

ECG 2 - Mathématiques appliquées

Lilian Saligue

2024-2025

Table des matières

I	Troisième semestre	1
1	Suites récurrentes	3
1.1	Rappel : Les suites récurrentes de première année.	3
1.1.1	Les suites récurrentes linéaires d'ordre 1	3
1.1.2	Les suites récurrentes linéaires d'ordre 2	4
1.2	Le cas $u_{n+1} = f(u_n)$	5
1.2.1	Existence de la suite	6
1.2.2	Le cas particulier où f est croissante	7
1.2.3	Étude de la convergence	8
1.2.4	Récapitulatif de la méthode et exemple complet	10
2	Intégrales : Révisions et compléments	15
2.1	Rappel de première année	15
2.1.1	Définitions et règles de calculs	15
2.1.2	Outil de calculs	17
2.2	Intégrales impropres	18
2.2.1	Définitions	18
2.2.2	Exemples fondamentaux	20
2.2.3	Règles de calculs	23
2.2.4	Intégrales doublement impropres	26
3	Espaces vectoriels	29
3.1	Construction et définition	29
3.1.1	Loi de composition interne et externe	29
3.1.2	Définitions et exemples fondamentales	32
3.2	Sous-espaces vectoriels	34
3.2.1	Sous-espaces vectoriels engendrés et notion de rang	36
3.3	Familles libres, génératrices et bases	39
3.3.1	Familles génératrices	39
3.3.2	Familles libres	42
3.3.3	Bases	44

4	Statistiques bivariées	49
4.1	Rappel sur les statistiques univariées	49
4.2	Définitions	50
4.3	Covariance et corrélation	52
4.4	Régression et méthode des moindres carrés	54
5	Comparaisons de suites réelles	59
5.1	Relation de négligeabilité	59
5.1.1	Définition et caractérisation	59
5.1.2	Règles de calculs	61
5.2	Relation d'équivalence	62
5.2.1	Définition et caractérisation	62
5.2.2	Règles de calculs	63
5.2.3	Méthodes de calculs	64
6	Séries : Révisions et compléments	67
6.1	Rappels	67
6.1.1	Définitions et propriétés générales	67
6.1.2	Séries usuelles	69
6.1.3	Convergence absolue	70
6.2	Séries à termes positifs	71
6.2.1	Série de Riemann	71
6.2.2	Utilisation des relations de comparaisons	73
7	Applications linéaires et endomorphismes	77
7.1	Définitions et règles de calculs	77
7.2	Notion d'image et de noyau	81
7.2.1	Noyau	81
7.2.2	Image	83
7.3	Utilisation des dimensions et des bases	84
7.3.1	Caractérisation d'une application linéaire	84
7.3.2	Notion de rang	85
7.3.3	Théorème du rang	87
7.4	Matrice associée à une application linéaire	88
7.4.1	Généralités	88
7.4.2	Parallélisme dans le calcul	93
7.4.3	Notion de changement de base	97
8	Révisions : Variables aléatoires discrètes	101
8.1	Généralités	101
8.2	Notion de moments	104

8.2.1	Espérance	104
8.2.2	Variance	106
8.3	Lois usuelles	108
8.3.1	Lois discrètes finis	109
8.3.2	Lois discrètes infinis	113
9	Couples de variables aléatoires	117
9.1	Loi d'un couple de variables aléatoires discrètes	117
9.1.1	Définitions	117
9.1.2	Aparté sur les séries doubles et un système complet d'évè- nements	118
9.1.3	Lois marginales et loi conditionnelles	120
9.2	Variable aléatoire du type $Z = g(X, Y)$	123
9.2.1	Définition	123
9.2.2	Espérance et théorème du transfert	125
9.2.3	Lois classiques	127
9.3	Notion de covariance	132
9.3.1	Définition et propriétés fondamentales	132
10	Réduction de matrices carrées	137
11	Comparaisons de fonctions réelles et développements limités	139
11.1	Relations de comparaisons	139
11.1.1	Notion de voisinage	139
11.1.2	Relation de négligeabilité	140
11.1.3	Relation d'équivalence	143
11.2	Développements limités	146
11.2.1	Ordre 0	146
11.2.2	Ordre 1	146
11.2.3	Ordre 2	147
11.2.4	Développements limités usuelles	148
11.3	Applications des développements limités	149
11.3.1	Aspect graphique	149
11.3.2	Application aux suites	151
12	Intégrales impropres de fonctions positives	153
12.1	Généralités	153
12.2	Inégalités	155
12.3	Négligeabilité	155
12.4	Équivalence	156
12.5	Convergence absolue	156

13	Systèmes différentielles linéaires	159
14	Suites de variables aléatoires discrètes	161
14.1	Généralités	161
14.1.1	Définitions	161
14.1.2	Indépendances	162
14.2	Lois usuelles	164
14.3	Moments	165
II	Quatrième semestre	167

Première partie
Troisième semestre

Chapitre 1

Suites récurrentes

De nombreuses suites réelles s'expriment de manières récursives, c'est-à-dire que le ou les premiers termes sont données et les autres termes s'obtiennent à partir des précédents. Nous avons en première année déjà étudiée certaines suites de ce type comme les bien connus suites arithmétiques, géométriques, arithmético-géométrique ou encore les suites récurrentes linéaires d'ordre 2. Nous avons également étudiée quelques suites récurrentes sous la forme plus générales $u_{n+1} = f(u_n)$ et ce chapitre à pour but de pousser l'étude de ces dernières.

1.1 Rappel : Les suites récurrentes de première année.

1.1.1 Les suites récurrentes linéaires d'ordre 1

Listons les résultats sur les suites arithmétiques, géométriques et arithmético-géométriques.

Définition 1.1.1. Soit $r \in \mathbb{R}$. On dit qu'une suite $(u_n)_{n \in \mathbb{N}}$ est arithmétique de raison r lorsque pour tout $n \in \mathbb{N}$, $u_{n+1} = u_n + r$.

Propriété 1.1.2. Soit $(u_n)_{n \in \mathbb{N}}$ une suite arithmétique réelle de raison r , on a alors

$$\forall n \in \mathbb{N}, u_n = u_0 + nr.$$

Définition 1.1.3. Soit $q \in \mathbb{R}$. On dit qu'une suite $(u_n)_{n \in \mathbb{N}}$ est géométrique de raison q lorsque pour tout $n \in \mathbb{N}$, $u_{n+1} = qu_n$.

Propriété 1.1.4. Soit $(u_n)_{n \in \mathbb{N}}$ une suite géométrique réelle de raison $q \neq 0$, on a alors

$$\forall n \in \mathbb{N}, u_n = u_0 q^n.$$

Définition 1.1.5. On dit que $(u_n)_{n \in \mathbb{N}}$ est une suite arithmético-géométrique lorsque qu'elle est définie par un premier terme u_0 et la relation de récurrence

$$\forall n \in \mathbb{N}, u_{n+1} = au_n + b,$$

où a et b sont deux réels.

Ces dernières suites sont les plus délicates à étudier et nous rappelons donc la méthodologie les concernant :

Méthode 1.1.6. Soit $(u_n)_{n \in \mathbb{N}}$ une suite arithmético-géométrique définie par la relation pour tout $n \in \mathbb{N}$, $u_{n+1} = au_n + b$ où $a \neq 1$.

1. On commence par déterminer le point fixe l de la relation de récurrence, c'est à dire la solution de l'équation $l = al + b$.
2. On définit la suite $(v_n)_{n \in \mathbb{N}}$ par pour tout $n \in \mathbb{N}$, $v_n = u_n - l$ et on montre que cette suite est géométrique de raison a et de premier terme $u_0 - l$. On a alors pour tout $n \in \mathbb{N}$, $v_n = v_0 a^n = (u_0 - l)a^n$.
3. On en déduit l'expression explicite de $(u_n)_{n \in \mathbb{N}}$ en remarquant que pour tout $n \in \mathbb{N}$, $v_n = u_n - l$ est équivalent à $u_n = v_n + l$. On a alors

$$\forall n \in \mathbb{N}, u_n = (u_0 - l)a^n + l.$$

1.1.2 Les suites récurrentes linéaires d'ordre 2

Rappelons la définition puis le théorème donnant la forme générale d'une telle suite

Définition 1.1.7. Soit α, β, a et b quatre réels et soit $(u_n)_{n \in \mathbb{N}}$ une suite de réels. On dit que $(u_n)_{n \in \mathbb{N}}$ est une suite récurrente linéaire d'ordre 2 lorsque elle est définie par

$$\begin{cases} u_0 = \alpha; \\ u_1 = \beta; \\ \forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n \end{cases}$$

Théorème 1.1.8. Soit $(u_n)_{n \in \mathbb{N}}$ une suite récurrente linéaire d'ordre 2 définie par la relation de récurrence pour tout $n \in \mathbb{N}$, $u_{n+2} = au_{n+1} + bu_n$. Lorsque l'équation caractéristique $x^2 - ax - b = 0$ a :

- deux solutions distinctes r_1 et r_2 (c'est-à-dire que le discriminant du trinôme est strictement positif) on a :

$$\exists (\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = \lambda r_1^n + \mu r_2^n;$$

- une unique solution r (c'est-à-dire que le discriminant du trinôme est nul)
on a :

$$\exists(\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = (\lambda + \mu n)r^n.$$

Remarque 1.1.9. La forme général dépendant de deux constantes, on n'oublie pas de les déterminer en utilisant un système à l'aide des valeurs de u_0 et de u_1 .

1.2 Le cas $u_{n+1} = f(u_n)$

Le principal cas d'étude de ce chapitre sont les suites réelles du type

$$\begin{cases} u_0 \in \mathbb{R}, \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

où f est une fonction réelle. Nous allons à présent expliciter la méthodologie pour travailler sur les suites de ce type.

Pour étudier cette suite, nous aurons besoin de plusieurs informations sur la fonction f , il est donc bien souvent recommandé de commencer par étudier complètement la fonction f .

Exemple 1.2.1. On va considérer l'exemple de la suite

$$\begin{cases} u_0 = 3 \\ \forall n \in \mathbb{N}, u_{n+1} = \frac{1}{2} \left(u_n + \frac{3}{u_n} \right) \end{cases}$$

En posant $f : \mathbb{R}^* \rightarrow \mathbb{R}$
 $x \mapsto \frac{1}{2} \left(x + \frac{3}{x} \right)$, on a que la relation de récurrence s'écrit bien $u_{n+1} = f(u_n)$.

Étudions cette fonction. On remarque qu'il s'agit de la somme d'un polynôme et de la fonction inverse donc f est de classe \mathcal{C}^∞ sur \mathbb{R}^* et pour tout $x \in \mathbb{R}^*$, on a $f'(x) = \frac{1}{2} \left(1 - \frac{3}{x^2} \right) = \frac{x^2-3}{2x^2} = \frac{(x-\sqrt{3})(x+\sqrt{3})}{2x^2}$. On en déduit le tableau de variation suivant :

x	$-\infty$	$-\sqrt{3}$	0	$\sqrt{3}$	$+\infty$	
$f'(x)$	$+$	0	$-$	$-$	0	$+$
f	$-\infty$	$-\sqrt{3}$	$-\infty$	$+\infty$	$\sqrt{3}$	$+\infty$

1.2.1 Existence de la suite

Pour la partie étude de suite d'une suite réelle définie par une relation de récurrence $u_{n+1} = f(u_n)$, la première chose à faire est de vérifier si la suite $(u_n)_{n \in \mathbb{N}}$ est bien définie. En général on couple cela avec l'appartenance de la suite à un intervalle qui nous aidera dans notre étude de suite.

Cela se fait en général par récurrence en utilisant la notion d'intervalle stable.

Définition 1.2.2. Soit I un intervalle de \mathbb{R} et soit f une fonction définie sur un domaine contenant I . On dit que I est stable par f si $f(I) \subset I$. Autrement dit I est stable par f si et seulement si pour tout $x \in I$, $f(x) \in I$.

Exemple 1.2.3. On considère l'exemple précédent et le tableau de variation nous indique que $[\sqrt{3}, +\infty[$ est un intervalle stable par f .

En effet pour tout $x \in [\sqrt{3}, +\infty[$, comme f est croissante sur $[\sqrt{3}, +\infty[$, on a $f(\sqrt{3}) \leq f(x)$ or $f(\sqrt{3}) = \frac{1}{2} \left(\sqrt{3} + \frac{3}{\sqrt{3}} \right) = \frac{1}{2}(\sqrt{3} + \sqrt{3}) = \frac{2\sqrt{3}}{2} = \sqrt{3}$. On a donc $\sqrt{3} \leq f(x)$ donc $f(x) \in [\sqrt{3}, +\infty[$. On a donc bien $f([\sqrt{3}, +\infty[) \subset [\sqrt{3}, +\infty[$.

Remarque 1.2.4. Dans l'immense majorité des cas, l'intervalle stable sera donné par l'énoncé ou sera l'ensemble de définition de f .

Cette notion est nécessaire dans le résultat suivant :

Propriété 1.2.5. Soit f une fonction réelle et soit I est un intervalle stable de f . Soit $(u_n)_{n \in \mathbb{N}}$ une suite définie par $\begin{cases} u_0 \in I, \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$. On a alors que pour tout $n \in \mathbb{N}$, $u_n \in I$.

Remarque 1.2.6. On n'utilisera pas ce résultat en pratique car on démontrera par récurrence ce fait dans chacune de nos études.

Exemple 1.2.7. On considère l'exemple précédent et montrons par récurrence que pour tout $n \in \mathbb{N}$, $u_n \in [\sqrt{3}, +\infty[$:

- **Initialisation :** On a $u_0 = 3 \geq \sqrt{3}$ donc on a bien $u_0 \in [\sqrt{3}, +\infty[$
- **Hérédité :** Soit $n \in \mathbb{N}$ fixé quelconque et supposons la propriété vraie au rang n et montrons là au rang $n + 1$, c'est-à-dire supposons que u_n existe et que $u_n \in [\sqrt{3}, +\infty[$ et montrons que u_{n+1} existe et que $u_{n+1} \in [\sqrt{3}, +\infty[$. Comme $u_n \in [\sqrt{3}, +\infty[$ et que $[\sqrt{3}, +\infty[$ est stable par f , on a que $f(u_n)$ existe bien et que $f(u_n) \in [\sqrt{3}, +\infty[$ soit u_{n+1} existe et $u_{n+1} \in [\sqrt{3}, +\infty[$ ce qui prouve notre hérédité.

- **Conclusion :** Notre propriété est bien initialisée et est héréditaire à partir du rang 0 donc par le principe de récurrence on a prouvé que $(u_n)_{n \in \mathbb{N}}$ existe bien et que pour tout $n \in \mathbb{N}$, $u_n \in [\sqrt{3}, +\infty[$.

