidienne dans $\mathbb{R}^n$ et norme). *Soient M_1 et M_2 deux points dans $\mathbb{R}^n$ de coordonnées respectives $x = (x_1, x_2, \cdots, x_n)$ et $a = (a_1, a_2, \cdots, a_n)$. La distance euclidienne de M_1 à M_2 est*

$$d(M_1, M_2) = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots (x_n - a_n)^2} = \sqrt{\sum_{k=1}^{n} (x_k - a_k)^2}.$$

Pour $h = (h_1, h_2, \cdots, h_n) \in \mathbb{R}^n$*, la norme euclidienne de* h *est*

$$\|h\| = \sqrt{h_1^2 + h_2^2 + \cdots h_n^2} = \sqrt{\sum_{k=1}^{n} h_k^2}.$$

Définition 5.8 (Boules de rayon R). *Soit* $x_0 \in \mathbb{R}^n$*. On appelle boule de centre* x_0 *et de rayon* R*, que l'on note* $B_R(x_0)$*, l'ensemble des points de* $\mathbb{R}^n$ *qui sont à une distance euclidienne au plus* R *de* x_0*:*

$$B_R(x_0) = \{\ x \in \mathbb{R}^n \mid d(x, x_0) \leq R\}.$$

Remarque 5.9. *Dans* $\mathbb{R}^3$*, la boule de centre* $a \in \mathbb{R}^n$ *et de rayon* R *est la boule usuelle (contenue dans une sphère). Dans* $\mathbb{R}^2$*, c'est le disque de centre* a *et de rayon* R*.*

On peut maintenant donner la définition de limite en un point pour une fonction de plusieurs variables.

Définition 5.10 (Limite). *Soit* $F : \mathbb{R}^n \to \mathbb{R}$*, et soit* $a \in \mathbb{R}^n$*.*

On dit que f *admet une limite finie* $\ell \in \mathbb{R}$ *quand* x *tend vers* a*, si:*

Pour tout intervalle $]\ell - \epsilon, \ell + \epsilon[\subset \mathbb{R}$*, il existe* $\delta > 0$ *tel que*

$$x \in B_\delta(a) \Rightarrow f(x) \in]\ell - \epsilon,\ \ell + \epsilon[.$$

Remarque 5.11. *Bien que fondamentale en analyse, nous n'utiliserons pas cette définition de la limite pour montrer la continuité des fonctions. Au lieu de cela, nous travaillerons avec des fonctions "standards" dont nous connaîtrons les propriétés de continuité, ou dont nous déduirons les propriétés de continuité par des théorèmes usuels (somme, produit, quotient ou composition de fonctions continues).*

Théorème 5.12. *Soient* f *et* g *deux fonctions de* $\mathbb{R}^n \to \mathbb{R}$*, continues en* $a \in X$*. On a:*

(1) Pour tout λ *et* μ *dans* $\mathbb{R}$*,* $\lambda f + \mu g$ *est continue en* a*.*
(2) La fonction $f\,g$ *est continue en* a*.*
(3) Si $g(a) \neq 0$*, alors* $\dfrac{f}{g}$ *est continue en* a*.*

Théorème 5.13. *Soient* $f : \mathbb{R}^n \to \mathbb{R}$ *et* $g : \mathbb{R} \to \mathbb{R}$*. Si* f *est continue en* a *et si* g *est continue en* $f(a)$*, alors* $g \circ f$ *est continue en* a*.*

5.3. **Rappels fondamentaux sur les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.** Dans ce paragraphe, on rappelle la notion de dérivée d'une fonction de la variable réelle (fonction de $\mathbb{R}$ dans $\mathbb{R}$). Cela nous servira de comparaison, jusqu' un certain point, avec la notion de différentielle pour les fonctions de plusieurs variables (fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$).

Définition 5.14 (Dérivée en un point). *Soit $f : \mathbb{R} \to \mathbb{R}$. Alors f est dérivable en a si la limite suivante existe*

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a}. \tag{6}$$

On note alors $f'(a)$ cette limite.

Remarque 5.15. *o) Comme le dénominateur $x-a$ tend vers zéro quand x tend vers a, pour que cette limite existe, il faut au moins, même si cela ne suffit pas forcément, que le numérateur tende vers zéro, c'est à dire que $f(x)$ tende vers $f(a)$ quand x tend vers a, ce qui signifie que la fonction f est continue en a. On retrouve le résultat connu suivant: si f est dérivable en un point a alors nécessairement f est continue en a. La réciproque est fausse.*

i) La quantité $\frac{f(x)-f(a)}{x-a}$ est appelée taux d'accroissement entre x et a. Elle mesure la pente de la droite qui relie les points A de coordonées $(a, f(a))$ et M de coordonnées $(x, f(x))$.

ii) En réécrivant $x = a + h$, on obtient $x - a = h$ et donc:

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}. \tag{7}$$

iii) Une remarque importante pour la suite est que si f est dérivable en a, alors, la fonction $\epsilon(h) = \frac{f(a+h)-f(a)}{h} - f'(a)$ tend vers 0 quand h tend vers 0. On peut donc écrire, à partir de (7):

$$f(a+h) = f(a) + f'(a) \times h \;+\; h\epsilon(h) \qquad \text{avec } \lim_{h\to 0} \epsilon(h) = 0 \tag{8}$$

Notez bien que pour a fixé, l'application $h \mapsto f'(a) \times h$ est une application linéaire. L'égalité (8) s'interprète donc comme: "près du point a, l'application f s'écrit comme sa valeur au point a (c'est $f(a)$) + une application linéaire (c'est $f'(a) \times h$) + un reste "petit" devant l'application linéaire (c'est $h\epsilon(h)$).

iv) En revenant à la notation avec x au lieu de h dans (8), on obtient

$$f(x) = \underbrace{f(a) + f'(a) \times (x-a)}_{\text{donne l'équation de la droite tangente en } a} + \;(x-a)\epsilon(x-a) \text{ avec } \lim_{x\to a} \epsilon(x-a) = 0 \tag{9}$$

L'équation de la droite tangente à $\mathcal{C}_f$ en a est: $y = f(a) + f'(a) \times (x-a)$.

iv) Dans la limite où x tend vers a, c'est à dire dans la limite où h tend vers 0, la droite qui relie les points A et M se rapproche de la droite tangente à $\mathcal{C}_f$ en a. La

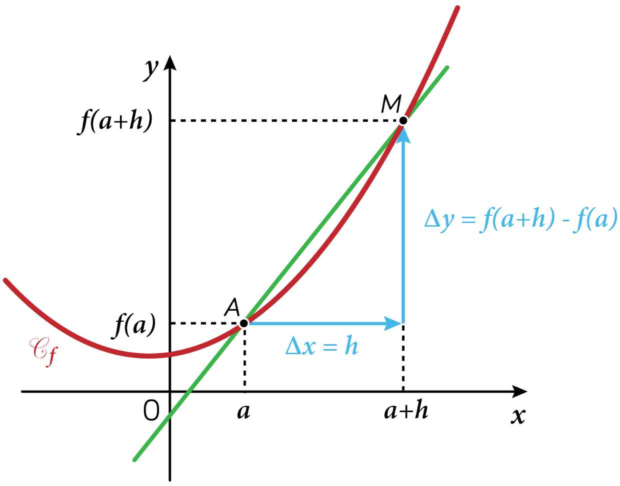

FIGURE 4. Taux d'accroissement

valeur de a étant donnée, a, $f(a)$ et $f'(a)$ sont connues. Les inconnues de l'équation $y = f(a) + f'(a) \times (x - a)$ sont x et y. L'ensemble des couples (x, y) solutions de cette équation sont représentés par une droite dans le plan.

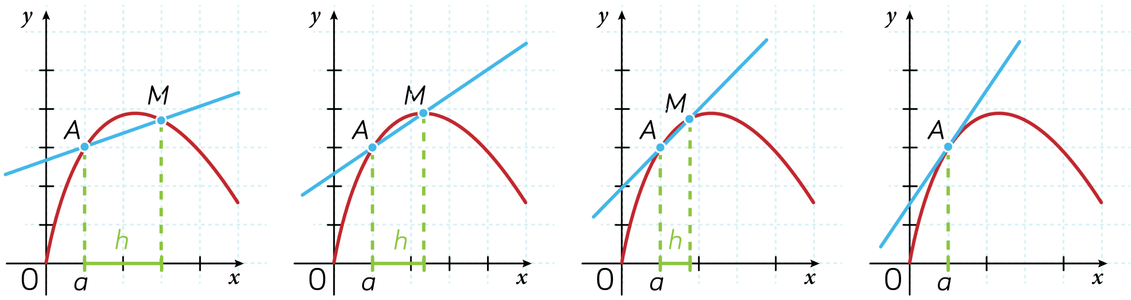

FIGURE 5. Tangente et dérivée

Le taux d'accroissement $\frac{f(a+h)-f(a)}{h}$ *donne la pente de la droite qui relie les points* $M\begin{pmatrix} a+h \\ f(a+h) \end{pmatrix}$ et $A\begin{pmatrix} a \\ f(a) \end{pmatrix}$. *La pente de la droite tangente à $\mathcal{C}_f$ en A vaut $f'(a)$ qui est la limite quand $h \to 0$ du taux d'accroissement* $\frac{f(a+h)-f(a)}{h}$.

DÉFINITION 5.16. *Soit $f : \mathbb{R} \to \mathbb{R}$. Alors f est dérivable sur $I \subset \mathbb{R}$ si f est dérivable en tout point de I.*

On a les théorèmes usuels suivants:

Théorème 5.17. *Soit $f : \mathbb{R} \to \mathbb{R}$. Si f est dérivable en a alors f est continue en a*

Remarque 5.18. *En général, la réciproque est fausse. Donc pour qu'une fonction soit dérivable en un point a, cela ne suffit pas de dire qu'elle est continue en a.*

Théorème 5.19. *Soient f et g deux fonctions de $\mathbb{R} \to \mathbb{R}$, dérivables en $a \in \mathbb{R}$. On a:*

(1) Pour tout λ et μ dans $\mathbb{R}$, $\lambda f + \mu g$ est dérivable en x_0 et $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$.
(2) La fonction $f\,g$ est dérivable en a et $(f\,g)'(a) = f'(a)g(a) + f(a)g'(a)$.
(3) Si $g(a) \neq 0$, alors $\dfrac{f}{g}$ est dérivable en a et

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{g^2(a)}$$

Théorème 5.20. *Soient $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$. Si f est dérivable en a et si g est dérivable en $f(a)$, alors $f \circ g$ est dérivable en a et*

$$(f \circ g)'(a) = g'(a)\, f'(g(a)).$$

5.4. **Différentielle - Dérivées partielles.** Dans ce paragraphe, nous allons définir la différentielle d'une fonction de plusieurs variables réelles, qui généralise la notion de dérivée de fonction de $\mathbb{R}$ dans $\mathbb{R}$.

Pour définir la différentielle en un point $A \in \mathbb{R}^n$ d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$, on ne peut pas reprendre l'expression (7) $\lim_{h\to 0} \frac{f(A+h)-f(A)}{h}$, parce-que cela n'aurait pas de sens de diviser par h qui est maintenant un élément de $\mathbb{R}^n$.

Définition 5.21 (Différentielle). *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables réelles. On dit que f est différentiable en $A = (a_1, a_2, \cdots, a_n)$ s'il existe une application linéaire $L_A(h)$ de $\mathbb{R}^n$ dans $\mathbb{R}$ telle que*

$$\lim_{h\to 0_{\mathbb{R}^n}} \frac{f(A+h) - f(A) - L_A(h)}{\|h\|} = 0$$

L'application linéaire L_A sera notée $\mathrm{d}f(A)$ et on l'appelle différentielle de f en A. On écrira $L_A(h) = \mathrm{d}f(A).h$

Remarque 5.22. *i) L'analogie avec le cas d'une seule variable réelle est qu'au lieu d'avoir*

$$f(a+h) = f(a) + f'(a) \times h + h\epsilon(h)$$

où $\epsilon(h)$ est une fonction qui tend vers 0 quand h tend vers 0, on a

$$f(A+h) = f(A) + \mathrm{d}f(A).h + \|h\|\epsilon(\|h\|). \tag{10}$$

Dans le cas $\mathbb{R}^2$*, si on remplace* $h = (h_1, h_2) \in \mathbb{R}^2$ *par* $x - A \in \mathbb{R}^2$*, c'est à dire on pose* $x = (x_1, x_2) = A + h = (a_1 + h_1, a_2 + h_2)$*, et si on pose* $y = f(x) \in \mathbb{R}$ *dans* (10)*, on obtient:*

$$y = f(A) + \mathrm{d}f(A).(x - A)$$

qui donnera l'équation du plan tangent au graphe de f *au point* $(A, f(A))$*.*

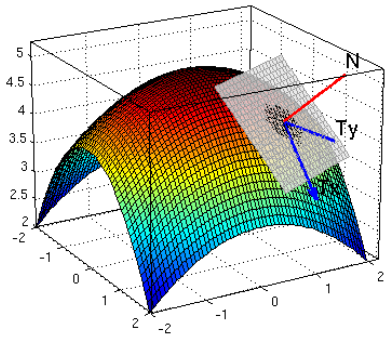

Figure 6. Plan tangent dans $\mathbb{R}^2$

ii) Les notations $df(A)$ *et* $df(A).h$ *peuvent sembler ici sibyllines. Mais vous connaissez déjà ces objets mathématiques. La notation* $df(A).h$ *n'est rien d'autre qu'une expression de la forme* (5)*:*

$$df(A).h = \alpha_1 h_1 + \alpha_2 h_2 + \cdots + \alpha_n h_n$$

où les variables sont $h_1, h_2, \cdots, h_n$*, (au lieu de* $x_1, x_2, \cdots, x_n$ *dans* (5)*) et* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *sont des constantes. L'expression* $df(A)$ *est alors l'application linéaire de* $\mathbb{R}^n$ *dans* $\mathbb{R}$ *qui à tout* $h = (h_1, h_2, \cdots, h_n)$ *associe le réel* $df(A).h = \alpha_1 h_1 + \alpha_2 h_2 + \cdots + \alpha_n h_n$*. L'application* $df(A)$ *est donc le même objet que l'application linéaire* F *dans* (5)*.*

Dans ce qui suit, nous allons voir comment déterminer facilement l'application linéaire $\mathrm{d}f(A)$.

Définition 5.23 (Dérivées partielles). *Soit* $f : \mathbb{R}^n \to \mathbb{R}$*. On appelle dérivée partielle de la fonction* $f(x) = f(x_1, x_2, \cdots, x_n)$ *par rapport à la i-ème variable, en* $A = (a_1, a_2, \cdots, a_n)$*, la quantité*

$$\lim_{h \to 0} \frac{f(a_1, a_2, \cdots, a_i + h, \cdots, a_n) - f(a_1, a_2, \cdots, a_i, \cdots, a_n)}{h}$$

On notera cette quantité

$$\frac{\partial f}{\partial x_i}(a_1, a_2, \cdots, a_n)$$

La fonction dérivée partielle, qui est une fonction des n variables $x_1, x_2, \cdots x_n$ est notée

$$\frac{\partial f}{\partial x_i} \quad ou \quad \frac{\partial f}{\partial x_i}(x_1, x_2, \cdots, x_n)$$

Remarque 5.24. *Dans le cas $\mathbb{R}^2$, on a donc deux fonctions dérivées partielles:*

$$\frac{\partial f}{\partial x_1}(a_1, a_2) = \lim_{h \to 0} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}$$

et

$$\frac{\partial f}{\partial x_2}(a_1, a_2) = \lim_{h \to 0} \frac{f(a_1, a_2 + h) - f(a_1, a_2)}{h}$$

La première dérivée partielle donne la pente de la droite tangente au graphe de f dans la direction x. L'accroissement est uniquement dans la première variable: on passe de $x_1 = a_1$ à $x_1 = a_1 + h$. La valeur de x_2 reste fixée à $x_2 = a_2$ (on dit parfois "gelée").

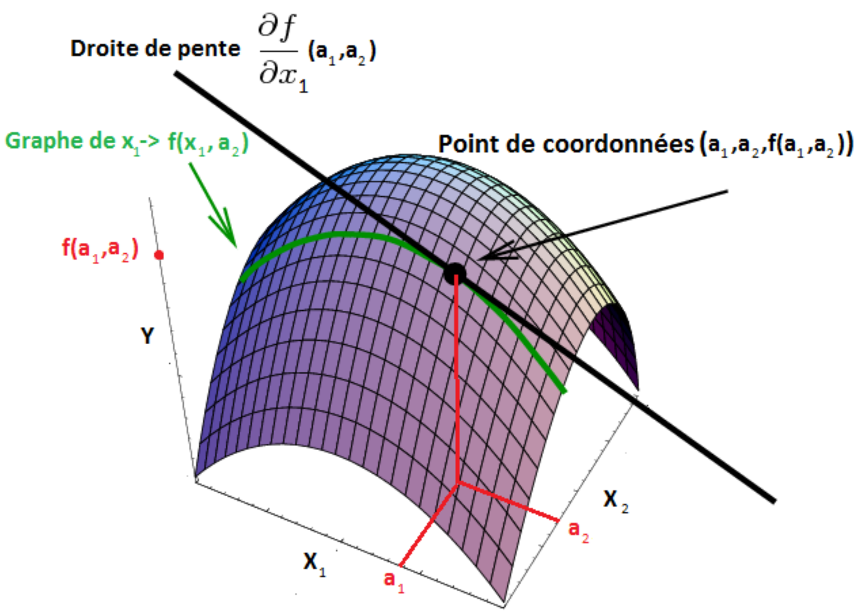

Figure 7. Dérivées partielles - droite tangente dans $\mathbb{R}^2$

Dans la Figure 7, la courbe en vert est celle obtenue lorsqu'on fixe x_2 à la valeur a_2, et que l'on fait varier x_1. C'est donc le tracé de la courbe de la fonction $g : \mathbb{R} \to \mathbb{R}$, $g(x_1) = f(x_1, x_2)$. La droite en noir est alors la tangente à cette courbe verte, au point a_1 . La pente de cette tangente est la dérivée de la fonction $g(\cdot)$ au point a_1:

$$\begin{aligned} g'(a_1) &= \lim_{h\to 0} \frac{g(a_1+h) - g(a_1)}{h} \\ &= \lim_{h\to 0} \frac{f(a_1+h, a_2) - f(a_1, a_2)}{h} \\ &= \frac{\partial f}{\partial x_1}(a_1, a_2). \end{aligned}$$

On peut faire le même raisonnement si on fixe la valeur de x_1 à a_1, et si on fait varier la valeur de x_2 de a_2 à a_2+h. On trouvera alors la dérivée partielle par rapport à la variable x_2 en (a_1, a_2).