Remarque 1.2.8. L'argument dans l'hérédité est toujours le même, à savoir " $u_n \in I$ or comme I est stable par f , on a que $f(u_n)$ existe bien et que $f(u_n) \in I$ soit u_{n+1} existe et $u_{n+1} \in I$ ". Un argument calculatoire peut également intervenir mais cela revient au calcul de l'intervalle stable.

Remarque 1.2.9. Il arrive parfois que les premiers termes de la suite ne sont pas dans l'intervalle stable mais à partir du moment qu'un terme est dans un tel intervalle, tout les termes suivant ne pourront pas sortir de cette intervalle. On devra donc faire attention et parfois ne pas considérer les premiers termes de la suite.

Remarque 1.2.10. L'appartenance à un intervalle peut également fournir un encadrement de la suite qui pourra nous servir pour l'étude de la convergence.

Exemple 1.2.11. Le fait que pour tout $n \in \mathbb{N}$, $u_n \in [\sqrt{3}, +\infty[$ se traduit également par le fait que pour tout $n \in \mathbb{N}$, $\sqrt{3} \leq u_n$.

1.2.2 Le cas particulier où f est croissante

Un cas particulier est lorsque f est une fonction croissante sur l'intervalle I . En effet la croissance de la fonction à une conséquence sur la monotonie de la suite $(u_n)_{n \in \mathbb{N}}$.

Propriété 1.2.12. Soit f une fonction réelle croissante (resp. strictement croissante) sur un intervalle I stable. Soit $(u_n)_{n \in \mathbb{N}}$ une suite définie par

$$\begin{cases} u_0 \in I, \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

On a alors que la suite $(u_n)_{n \in \mathbb{N}}$ est monotone (resp. strictement monotone).

Remarque 1.2.13. Comme précédemment, on utilise pas ce résultat directement mais on montre le résultat par récurrence avec un argument d'hérédité du type "Si $u_n < u_{n+1}$ alors par croissance de f on a $f(u_n) < f(u_{n+1})$ soit $u_{n+1} < u_{n+2}$ ". On pourra se faire une idée du sens de variation grâce à la comparaison de u_0 et u_1 .

Exemple 1.2.14. On considère le même exemple que précédemment. Comme notre suite est intégralement dans l'intervalle $[\sqrt{3}, +\infty[$ on peut restreindre l'étude de f à cette intervalle là et par l'étude que l'on en a fait, f est croissante sur

cette intervalle. Le résultat précédent nous amène donc à penser que $(u_n)_{n \in \mathbb{N}}$ est monotone. Comme $u_1 = f(u_0) = f(3) = \frac{1}{2} \left(3 + \frac{3}{3}\right) = \frac{4}{2} = 2$, on va montrer que $(u_n)_{n \in \mathbb{N}}$ est décroissante en montrant par récurrence que pour tout $n \in \mathbb{N}$, $u_{n+1} \leq u_n$.

- **Initialisation :** Comme on vient de la calculer $u_1 \leq u_0$ donc la propriété est vraie au rang 0.
- **Hérédité :** Soit $n \in \mathbb{N}$ et supposons la propriété vraie au rang n et montrons là au rang $n + 1$, c'est-à-dire supposons que $u_{n+1} \leq u_n$ et montrons que $u_{n+2} \leq u_{n+1}$.
Comme $u_{n+1} \leq u_n$, que u_n et u_{n+1} sont des éléments de $[\sqrt{3}, +\infty[$ et que f est croissante sur $[\sqrt{3}, +\infty[$, on a $f(u_{n+1}) \leq f(u_n)$ soit $u_{n+2} \leq u_{n+1}$ ce qui prouve notre hérédité.
- **Conclusion :** La propriété est initialisée au rang 0 et est héréditaire à partir de ce rang donc par le principe de récurrence on a montré que pour tout $n \in \mathbb{N}$, $u_{n+1} \leq u_n$, autrement dit la suite $(u_n)_{n \in \mathbb{N}}$ est décroissante.

Remarque 1.2.15. Le cas d'une fonction f décroissante ne donne pas un résultat aussi simple d'utilisation. On devrait dans ce cas utiliser $f \circ f$ et remarquer que les sous-suites $(u_{2n})_{n \in \mathbb{N}}$ et $(u_{2n+1})_{n \in \mathbb{N}}$ sont chacune monotones et de sens de variations opposées.

1.2.3 Étude de la convergence

L'étude de la convergence se fait en deux étapes.

Il nous faut dans un premier temps justifier de la convergence de la suite, pour cela on utilise des arguments qui dépendront de chaque cas particuliers. Dans beaucoup de cas le théorème de la limite monotone nous sera fort utile.

Exemple 1.2.16. La suite $(u_n)_{n \in \mathbb{N}}$ est décroissante et est minorée par $\sqrt{3}$, elle est donc convergente vers un réel l .

Si la suite $(u_n)_{n \in \mathbb{N}}$ est convergente, on peut calculer sa limite à l'aide de la définition et du résultat suivant :

Définition 1.2.17. Soit E un ensemble et soit f une application d'un sous-ensemble de E dans E . On dit que $l \in E$ est un point fixe de f si $f(l) = l$.

Théorème 1.2.18. Soit f une fonction réelle continue. Soit $(u_n)_{n \in \mathbb{N}}$ une suite définie par $\begin{cases} u_0 \in I, \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$. Si $(u_n)_{n \in \mathbb{N}}$ est convergente, alors $(u_n)_{n \in \mathbb{N}}$ converge vers un point fixe de f .

Remarque 1.2.19. L'étude des points fixes d'une fonction f peut se faire à l'aide de l'étude de la fonction $x \mapsto f(x) - x$. Cette fonction peut de plus déjà avoir été étudiée dans le cadre de l'étude de la monotonie de la suite $(u_n)_{n \in \mathbb{N}}$.

Exemple 1.2.20. On conserve les données des exemples précédents et on pose $g : x \mapsto f(x) - x$. On a que g est une fonction dérivable sur \mathbb{R}^* en tant que différence de f et d'un polynôme qui sont toutes deux des fonctions dérivables sur \mathbb{R}^* . On a de plus pour tout $x \in \mathbb{R}^*$, $g'(x) = \frac{x^2-3}{2x^2} - 1 = \frac{x^2-3-2x^2}{2x^2} = -\frac{x^2+3}{2x^2} < 0$. On en déduit donc le tableau de variation suivant :

x	$-\infty$	0	$+\infty$
$g'(x)$	-		-
g	$+\infty$ ↘ $-\infty$	$+\infty$ ↘ $-\infty$	$-\infty$

Comme g est continue, strictement décroissant sur \mathbb{R}_+^* et que $0 \in] \lim_{x \rightarrow +\infty} f(x), \lim_{x \rightarrow 0^+} f(x)[= \mathbb{R}$, on a par le théorème des valeurs intermédiaire qu'il existe un unique réel $\alpha \in \mathbb{R}_+^*$ tel que $g(x) = 0$ ou encore tel que $f(x) = x$. On a déjà vu que $f(\sqrt{3}) = \sqrt{3}$ donc on a $\alpha = \sqrt{3}$.

Par le même procédé, on montre que $-\sqrt{3}$ est l'unique solution de $g(x) = 0$ sur \mathbb{R}_-^* .

On déduit ainsi que les points fixes de f sont $-\sqrt{3}$ et $\sqrt{3}$.

Remarque 1.2.21. On prendra garde au fait que l'existence d'un point fixe n'implique pas la convergence de la suite, on devra donc toujours montrer la convergence de la suite dans un premier temps. De même une fonction n'ayant pas toujours un unique point fixe, il faut donc faire attention à la cohérence des résultats pour choisir le bon point fixe en fonction de la suite convergente que l'on considère !

Remarque 1.2.22. Une suite du type $u_{n+1} = f(u_n)$ dont le premier terme est un point fixe est une suite constante.

Exemple 1.2.23. Les points fixes de f sont $-\sqrt{3}$ et $\sqrt{3}$ or la suite $(u_n)_{n \in \mathbb{N}}$ est à termes positifs donc $-\sqrt{3}$ qui est strictement négatif est à exclure. La seule possibilité est donc $\sqrt{3}$. On a donc

$$\lim_{n \rightarrow +\infty} u_n = \sqrt{3}.$$

La vitesse de convergence peut s'étudier grâce à divers outils. Le principal que nous utilisons est l'inégalité des accroissements finis vu en première année que nous rappelons à présent :

Théorème 1.2.24 (Inégalités des accroissements finis). *Soit f une fonction dérivable sur I . Si il existe $k \in \mathbb{R}_+$ tel pour tout $x \in I$, $|f'(x)| \leq k$ alors :*

$$\forall (a, b) \in I^2, |f(a) - f(b)| \leq k|a - b|.$$

Exemple 1.2.25. En reprenant une dernière fois l'exemple suivi, on a vu que pour tout $x \in [\sqrt{3}, +\infty[$, $f'(x) = \frac{1}{2} \left(\frac{x^2-3}{x^2} \right)$ or sur $[\sqrt{3}, +\infty[$ on a que $0 \leq x^2 - 3 < x^2$ donc en particulier $0 \leq \frac{x^2-3}{x^2} < 1$. On en déduit donc que $|f'(x)| = f'(x) \leq \frac{1}{2}$. Pour tout $n \in \mathbb{N}$, on a $u_n \in [\sqrt{3}, +\infty[$ (ainsi que $\sqrt{3} \in [\sqrt{3}, +\infty[$) donc on a par l'inégalité des accroissements finis $|f(u_n) - f(\sqrt{3})| \leq \frac{1}{2}|u_n - \sqrt{3}|$ soit

$$|u_{n+1} - \sqrt{3}| \leq \frac{1}{2}|u_n - \sqrt{3}|.$$

Montrons à présent par récurrence que pour tout $n \in \mathbb{N}$, $|u_n - \sqrt{3}| \leq \frac{5}{2^n}$.

- **Initialisation :** On a $|u_0 - \sqrt{3}| = |3 - \sqrt{3}| \leq 3 + \sqrt{3} \leq 5$ et $\frac{5}{2^0} = 5$. Comme $5 \leq 5$ notre initialisation est bien vérifiée.
- **Hérédité :** Soit $n \in \mathbb{N}$ fixé quelconque et supposons que notre inégalité est vraie au rang n et montrons là au rang $n + 1$, autrement dit supposons $|u_n - \sqrt{3}| \leq \frac{5}{2^n}$ et montrons $|u_{n+1} - \sqrt{3}| \leq \frac{5}{2^{n+1}}$. Par la question précédente, on a :

$$\begin{aligned} |u_{n+1} - \sqrt{3}| &\leq \frac{1}{2}|u_n - \sqrt{3}| \\ &\leq \frac{1}{2} \frac{5}{2^n} \text{ par hypothèse de récurrence} \\ &= \frac{5}{2^{n+1}} \end{aligned}$$

ce qui montre notre hérédité.

- **Conclusion :** La propriété est initialisé au rang 0 et est héréditaire à partir de ce rang, donc par le principe de récurrence on a montré que pour tout $n \in \mathbb{N}$, $|u_n - \sqrt{3}| \leq \frac{5}{2^n}$

1.2.4 Récapitulatif de la méthode et exemple complet

La méthodologie pour étudier une suite du type $u_{n+1} = f(u_n)$ est la même que pour toutes suites, à savoir, existence, étude de la monotonie, étude de la convergence et calcul éventuelle de la limite. Dans ce cas particulier, on a quelques outils et pistes classiques d'étude :

- Méthode 1.2.26.**
1. Étude complète de la fonction f (et éventuellement de $g : x \mapsto f(x) - x$ si cela est nécessaire par la suite).
 2. Détermination d'un intervalle stable.
 3. Existence de la suite par récurrence et localisation des termes.
 4. Étude de la monotonie de la suite (par exemple à l'aide du cas particulier de f est croissante ou de l'étude de $g : x \mapsto f(x) - x$)
 5. Étude de la convergence de la suite.
 6. Si la suite est convergente :
 - (a) Étude des points fixes à l'aide de $g : x \mapsto f(x) - x$.
 - (b) Conclusion sur la valeur de $\lim_{n \rightarrow +\infty} u_n$
 - (c) Éventuelle étude de la vitesse de convergence par l'inégalité des accroissements finis

Appliquons cette méthodologie à l'exemple suivant et en illustrant l'importance du premier terme de la suite :

Exemple 1.2.27. On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par

$$\begin{cases} u_0 \in \mathbb{R}_+ \\ \forall n \in \mathbb{N}, u_{n+1} = u_n^2 + \frac{3}{16} \end{cases}$$

On pose $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ et $g : \mathbb{R}_+ \rightarrow \mathbb{R}$. Il s'agit de deux fonctions polynômiales de classe \mathcal{C}^∞ sur leur ensemble de définition.

Dressons les tableaux de variations des deux fonctions. On a pour tout $x \in \mathbb{R}_+$, $f'(x) = 2x \geq 0$.

Pour g on observe que pour tout $x \in \mathbb{R}$, on a $g(x) = x^2 - x + \frac{3}{16}$, il s'agit d'un trinôme de discriminant $(-1)^2 - 4 \cdot 1 \cdot \frac{3}{16} = 1 - \frac{3}{4} = \frac{1}{4}$. Il admet donc deux racines qui sont $\frac{1-\frac{1}{2}}{2} = \frac{1}{4}$ et $\frac{1+\frac{1}{2}}{2} = \frac{3}{4}$ qui sont donc les deux points fixes de f .

On a de plus $g'(x) = 2x - 1$ qui est positif sur $[\frac{1}{2}, +\infty[$ et négatif sinon. On en déduit les tableaux de variations suivants :

x	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$+\infty$	
$f'(x)$		+	+	+	+	
f	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{7}{16}$	$\frac{3}{4}$	$+\infty$	
$g'(x)$		-	-	0	+	+
g	$\frac{3}{16}$	0	$-\frac{1}{16}$	0	$+\infty$	

On observe trois intervalles stables par f :

- $[0, \frac{1}{4}]$ car $f([0, \frac{1}{4}]) = [f(0), f(\frac{1}{4})] = [\frac{3}{16}, \frac{1}{4}] \subset [0, \frac{1}{4}]$ par croissance de f ;
- $[\frac{1}{4}, \frac{3}{4}]$ car $f([\frac{1}{4}, \frac{3}{4}]) = [f(\frac{1}{4}), f(\frac{3}{4})] = [\frac{1}{4}, \frac{3}{4}]$ par croissance de f .
- $[\frac{3}{4}, +\infty[$ car $f([\frac{3}{4}, +\infty[) = [f(\frac{3}{4}), \lim_{x \rightarrow +\infty} f(x)[= [\frac{3}{4}, +\infty[$ par croissance de f .

Dans ce cas là, ces trois intervalles recouvrent intégralement \mathbb{R}_+ et réalisons une disjonction des cas suivant les valeurs de u_0 .