Dans la Figure 8, la droite en bleu est la tangente à la courbe obtenue à partir de $f(x_1, x_2)$ en faisant varier x_1 près de a_1 et en fixant x_2 à la valeur a_2. La pente de cette droite bleue est la dérivée partielle de f en (a_1, a_2) par rapport à la première variable. La droite en rouge est la tangente à la courbe obtenue à partir de $f(x_1, x_2)$ en fixant x_1 à la valeur a_1 et en faisant varier x_2 près de a_2. La pente de cette droite rouge est la dérivée partielle de f en (a_1, a_2) par rapport à la seconde variable.

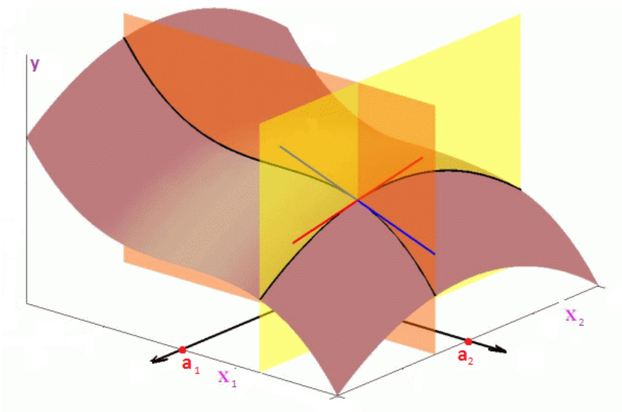

FIGURE 8. Dérivées partielles - droite tangente dans $\mathbb{R}^2$

Les deux droites engendrent un plan, qui est le plan tangent au graphe de f en (a_1, a_2) – on dit aussi au point $(a_1, a_2, f(a_1, a_2))$ – et dont l'équation permet de retrouver la différentielle de f en (a_1, a_2).

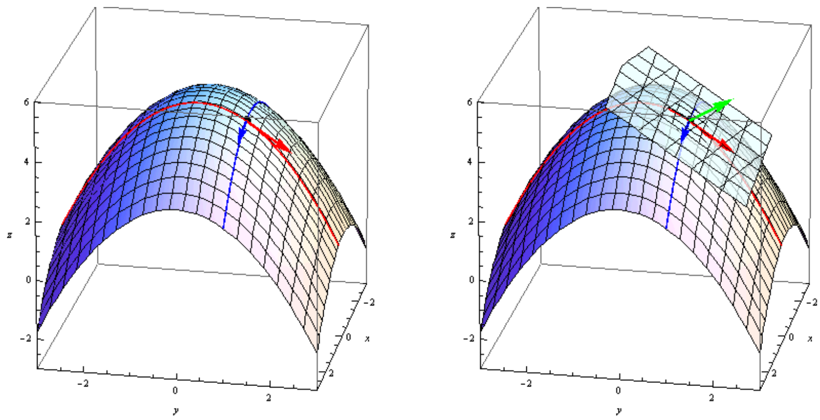

FIGURE 9. Plan tangent

PROPOSITION 5.25.
a) On suppose que f est différentiable en $a \in \mathbb{R}^n$. Alors toutes les dérivées partielles de f existent au point a et pour tout $h = (h_1, h_2, \cdots, h_n) \in \mathbb{R}^n$,

$$
\begin{aligned}
df(a).h &= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \times h_k \\
&= \frac{\partial f}{\partial x_1}(a) \times h_1 \; + \; \frac{\partial f}{\partial x_2}(a) \times h_2 \; + \; \cdots \; + \; \frac{\partial f}{\partial x_n}(a) \times h_n
\end{aligned} \tag{11}
$$

b) Si f est telle que toutes ses dérivées partielles en $A \in \mathbb{R}^n$ existent et sont continues, alors la différentielle de f en A existe et est donnée par (11).

REMARQUE 5.26. *Dans le cas $\mathbb{R}^2$, on obtient, pour $x = (x_1, x_2)$:*

$$
df(a).h = \frac{\partial f}{\partial x_1}(a) \times h_1 \; + \; \frac{\partial f}{\partial x_2}(a) \times h_2 \tag{12}
$$

On trouve aussi la notation suivante:

$$
\begin{aligned}
df &= \frac{\partial f}{\partial x_1} dx_1 \; + \; \frac{\partial f}{\partial x_2} dx_2 \\
&= f'_{x_1} dx_1 \; + \; f'_{x_2} dx_2.
\end{aligned}
$$

Ces écritures se généralisent aisément dans le cas $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$.

DÉFINITION 5.27 (Dérivées partielles d'ordre supérieur). *On définit les dérivées secondes suivantes*

$$
\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) \quad \text{et} \quad \frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right)
$$

De façon similaire, on définit les dérivées partielles d'ordre quelconque.

Exemples:
i) Soit $f(x_1,x_2)=x_1x_2^2+x_1^2$. On a $\dfrac{\partial f}{\partial x_1}(x_1,x_2)=x_2^2+2x_1$. Pour obtenir ce résultat, on a considéré dans l'expression $f(x_1,x_2)=x_1x_2^2+{x_1}^2$ que x_1 est la seule variable et que x_2 est un paramètre fixé.

De même, on obtient $\dfrac{\partial f}{\partial x_2}(x_1,x_2)=x_1\times 2x_2+0=2x_1x_2$.

Et on a:
$$\frac{\partial^2 f}{\partial x_1\partial x_2}=\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right)=\frac{\partial}{\partial x_1}\left(2x_1x_2\right)=2x_2$$
$$\frac{\partial^2 f}{\partial x_2\partial x_1}=\frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)=\frac{\partial}{\partial x_2}\left(x_2^2+2x_1\right)=2x_2$$
$$\frac{\partial^2 f}{\partial x_1^2}=\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right)=\frac{\partial}{\partial x_1}\left(x_2^2+2x_1\right)=2$$
$$\frac{\partial^2 f}{\partial x_2^2}=\frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_2}\right)=\frac{\partial}{\partial x_2}\left(2x_1x_2\right)=2x_1$$

ii) Soit $f(x,y)=\sin(xy)+y\mathrm{e}^{2x+y}$. On obtient $\dfrac{\partial f}{\partial x}(x,y)=y\cos(xy)+2y\mathrm{e}^{2x+y}$ et $\dfrac{\partial f}{\partial y}(x,y)=x\cos(xy)+\mathrm{e}^{2x+y}+y\mathrm{e}^{2x+y}$.

Remarque 5.28. *Dans l'exemple i) ci-dessus, on constate que* $\frac{\partial^2 f}{\partial x_1\partial x_2}=\frac{\partial^2 f}{\partial x_2\partial x_1}$. *On peut montrer que c'est toujours vrai si f vérifie certaines conditions.*

Théorème 5.29 (Théorème de Schwarz). *Soit $f: X\subset\mathbb{R}^n\to\mathbb{R}$ (et X un ouvert). Si les dérivées partielles d'ordre 2 de f existent et sont continues, alors*
$$\frac{\partial^2 f}{\partial x_i\partial x_j}=\frac{\partial^2 f}{\partial x_j\partial x_i}$$

Définition 5.30 (Matrice Hessienne). *Soit $f:\mathbb{R}^n\to\mathbb{R}$, telle que les dérivées secondes soient toutes bien définies. On appelle matrice Hessienne de f, notée $H(f)$ la matrice carrée d'ordre n dont le coefficient d'indice i et j est égal à $\frac{\partial^2 f}{\partial x_i\partial x_j}$. On a donc*
$$H(f)=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}$$
On notera la matrice Hessienne de f au point $x=a$: $H(f)_{|x=a}$.

Exemple: Pour $f(x,y) = xy^2 + x^2$, on a

$$H(f) = \begin{pmatrix} 2 & 2y \\ 2y & 2x \end{pmatrix}$$

et

$$H(f)_{|(x,y)=(2,-1)} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$$

5.5. **Fonctions implicites.** Nous allons voir dans ce paragraphe la notion de fonction implicite. On se limitera au cas où la fonction implicite est donnée à partir d'une fonction de deux variables.

Quand on se donne une fonction $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto y = f(x)$, on peut l'étudier directement comme fonction de la variable x. Dans ce cas, y est donnée *explicitement* par rapport à x. C'est le cas usuel que vous avez étudié au semestre 1. Il peut arriver que la façon dont varie y en fonction de x ne soit pas donnée directement, mais par une relation plus compliquée qui lie les variables x et y.

Exemple: Soit la fonction $F : (x,y) \mapsto F(x,y) = \cos y + x - \frac{1}{2}$. Alors, pour $x_0 = 0$ et $y_0 = \frac{\pi}{3}$, on a $F(x_0,y_0) = \cos(\frac{\pi}{3}) + 0 - \frac{1}{2} = 0$. Pour tout x dans $[-\frac{1}{2}, \frac{3}{2}]$, on a $\frac{1}{2} - x$ qui est dans l'intervalle $[-1,1]$, et donc pour tout x fixé dans $[-\frac{1}{2}, \frac{3}{2}]$, l'équation d'inconnue y: $\cos(y) + x - \frac{1}{2} = 0$, i.e., $\cos(y) = \frac{1}{2} - x$, admet bien une solution et une seule y dans $[0,\pi]$. Pour x donné dans $[-\frac{1}{2}, \frac{3}{2}]$, on a donc y qui est donnée de façon unique dans $[0,\pi]$.

Plus généralement, on sera amené à considérer des fonctions implicites lorsque l'on se donne une fonction de 2 variables $F(x,y)$, telle que $F(x_0,y_0) = a_0$ en un point (x_0,y_0), et qu'on cherche, pour toute valeur de x donnée "près" de x_0, la (ou les) valeurs de y "près" de y_0 telles que $F(x,y) = a_0$. Si on observe la figure 10, et qu'on se fixe un point (x_0,y_0) et la valeur associée $F(x_0,y_0)$, alors on peut visualiser pour x près de x_0, la valeur de y telle que $F(x,y) = F(x_0,y_0)$, i.e., on se déplace sur une courbe qui passe par (x_0,y_0) et qui reste à la même hauteur (altitude, côte) $F(x_0,y_0)$. Dans ces deux cas, on peut bien pour chaque x trouver un y associé, c'est à dire on trouve $y = f(x)$ en fonction de x. Cependant, en général, on n'obtient pas d'expression *explicite* pour f.

Remarquez aussi que si on prend $F(x,y)$ représentée par la figure 10 et qu'on choisit (x_0,y_0) telle que $F(x_0,y_0)$ soit à un maximum, alors, quand x varie près de x_0, on ne trouve pas de valeur y_0 telle que $F(x,y)$ reste à la valeur maximum. On pourrait donner aussi un exemple pour lequel, pour chaque x proche de x_0, il existe plusieurs valeurs associées y telles que $F(x,y) = F(x_0,y_0)$.

Une fonction implicite n'est donc pas toujours bien définie.

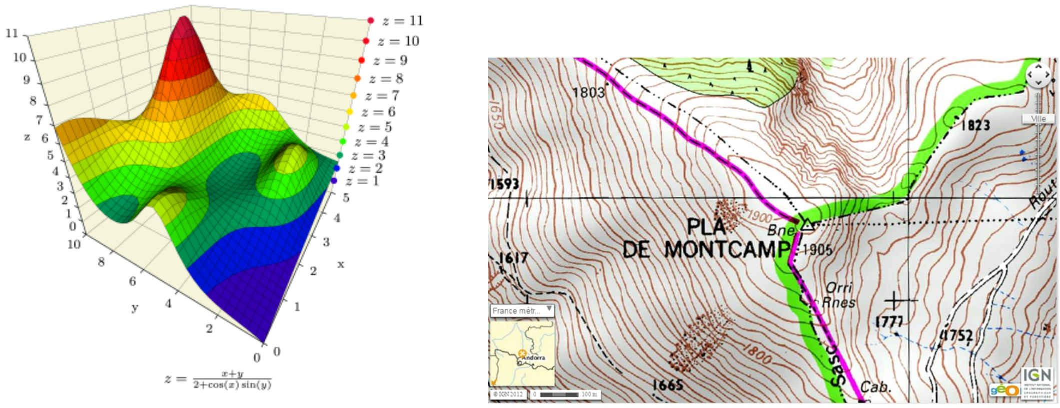

FIGURE 10. Fonctions implicites

Dans ce qui suit, nous allons donner des conditions pour qu'une fonction implicite donnée à partir d'une fonction de deux variables soit bien définie.

THÉORÈME 5.31. *Soit* $F : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto F(x, y)$, *et soit* $(x_0, y_0) \in \mathbb{R}^2$ *tel que* $F(x_0, y_0) = 0$.

On suppose que F *est continue sur un voisinage de* (x_0, y_0) *et on suppose de plus que* F *admet des dérivées partielles* $\frac{\partial F}{\partial x}$ *et* $\frac{\partial F}{\partial y}$ *continues sur un voisinage de* (x_0, y_0). *On suppose aussi que* $\frac{\partial F}{\partial y}$ *ne s'annule pas en* (x_0, y_0).

Alors il existe une fonction $f : x \mapsto y = f(x)$, *unique, continue, définie sur un voisinage* $]x_0-\epsilon, x_0+\epsilon[$ *de* x_0, *telle que pour tout* $x \in]x_0-\epsilon, x_0+\epsilon[$, *on a* $F(x, f(x)) = 0$.

REMARQUE. *La proposition précédente reste valable si on a* $F(x_0, y_0) = a \neq 0$ *et que la fonction implicite* f *est donnée par l'équation* $F(x, f(x)) = a$.

Alors qu'une fonction implicite f, même si elle existe, n'a pas forcément d'expression explicite en x, il est possible de déterminer l'expression de sa dérivée au point x_0. Ceci ne suffit pas à reconstruire la fonction f par intégration puisqu'on n'a seulement la valeur de sa dérivée en un point, et pas en tout x.

PROPOSITION 5.32. *Sous les hypothèses du théorème précédent, la fonction implicite* f *définie par l'équation* $F(x, f(x)) = 0$, *est dérivable en* x_0 *et*

$$f'(x_0) = -\frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}.$$

Exemple: Soit $F(x, y) = e^{xy} - x^3 + y$. Soit $(x_0, y_0) = (1, 0)$. Alors $F((x_0, y_0)) = F((1, 0)) = 1 - 1 + 0 = 0$. La fonction F est bien continue au voisinage de $(1, 0)$, et admet des dérivées partielle

$$\frac{\partial F}{\partial x} = ye^{xy} - 3x^2, \quad \frac{\partial F}{\partial y} = xe^{xy} + 1$$

qui sont continues au voisinage de $(1,0)$. De plus

$$\frac{\partial F}{\partial y}(x_0, y_0) = x_0 e^{x_0 y_0} + 1 = 2 \neq 0$$

On peut donc appliquer le théorème des fonctions implicites: il existe une fonction $f :]x_0 - \epsilon, x_0 + \epsilon[\to \mathbb{R}$, $x \mapsto y = f(x)$ telle que pour tout $x \in]x_0 - \epsilon, x_0 + \epsilon[$, on a $F(x, f(x)) = 0$.

On a de plus

$$f'(x_0) = f'(1) = -\frac{y_0 e^{x_0 y_0} - 3x_0^2}{x_0 e^{x_0 y_0} + 1} = -\frac{-3}{2} = \frac{3}{2}.$$

6. Calcul d'extrema libres

6.1. **Formule de Taylor.** Dans ce paragraphe, nous allons donner une technique de calcul qui permet d'approximer en un point $a = (x_1, x_2, \cdots, x_n)$ de $\mathbb{R}^n$ donné une fonction de plusieurs variables $F : \mathbb{R}^n \to \mathbb{R}$ par un polynôme de degré $k = 1$ ou 2 dans les variables $x_1, x_2, \cdots x_n$. L'intérêt d'une telle approximation est de pouvoir étudier plus facilement, près du point a, le comportement de la fonction F, en étudiant le polynôme de degré 1 ou 2 qui approxime la fonction F.

Dans le cas d'une fonction f d'une seule variable $x \in \mathbb{R}$, on a déjà vu qu'on pouvait approximer en un point a une fonction par un polynôme de degré $k = 1$ (équation de droite). Voir les équations (8) et (9). D'un point de vue graphique, cela revient à approximer la courbe de f près de a par une droite.

Comme on l'a déjà vu, ce type d'approximation n'est valable que si x est "proche" de a. Quand x s'éloigne de a, il y a toutes les chances que cette approximation n'ait plus aucun sens, ce qui revient à dire, dans les formules (8) et (9), que la fonction $\epsilon(h)$ ne tend pas forcément vers 0 quand h tend vers une autre valeur que 0.