1. Si $u_0 \in [0, \frac{1}{4}]$. On montre par récurrence que pour tout $n \in \mathbb{N}$, $u_n \in [0, \frac{1}{4}]$:

- Initialisation : L'hypothèse sur u_0 est exactement notre initialisation.
- Hérédité : Soit $n \in \mathbb{N}$ fixé quelconque et supposons la propriété vraie au rang n et montrons là au rang $n + 1$, c'est-à-dire supposons que u_n existe et que $u_n \in [0, \frac{1}{4}]$ et montrons que u_{n+1} existe et que $u_{n+1} \in [0, \frac{1}{4}]$.

Comme $u_n \in [0, \frac{1}{4}]$ et que $[0, \frac{1}{4}]$ est stable par f , on a que $f(u_n)$ existe bien et que $f(u_n) \in [0, \frac{1}{4}]$ soit u_{n+1} existe et $u_{n+1} \in [0, \frac{1}{4}]$ ce qui prouve notre hérédité.

- Conclusion : Notre propriété est bien initialisé et est héréditaire à partir du rang 0 donc par le principe de récurrence on a prouvé que $(u_n)_{n \in \mathbb{N}}$ existe bien et que pour tout $n \in \mathbb{N}$, $u_n \in [0, \frac{1}{4}]$.

Le tableau de variation de g nous apprend que pour tout $x \in [0, \frac{1}{4}]$, $g(x) \geq 0$ soit $f(x) \geq x$ or comme pour tout $n \in \mathbb{N}$, $u_n \in [0, \frac{1}{4}]$, on peut appliquer cette inégalité en u_n et l'on a $f(u_n) \geq u_n$ soit $u_{n+1} \geq u_n$ et donc la suite

$(u_n)_{n \in \mathbb{N}}$ est croissante.

La suite $(u_n)_{n \in \mathbb{N}}$ est croissante et majorée par $\frac{1}{4}$, elle est donc convergente vers un réel l qui, comme f est continue est un point fixe de f à savoir $\frac{1}{4}$ ou $\frac{3}{4}$. Comme pour tout $n \in \mathbb{N}$, $u_n \in [0, \frac{1}{4}]$ le point fixe $\frac{3}{4}$ est exclu ce qui impose que

$$\lim_{n \rightarrow +\infty} u_n = \frac{1}{4}.$$

2. Si $u_0 \in [\frac{1}{4}, \frac{3}{4}]$. On montre par récurrence que pour tout $n \in \mathbb{N}$, $u_n \in [\frac{1}{4}, \frac{3}{4}]$:

- Initialisation : L'hypothèse sur u_0 est exactement notre initialisation.
- Hérédité : Soit $n \in \mathbb{N}$ fixé quelconque et supposons la propriété vraie au rang n et montrons là au rang $n + 1$, c'est-à-dire supposons que u_n existe et que $u_n \in [\frac{1}{4}, \frac{3}{4}]$ et montrons que u_{n+1} existe et que $u_{n+1} \in [\frac{1}{4}, \frac{3}{4}]$.
Comme $u_n \in [\frac{1}{4}, \frac{3}{4}]$ et que $[\frac{1}{4}, \frac{3}{4}]$ est stable par f , on a que $f(u_n)$ existe bien et que $f(u_n) \in [\frac{1}{4}, \frac{3}{4}]$ soit u_{n+1} existe et $u_{n+1} \in [\frac{1}{4}, \frac{3}{4}]$ ce qui prouve notre hérédité.
- Conclusion : Notre propriété est bien initialisé et est héréditaire à partir du rang 0 donc par le principe de récurrence on a prouvé que $(u_n)_{n \in \mathbb{N}}$ existe bien et que pour tout $n \in \mathbb{N}$, $u_n \in [\frac{1}{4}, \frac{3}{4}]$.

Le tableau de variation de g nous apprend que pour tout $x \in [\frac{1}{4}, \frac{3}{4}]$, $g(x) \leq 0$ soit $f(x) \leq x$ or comme pour tout $n \in \mathbb{N}$, $u_n \in [\frac{1}{4}, \frac{3}{4}]$, on peut appliquer cette inégalité en u_n et l'on a $f(u_n) \leq u_n$ soit $u_{n+1} \leq u_n$ et donc la suite $(u_n)_{n \in \mathbb{N}}$ est décroissante.

La suite $(u_n)_{n \in \mathbb{N}}$ est décroissante et minorée par $\frac{1}{4}$, elle est donc convergente vers un réel l qui, comme f est continue est un point fixe de f à savoir $\frac{1}{4}$ ou $\frac{3}{4}$. Si $u_0 = \frac{3}{4}$ alors $(u_n)_{n \in \mathbb{N}}$ est la suite constante égale à $\frac{3}{4}$ et

$$\lim_{n \rightarrow +\infty} u_n = \frac{3}{4}.$$

Si $u_0 \in [\frac{1}{4}, \frac{3}{4}[$ alors par décroissance, on a $u_n \leq u_0 < \frac{3}{4}$ ce qui par passage à la limite nous donne $l \leq u_0 < \frac{3}{4}$. Cela exclu $\frac{3}{4}$ des valeurs possibles pour la limite et on a donc

$$\lim_{n \rightarrow +\infty} u_n = \frac{1}{4}.$$

3. Si $u_0 \in [\frac{3}{4}, +\infty[$. On montre par récurrence que pour tout $n \in \mathbb{N}$, $u_n \in [\frac{3}{4}, +\infty[$:

- Initialisation : L'hypothèse sur u_0 est exactement notre initialisation.
- Hérédité : Soit $n \in \mathbb{N}$ fixé quelconque et supposons la propriété vraie au rang n et montrons là au rang $n + 1$, c'est-à-dire supposons que u_n existe et que $u_n \in \left[\frac{3}{4}, +\infty\right[$ et montrons que u_{n+1} existe et que $u_{n+1} \in \left[\frac{3}{4}, +\infty\right[$.
Comme $u_n \in \left[\frac{3}{4}, +\infty\right[$ et que $\left[\frac{3}{4}, +\infty\right[$ est stable par f , on a que $f(u_n)$ existe bien et que $f(u_n) \in \left[\frac{3}{4}, +\infty\right[$ soit u_{n+1} existe et $u_{n+1} \in \left[\frac{3}{4}, +\infty\right[$ ce qui prouve notre hérédité.
- Conclusion : Notre propriété est bien initialisé et est héréditaire à partir du rang 0 donc par le principe de récurrence on a prouvé que $(u_n)_{n \in \mathbb{N}}$ existe bien et que pour tout $n \in \mathbb{N}$, $u_n \in \left[\frac{3}{4}, +\infty\right[$.

Le tableau de variation de g nous apprend que pour tout $x \in \left[\frac{3}{4}, +\infty\right[$, $g(x) \geq 0$ soit $f(x) \geq x$ or comme pour tout $n \in \mathbb{N}$, $u_n \in \left[\frac{3}{4}, +\infty\right[$, on peut appliquer cette inégalité en u_n et l'on a $f(u_n) \geq u_n$ soit $u_{n+1} \geq u_n$ et donc la suite $(u_n)_{n \in \mathbb{N}}$ est croissante.

Si la suite $(u_n)_{n \in \mathbb{N}}$ est convergente alors, elle est donc convergente vers un réel l qui, comme f est continue est un point fixe de f à savoir $\frac{1}{4}$ ou $\frac{3}{4}$.

Si $u_0 = \frac{3}{4}$, alors $(u_n)_{n \in \mathbb{N}}$ est la suite constante égale à $\frac{3}{4}$ et

$$\lim_{n \rightarrow +\infty} u_n = \frac{3}{4}.$$

Si $u_0 > \frac{3}{4}$ alors par croissance, on a $u_n \geq u_0 > \frac{3}{4}$ ce qui par passage à la limite nous donne $l \geq u_0 > \frac{3}{4}$ donc la suite convergerait vers un point fixe de f strictement plus grand que $\frac{3}{4}$ or il n'existe pas de tel point fixe donc c'est absurde. Lorsque $u_0 > \frac{3}{4}$ la suite $(u_n)_{n \in \mathbb{N}}$ diverge, soit

$$\lim_{n \rightarrow +\infty} u_n = +\infty.$$

Pour récapituler, on a que :

- Si $u_0 \in \left[0, \frac{3}{4}\right[$ la suite $(u_n)_{n \in \mathbb{N}}$ converge et $\lim_{n \rightarrow +\infty} u_n = \frac{1}{4}$;
- Si $u_0 = \frac{3}{4}$ la suite $(u_n)_{n \in \mathbb{N}}$ converge et $\lim_{n \rightarrow +\infty} u_n = \frac{3}{4}$ (c'est une suite constante);
- Si $u_0 > \frac{3}{4}$, la suite $(u_n)_{n \in \mathbb{N}}$ diverge vers $+\infty$.

Chapitre 2

Intégrales : Révisions et compléments

L'intégration est un outil incontournable de l'analyse et qui a de nombreuses applications que nous verrons dans de nombreux chapitres, notamment en probabilités avec les variables aléatoires continues.

Nous allons dans ce chapitre revenir sur l'intégration d'une fonction continue sur un segment vue en première année avant d'étendre cette notion à des fonctions sur des intervalles du type $[a, +\infty[$.

2.1 Rappel de première année

2.1.1 Définitions et règles de calculs

Le calcul intégral est fortement lié à la notion de primitive dont nous devons connaître les primitives usuelles par cœur, c'est pourquoi nous rappelons la définition et le tableau suivant :

Définition 2.1.1. Soit f une fonction continue sur $[a, b]$. On dit que F est UNE primitive de f sur $[a, b]$ si F est dérivable sur $[a, b]$ et que $F' = f$.

Remarque 2.1.2. Il a également été vu en première année que toutes fonctions continues sur un segment admet des primitives d'où l'importance de la notion de continuité. Une conséquence directe de cela est qu'une fonction qui est une primitive d'une autre fonction est automatiquement au moins de classe \mathcal{C}^1 .

Exemples 2.1.3. Soit $\lambda \in \mathbb{R}$ et soit $k \in \mathbb{R}$. On regroupe dans le tableau suivant l'ensemble des primitives usuelles qu'il faut connaître par cœur !

Fonction	Où	<u>UNE</u> primitive
0	\mathbb{R}	k
λ	\mathbb{R}	$\lambda x + k$
x	\mathbb{R}	$\frac{1}{2}x^2 + k$
$x^n, n \in \mathbb{N}$	\mathbb{R}	$\frac{x^{n+1}}{n+1} + k$
$\frac{1}{x}$	\mathbb{R}_+^* ou \mathbb{R}_-^*	$\ln x + k$
$\frac{1}{x^n}, n \geq 2$	\mathbb{R}_+^* ou \mathbb{R}_-^*	$-\frac{1}{(n-1)x^{n-1}} + k$
\sqrt{x}	\mathbb{R}_+^*	$\frac{2}{3}x^{\frac{3}{2}} + k$
$\frac{1}{\sqrt{x}}$	\mathbb{R}_+^*	$2\sqrt{x} + k$
$x^\alpha, \alpha \neq -1$	\mathbb{R}_+^*	$\frac{1}{\alpha+1}x^{\alpha+1} + k$
$\ln(x)$	\mathbb{R}_+^*	$x \ln(x) - x + k$
e^x	\mathbb{R}	$e^x + k$

On suppose à présent que u est une fonction de classe \mathcal{C}^1 sur un intervalle compatible avec les compositions suivantes :

Fonction	<u>UNE</u> primitive
$u'u$	$\frac{1}{2}u^2 + k$
$u'u^n$	$\frac{1}{n+1}u^{n+1} + k$
$u'u^\alpha, \alpha \neq -1$	$\frac{1}{\alpha+1}u^{\alpha+1} + k$
$\frac{u'}{u}$	$\ln(u) + k$
$\frac{u'}{\sqrt{u}}$	$2\sqrt{u} + k$
$u'e^u$	$e^u + k$

Les primitives nous permettent donc de calculer les intégrales de fonctions continues sur un segment de la manière suivante :

Théorème 2.1.4. Soit f une fonction continue sur I et soit F une primitive de f . On a donc pour tout $(a, b) \in I^2$,

$$\int_a^b f(t) dt = [F(t)]_a^b = F(b) - F(a).$$

2.1.2 Outil de calculs

Pour calculer des intégrales sur un segment nous avons un certain nombre de propriétés fort utiles pour simplifier nos calculs :

Propriété 2.1.5. Soit f et g deux fonctions continues sur un intervalle $[a, b]$, soit $c \in [a, b]$ et soit $\lambda \in \mathbb{R}$. On a alors :

- $\int_a^b f(t) + g(t) dt = \int_a^b f(t) dt + \int_a^b g(t) dt$;
- $\int_a^b \lambda f(t) dt = \lambda \int_a^b f(t) dt$.
- (Relation de Chasles) $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$.
- Si $a \leq b$ et que pour tout $x \in [a, b]$, $f(x) \leq g(x)$ alors $\int_a^b f(t) dt \leq \int_a^b g(t) dt$.

Dans le cas de fonctions positives on a le résultat suivant :

Propriété 2.1.6 (Positivité de l'intégrale). Soit f une fonction continue positive sur $[a, b]$ avec $a \leq b$. On a donc

$$0 \leq \int_a^b f(t) dt.$$

Remarque 2.1.7. Le cas des fonctions positives sera étudié plus en détails à la fin de ce chapitre avec la notion de convergence absolue.

Concluons ces rappels sur les intégrales de fonctions continues sur un segment par deux techniques de calculs puissantes :

Théorème 2.1.8 (Changement de variables). Soit f une fonction continue sur I et g une fonction de classe \mathcal{C}^1 sur $[a, b]$ à valeurs dans I . On a alors

$$\int_a^b f(g(x))g'(x) dx = \int_{g(a)}^{g(b)} f(t) dt.$$

Théorème 2.1.9 (Intégration par parties). Soit u et v deux fonctions de classe \mathcal{C}^1 sur $[a, b]$ alors

$$\int_a^b u'(t)v(t) dt = [u(t)v(t)]_a^b - \int_a^b u(t)v'(t) dt.$$

2.2 Intégrales impropres

Nous allons à présent nous intéresser à la notion d'intégrale d'une fonction non plus sur un segment mais sur un intervalle infini. Pour cela la notion de limite va intervenir et nous redoublerons d'attention pour ne pas utiliser de résultats sur ces intégrales sans vérifier les convergences des limites concernées.

2.2.1 Définitions

Définition 2.2.1. Soit f une fonction continue sur un intervalle de la forme $[a, +\infty[$ où $a \in \mathbb{R}$.