On commence par donner la formule de Taylor pour une fonction d'une variable, dans le cas $k = 2$, ce qui revient à dire qu'on donne un moyen de calculer un polynôme de degr{e 2 qui approxime une fonction donnée en un point donné.

Theorem 6.1 (Formule de Taylor à l'ordre 2 pour une fonction d'une variable). *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue, deux fois dérivable et à dérivée seconde continue sur un intervalle I centré en a, alors, pour tout h tel que $a + h \in I$, on a*

$$f(a+h) = \underbrace{f(a) + hf'(a) + \frac{h^2}{2} f''(a)}_{\textit{Polynôme de degré 2 dans la variable } h} + h^2 \epsilon(h),$$

où $\lim_{h\to 0} \epsilon(h) = 0.$

Ceci s'écrit aussi:

Pour tout $x \in I$,

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2} f''(a) + (x-a)^2\epsilon(x-a),$$

où $\lim_{x\to a} \epsilon(x-a) = 0.$

Remarque. *L'approximation en un point a d'une fonction f par un polynôme de degré 2, dont la courbe représentative est une parabole, est plus précise que l'approximation par un polynôme de degré 1, dont la courbe représentative est une droite, qui est la droite tangente à f en a.*

Pour des fonctions de plusieurs variables, on a déjà vu le développement de Taylor à l'ordre $k = 1$. Il est donné par les équations (10), où la différentielle est donnée par les dérivées partielles dans (12).

Cela donne donc le théorème suivant

Théorème 6.2 (Développement de Taylor à l'ordre 1 pour une fonction de plusieurs variables). *Soit F une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, dont toutes les dérivées partielles premières sont définies et continues au voisinage d'un point $a \in \mathbb{R}^n$. Alors on a pour $h \in \mathbb{R}^n$,*

$$\begin{aligned} F(a+h) &= \underbrace{F(a) \; + \; \frac{\partial F}{\partial x_1}(a)h_1 + \frac{\partial F}{\partial x_2}(a)h_2 + \cdots + \frac{\partial F}{\partial x_n}(a)h_n}_{\textit{Polynôme d'ordre } 1 \textit{ dans les variables } h_i} + \|h\|\epsilon(h) \\ &= F(a) \; + \; \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(a)h_j + \|h\|\epsilon(h) \end{aligned}$$

De même, on a le résultat suivant:

Théorème 6.3 (Développement de Taylor à l'ordre 2 pour une fonction de plusieurs variables). *Soit F une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, dont toutes les dérivées partielles premières et secondes sont définies et continues au voisinage d'un point $a \in \mathbb{R}^n$. Alors on a pour $h \in \mathbb{R}^n$,*

$$F(a+h) = \underbrace{F(a) \; + \; \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(a)h_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k}(a)h_j\, h_k}_{\textit{Polynôme d'ordre } 2 \textit{ dans les variables } h_i} + \|h\|^2\epsilon(h)$$

Exemple: i) Cas d'une fonction d'une seule variable. Soit $f(x) = \frac{1}{1-x}$. Calculons le développement de Taylor à l'ordre 2 en $a = 0$. On commence par calculer la dérivée première et la dérivée seconde. On a

$$f(0) = 1, \quad f'(x) = \frac{1}{(1-x)^2}, \quad f'(0) = 1, \quad f''(x) = \frac{2(1-x)}{(1-x)^4}, \quad f''(0) = 2$$

La fonction f est bien deux fois fois dérivable et à dérivée second continue au voisinage de 0. On a donc le développement suivant:

$$\begin{aligned}\frac{1}{1-h} &= 1 + h \times 1 + \frac{h^2}{2} \times 2 + h^2\epsilon(h)\\ &= 1 + h + h^2 + h^2\epsilon(h)\end{aligned}$$

ii) Cas d'une fonction d'une seule variable. Soit $g(x) = \sin(x)$. Calculons le développement de Taylor à l'ordre 2 en $a = 0$ de cette fonction. On commence par calculer les dérivées d'ordre 1 et 2. On a

$$g(0) = 1,\ g'(x) = \cos(x),\ g'(0) = \cos(0) = 1,\ g''(x) = -\sin(x),\ g''(0) = -\sin(0) = 0.$$

La fonction g est bien deux fois fois dérivable et à dérivée second continue au voisinage de 0. On a donc le développement suivant:

$$\begin{aligned}\sin(0+h) = \sin(h) &= \sin(0) + h\cos(0) + \frac{h^2}{2}(-\sin(0)) + h^2\epsilon(h)\\ &= 0 + h + 0 + h^2\epsilon(h) = h + h^2\epsilon(h).\end{aligned}$$

ii) Cas d'une fonction de deux variables. Soit $F(x, y) = e^{xy} - x^3 + y$. Calculons le développement de Taylor à l'ordre 2 au point $a = (1, 0)$. Pour cela on va calculer les dérivées partielles d'ordre 1 et d'ordre 2. On a déjà calculé les dérivées d'ordre 1:

$$\frac{\partial F}{\partial x} = ye^{xy} - 3x^2, \quad , \frac{\partial F}{\partial x}(1,0) = -3, \quad \frac{\partial F}{\partial y} = xe^{xy} + 1, \quad , \frac{\partial F}{\partial y}(1,0) = 2$$

et on calcule les dérivées partielles d'ordre 2:

$$\frac{\partial^2 F}{\partial x^2} = \frac{\partial F}{\partial x} ye^{xy} - 3x^2 = y^2e^{xy} - 6x,\ \frac{\partial^2 F}{\partial x^2}(1,0) = -6$$

$$\frac{\partial^2 F}{\partial y^2} = \frac{\partial F}{\partial y} xe^{xy} + 1 = x^2e^{xy},\ \frac{\partial^2 F}{\partial y^2}(1,0) = 1$$

$$\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial F}{\partial y} ye^{xy} - 3x^2 = e^{xy} + xye^{xy},\ \frac{\partial^2 F}{\partial y \partial x}(1,0) = 1$$

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial F}{\partial x} xe^{xy} + 1 = e^{xy} + xye^{xy},\ \frac{\partial^2 F}{\partial x \partial y}(1,0) = 1$$

On obtient donc le développement de Taylor suivant

$$\begin{aligned} & e^{(1+h_1)h_2} - (1+h_1)^3 + h_2 \\ & = 0 - 3h_1 + 2h_2 + \frac{1}{2}\left((-6)\times h_1^2 + 1\times h_2^2 + 1\times h_1\,h_2 + 1\times h_2\,h_1\right) + \|h\|^2\epsilon(h) \\ & = \left(-3h_1 + 2h_2 - 3h_1^2 + \frac{1}{2}h_2^2 + 2h_1\,h_2\right) + \|h\|^2\epsilon(h) \end{aligned}$$

6.2. **Recherche d'extrema de fonctions de plusieurs variables.** Pour trouver les extrema (maximum, minimum) d'une fonction numérique de plusieurs variables, on va procéder, pour certains aspects, de manière analogue au cas de fonctions d'une seule variable. Cependant, il ne sera pas possible d'utiliser un tableau de variation comme dans le cas de fonction d'une seule variable. Par contre, comme dans le cas de fonction d'une seule variable, on pourra utiliser des propriétés de concavité ou de convexité de la fonction.

Définition 6.4 (Extremum local). *Soit F une fonction de $\mathbb{R}^n$ définie sur un ensemble Ω. Soit $a \in \mathbb{R}^n$ un point intérieur à Ω (dans les cas qui nous intéresse, cela signifie que a n'est pas sur le bord du domaine; dans le cas général, cela signifie qu'il existe une boule centrée en a entièrement contenue dans Ω).*

Le point a est un maximum local de F s'il existe un voisinage $\mathcal{V}_a$ du point a tel que pour tout $x \in \mathcal{V}_a$, on a $F(x) - F(a) \leq 0$.

Le point a est un maximum global si pour tout $x \in \Omega$ on a $F(x) - F(a) \leq 0$.

Le point a est un minimum local de F s'il existe un voisinage $\mathcal{V}_a$ du point a tel que pour tout $x \in \mathcal{V}_a$, on a $F(x) - F(a) \geq 0$.

Le point a est un minimum global si pour tout $x \in \Omega$ on a $F(x) - F(a) \geq 0$.

On appelle extremum local, un maximum ou un minimum local.

Pour les fonctions (dérivables) d'une seule variable réelle, un extrema local qui ne sont pas sur le bord du domaine est tel que la dérivée de la fonction s'annule en ce point. On a une propriété similaire dans le cas de fonctions de plusieurs variables.

Définition 6.5. *Soit $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, qui admet des dérivées partielles en tout point de Ω. Soit a un point intérieur à Ω. On dira que a est un point critique ou un point stationnaire si*

$$\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0.$$

Proposition 6.6. *Soit $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ une application qui admet des dérivées partielles continues. Soit $a \in \Omega$ tel que a est un point intérieur de Ω. Si a est un extremum local, alors toutes les dérivées partielles de F s'annulent au point a, c'est à dire que a est un point critique de F.*

Remarque. *i) Pour rechercher les extrema locaux, on commencera donc par calculer les dérivées partielles. Ceci nous permettra de déterminer les points critiques. C'est*

parmi ces points critiques que l'on trouvera les extrema locaux. Les points critiques sont donc les candidats pour être des extremum locaux; et ce sont les seuls candidats si Ω est un ouvert, ce qui réduit alors considérablement le champs de recherche des extrémas.
ii) ⚠ *Un point critique n'est pas toujours un minimum ou un maximum local. Voir par exemple la figure 11.*

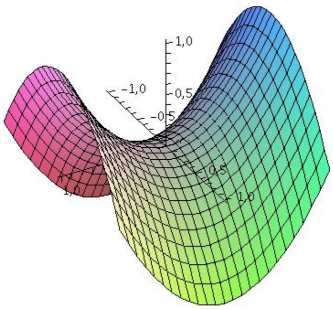

FIGURE 11. Point selle ou Point col

Il nous reste maintenant à trouver une propriété qui va caractériser parmi les points critiques ceux qui sont des extremas. Pour cela, on va avoir besoin de calculer les dérivées partielles secondes.

En utilisant la formule de Taylor vue dans le Théorème 6.3, on a pour tout point critique a

$$F(a+h)-F(a) = \underbrace{\sum_{j=1}^{n}\frac{\partial F}{\partial x_j}(a)h_j}_{=0 \text{ car } a \text{ pt critique}} + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 F}{\partial x_j \partial x_k}(a)h_j\, h_k + \underbrace{\|h\|^2\epsilon(h)}_{\text{terme négligeable}}$$

Donc, pour h "petit", c'est à dire pour $a+h$ au voisinage de a, le signe de $F(a+h)-F(a)$ est donné par le signe de

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 F}{\partial x_j \partial x_k}(a)h_j\, h_k \tag{13}$$

Pour déterminer le signe de l'expression (13) quand $\|h\|$ est proche de zéro, on va utiliser la définition suivante.

DÉFINITION 6.7. *Soit $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ une matrice carrée d'ordre n. On dit que A est une matrice définie positive si pour tout vecteur colonne X de $\mathbb{R}^n$ qui ne soit pas le vecteur nul, on a*

$${}^tXAX > 0.$$

Soit $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ une matrice carrée d'ordre n. On dit que A est une matrice définie négative si pour tout vecteur colonne X de $\mathbb{R}^n$ qui ne soit pas le vecteur nul, on a

$$^tXAX < 0.$$

Il existe une caractérisation pratique du caractère défini positif (ou défini négatif) d'une matrice :

Proposition 6.8 (Critère de Sylvester). *Soit $A = (a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq n}} \in \mathfrak{M}_{n,n}(\mathbb{R})$ une matrice symétrique à coefficients réels. On note Δ_p le déterminant $p \times p$ de la matrice A tronquée à la p-ème ligne et p-ème colonne en haut à gauche:*

$$\Delta_p = \det\left((a_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq p}}\right).$$

A est définie positive *si et seulement si*

$$\Delta_1 > 0,\ \ \Delta_2 > 0, \Delta_3 > 0, \cdots, \Delta_n > 0.$$

A est définie négative *si et seulement si $(-A)$ est définie positive si et seulement si*

$$\Delta_1 < 0,\ \ \Delta_2 > 0, \Delta_3 < 0, \Delta_4 > 0, \cdots, .$$

Remarque. *Pour une matrice carrée A donnée, il est possible qu'elle ne soit ni définie positive, ni définie négative.*

Théorème 6.9 (Conditions suffisantes d'existence d'extremum local). *Soit F une application d'un ensemble $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}$ dont les dérivées partielles secondes sont bien définies et sont continues.*

Si a est un point intérieur de Ω tel que $\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0$ et si la Hessienne de F au point a, $H(F)_{|x=a}$, est définie positive alors a est un minimum local de F.
Si a est un point intérieur de Ω tel que $\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0$ et si la Hessienne de F au point a, $H(F)_{|x=a}$, est définie négative alors a est un maximum local de F.

Remarque. *Si a est un point critique, i.e., $\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0$, mais que la Hessienne de f au point a n'est ni définie positive, ni définie négative, alors on ne peut rien conclure pour le point a.*

Méthode pratique de recherche d'extrema locaux:
• Si F est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$:
i) On calcule les dérivées partielles et les dérivées partielles secondes de F, et on vérifie qu'elle sont continues (ce point sera admis).

ii) On recherche les points critiques $a \in \mathbb{R}^n$, i.e., les points $a \in \mathbb{R}^n$ tels que

$$\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0.$$

iii) On calcule la Hessienne de f en chaque point critique, et on vérifie par le critère de Sylvestre si la Hessienne est définie positive (on aura alors un minimum local) ou définie négative (on aura alors un maximum local). Si elle n'est ni l'un ni l'autre, on ne peut pas conclure.

• Si F est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ (cas $n = 2$): On applique aussi la méthode précédente. Comme c'est un cas plus simple, on peut aussi discuter des points cols (ou points selles).
i) On calcule les dérivées partielles et des dérivées partielles secondes, et on vérifie qu'elle sont continues (ce point sera admis).
ii) On recherche les points critiques $a \in \mathbb{R}^n$, i.e., les points $a \in \mathbb{R}^n$ tels que

$$\frac{\partial F}{\partial x_1}(a) = \frac{\partial F}{\partial x_2}(a) = \cdots = \frac{\partial F}{\partial x_n}(a) = 0.$$

iii) Si a est un point critique, la Hessienne de F au point a est une matrice 2×2 de la forme

$$H(F)_{|x=a} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}.$$

- ◇ Si $rt - s^2 > 0$ et $r > 0$ (ou $t > 0$), le point a correspond à un minimum local.
- ◇ Si $rt - s^2 > 0$ et $r < 0$ (ou $t < 0$), le point a correspond à un maximum local
- ◇ Si $rt - s^2 < 0$ alors a est un point col (on dit aussi point selle).
- ◇ Si $rt - s^2 = 0$, on ne peut pas conclure. Il faut alors écrire la formule de Taylor à l'ordre supérieur, ce qui ne sera pas au programme de cette année.

7. Calcul d'extrema liés

Dans cette section, nous nous intéresserons au calcul d'extrema de fonctions sous contraintes. Comme dans le chapitre précédent, il s'agit de trouver des extrema de fonctions de plusieurs variables, mais cette fois, en imposant des contraintes sur les variables.

En micro-économie, ces types de problèmes sont très courants.

Exemple. Un consommateur cherche à maximiser son utilité, qui est une mesure de sa satisfaction, de son bien être, en consommant dans un panier de biens, mais en respectant une contrainte de budget. La modélisation est la suivante. On considère qu'il existe n biens différents $b_1, b_2, \cdots, b_n$, et dont le prix unitaire est noté p_i pour le bien b_i. Si on note $x_1, x_2, \cdots, x_n$ les quantités achetées des biens $b_1, b_2, \cdots, b_n$, et $U(x_1, x_2, \cdots, x_n)$ la fonction utilité, alors le problème à résoudre est

$$(14) \qquad \begin{cases} \max_{x_1,x_2,\cdots x_n} U(x_1, x_2, \cdots, x_n) & \text{optimisation} \\ p_1x_1 + p_2x_2 + \cdots + p_nx_n \leq I & \text{contrainte d'égalité} \\ x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0 & \text{contraintes d'inégalité} \end{cases}$$

où I est le budget que s'autoriser le consommateur.

Dans ce problème, si on note $\mathcal{A}$ l'ensemble des n-uplets $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ tels que $p_1x_1 + p_2x_2 + \cdots + p_nx_n \leq I$ et $x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0$, alors le problème de maximisation sous contrainte (14) revient à maximiser la fonction $U(x_1, x_2, \cdots, x_n)$ pour $\mathbf{x} \in A$.

Exemple. Une entreprise cherche à maximiser ses profits, tout en respectant une contrainte sur les coûts d'investissement et de fonctionnement.

Pour ce chapitre, nous nous restreindrons au cas où il y a une seule fonction réelle de deux variables $f(x_1, x_2)$ dont on cherchera les extremas, et avec une seule contrainte qui sera soit de type égalité $g(x_1, x_2) = 0$.