- Si la limite $\lim_{x \rightarrow +\infty} \int_a^x f(t) dt$ existe et est finie alors on dit que l'intégrale impropre en $+\infty$ $\int_a^{+\infty} f(t) dt$ est convergente et sa valeur est la valeur de la limite trouvée.
- Si la limite n'existe pas ou est divergente on dit alors que l'intégrale impropre $\int_a^{+\infty} f(t) dt$ est divergente.

Remarque 2.2.2. On parle parfois d'intégrale généralisée au lieu d'intégrale impropre.

On peut donc établir la méthodologie suivante pour étudier une intégrale impropre en $+\infty$:

- Méthode 2.2.3.**
1. On vérifie la régularité de la fonction f sur l'intervalle $[a, +\infty[$;
 2. On pose un $x \in [a, +\infty[$ quelconque et l'on considère le segment $[a, x]$. On étudie l'intégrale sur un segment $\int_a^x f(t) dt$ en utilisant les outils de première année ;
 3. On étudie la limite $\lim_{x \rightarrow +\infty} \int_a^x f(t) dt$;
 4. On conclut quant à la nature de l'intégrale impropre.

Exemples 2.2.4. • Étudions $\int_1^{+\infty} \frac{dt}{t^2}$.

On observe que $t \mapsto \frac{1}{t^2}$ est continue sur $[1, +\infty[$ en tant que quotient d'un polynôme ne s'annulant pas sur cet intervalle.

Soit $x \in [1, +\infty[$ et l'on a

$$\begin{aligned} \int_1^x \frac{dt}{t^2} &= \left[\frac{-1}{2-1} \frac{1}{t^{2-1}} \right]_1^x = \left[\frac{-1}{t} \right]_1^x \\ &= \frac{-1}{x} - \frac{-1}{1} = 1 - \frac{1}{x} \end{aligned}$$

et comme $\lim_{x \rightarrow +\infty} 1 - \frac{1}{x} = 1$ on a que l'intégrale impropre $\int_1^{+\infty} \frac{dt}{t^2}$ est convergente et

$$\int_1^{+\infty} \frac{dt}{t^2} = 1.$$

- Étudions $\int_0^{+\infty} e^t dt$.

L'exponentielle est continue sur \mathbb{R} donc également sur $[0, +\infty[$. Soit $x \in [0, +\infty[$ et l'on a

$$\int_0^x e^t dt = [e^t]_0^x = e^x - 1$$

et comme $\lim_{x \rightarrow +\infty} e^x - 1 = +\infty$ on a que l'intégrale impropre $\int_0^{+\infty} e^t dt$ est divergente.

On considère parfois des intégrales impropres en $-\infty$ selon le même modèle de définition :

Définition 2.2.5. Soit f une fonction continue sur un intervalle de la forme $] -\infty, b]$ où $b \in \mathbb{R}$.

- Si la limite $\lim_{x \rightarrow -\infty} \int_x^b f(t) dt$ existe et est finie alors on dit que l'intégrale impropre en $-\infty$ $\int_{-\infty}^b f(t) dt$ est convergente et sa valeur est la valeur de la limite trouvée.
- Si la limite n'existe pas ou est divergente on dit alors que l'intégrale impropre $\int_{-\infty}^b f(t) dt$ est divergente.

Exemple 2.2.6. Étudions $\int_{-\infty}^{-2} \frac{1}{t} dt$.

La fonction inverse est continue sur \mathbb{R}^* donc également sur $] -\infty, -2]$. Soit $x \in] -\infty, -2]$ et l'on a

$$\int_x^{-2} \frac{1}{t} dt = [\ln(|t|)]_x^{-2} = \ln(|-2|) - \ln(|x|) = \ln(2) - \ln(-x)$$

et comme $\lim_{x \rightarrow -\infty} \ln(2) - \ln(-x) = +\infty$ on a que l'intégrale impropre $\int_{-\infty}^{-2} \frac{1}{t} dt$ est divergente.

Remarque 2.2.7. La méthodologie employée ici est l'adaptation direct de celle utilisée pour les intégrales impropres en $+\infty$.

2.2.2 Exemples fondamentaux

Deux familles d'exemples fondamentaux sont à connaître et maîtriser absolument !

Nous verrons dans un prochain chapitre que beaucoup de convergences d'intégrales impropres dépendront de ces exemples de références (et même la convergence de certaines séries numériques).

2.2.2.1 L'intégrale d'exponentielle

Théorème 2.2.8. Soit $\alpha \in \mathbb{R}$, on a alors que l'intégrale impropre $\int_0^{+\infty} e^{-\alpha t} dt$ est :

- convergente si $\alpha > 0$ et dans ce cas $\int_0^{+\infty} e^{-\alpha t} dt = \frac{1}{\alpha}$.
- divergente si $\alpha \leq 0$.

Démonstration. La fonction $t \mapsto e^{-\alpha t}$ est de classe \mathcal{C}^∞ sur \mathbb{R} en tant que composée d'un polynôme et de la fonction exponentielle. A fortiori elle est donc continue sur $[0, +\infty[$.

Soit $x \in [0, +\infty[$ et calculons $\int_0^x e^{-\alpha t} dt$.

- Si $\alpha = 0$, alors on a

$$\int_0^x e^{-\alpha t} dt = \int_0^x 1 dt = [t]_0^x = x$$

;

- Si $\alpha \neq 0$, alors on a

$$\begin{aligned} \int_0^x e^{-\alpha t} dt &= \left[\frac{1}{-\alpha} e^{-\alpha t} \right]_0^x \\ &= \frac{-e^{-\alpha x}}{\alpha} - \frac{-e^0}{\alpha} = \frac{1 - e^{-\alpha x}}{\alpha} \end{aligned}$$

On peut donc à présent calculer les limites des intégrales que nous venons de calculer :

- Si $\alpha = 0$ alors

$$\lim_{x \rightarrow +\infty} \int_0^x e^{-\alpha t} dt = \lim_{x \rightarrow +\infty} x = +\infty,$$

donc l'intégrale impropre est divergente dans ce cas.

- Si $\alpha < 0$ alors

$$\lim_{x \rightarrow +\infty} \int_0^x e^{-\alpha t} dt = \lim_{x \rightarrow +\infty} \frac{1 - e^{-\alpha x}}{\alpha} = +\infty,$$

donc l'intégrale impropre est divergente dans ce cas.

- Si $\alpha > 0$ alors

$$\lim_{x \rightarrow +\infty} \int_0^x e^{-\alpha t} dt = \lim_{x \rightarrow +\infty} \frac{1 - e^{-\alpha x}}{\alpha} = \frac{1 - 0}{\alpha} = \frac{1}{\alpha},$$

donc l'intégrale impropre est convergente dans ce cas et l'on a

$$\int_0^{+\infty} e^{-\alpha t} dt = \frac{1}{\alpha}.$$

□

Exemples 2.2.9. 1. On a vu dans un exemple précédent que $\int_0^{+\infty} e^t dt$ est divergente, ce qui correspond au cas $\alpha = -1$.

2. On a que $\int_0^{+\infty} e^{-\frac{t}{2}} dt$ est convergente car il s'agit d'une exponentielle avec $\alpha = \frac{1}{2}$. On a de plus $\int_0^{+\infty} e^{-\frac{t}{2}} dt = \frac{1}{\frac{1}{2}} = 2$.

2.2.2.2 Les intégrales de Riemann

La seconde famille d'exemples fondamentaux sont les intégrales de Riemann :

Définition 2.2.10. On appelle intégrale de Riemann en $+\infty$ les intégrales impropres du type $\int_c^{+\infty} \frac{1}{t^\alpha} dt$ ou $\alpha \in \mathbb{R}$ et $c \in \mathbb{R}_+^*$.

Théorème 2.2.11. L'intégrale de Riemann $\int_c^{+\infty} \frac{1}{t^\alpha} dt$ est convergente si et seulement si $\alpha > 1$.

Démonstration. Soit $t \mapsto \frac{1}{t^\alpha}$ est de classe \mathcal{C}^∞ sur \mathbb{R}_+^* en tant qu'inverse d'une fonction puissance (on pourrait également dire qu'il s'agit de la fonction puissance $-\alpha$). Comme $c > 0$ par hypothèse, on a en particulier que cette fonction est continue sur $[c, +\infty[$.

Soit $x \in [c, +\infty[$ et calculons $\int_c^x \frac{1}{t^\alpha} dt$.

- Si $\alpha \neq 1$, on a alors

$$\begin{aligned} \int_c^x \frac{1}{t^\alpha} dt &= \int_c^x t^{-\alpha} dt = \left[\frac{1}{-\alpha + 1} t^{-\alpha+1} \right]_c^x \\ &= \left[\frac{t^{1-\alpha}}{1-\alpha} \right]_c^x = \frac{x^{1-\alpha}}{1-\alpha} - \frac{c^{1-\alpha}}{1-\alpha} \\ &= \frac{x^{1-\alpha} - c^{1-\alpha}}{1-\alpha} \end{aligned}$$

- Si $\alpha = 1$, on a alors

$$\begin{aligned}\int_c^x \frac{1}{t^\alpha} dt &= \int_c^x \frac{1}{t} dt = [\ln(t)]_c^x \\ &= \ln(x) - \ln(c)\end{aligned}$$

On peut donc à présent calculer les limites des intégrales que nous venons de calculer :

- Si $\alpha < 1$, alors $1 - \alpha > 0$ et on a

$$\lim_{x \rightarrow +\infty} \int_c^x \frac{1}{t^\alpha} dt = \lim_{x \rightarrow +\infty} \frac{x^{1-\alpha} - c^{1-\alpha}}{1-\alpha} = +\infty,$$

donc l'intégrale impropre est divergente dans ce cas ;

- Si $\alpha > 1$, alors $1 - \alpha < 0$ et on a

$$\lim_{x \rightarrow +\infty} \int_c^x \frac{1}{t^\alpha} dt = \lim_{x \rightarrow +\infty} \frac{x^{1-\alpha} - c^{1-\alpha}}{1-\alpha} = \frac{0 - c^{1-\alpha}}{1-\alpha},$$

donc l'intégrale impropre est convergente ;

- Si $\alpha = 1$, alors

$$\lim_{x \rightarrow +\infty} \int_c^x \frac{1}{t^\alpha} dt = \lim_{x \rightarrow +\infty} \ln(x) - \ln(c) = +\infty,$$

donc l'intégrale impropre est divergente dans ce cas.

On en déduit donc que l'intégrale impropre $\int_c^\infty \frac{1}{t^\alpha} dt$ converge si $\alpha > 1$ et diverge si $\alpha \leq 1$.

□

Remarque 2.2.12. Dans le cas classique où $c = 1$, la valeur dans le cas de convergence (c'est-à-dire lorsque $\alpha > 1$) est à connaître. Par les calculs précédents, on a

$$\int_1^\infty \frac{1}{t^\alpha} dt = \frac{1}{\alpha - 1}.$$

Exemples 2.2.13. Dans les exemples précédents, on a déjà vu deux intégrales de Riemann. $\int_1^{+\infty} \frac{dt}{t}$ est divergente puisque ici $\alpha = 1$ et $\int_1^{+\infty} \frac{dt}{t^2}$ est convergente puisque ici $\alpha = 2$.

2.2.3 Règles de calculs

Comme dans le cas de l'intégrale sur un segment, on peut utiliser des règles de calculs pour simplifier nos calculs pour les intégrales impropres à conditions de bien respecter les convergences.

Toutes les propriétés vont être énoncées pour les intégrales impropres en $+\infty$ mais sont adaptable aux intégrales impropres en $-\infty$.

Propriété 2.2.14 (Linéarité de l'intégrale). *Soit f et g deux fonctions continues sur un intervalle $[a, +\infty[$ et soit $\lambda \in \mathbb{R}$. On a alors si $\int_a^{+\infty} f(t) dt$ et $\int_a^{+\infty} g(t) dt$ convergent :*

- $\int_a^{+\infty} f(t) + g(t) dt = \int_a^{+\infty} f(t) dt + \int_a^{+\infty} g(t) dt;$
- $\int_a^{+\infty} \lambda f(t) dt = \lambda \int_a^{+\infty} f(t) dt.$

Remarque 2.2.15. La réciproque est fautive, en effet $\int_1^{+\infty} \frac{1}{t} dt$ et $\int_1^{+\infty} -\frac{1}{t} dt$ sont des intégrales impropres divergentes or $\int_1^{+\infty} \frac{1}{t} - \frac{1}{t} dt = \int_1^{+\infty} 0 dt$ est convergente est vaut 0.

Exemple 2.2.16. Étudions $\int_0^{+\infty} e^{-3t} - 4e^{-t} dt$. Comme 3 et 1 sont strictement positifs, on a que $\int_0^{+\infty} e^{-3t} dt$ et $\int_0^{+\infty} e^{-t} dt$ sont convergentes et donc par linéarité on a que $\int_0^{+\infty} e^{-3t} - 4e^{-t} dt$ est convergente et l'on a :

$$\begin{aligned} \int_0^{+\infty} e^{-3t} - 4e^{-t} dt &= \int_0^{+\infty} e^{-3t} dt - 4 \int_0^{+\infty} e^{-t} dt \\ &= \frac{1}{3} - 4 \frac{1}{1} = \frac{1}{3} - 4 = -\frac{11}{3} \end{aligned}$$

La relation de Chasles peut également s'utiliser de la manière suivante pour les intégrales impropres :

Propriété 2.2.17 (Relation de Chasles). *Soit f une fonction continue sur $[a, +\infty[$ tel que $\int_a^{+\infty} f(t) dt$ soit une intégrale impropre convergente. Soit $c \in [a, +\infty[$ alors $\int_c^{+\infty} f(t) dt$ est convergente et on a :*

$$\int_a^{+\infty} f(t) dt = \int_a^c f(t) dt + \int_c^{+\infty} f(t) dt.$$

Remarque 2.2.18. Si l'on ne sait pas que $\int_a^{+\infty} f(t) dt$ est convergente mais que l'on a $\int_c^{+\infty} f(t) dt$ qui l'est, alors la continuité de f sur le segment $[a, c]$ permet de conclure à la convergence de $\int_a^{+\infty} f(t) dt$.

Exemple 2.2.19. Étudions $\int_1^{+\infty} \frac{1}{t^7} + e^{-6t} dt$.