Définition 7.1. *Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$. Soit $\mathcal{A} \subset \mathbb{R}^2$ un ensemble sur lequel la fonction f est définie.*
On dit que f admet un maximum local en $\mathbf{a} \in \mathbb{R}^2$ sous la contrainte $\mathbf{x} \in \mathcal{A}$, si il existe $R > 0$, tel que pour tout $x \in \mathcal{A} \cap B_R(\mathbf{a})$

$$f(\mathbf{x}) \leq f(\mathbf{a})$$

On dit que f admet un minimum local en $\mathbf{a} \in \mathbb{R}^2$ sous la contrainte $\mathbf{x} \in \mathcal{A}$, si il existe $R > 0$, tel que pour tout $\mathbf{x} \in \mathcal{A} \cap B_R(\mathbf{a})$

$$f(\mathbf{x}) \geq f(\mathbf{a})$$

Remarque 7.2. *Si la contrainte s'exprime sous la forme $g(\mathbf{x}) = \alpha$, alors $\mathcal{A}$ dans la définition précédente est l'ensemble des valeurs de $\mathbf{x} \in \mathbb{R}^2$ telles que $\mathbf{x} \in \mathcal{U}$ et $g(\mathbf{x}) = \alpha$. Si la contrainte s'exprime sous la forme $g(\mathbf{x}) \leq \alpha$, alors $\mathcal{A}$ dans la définition précédente est l'ensemble des valeurs de $\mathbf{x} \in \mathbb{R}^2$ telles que $\mathbf{x} \in \mathcal{U}$ et $g(\mathbf{x}) \leq \alpha$.*

On considère $f : \mathcal{U} \to \mathbb{R}^2$, et $g : \mathcal{U} \to \mathbb{R}^2$ où $\mathcal{U}$ est un ensemble ouvert de $\mathbb{R}^2$. Le problème à étudier consiste à trouver la valeur maximum sur $\mathcal{U}$ de f sous la contrainte $g(\mathbf{x}) = \alpha$, où $\mathbf{x}$ désigne le couple (x_1, x_2).

$$(15) \qquad \begin{cases} \displaystyle\max_{(x_1,x_2)\in\mathcal{U}} f(x_1,x_2) & \text{(optimisation)} \\ \\ g(x_1,x_2) = \alpha & \text{(contrainte)} \end{cases}$$

Définition 7.3. *On dit que la contrainte $g(x) = \alpha$ est qualifiée au point $a \in \mathbb{R}^2$, si $g(a) = \alpha$ et si les dérivées partielles de g en a ne sont pas toutes nulles à la fois:*

$$\begin{pmatrix} \frac{\partial g}{\partial x_1}(a) \\ \frac{\partial g}{\partial x_2}(a) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Proposition 7.4 (Condition nécessaire du 1er ordre).
On considère le problème de maximisation sous contrainte donné par (15).
On suppose:

i) L'ensemble $\mathcal{U}$ est un ouvert, et les fonctions f et g de $\mathcal{U} \subset \mathbb{R}^2$ dans $\mathbb{R}$ sont continues, et admettent des dérivées partielles continues sur $\mathcal{U}$.
ii) $\mathbf{a} = (a_1, a_2)$ est un élément de $\mathcal{U}$ qui est tel que la contrainte soit qualifiée en $\mathbf{a}$.

Sous ces hypothèses i) et ii) on a:
Si f admet un extremum local en $\mathbf{a}$ sous la contrainte $g(\mathbf{x}) = \alpha$, alors il existe un nombre réel λ tel que

$$(16) \qquad \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) \\ \frac{\partial f}{\partial x_2}(\mathbf{a}) \end{pmatrix} = \lambda \begin{pmatrix} \frac{\partial g}{\partial x_1}(\mathbf{a}) \\ \frac{\partial g}{\partial x_2}(\mathbf{a}) \end{pmatrix}$$

Remarque 7.5. *Ce résultat nous dit que les extremas locaux de la fonction $f(\cdot,\cdot)$ sont soit des points tels que $g(\mathbf{x}) = \alpha$ et où la contrainte n'est pas qualifiée, soit des points tels que $g(\mathbf{x}) = \alpha$ où la contrainte est qualifiée, et dans ce cas ils vérifient l'équation* (16) *pour un certain λ.*

DÉFINITION 7.6.
i) Un réel λ qui vérifie l'équation (16) est appelée un multiplicateur de Lagrange.
ii) On appelle lagrangien du problème (15) la fonction

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda\big(g(\mathbf{x}) - \alpha\big).$$

L'équation (16) est alors équivalente à

$$(\nabla_{\mathbf{x}}\mathcal{L})(\mathbf{a}, \lambda) := \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial x_1}(\mathbf{a}, \lambda) \\ \frac{\partial \mathcal{L}}{\partial x_2}(\mathbf{a}, \lambda) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) - \lambda \frac{\partial g}{\partial x_1}(\mathbf{a}) \\ \frac{\partial f}{\partial x_2}(\mathbf{a}) - \lambda \frac{\partial g}{\partial x_2}(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

iii) On appelle point critique ou point stationnaire du Lagrangien $\mathcal{L}$ tout triplet $(\mathbf{a}, \lambda) = (a_1, a_2, \lambda)$ qui vérifie

$$(\nabla\mathcal{L})(\mathbf{a}, \lambda) := \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial x_1}(\mathbf{a}, \lambda) \\ \frac{\partial \mathcal{L}}{\partial x_2}(\mathbf{a}, \lambda) \\ \frac{\partial \mathcal{L}}{\partial \lambda}(\mathbf{a}, \lambda) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) - \lambda \frac{\partial g}{\partial x_1}(\mathbf{a}) \\ \frac{\partial f}{\partial x_2}(\mathbf{a}) - \lambda \frac{\partial g}{\partial x_2}(\mathbf{a}) \\ g(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{17}$$

REMARQUE 7.7. *i) Puisque $\mathcal{L}(x, y, \lambda) = f(\mathbf{x}) - \lambda(g(\mathbf{x}) - \alpha)$, on obtient $\frac{\partial \mathcal{L}}{\partial \lambda} = g$. C'est cette égalité que l'on retrouve dans (17). Cela signifie que (17) est équivalente à (16) sous la contrainte $g(\mathbf{x}) = \alpha$. C'est pourquoi dans certains ouvrages on appelle points critiques du lagrangien uniquement les points $\mathbf{a} = (a_1, a_2)$ qui vérifient (16), et qu'on étudie parmi ces points critiques ceux qui vérifient $g(\mathbf{x}) - \alpha = 0$.*
ii) Comme la contrainte $g(\mathbf{x}) = \alpha$ s'écrit aussi bien $g(\mathbf{x}) - \alpha = 0$ et $\alpha - g(\mathbf{x}) = 0$, le signe devant λ dans le lagrangien n'a aucune importance, et on peut tout aussi bien considérer le lagrangien $f(\mathbf{x}) + \lambda(g(\mathbf{x}) - \alpha)$ au lieu de $f(\mathbf{x}) - \lambda(g(\mathbf{x}) - \alpha)$.

DÉFINITION 7.8 (Hessienne bordée). *La matrice hessienne bordée du lagrangien $\mathcal{L}$ en $(\mathbf{a}, \lambda) = (a_1, a_2, \lambda)$ est*

$$H(\mathcal{L})_{|(a_1,a_2,\lambda)} = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x_1^2}(a_1, a_2, \lambda) & \frac{\partial^2 \mathcal{L}}{\partial x_2 \partial x_1}(a_1, a_2, \lambda) & \frac{\partial g}{\partial x_1}(a_1, a_2) \\ \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_2}(a_1, a_2, \lambda) & \frac{\partial^2 \mathcal{L}}{\partial x_2^2}(a_1, a_2, \lambda) & \frac{\partial g}{\partial x_2}(a_1, a_2) \\ \frac{\partial g}{\partial x_1}(a_1, a_2) & \frac{\partial g}{\partial x_2}(a_1, a_2) & 0 \end{pmatrix}$$

C'est la matrice hessienne de $\mathcal{L}$ dans les trois variables x_1, x_2, λ.

PROPOSITION 7.9 (Condition suffisante du 2ème ordre). *Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^2$, et soient f et g deux applications de $\mathcal{U}$ dans $\mathbb{R}$, continues, qui admettent des dérivées partielles d'ordre un et d'ordre deux continues.*

Soit (a_1, a_2, λ) un point critique du lagrangien $\mathcal{L}$, c'est à dire $(\nabla_{\mathbf{x},\lambda}\mathcal{L})(a_1, a_2, \lambda) = 0_{\mathbb{R}^2}$. Alors

- *Si le déterminant de la hessienne bordée $H(\mathcal{L})_{|(a_1,a_2,\lambda)}$ est strictement négatif, alors (a_1, a_2) est un minimum local du problème avec contrainte.*
- *Si le déterminant de la hessienne bordée $H(\mathcal{L})_{|(a_1,a_2,\lambda)}$ est strictement positif, alors (a_1, a_2) est un maximum local du problème avec contrainte.*
- *Si le déterminant de la hessienne bordée $H(\mathcal{L})_{|(a_1,a_2,\lambda)}$ est nul, on ne peut rien dire.*

Méthodologie:

- On cherche les points stationnaires (points critiques) du Lagrangien $\mathcal{L}$, c'est à dire, on cherche les $(\mathbf{a}, \lambda) = (a_1, a_2, \lambda)$ tels que