On reconnaît une intégrale de Riemann de paramètre $\alpha = 7 > 1$ donc $\int_1^{+\infty} \frac{dt}{t^7}$ est convergente. De plus comme $\int_0^{+\infty} e^{-6t} dt$ est convergente en tant qu'intégrale d'exponentielle où $\alpha > 0$ donc par relation de Chasles $\int_1^{+\infty} e^{-6t} dt$ est également convergente et

$$\int_0^{+\infty} e^{-6t} dt = \int_0^1 e^{-6t} dt + \int_1^{+\infty} e^{-6t} dt,$$

soit

$$\begin{aligned} \int_1^{+\infty} e^{-6t} dt &= \int_0^{+\infty} e^{-6t} dt - \int_0^1 e^{-6t} dt \\ &= \frac{1}{6} - \left[\frac{-1}{6} e^{-6t} \right]_0^1 \\ &= \frac{1}{6} - \left(\frac{-1}{6} e^{-6} - \frac{-1}{6} e^0 \right) \\ &= \frac{1}{6} + \frac{e^{-6}}{6} - \frac{1}{6} = \frac{e^{-6}}{6} \end{aligned}$$

On en déduit donc par linéarité que

$$\begin{aligned} \int_1^{+\infty} \frac{1}{t^7} + e^{-6t} dt &= \int_1^{+\infty} \frac{1}{t^7} dt + \int_1^{+\infty} e^{-6t} dt \\ &= \frac{1}{7-1} + \frac{e^{-6}}{6} = \frac{1+e^{-6}}{6} \end{aligned}$$

La positivité de la fonction a également les mêmes conséquences pour les intégrales impropres que pour les intégrales sur des segments :

Propriété 2.2.20 (Positivité de l'intégrale). *Soit f une fonction continue positive sur $[a, +\infty[$ tel que $\int_a^{+\infty} f(t) dt$ est convergente. On a donc*

$$0 \leq \int_a^{+\infty} f(t) dt.$$

Exemple 2.2.21. Soit $\alpha > 1$, $[1, +\infty[\rightarrow \mathbb{R}$ est continue et est positive sur

$$t \mapsto \frac{1}{t^\alpha}$$

son ensemble de définition et comme l'intégrale de Riemann $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ est convergente par hypothèse sur α , on a par positivité de l'intégrale que $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ est positive.

Cela nous donne donc le résultat suivant pour comparer les intégrales impropres de deux fonctions :

Corollaire 2.2.22. Soit f et g deux fonctions continues sur $[a, +\infty[$ telles que pour tout $x \in [a, +\infty[$, $f(x) \leq g(x)$. Si $\int_a^{+\infty} f(t) dt$ et $\int_a^{+\infty} g(t) dt$ sont convergentes, on a alors

$$\int_a^{+\infty} f(t) dt \leq \int_a^{+\infty} g(t) dt.$$

Exemple 2.2.23. On a pour tout $t \in [1, +\infty[$, $\frac{1}{t^4} \leq \frac{1}{t^2}$ et que les intégrales associées sont convergentes en tant qu'intégrales de Riemann et on a donc

$$\int_1^{+\infty} \frac{1}{t^4} dt \leq \int_1^{+\infty} \frac{1}{t^2} dt.$$

Théorème 2.2.24. Soit f une fonction continue positive sur $[a, +\infty[$ tel que $\int_a^{+\infty} f(t) dt$ est convergente. On a alors que $\int_a^{+\infty} f(t) dt = 0$ si et seulement si f est la fonction nulle sur $[a, b]$.

Remarque 2.2.25. L'intégration par parties et le changement de variables sur des intégrales impropres sont hors programme. En effet elles soulèvent quelques problèmes théoriques que nous ne pouvons pas traiter. On contournera ce problème par l'utilisation de ces techniques sur les intégrales sur les segments $[a, x]$ puis on calculera les limites lorsque cela est nécessaires.

Exemple 2.2.26. Étudions $\int_0^{+\infty} te^{-t} dt$.

$t \mapsto te^{-t}$ est continue sur $[0, +\infty[$ en tant que produit de fonctions usuelles continues. Soit $x \in [0, +\infty[$ et calculons $\int_0^x te^{-t} dt$ à l'aide d'une intégration par partie (les fonctions considérées sont \mathcal{C}^1 en tant que polynôme et exponentielle de polynôme.)

$$\begin{aligned} \int_0^x te^{-t} dt &= [-te^{-t}]_0^x - \int_0^x 1 \cdot (-e^{-t}) dt \\ &= -xe^{-x} + 0 \cdot e^0 - [e^{-t}]_0^x \\ &= -xe^{-x} - (e^{-x} - e^0) \\ &= 1 - xe^{-x} - e^{-x} \end{aligned}$$

Par croissance comparée on a que $\lim_{x \rightarrow +\infty} xe^{-x} = 0$ donc on a

$$\lim_{x \rightarrow +\infty} \int_0^x te^{-t} dt = 1 - 0 - 0 = 1.$$

L'intégrale $\int_0^{+\infty} te^{-t} dt$ est donc convergente et l'on a

$$\int_0^{+\infty} te^{-t} dt = 1.$$

2.2.4 Intégrales doublement impropres

Une question naturelle que l'on peut se poser est la question d'une intégrale ayant comme bornes $-\infty$ et $+\infty$. Il s'agit des intégrales doublement impropres.

Définition 2.2.27. Soit f une fonction continue sur \mathbb{R} . Si il existe un réel c tel que les intégrales impropres $\int_{-\infty}^c f(t) dt$ et $\int_c^{+\infty} f(t) dt$ sont convergentes, alors l'intégrale doublement impropre $\int_{-\infty}^{+\infty} f(t) dt$ est convergente et l'on a

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{-\infty}^c f(t) dt + \int_c^{+\infty} f(t) dt.$$

Sinon, on dit qu'elle diverge.

Remarques 2.2.28. • Si une intégrale doublement impropre converge tout les réels c donneront le résultat par relation de Chasles.

- Une erreur classique est de considérer la limite de $\int_{-x}^x f(t) dt$ pour prouver la convergence d'une intégrale doublement impropre, cela est totalement faux. Un effet de compensation peut avoir lieu entre les deux "cotés" de l'intégrale qui masquera la divergence potentielle de l'intégrale doublement impropre.
- Lorsque la fonction f est paire ou impaire, l'usage de 0 pour c peut nous aider pour limiter le travail sur les deux intégrales impropres à étudier.

Exemples 2.2.29. 1. On pose $f : \mathbb{R} \rightarrow \mathbb{R}$. f est une fonction continue et l'on va étudier l'intégrale doublement impropre $\int_{-\infty}^{+\infty} f(t) dt$ en étudions les intégrales impropres $\int_{-\infty}^0 f(t) dt$ et $\int_0^{+\infty} f(t) dt$.

$x \mapsto xe^{-|x|}$

On observe que f est impaire (c'est-à-dire que $f(-t) = -f(t)$ pour tout $t \in \mathbb{R}$) et donc en particulier pour tout $x \in \mathbb{R}_+$, $\int_{-x}^0 f(t) dt = -\int_0^x f(t) dt$ (par changement de variable classique) et donc $\int_{-\infty}^0 f(t) dt$ converge si et seulement si $\int_0^{+\infty} f(t) dt = \int_0^{+\infty} te^{-t} dt$ est convergente, ce qui est le cas par l'exemple précédent.

Comme $\int_0^{+\infty} f(t) dt$ est convergente, on a que $\int_{-\infty}^0 f(t) dt$ l'est également donc par définition, l'intégrale doublement impropre $\int_{-\infty}^{+\infty} f(t) dt$ est bien convergente.

On a de plus

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^{+\infty} f(t) dt$$

$$\begin{aligned}
&= \lim_{x \rightarrow +\infty} \int_{-x}^0 f(t) dt + \int_0^{+\infty} f(t) dt \\
&= - \lim_{x \rightarrow +\infty} \int_0^x f(t) dt + \int_0^{+\infty} f(t) dt \\
&= - \int_0^{+\infty} f(t) dt + \int_0^{+\infty} f(t) dt \\
&= 0
\end{aligned}$$

2. On considère la fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ et étudions l'intégrale doublement impropre $\int_{-\infty}^{+\infty} g(t) dt$.

ment impropre $\int_{-\infty}^{+\infty} g(t) dt$.

Soit $c \in \mathbb{R}$ et observons $\int_c^{+\infty} g(t) dt$. On a pour tout $x \in [c, +\infty[$,

$$\int_c^x g(t) dt = \int_c^x t^3 dt = \left[\frac{t^4}{4} \right]_c^x = \frac{x^4 - c^4}{4}$$

or $\lim_{x \rightarrow +\infty} \int_c^x g(t) dt = +\infty$ donc l'intégrale $\int_c^{+\infty} g(t) dt$ est divergente donc

l'intégrale doublement impropre $\int_{-\infty}^{+\infty} g(t) dt$ l'est également.

On prendra donc attention à de ne pas aller trop vite car

$$\begin{aligned}
\int_{-x}^x g(t) dt &= \left[\frac{t^4}{4} \right]_{-x}^x = \frac{x^4 - (-x)^4}{4} \\
&= \frac{x^4 - x^4}{4} = 0
\end{aligned}$$

et l'on pourra être tenté de dire que l'intégrale doublement impropre converge, ce qui est faux, comme nous l'avons prouvé précédemment !

Chapitre 3

Espaces vectoriels

En première année, uniquement l'espace vectoriel \mathbb{R}^n et ces sous-espaces vectoriels ont été abordées. Nous allons cette année étudier deux nouvelles familles d'espaces vectoriels. L'objectif est de fournir tout le nécessaire à l'étude des applications linéaires et les bases pour la réduction des matrices carrées qui est l'objectif du programme d'algèbre de l'année.

Contrairement à la première année, nous ne mettrons plus de flèches sur nos vecteurs. Nous devons donc avoir conscience de quels éléments sont des vecteurs et de quels éléments sont des scalaires.

3.1 Construction et définition

3.1.1 Loi de composition interne et externe

Pour définir un espace vectoriel, nous avons besoin de définir ce que l'on appelle une loi de composition interne et une loi de composition externe qui nous permettront d'additionner des vecteurs entre eux et les multiplier par des coefficients réels.

Définitions 3.1.1. Soit E un ensemble.

1. Soit $+ : E \times E \rightarrow E$. On dit que $+$ est une loi de composition

$$(x, y) \mapsto x + y$$

interne sur E lorsque :

- $\forall (x, y) \in E^2, x + y = y + x$;
- $\forall (x, y, z) \in E^3, (x + y) + z = x + (y + z)$
- $\exists 0_E \in E, \forall x \in E, x + 0_E = 0_E + x = x$

- $\forall x \in E, \exists y \in E, x + y = 0_E$. On note alors $-x$ cet élément.

2. Soit $\cdot : \mathbb{R} \times E \rightarrow E$. On dit que \cdot est une loi de composition ex-

$$(\lambda, x) \mapsto \lambda.x$$

terne sur E lorsque :

- $\forall (\lambda, \mu) \in \mathbb{R}^2, \forall x \in E, \lambda.(\mu.x) = (\lambda\mu).x$;
- $\forall \lambda \in \mathbb{R}, \forall (x, y) \in E^2, \lambda.(x + y) = \lambda.x + \lambda.y$
- $\forall (\lambda, \mu) \in \mathbb{R}^2, \forall x \in E, (\lambda + \mu).x = \lambda.x + \mu.x$;
- $\forall x \in E, 1.x = x$

Exemples 3.1.2. 1. Sur \mathbb{R}^n l'addition

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$((x_1, \dots, x_n), (y_1, \dots, y_n)) \mapsto (x_1 + y_1, \dots, x_n + y_n)$$

et la multiplication par un réel

$$\cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$(\lambda, (x_1, \dots, x_n)) \mapsto (\lambda x_1, \dots, \lambda x_n)$$

sont respectivement une loi de composition interne et une loi de composition externe sur \mathbb{R}^n .

2. Sur $\mathbb{R}_n[x]$ l'addition

$$+ : \mathbb{R}_n[x] \times \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$$

$$((P, Q)) \mapsto P + Q$$

et la multiplication par un réel

$$\cdot : \mathbb{R} \times \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$$

$$(\lambda, P) \mapsto \lambda.P$$

sont respectivement une loi de composition interne et une loi de composition externe sur $\mathbb{R}_n[x]$.

3. Sur $\mathcal{M}_{n,p}(\mathbb{R})$ l'addition

$$\begin{aligned} + & : \mathcal{M}_{n,p}(\mathbb{R}) \times \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathcal{M}_{n,p}(\mathbb{R}) \\ & (A, B) \quad \mapsto \quad A + B \end{aligned}$$

et la multiplication par un réel

$$\begin{aligned} \cdot & : \mathbb{R} \times \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathcal{M}_{n,p}(\mathbb{R}) \\ & (\lambda, A) \quad \mapsto \quad \lambda A \end{aligned}$$

sont respectivement une loi de composition interne et une loi de composition externe sur $\mathcal{M}_{n,p}(\mathbb{R})$.

Une notion fondamentale est celle de combinaison linéaire :

Définition 3.1.3. Soit $(E, +, \cdot)$ un ensemble muni d'une loi de composition interne et d'une loi composition externe. Soit $p \in \mathbb{N}^*$, soit $x \in E$ et soit $(x_1, \dots, x_p) \in E^p$. On dit que x est une combinaison linéaire de x_1, \dots, x_p si il existe p réels $\lambda_1, \dots, \lambda_p$ tels que

$$x = \lambda_1 x_1 + \dots + \lambda_p x_p = \sum_{i=1}^p \lambda_i x_i.$$

Exemples 3.1.4. • Dans $\mathcal{M}_2(\mathbb{R})$, la matrice $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ est une combinaison linéaire des matrices $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$ car on a par exemple

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

ou encore

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} = 2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + (-1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Dans $\mathbb{R}_2[x]$, $x^2 - 1$ n'est pas une combinaison linéaire de $x - 1$ et $x + 1$ car toutes combinaisons linéaires de $x - 1$ et $x + 1$ n'est qu'au plus de degré 1. Un produit de vecteurs (ici de polynômes) n'est PAS une combinaison linéaire!

3.1.2 Définitions et exemples fondamentales

Définition 3.1.5. Soit $(E, +, \cdot)$ un ensemble muni d'une loi de composition interne et d'une loi composition externe. Soit $n \in \mathbb{N}$. On dit que E est un espace vectoriel réel de dimension n si il existe une bijection ψ entre E et \mathbb{R}^n qui préserve les combinaisons linéaires, autrement dit si il existe une bijection ψ entre E et \mathbb{R}^n telle que pour tout $p \in \mathbb{N}^*$, pour tout $(x_1, \dots, x_p) \in E^p$ et pour tout $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$, on a

$$\psi \left(\sum_{i=1}^p \lambda_i x_i \right) = \sum_{i=1}^p \lambda_i \psi(x_i).$$

Remarque 3.1.6. Il est possible de montrer qu'un ensemble est un espace vectoriel uniquement si il préserve les combinaisons linéaires de deux éléments puis en retrouvant la définition par une récurrence.

Remarque 3.1.7. En conservant les notation de la définition précédente, on a que $\psi(0_E) = 0_{\mathbb{R}^n}$, soit que l'élément neutre de E s'envoie sur le neutre de \mathbb{R}^n .

Notation 3.1.8. • On note $\dim(E)$ la dimension de l'espace vectoriel réel E .