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_1}(\mathbf{a}) = 0 \\[2ex] \dfrac{\partial \mathcal{L}}{\partial x_2}(\mathbf{a}) = 0 \\[2ex] \dfrac{\partial \mathcal{L}}{\partial \lambda}(\mathbf{a}) = 0 \end{cases} \quad \text{c'est à dire} \quad \begin{cases} \dfrac{\partial f}{\partial x_1}(\mathbf{a}) - \lambda \dfrac{\partial g}{\partial x_1}(\mathbf{a}) = 0 \\[2ex] \dfrac{\partial f}{\partial x_2}(\mathbf{a}) - \lambda \dfrac{\partial g}{\partial x_2}(\mathbf{a}) = 0 \\[2ex] g(\mathbf{a}) - \alpha = 0 \end{cases}$$

- En chaque point stationnaire $(\mathbf{a}, \lambda) = (a_1, a_2, \lambda)$ trouvé, on calcule la matrice hessienne bordée.
  - Si son déterminant est strictement positif, alors $\mathbf{a}$ est un maximum local
  - Si son déterminant est strictement négatif, alors $\mathbf{a}$ est un minimum local
  - Si son déterminant est nul, on ne peut pas conclure
- Éventuellement, on discute séparément les points $\mathbf{x}$ tels que $g(\mathbf{x}) = \alpha$ et $\nabla g(\mathbf{x}) = 0_{\mathbb{R}^2}$ (i.e., là où la contrainte n'est pas qualifiée en $\mathbf{x}$).

Exemple: On cherche les extremas locaux de la fonction $f(x, y) = x^2 + y^2$ sous la contrainte $g(x, y) = x^2y - 16 = 0$.

Le lagrangien associé est $\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda g(x, y) = x^2 + y^2 - \lambda(x^2y - 16)$.

• On recherche les points stationnaires du lagrangien:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 0 \\ \frac{\partial \mathcal{L}}{\partial y} = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x - 2\lambda xy = 0 \\ 2y - \lambda x^2 = 0 \\ x^2y - 16 = 0 \end{cases} \Leftrightarrow \begin{cases} x(1 - \lambda y) = 0 \\ 2y = \lambda x^2 = 0 \\ x^2y = 16 \end{cases}$$

La première équation donne $x = 0$ ou $(1 - \lambda y) = 0$. Le cas $x = 0$ est exclu car la dernière équation deviendrait $0 = 16$. Le cas $(1 - \lambda y) = 0$ donne $y = 1/\lambda$. La dernière équation donne $x^2 = 16/y$. On reporte ces deux résultats dans la deuxième équation et cela donne $2y - \frac{1}{y}\frac{16}{y} = 0$, et donc $y^3 = 8$, c'est à dire $y = 2$. On obtient donc, puisque $y = 1/\lambda$, que $\lambda = 1/2$. Et cela donne $x^2 = 8$, i.e., $x = \pm\sqrt{8} = \pm 2\sqrt{2}$. On a donc deux points critiques pour le lagrangien:

$$(2\sqrt{2}; 2; \frac{1}{2}) \quad \text{et} \quad (-2\sqrt{2}; 2; \frac{1}{2}).$$

• On calcule la hessienne bordée

$$H(\mathcal{L}) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial g}{\partial x} \\ \frac{\partial^2 \mathcal{L}}{\partial x \partial y} & \frac{\partial^2 \mathcal{L}}{\partial y^2} & \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & 0 \end{pmatrix} = \begin{pmatrix} 2 - 2\lambda y & -2\lambda x & 2xy \\ -2\lambda x & 2 & x^2 \\ 2xy & x^2 & 0 \end{pmatrix}$$

puis on l'évalue en chacun des points critiques:

$$H(\mathcal{L})_{|(2\sqrt{2};2;\frac{1}{2})} = \begin{pmatrix} 0 & -2\sqrt{2} & 8\sqrt{2} \\ -2\sqrt{2} & 2 & 8 \\ 8\sqrt{2} & 8 & 0 \end{pmatrix} \text{ et } H(\mathcal{L})_{|(-2\sqrt{2};2;\frac{1}{2})} = \begin{pmatrix} 0 & 2\sqrt{2} & -8\sqrt{2} \\ 2\sqrt{2} & 2 & 8 \\ -8\sqrt{2} & 8 & 0 \end{pmatrix}$$

Dans les deux cas, le déterminant vaut -768, et on conclut donc que les deux points $(2\sqrt{2}; 2)$ et $(-2\sqrt{2}; 2)$ sont deux minima locaux, qui correspondent chacun au même multiplicateur de Lagrange $\lambda = 1/2$.

• On cherche enfin les points qui ne sont pas qualifiés, c'est à dire les couples (x, y) telque la contrainte $g(x, y) = 0$ soit vérifiée, mais $(\partial g)/(\partial x) = (\partial g)/(\partial y) = 0$. On résout donc le système

$$\begin{cases} g(x, y) = x^2 y - 16 = 0 \\ \frac{\partial g}{\partial x} = 2xy = 0 \\ \frac{\partial g}{\partial y} = x^2 = 0 \end{cases}$$

Ce système n'a pas de solution. Il n'y a donc pas de discussion supplémentaire.

Les seuls extremas locaux sont donc les deux minimas locaux $(2\sqrt{2}; 2)$ et $(-2\sqrt{2}; 2)$

Remarque 7.10. *i) Les résultats précédents se généralisent au cas $\mathbb{R}^n$ et aussi au cas de p contraintes $g_1(x) = \alpha_1, g_2(x) = \alpha_2, \cdots, g_p(x) = \alpha_p$. On a alors p multiplicateurs de Lagrange $\lambda_1, \lambda_2, \cdots, \lambda_p$. Et la "qualification des contraintes" est une condition sur la matrice jacobienne des contraintes, qui est la matrice des dérivées partielles:*

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial x_1}(a) & \cdots & \frac{\partial g_p}{\partial x_n}(a) \end{pmatrix}$$

La contrainte est qualifiée si le rang de cette matrice est maximum, c'est à dire si on peut trouver une matrice carrée de taille maximum $\inf\{p, n\}$ *extraite de la matrice jacobienne, dont le déterminant est non nul.*
ii) La proposition 7.9 est spécifique au cas $\mathbb{R}^2$*, mais il existe un énoncé qui se généralise (de façon non triviale) au cas* $\mathbb{R}^n$ *et avec* p *contraintes d'égalité.*

PROPOSITION 7.11. *Dans le problème* (15)*, pour* α *étant une valeur donnée pour la contrainte, le multiplicateur de Lagrange* λ *associé à un extremum* $\mathbf{a}$ *(il peut y avoir plusieurs extrema) mesure le taux d'accroissement de la valeur de l'extremum près de* $\mathbf{a}$ *lorsqu'on varie la valeur de* α *de la contrainte dans le problème* (15)*. En d'autres termes, le multiplicateur de Lagrange mesure l'effet marginal sur un maximum local (ou un minimum local) pour un déplacement de la contrainte près de la valeur* α*.*

Nous n'aurons pas le temps de traiter cette année le cas d'une contrainte sous forme d'inégalité. Ce problème est traité par exemple dans les ouvrages [1, 2].

A savoir et savoir faire pour les contrôles de connaissances: Analyse

Au moins 80% de la partie de l'examen qui portera sur l'analyse (chapitres 5 à 7) concernera les points suivants:

- Pour le Chapitre 5
  - Savoir déterminer le domaine de définition d'une fonction d'une variable (programme du semestre 1)
  - Connaître les dérivées usuelles de fonctions d'une seule variable: voir tableau donné en TD (programme du semestre 1).
  - Savoir calculer la composée de deux fonctions d'une seule variable, et savoir calculer sa dérivée (programme du semestre 1)
  - Déterminer le domaine de définition d'une fonction de plusieurs variables à valeurs dans $\mathbb{R}$.
  - Savoir calculer des dérivées partielles de fonctions de plusieurs variables à valeurs dans $\mathbb{R}$.
  - La partie fonction implicite n'est pas exigible (ce qui ne vous dispense pas de connaître le cours dont vous aurez besoin dans le futur)
- Pour le Chapitre 6: extrema libres
  - Savoir donner le domaine de définition de la fonction à étudier

- Connaître la définition de point critique d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. Savoir déterminer les points critiques d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ (dans les cas $n = 2$ ou $n = 3$).
- Calcul de hessienne
- Déterminer la nature des points critiques dans les cas non dégénérés pour des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ et de $\mathbb{R}^3$ dans $\mathbb{R}$.

• Pour le Chapitre 7: extrema liés
  - Aucune connaissance exigible sur ce chapitre 7 en examen final et en examen de seconde session. Pour le second contrôle continu de TD, voir directement avec les enseignants concernés.

References

[1] Thierry Lafay, Mathématiques pour les économistes, Édition Archétype82 (2013)
[2] Stéphane Rossignol, Mathématiques en économie-gestion, Édition Dunod (2015).