- On appelle les éléments de E les vecteurs de E et les coefficients réels λ sont appelées les scalaires. Les scalaires se placent toujours devant un vecteur!
- On note en général $+$ la loi de composition interne dans E et dans \mathbb{R}^n , on prendra garde à ne pas faire de confusion entre les divers éléments.
- On note en général λx et non pas $\lambda \cdot x$. On prendra garde à ne pas oublier également la présence de la loi de composition externe.

Remarque 3.1.9. La définition que l'on donne limite en réalité les espaces vectoriels aux espaces vectoriels de dimensions finis mais il existe en réalité des espaces vectoriels de dimension infini qui sont hors programme.

Donnons quelques règles de calculs dans les espaces vectoriels :

Propriété 3.1.10. Soit E un espace vectoriel réel de dimension finie. On a alors pour tout $x \in E$ et pour tout $\lambda \in \mathbb{R}$:

- $0x = 0_E$
- $\lambda 0_E = 0_E$
- $(-1).x = -x$
- $\lambda x = 0_E$ si et seulement si ($\lambda = 0$ ou $x = 0_E$).

Citons à présents les trois exemples de références que nous utiliserons tout au long de l'année :

- Théorème 3.1.11.** 1. Soit $n \in \mathbb{N}$, \mathbb{R}^n est un espace vectoriel réel de dimension n .
2. Soit $n \in \mathbb{N}$, $\mathbb{R}_n[x]$ est un espace vectoriel réel de dimension $n + 1$.
3. Soit $(n, p) \in (\mathbb{N}^*)^2$, $\mathcal{M}_{n,p}(\mathbb{R})$ est un espace vectoriel réel de dimension np .

Démonstration. 1. \mathbb{R}^n est trivialement en bijection avec lui même par l'application identité qui conserve bien les combinaisons linéaires.

2. On pose $f : \mathbb{R}_n[x] \rightarrow \mathbb{R}^{n+1}$. Il s'agit d'une bijection car :
- $$\sum_{i=0}^n a_i x^i \mapsto (a_0, a_1, \dots, a_n)$$

- f est injective car si $f\left(\sum_{i=0}^n a_i x^i\right) = f\left(\sum_{i=0}^n b_i x^i\right)$, on a alors $(a_0, \dots, a_n) = (b_0, \dots, b_n)$ soit pour tout $i \in \llbracket 0, n \rrbracket$, $a_i = b_i$, soit $\sum_{i=0}^n a_i x^i = \sum_{i=0}^n b_i x^i$.
- f est surjective car pour tout $(a_0, \dots, a_n) \in \mathbb{R}^{n+1}$, $\sum_{i=0}^n a_i x^i$ est un polynôme de $\mathbb{R}_n[x]$ et $f\left(\sum_{i=0}^n a_i x^i\right) = (a_0, \dots, a_n)$.

De plus pour tout $(P, Q) \in \mathbb{R}_n[x]^2$ et pour tout $(\lambda, \mu) \in \mathbb{R}^2$, on a $f(\lambda P + \mu Q) = \lambda f(P) + \mu f(Q)$. En effet, on pose $P = \sum_{i=0}^n a_i x^i$ et $Q = \sum_{i=0}^n b_i x^i$. On

a donc $\lambda P + \mu Q = \sum_{i=0}^n (\lambda a_i + \mu b_i) x^i$ donc

$$\begin{aligned} f(\lambda P + \mu Q) &= (\lambda a_0 + \mu b_0, \dots, \lambda a_n + \mu b_n) \\ &= \lambda(a_0, \dots, a_n) + \mu(b_0, \dots, b_n) \\ &= \lambda f(P) + \mu f(Q) \end{aligned}$$

On en déduit par récurrence que f conserve toutes les combinaisons linéaires.

$\mathbb{R}_n[x]$ est donc bien un espace vectoriel réel de dimension $n + 1$.

3. On pose $f : \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathbb{R}^{np}$. Il s'agit
 $(a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}} \mapsto (a_{1,1}, a_{1,2}, \dots, a_{1,p}, a_{2,1}, \dots, a_{n,p})$
d'une bijection car :

- f est injective car si $f \left((a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}} \right) = f \left((b_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}} \right)$, on a alors pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, $a_{i,j} = b_{i,j}$, soit l'égalité entre les deux matrices recherchées.
- f est surjective car pour tout $(a_{1,1}, a_{1,2}, \dots, a_{1,p}, a_{2,1}, \dots, a_{n,p}) \in \mathbb{R}^{np}$, $(a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}}$ est une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ et

$$f \left((a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}} \right) = (a_{1,1}, a_{1,2}, \dots, a_{1,p}, a_{2,1}, \dots, a_{n,p}).$$

De plus pour tout $(A, B) \in \mathcal{M}_{n,p}(\mathbb{R})^2$ et pour tout $(\lambda, \mu) \in \mathbb{R}^2$, on a $f(\lambda A + \mu B) = \lambda f(A) + \mu f(B)$. En effet, on pose $A = (a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}}$ et $B = (b_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}}$. On a donc $\lambda A + \mu B = (\lambda a_{i,j} + \mu b_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, n \rrbracket}}$ donc

$$\begin{aligned} f(\lambda A + \mu B) &= (\lambda a_{1,1} + \mu b_{1,1}, \lambda a_{1,2} + \mu b_{1,2}, \dots, \lambda a_{1,p} + \mu b_{1,p}, \lambda a_{2,1} + \mu b_{2,1}, \dots, \lambda a_{n,p} + \mu b_{n,p}) \\ &= \lambda(a_{1,1}, a_{1,2}, \dots, a_{1,p}, a_{2,1}, \dots, a_{n,p}) + \mu(b_{1,1}, b_{1,2}, \dots, b_{1,p}, b_{2,1}, \dots, b_{n,p}) \\ &= \lambda f(A) + \mu f(B) \end{aligned}$$

On en déduit par récurrence que f conserve toutes les combinaisons linéaires.

$\mathcal{M}_{n,p}(\mathbb{R})$ est donc bien un espace vectoriel réel de dimension np . □

3.2 Sous-espaces vectoriels

3.2.0.1 Généralités

Une notion indispensable associée aux espaces vectoriels est la notion de sous-espaces vectoriels :

Définition 3.2.1. Soit E un espace vectoriel réel de dimension finie. Soit F un ensemble. On a que F est un sous-espace vectoriel de E lorsque F vérifie :

1. $F \subset E$

2. $F \neq \emptyset$
3. $\forall (x, y) \in F^2, \forall \lambda \in \mathbb{R}, \lambda x + y \in F$.

Remarques 3.2.2. 1. Il est possible de remplacer $F \neq \emptyset$ par $0_E \in F$ car un sous-espace vectoriel contient obligatoirement le neutre de E . De plus montrer que le neutre est dans F est en général très simple.

2. Il est possible de remplacer l'axiome

$$\forall (x, y) \in F, \forall \lambda \in \mathbb{R}, \lambda x + y \in F$$

par

$$\forall (x, y) \in F^2, x + y \in F \quad \text{et} \quad \forall x \in F, \forall \lambda \in \mathbb{R}, \lambda x \in F.$$

3. Cette définition concorde bien avec la caractérisation des sous-espaces vectoriels de \mathbb{R}^n donnée en première année.
4. $\{0_E\}$ et E sont toujours des sous-espaces vectoriels de E .

Exemples 3.2.3. 1. Montrons que $\mathbb{R}_1[x]$ est un sous espace vectoriel de $\mathbb{R}_2[x]$.

- Un polynôme de degré au plus 1 est automatiquement de degré au plus 2 donc on a bien $\mathbb{R}_1[x] \subset \mathbb{R}_2[x]$;
- Le polynôme nul est bien un élément de $\mathbb{R}_1[x]$ donc $\mathbb{R}_1[x] \neq \emptyset$;
- Soit $(P, Q) \in \mathbb{R}_1[x]^2$ et soit $\lambda \in \mathbb{R}$ et montrons que $\lambda P(x) + Q(x) \in \mathbb{R}_1[x]$. On considère $(a_0, a_1, b_0, b_1) \in \mathbb{R}^4$ tel que $P(x) = a_0 + a_1x$ et $Q(x) = b_0 + b_1x$. On a donc

$$\lambda P(x) + Q(x) = \lambda(a_0 + a_1x) + b_0 + b_1x = (\lambda a_0 + b_0) + (\lambda a_1 + b_1)x \in \mathbb{R}_1[x].$$

$\mathbb{R}_1[x]$ est donc bien un sous-espace vectoriel de $\mathbb{R}_2[x]$.

2. On note \mathcal{D} l'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{R})$. Montrons qu'il s'agit d'un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$.
 - \mathcal{D} est composé de matrices de $\mathcal{M}_n(\mathbb{R})$ donc on a $\mathcal{D} \subset \mathcal{M}_n(\mathbb{R})$;
 - La matrice nulle 0_n est une matrice diagonale donc $0_n \in \mathcal{D}$ donc $\mathcal{D} \neq \emptyset$;
 - Soit $(A, B) \in \mathcal{D}^2$ et soit $\lambda \in \mathbb{R}$ et montrons que $\lambda A + B$ est diagonale. La matrice λA est toujours diagonale et la somme de deux matrices diagonales étant également diagonale, on a bien que $\lambda A + B$ est une matrice diagonale, soit que $\lambda A + B \in \mathcal{D}$.

\mathcal{D} est donc bien un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$.

Remarque 3.2.4. En réalité $\mathbb{R}_k[x]$ est un sous espace vectoriel de $\mathbb{R}_n[x]$ dès que $k \leq n$. Cela sera montré en TD.

Propriété 3.2.5. *Un sous-espace vectoriel réel F d'un espace vectoriel réel E de dimension finie est un espace vectoriel réel de dimension finie. On a de plus*

$$\dim(F) \leq \dim(E)$$

On a également $\dim(F) = \dim(E)$ si et seulement si $F = E$.

Méthode 3.2.6. Pour montrer qu'un ensemble est un espace vectoriel, on montre en général qu'il s'agit d'un sous-espace vectoriel d'un espace de référence.

Méthode 3.2.7. Pour montrer l'égalité entre deux espaces vectoriels de dimensions finies, on montre en général une seule inclusion et l'égalité des dimensions.

3.2.1 Sous-espaces vectoriels engendrés et notion de rang

Le principal type de sous-espaces vectoriels que nous étudierons seront les sous-espaces vectoriels engendrés. En effet nous essaierons de déterminer tout nos sous-espaces comme des sous-espaces vectoriels engendrés pour simplifier grandement leurs études.

Définition 3.2.8. Soit E un espace vectoriel réel, soit $p \in \mathbb{N}^*$ et soit $(e_1, \dots, e_p) \in E^p$. L'ensemble des combinaisons linéaires de e_1, \dots, e_p est appelé sous-espace vectoriel engendré par la famille $\{e_1, \dots, e_p\}$. On note ce dernier $Vect(e_1, \dots, e_p)$. On dit que la famille $\{e_1, \dots, e_p\}$ est une famille génératrice de $Vect(e_1, \dots, e_p)$.

Remarque 3.2.9. On a $Vect(e_1, \dots, e_p) = \left\{ \sum_{i=1}^p \lambda_i e_i \mid (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p \right\}$.

Exemples 3.2.10. • On considère l'ensemble des matrices symétriques de $\mathcal{M}_2(\mathbb{R})$.

$$\begin{aligned} \mathcal{S}_2(\mathbb{R}) &= \left\{ \begin{pmatrix} a & c \\ c & b \end{pmatrix} \mid (a, b, c) \in \mathbb{R}^3 \right\} \\ &= \left\{ a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + c \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mid (a, b, c) \in \mathbb{R}^3 \right\} \\ &= Vect \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \end{aligned}$$

- On considère F l'ensemble des polynômes de $\mathbb{R}_2[x]$ qui s'annule en 1. On a alors

$$\begin{aligned} F &= \{ax^2 + bx + c \mid a + b + c = 0\} \\ &= \{ax^2 + bx - a - b \mid (a, b) \in \mathbb{R}^2\} \\ &= \{a(x^2 - 1) + b(x - 1) \mid (a, b) \in \mathbb{R}^2\} \\ &= \text{Vect}(x^2 - 1, x - 1) \end{aligned}$$

Propriété 3.2.11. *Un sous-espace vectoriel engendré est un sous-espace vectoriel.*

Remarque 3.2.12. Cela nous fournit une autre manière de montrer qu'un ensemble est un sous-espace vectoriel, en arrivant à l'écrire comme un sous-espace vectoriel engendré.

Exemples 3.2.13. En considérant les deux exemples précédents, on a que les matrices symétriques de $\mathcal{M}_2(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{M}_2(\mathbb{R})$.

De même, les polynômes s'annulant en 1 de degré au plus 2 forme un sous-espace vectoriel de $\mathbb{R}_2[x]$.

Un résultat utile est également le suivant :

Propriété 3.2.14. *Soit E un espace vectoriel réel et soit F un sous-espace vectoriel de E . Soit $(e_1, \dots, e_p) \in E^p$ des vecteurs tels que pour tout $i \in \llbracket 1, p \rrbracket$, $e_i \in F$. On a alors*

$$\text{Vect}(e_1, \dots, e_p) \subset F.$$

On pourrait penser que la dimension d'un sous-espace vectoriel engendré est le cardinal de la famille qui l'engendre mais c'est faux!

Définition 3.2.15. Soit E un espace vectoriel réel, soit $p \in \mathbb{N}^*$ et soit $(e_1, \dots, e_p) \in E^p$. On appelle rang de la famille $\{e_1, \dots, e_p\}$ la dimension du sous-espace vectoriel $\text{Vect}(e_1, \dots, e_p)$ que l'on note $rg(e_1, \dots, e_p)$.

Remarques 3.2.16. • Le rang d'une famille de vecteurs ne peut donc jamais dépasser le cardinal de la famille, ni la dimension de l'espace vectoriel E .

- On rappelle que le rang d'une matrice est le rang de ces colonnes ou de ces lignes vu comme famille de vecteurs d'un espace \mathbb{R}^n . On prendra donc garde à ne pas confondre le rang d'une matrice et le rang d'une famille de matrices où les matrices sont des vecteurs d'un espace $\mathcal{M}_{n,p}(\mathbb{R})$. On reviendra plus en détails sur le rang d'une matrice dans le chapitre sur les applications linéaires.

Exemple 3.2.17. On pose $F = \text{Vect}(x^2, x^2 - x, x - 2, 2)$ un sous-espace vectoriel de $\mathbb{R}_2[x]$ or pour tout $ax^2 + bx + c \in \mathbb{R}_2[x]$, on a

$$ax^2 + bx + c = a.x^2 + 0(x^2 - x) + b(x - 2) + \left(\frac{1}{2}c + b\right) 2.$$

donc $ax^2 + bx + c \in F$. On a alors $\mathbb{R}_2[x] \subset F$ et comme $F \subset \mathbb{R}_2[x]$, on a $F = \mathbb{R}_2[x]$. F est donc un espace vectoriel de dimension 3 donc

$$\text{rg}(x^2, x^2 - x, x - 2, 2) = 3.$$

Pour travailler sur les sous-espaces vectoriels engendrés ou calculer un rang les règles de calculs suivantes sont fortes utiles :

Propriétés 3.2.18. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. On a alors :

1. Soit $x \in E$. Si x est une combinaison linéaire de e_1, \dots, e_p , on a alors

$$\text{Vect}(e_1, \dots, e_p, x) = \text{Vect}(e_1, \dots, e_p);$$

2. Pour tout $(i, j) \in \llbracket 1, p \rrbracket^2$. On a

$$\text{Vect}(e_1, \dots, e_i, \dots, e_p) = \text{Vect}(e_1, \dots, e_i + e_j, \dots, e_p);$$

3. Soit $\lambda_1, \dots, \lambda_p$ p réels tous non nuls, on a alors

$$\text{Vect}(e_1, \dots, e_p) = \text{Vect}(\lambda_1 e_1, \dots, \lambda_p e_p).$$

Remarque 3.2.19. La première propriété implique que $\text{Vect}(e_1, \dots, e_p, 0_E) = \text{Vect}(e_1, \dots, e_p)$.

Exemple 3.2.20. On considère l'exemple précédent et on a

$$\begin{aligned} \text{Vect}(x^2, x^2 - x, x - 2, 2) &= \text{Vect}(x^2, x^2 - x, x - 2 + 2, 2) = \text{Vect}(x^2, x^2 - x, x, 2) \\ &= \text{Vect}\left(x^2, x^2 - x, x, \frac{1}{2}2\right) = \text{Vect}(x^2, x^2 - x, x, 1) \\ &= \text{Vect}(x^2, x^2 - x + x, x, 1) = \text{Vect}(x^2, x^2, x, 1) \\ &= \text{Vect}(x^2, x^2 - x^2, x, 1) = \text{Vect}(x^2, 0, x, 1) \\ &= \text{Vect}(x^2, x, 1) = \mathbb{R}_2[x] \end{aligned}$$

Ces opérations sur les sous-espaces vectoriels engendrés nous permettent également d'en déduire les opérations suivantes sur les rang

Propriétés 3.2.21. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. On a alors :

1. Soit $x \in E$. Si x est une combinaison linéaire de e_1, \dots, e_p , on a alors

$$rg(e_1, \dots, e_p, x) = rg(e_1, \dots, e_p);$$

2. Pour tout $(i, j) \in \llbracket 1, p \rrbracket^2$. On a

$$rg(e_1, \dots, e_i, \dots, e_p) = rg(e_1, \dots, e_i + e_j, \dots, e_p);$$

3. Soit $\lambda_1, \dots, \lambda_p$ p réels tous non nuls, on a alors

$$rg(e_1, \dots, e_p) = rg(\lambda_1 e_1, \dots, \lambda_p e_p).$$

Remarque 3.2.22. De même, la première propriété implique que $rg(e_1, \dots, e_p, 0_E) = rg(e_1, \dots, e_p)$.

3.3 Familles libres, génératrices et bases

Pour une famille de vecteurs, on recherche en général le fait qu'elle ait une des deux propriétés suivantes : être génératrice ou être libre.

Le cas où une famille de vecteurs a les deux caractéristiques, on est dans le cas d'une base qui joue un rôle central en algèbre. La détermination de ces dernières va être un des objectifs principaux de beaucoup d'exercices d'algèbres.

3.3.1 Familles génératrices

Définition 3.3.1. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. On dit que la famille $\{e_1, \dots, e_p\}$ est génératrice de E lorsque que tout vecteur de E peut s'écrire comme une combinaison linéaire de e_1, \dots, e_p .

Autrement dit la famille $\{e_1, \dots, e_p\}$ est génératrice de E lorsque

$$E = Vect(e_1, \dots, e_p).$$

Remarque 3.3.2. Une famille de vecteurs est générateur d'un espace vectoriel, on doit toujours le préciser sous peine que cela manque de sens !

Remarque 3.3.3. Par définition le rang d'une famille génératrice $\{e_1, \dots, e_p\}$ d'un espace vectoriel réel E de dimension n est toujours n .

Méthode 3.3.4. Pour montrer qu'une famille de vecteurs $\{e_1, \dots, e_p\}$ est génératrice d'un espace vectoriel E , on considère un vecteur x quelconque de E et on doit déterminer des réels $(\lambda_1, \dots, \lambda_p)$ tel que

$$x = \lambda_1 e_1 + \dots + \lambda_p e_p.$$

Cela revient à déterminer au moins une solution d'un système.

Exemples 3.3.5. 1. Dans un exemple précédent, on a vu que pour tout $ax^2 + bx + c \in \mathbb{R}_2[x]$, on a

$$ax^2 + bx + c = a.x^2 + 0(x^2 - x) + b(x - 2) + \left(\frac{1}{2}c + b\right) 2.$$

Cette possibilité n'est pas unique mais suffit à justifier que la famille $\{x^2, x^2 - x, x - 2, 2\}$ est génératrice de $\mathbb{R}_2[x]$.

2. On considère la famille $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ de $\mathcal{M}_2(\mathbb{R})$.

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et résolvons l'équation vectoriel d'inconnues réelles

$\lambda_1, \lambda_2, \lambda_3$ et λ_4 :

$$A = \lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Cette équation est équivalente au système suivant :

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = a \\ \lambda_2 = b \\ \lambda_3 = c \\ \lambda_4 = d \end{cases} \Leftrightarrow \begin{cases} b + c + d + \lambda_1 = a \\ \lambda_2 = b \\ \lambda_3 = c \\ \lambda_4 = d \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = a - b - c - d \\ \lambda_2 = b \\ \lambda_3 = c \\ \lambda_4 = d \end{cases}$$

Ce système admet donc au moins une solution (dans ce cas une unique solution) donc la famille $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ est génératrice dans $\mathcal{M}_2(\mathbb{R})$.

Le caractère générateur est fortement liée aux sous-espaces vectoriels engendrée, on en déduit donc les propriétés similaires :

Propriété 3.3.6. Soit $\{e_1, \dots, e_p\}$ une famille génératrice d'un espace vectoriel E . On a alors toujours une famille génératrice de E lorsque l'on :

- Change l'ordre des vecteurs ;
- Ajoute un nouveau vecteur à la famille ;
- Multiplie un vecteur par un scalaire non nul ;
- Additionner un vecteur à un autre ;
- Retire un vecteur qui est combinaison linéaire des autres.

Remarque 3.3.7. En particulier, si la famille génératrice contient le vecteur nul, il est automatiquement combinaison linéaire des autres et l'on peut donc le retirer.

Exemple 3.3.8. La famille $\{(1, 2, 3), (3, -2, 1), (1, 0, 1)\}$ est génératrice de $F = Vect((1, 2, 3), (3, -2, 1), (1, 0, 1))$ or comme $(1, 2, 3) + (3, -2, 1) = (4, 0, 4)$ donc $\{(1, 2, 3), (4, 0, 4), (1, 0, 1)\}$ est génératrice de F . Comme $(4, 0, 4) = 4(1, 0, 1)$ on a donc

$$\{(1, 2, 3), (1, 0, 1)\}$$

est une famille génératrice de F .

Le dernier résultat qui nous intéresse est le lien entre le cardinal d'une telle famille et la dimension de l'espace :

Propriété 3.3.9. Soit E un espace vectoriel de dimension n et soit $\{e_1, \dots, e_p\}$ une famille génératrice de E . On a donc

$$n \leq p.$$

Remarque 3.3.10. On peut utiliser ce résultat par sa contraposée afin d'établir qu'une famille n'est pas génératrice d'un espace vectoriel donnée.

Exemple 3.3.11. La famille $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$ n'est pas génératrice dans $\mathcal{M}_2(\mathbb{R})$ car elle ne contient que trois vecteurs et que $\mathcal{M}_2(\mathbb{R})$ est de dimension 4.

3.3.2 Familles libres

Contrairement aux familles génératrices qui justifient de l'existence d'au moins une écriture de chaque vecteur comme combinaison linéaire, on s'intéresse à présent aux familles libres qui nous donne l'existence d'au plus une seule écriture comme combinaison linéaire.

Définition 3.3.12. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. La famille $\{e_1, \dots, e_p\}$ est une famille libre de vecteurs si aucun vecteur de ne peut s'écrire comme une combinaison linéaire des autres.

Dans le cas contraire on dit que la famille est liée.

Remarques 3.3.13. • Contrairement aux familles génératrices, ces caractéristiques est totalement dépendantes des vecteurs eux mêmes.

- Comme le vecteur nul est toujours combinaison linéaire de n'importe quelle vecteur, pour qu'une famille soit libre tout les vecteurs doivent être non nuls !
- Dans le cas d'une famille à deux éléments, au lieu de parlé d'une famille liée, on parlera de deux vecteurs colinéaires.

En pratique on caractérise les familles libres par le résultat suivant :

Théorème 3.3.14. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. La famille $\{e_1, \dots, e_p\}$ est une famille libre si et seulement si l'égalité vectoriel $\lambda_1 e_1 + \dots + \lambda_p e_p = 0$ implique que les réels $\lambda_1, \dots, \lambda_p$ soient tous nuls.

Ce théorème nous indique la méthode pour prouver qu'une famille est libre :

Méthode 3.3.15. Soit E un espace vectoriel réel et soit $(e_1, \dots, e_p) \in E^p$. Pour montrer que la famille $\{e_1, \dots, e_p\}$ est libre, on cherche donc à résoudre l'équation $\lambda_1 e_1 + \dots + \lambda_p e_p = 0$ qui nous amène à un système. On doit trouver que le système admet une **UNIQUE** solution pour obtenir la liberté qui sera $\lambda_1 = \lambda_2 = \dots = \lambda_p = 0$.

Exemples 3.3.16. 1. La famille $\{x + 1, x + 2\}$ est libre dans $\mathbb{R}_1[x]$. En effet pour $(\lambda, \mu) \in \mathbb{R}^2$ tel que $\lambda(x + 1) + \mu(x + 2) = 0$, on a

$$(\lambda + \mu)x + \lambda + 2\mu = 0 \Leftrightarrow \begin{cases} \lambda + \mu = 0 \\ \lambda + 2\mu = 0 \end{cases} \Leftrightarrow \lambda = \mu = 0,$$

par unicité des coefficients d'un polynômes.

2. La famille $\{(1, 2, 3), (1, 1, -1), (1, 4, 11)\}$ n'est pas libre. En effet pour $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ tel que $\lambda_1(1, 2, 3) + \lambda_2(1, 4, 11) + \lambda_3(1, 4, 11) = (0, 0, 0)$, on a que $(\lambda_1, \lambda_2, \lambda_3)$ est solution du système suivant :

$$\begin{aligned} & \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + 4\lambda_3 = 0 \\ 3\lambda_1 - \lambda_2 + 11\lambda_3 = 0 \end{cases} \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ -\lambda_2 + 2\lambda_3 = 0 \\ 3\lambda_1 - \lambda_2 + 11\lambda_3 = 0 \end{cases} \\ & \xrightarrow{L_3 \leftarrow L_3 - 3L_1} \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ -\lambda_2 + 2\lambda_3 = 0 \\ -4\lambda_2 + 8\lambda_3 = 0 \end{cases} \\ & \xrightarrow{L_3 \leftarrow L_3 - 4L_2} \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ -\lambda_2 + 2\lambda_3 = 0 \\ 0 = 0 \end{cases} \\ & \Leftrightarrow \begin{cases} \lambda_1 + 3\lambda_3 = 0 \\ \lambda_2 = 2\lambda_3 \end{cases} \\ & \Leftrightarrow \begin{cases} \lambda_1 = -3\lambda_3 \\ \lambda_2 = 2\lambda_3 \end{cases} \end{aligned}$$

Ce système admettant une infinité de solution, la famille n'est pas libre. De plus pour $\lambda_3 = 1$, on obtient $\lambda_1 = -3$ et $\lambda_2 = 2$ ce qui nous donne

$$-3(1, 2, 3) + 2(1, 1, -1) + (1, 4, 11) = (0, 0, 0)$$

soit

$$(1, 4, 11) = 3(1, 2, 3) - 2(1, 1, -1).$$

La série de remarque suivante est à connaître parfaitement :

- Remarques 3.3.17.**
- Une famille constitué d'un seul vecteur est libre si et seulement si ce vecteur est non nul ;
 - Une famille constitué de deux vecteurs est libre si et seulement si les deux vecteurs ne sont pas colinéaire ;
 - Si le vecteur nul est dans une famille alors cette famille est liée ;
 - Si une famille contient deux fois le même vecteur alors elle est liée ;
 - Une famille est liée si et seulement si il existe un vecteur qui est combinaison linéaire des autres ;

Une famille libre se transforme moins en une autre famille libre qu'une famille génératrice en une nouvelle famille génératrice :

Propriété 3.3.18. Soit $\{e_1, \dots, e_p\}$ une famille libre d'un espace vectoriel E . On a alors toujours une famille libre de E lorsque l'on :

- Change l'ordre des vecteurs ;
- Multiplie un vecteur par un scalaire non nul ;
- Ajoute un vecteur qui n'est **PAS** combinaison linéaire des autres ;
- Une sous-famille extraite est une famille libre.

Dans l'espace vectoriel $\mathbb{R}_n[x]$, on a la possibilité de construire des familles libres simplement grâce à ce critère :

Propriété 3.3.19. Soit (P_1, \dots, P_p) une famille de polynômes non nuls de $\mathbb{R}_n[x]$. Si les polynômes sont de degrés deux à deux distincts alors la famille $\{P_1, \dots, P_p\}$ est libre. On parle alors d'une famille échelonnée de polynômes.

Exemple 3.3.20. La famille $\{x^i\}_{i \in \llbracket 0, n \rrbracket}$ est échelonnée dans $\mathbb{R}_n[x]$, elle forme donc une famille libre de $\mathbb{R}_n[x]$.

Le dernier résultat qui nous intéresse est le lien entre le cardinal d'une telle famille et la dimension de l'espace :

Propriété 3.3.21. Soit E un espace vectoriel de dimension n et soit $\{e_1, \dots, e_p\}$ une famille libre de E . On a donc

$$p \leq n.$$

Remarque 3.3.22. On peut utiliser ce résultat par sa contraposée afin d'établir qu'une famille n'est pas libre dans un espace vectoriel donnée.

Exemple 3.3.23. La famille

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

de \mathbb{R}^4 n'est pas libre car elle est composée de 6 vecteurs et que la dimension de \mathbb{R}^4 est 4 qui est strictement inférieur à 6.

3.3.3 Bases

3.3.3.1 Définitions

Définition 3.3.24. Soit E un espace vectoriel réel. On dit que famille $\{e_1, \dots, e_p\}$ est une base de E lorsque la famille est libre et génératrice de E .

Remarque 3.3.25. Si une famille $\{e_1, \dots, e_p\}$ est libre, alors il s'agit d'une base du sous-espace vectoriel $Vect(e_1, \dots, e_p)$.

Les inégalités précédentes sur le cardinal d'une famille libre ou génératrice en fonction de la dimension nous donnent le résultat suivant :

Propriété 3.3.26. Soit E un espace vectoriel de dimension n et soit $\{e_1, \dots, e_p\}$ une base de E . On a alors $p = n$.

Démonstration. Comme la famille est libre, on a $p \leq n$ et comme elle est génératrice de E , on a également $n \leq p$ donc $p = n$. \square

Remarque 3.3.27. Toutes les bases d'un même espace vectoriel ont donc le même cardinal !

Lemme 3.3.28. Soit E un espace vectoriel réel de dimension n et soit $\{e_1, \dots, e_n\}$ une base de E . Soit $x \in E$, il existe un unique n -uplet de réels $(\lambda_1, \dots, \lambda_n)$ tel que

$$x = \lambda_1 e_1 + \dots + \lambda_n e_n.$$

Les coefficients $\lambda_1, \dots, \lambda_n$ sont appelés les coordonnées de x dans la base $\{e_1, \dots, e_n\}$.

Remarque 3.3.29. Cette écriture unique provient de l'existence d'une combinaison linéaire par le caractère générateur de la base et l'unicité par le caractère libre.

Méthode 3.3.30. Pour trouver les coordonnées dans une base, on pose l'équation d'inconnue $(\lambda_1, \dots, \lambda_n)$ et on résout $x = \lambda_1 e_1 + \dots + \lambda_n e_n$.

Résoudre cette équation et obtenir une unique solution permet également de prouver que la famille $\{e_1, \dots, e_n\}$ est une base.

En pratique pour déterminer si une famille est une base lorsque que l'on a pas besoin des coordonnées dans cette base, on utilise le théorème suivant :

Théorème 3.3.31. Soit E un espace vectoriel de dimension n .

1. Soit \mathcal{G} une famille génératrice de E . Si $Card(\mathcal{G}) = n$, alors \mathcal{G} est une base de E .
2. Soit \mathcal{L} une famille libre de E . Si $Card(\mathcal{L}) = n$, alors \mathcal{L} est une base de E .

Exemples 3.3.32. 1. On a vu dans un exemple précédent que la famille

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

est génératrice dans $\mathcal{M}_2(\mathbb{R})$ or cette famille est de cardinal 4 et l'espace vec-

toriel $\mathcal{M}_2(\mathbb{R})$ est de dimension 4. La famille $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$

est donc une base de $\mathcal{M}_2(\mathbb{R})$.

2. La famille $\{1, x + 1, x^2 + x + 1, x^3 + x^2 + x + 1, x^4 + x^3 + x^2 + x + 1\}$ est libre en tant que famille de polynômes échelonnées et est de cardinal 5 dans l'espace vectoriel $\mathbb{R}_4[x]$ qui est de dimension 5 donc $\{1, x + 1, x^2 + x + 1, x^3 + x^2 + x + 1, x^4 + x^3 + x^2 + x + 1\}$ est une base de $\mathbb{R}_4[x]$.

3.3.3.2 Bases canoniques

On appelle bases canoniques les bases de références des espaces vectoriels de références. En général si l'on ne précise pas la base associée à un espace vectoriel, la base que l'on considère est la base canonique.

Définition 3.3.33. La base canonique de \mathbb{R}^n est la famille $\{e_1, \dots, e_n\}$ où pour tout $i \in \llbracket 1, n \rrbracket$, $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ où l'unique 1 est en i ème position.

Pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on a

$$x = \sum_{i=1}^n x_i e_i.$$

Définition 3.3.34. La base canonique de $\mathbb{R}_n[x]$ est la famille $\{1, x, \dots, x^n\}$.

Pour tout $P = a_0 + a_1x + \dots + a_nx^n \in \mathbb{R}_n[x]$, on a

$$P = \sum_{i=0}^n a_i x^i.$$

Remarque 3.3.35. Comme $\{1, x, \dots, x^n\}$ est une base, on retrouve bien le résultat bien connu que deux polynômes sont égaux si et seulement si tout leurs coefficients sont égaux.

Définition 3.3.36. La base canonique de $\mathcal{M}_{n,p}(\mathbb{R})$ est la famille $\{E_{i,j}\}_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, p \rrbracket}}$ où

pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, $E_{i,j}$ désigne la matrice où tous les coefficients sont nul sauf le coefficient à la i -ème ligne et en j -ème colonne qui vaut 1.

Pour tout $M = (a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, p \rrbracket}} \in \mathcal{M}_{n,p}(\mathbb{R})$, on a

$$M = \sum_{i=1}^n \sum_{j=1}^p a_{i,j} E_{i,j}.$$

Exemple 3.3.37. On se place dans l'espace vectoriel $\mathcal{M}_{2,3}(\mathbb{R})$. Il s'agit d'un espace de dimension 6 dont la base canonique $(E_{1,1}, E_{1,2}, E_{1,3}, E_{2,1}, E_{2,2}, E_{2,3})$ est

$$\left(\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right).$$

La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ s'écrit dans cette base

$$A = E_{1,1} + 2E_{1,2} + 3E_{1,3} + 4E_{2,1} + 5E_{2,2} + 6E_{2,3}$$

Chapitre 4

Statistiques bivariées

En première année les statistiques ont été vu comme un moyen d'extraire des informations globales d'un ensemble épars de données. En général les données ne dépendent pas que d'un seul paramètre et dans ce chapitre on s'intéressera à des ensembles de données concernant deux traits. On développera les outils permettant de déterminer si il y a un certain type de corrélation entre les deux traits et comment l'estimer.

4.1 Rappel sur les statistiques univariées

Rappelons les principales définitions des statistiques univariées vues en première années.

- Définitions 4.1.1.**
- On appelle population l'ensemble des objets sur lesquels notre étude statistique à lieu. On note en général Ω la population.
 - On appelle individu un élément de la population étudiée. On note en général ω un individu.
 - Lorsqu'on ne considère qu'une partie de la population, on considère alors un échantillon.
 - Une variable statistique est un caractère de la population que l'on étudie.
 - Si le caractère correspond à une valeur numérique, on parle de variable quantitative, sinon on parle de variable statistique qualitative.
 - L'ensemble des valeurs que peut prendre une variable statistique X sur la population Ω est en général noté $X(\Omega)$.

On ne traitera que les cas de variables statistiques discrètes (voir finis).

Définitions 4.1.2. Soit X une variable statistique finie de taille N associée à $(x_i, n_i)_{i \in \llbracket 1, n \rrbracket}$ (où n_i désigne l'effectif associée à la modalité x_i).

- On appelle moyenne et on note \bar{x} la quantité

$$\bar{x} := \frac{1}{N} \sum_{i=1}^n n_i x_i.$$

- On appelle variance et on note s_X^2 la quantité

$$s_X^2 := \frac{1}{N} \sum_{i=1}^n n_i (x_i - \bar{x})^2.$$

- On appelle écart type la quantité $s_X = \sqrt{s_X^2}$.

En pratique pour calculer la variance, on utilise la formule suivante :

Théorème 4.1.3 (Formule de Koenig-Huygens). *Soit X une variable statistique quantitative et soit Y la variable statistique définie par $Y = X^2$. On a alors*

$$s_X^2 = \bar{y} - \bar{x}^2.$$

4.2 Définitions

Définition 4.2.1. Soient $\{\omega_1, \omega_2, \dots, \omega_n\}$ un échantillon d'une population Ω et $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ deux séries statistiques représentant les valeurs prises par deux caractères X et Y sur cet échantillon où pour tout $i \in \llbracket 1, n \rrbracket$, $x_i = X(\omega_i)$ et $y_i = Y(\omega_i)$.

On appelle série statistique double la donnée de la liste $((x_1, y_1), \dots, (x_n, y_n))$.

Remarque 4.2.2. Chaque couple est associée à un seul individu de l'échantillon puisque que pour tout $i \in \llbracket 1, n \rrbracket$, $(x_i, y_i) = (X(\omega_i), Y(\omega_i))$.

Exemple 4.2.3. On considère l'espérance de vie à la naissance des femmes et des hommes entre 2000 et 2007 :

Année	2000	2001	2002	2003	2004	2005	2006	2007
Femmes	82.8	82.9	83	82.9	83.9	83.8	84.2	84.5
Hommes	75.3	75.5	75.8	75.9	76.8	76.8	77.2	77.6

Ici Ω correspond aux années, X est le caractère décrivant l'espérance de vie à la naissance des femmes et Y décrivant celles des hommes.

Définition 4.2.4. Lorsque que l'on cherche à savoir si le caractère X est une cause du caractère Y , on appelle X la variable explicative et Y la variable à expliquer.

Exemple 4.2.5. On va chercher si on peut donner un modèle permettant d'exprimer l'espérance de vie des hommes à la naissance en fonction de celui des femmes. X est donc bien la variable explicative et Y la variable à expliquer dans ce cas.

Une manière classique de représenter les données graphiquement est de s'aider d'un nuage de point :

Définitions 4.2.6. On appelle nuage de points associé à la série statistique double $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ l'ensemble des points du plan M_i de coordonnée (x_i, y_i) . On appelle le point (\bar{x}, \bar{y}) le point moyen de la série statistique double où \bar{x} et \bar{y} désigne respectivement les moyennes des séries statistiques univariées (x_1, \dots, x_n) et (y_1, \dots, y_n) .

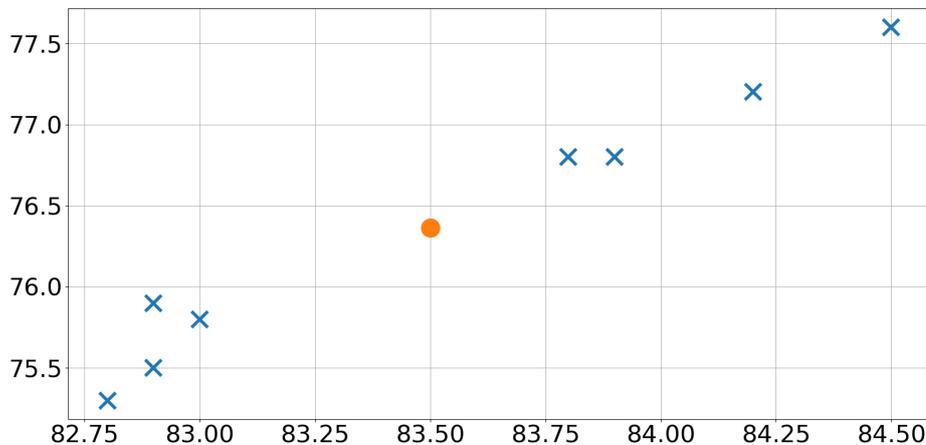
Exemple 4.2.7. On considère le même exemple que précédemment . On a donc

$$x = (82.8, 82.9, 83, 82.9, 83.9, 83.8, 84.2, 84.5)$$

et

$$y = (75.3, 75.5, 75.8, 75.9, 76.8, 76.8, 77.2, 77.6).$$

Par un outil de calcul on a $\bar{x} = 83.5$ et $\bar{y} = 76.3625$. On en déduit le nuage de point suivant où le point moyen est représenté par un cercle orange :



4.3 Covariance et corrélation

Pour étudier la répartition des valeurs d'une série statistique la variance est un outil puissant et de manière similaire on peut définir la notion de covariance pour étudier deux caractères conjointement.

Définition 4.3.1. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double et soit \bar{x} et \bar{y} les moyennes des séries statistiques univariées $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. On appelle covariance empirique

$$s_{x,y} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

On a la formule suivante qui permet de simplifier le calcul :

Théorème 4.3.2 (Formule de Koenig-Huygens). Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double et soit \bar{x} et \bar{y} les moyennes des séries statistiques univariées $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. On a alors

$$s_{x,y} = \left(\frac{1}{n} \sum_{i=1}^n x_i y_i \right) - \bar{x} \bar{y}.$$

Démonstration. On rappelle que $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ et $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. On a :

$$\begin{aligned} s_{x,y} &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x} y_i - x_i \bar{y} + \bar{x} \bar{y}) \\ &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \frac{\bar{x}}{n} \sum_{i=1}^n y_i - \frac{\bar{y}}{n} \sum_{i=1}^n x_i + \frac{\bar{x} \bar{y}}{n} \sum_{i=1}^n 1 \\ &= \left(\frac{1}{n} \sum_{i=1}^n x_i y_i \right) - \bar{x} \bar{y} - \bar{x} \bar{y} + \frac{n \bar{x} \bar{y}}{n} \\ &= \left(\frac{1}{n} \sum_{i=1}^n x_i y_i \right) - \bar{x} \bar{y} \end{aligned}$$

□

Exemple 4.3.3. En considérant toujours le même ensemble de données, on trouve que $s_{x,y} \simeq 0.494$.

Définition 4.3.4. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double et soit s_x et s_y les écarts types des séries statistiques univariées $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. Si s_x et s_y sont tout deux non nuls, on appelle coefficient de corrélation linéaire de la série statistique double le réel

$$r_{x,y} = \frac{s_{x,y}}{s_x s_y}.$$

Exemple 4.3.5. On a, par outils de calculs, que $s_x \simeq 0.632$ et $s_y \simeq 0.792$. On trouve donc que $r_{x,y} \simeq 0.986$.

Propriété 4.3.6. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double et soit s_x et s_y les écarts types et des séries statistiques univariées $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$. Si s_x et s_y sont tout deux non nuls, on a alors

- $|s_{x,y}| \leq s_x s_y$
- $|r_{x,y}| \leq 1$

Démonstration. On pose $f : \mathbb{R} \rightarrow \mathbb{R}$ et on va étudier cette fonction. On a déjà pour tout $\lambda \in \mathbb{R}$ que

$$f(\lambda) = \frac{1}{n} \sum_{i=1}^n (x_i + \lambda y_i - \overline{x + \lambda y})(x_i + \lambda y_i - \overline{x + \lambda y}) = \frac{1}{n} \sum_{i=1}^n (x_i + \lambda y_i - \overline{x + \lambda y})^2,$$

donc $f(\lambda) \geq 0$ en tant que sommes de carrés. De même on observe que

$$\begin{aligned} f(\lambda) &= \frac{1}{n} \sum_{i=1}^n (x_i + \lambda y_i - \overline{x + \lambda y})^2 \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x} + \lambda(y_i - \bar{y}))^2 \\ &= \frac{1}{n} \sum_{i=1}^n [(x_i - \bar{x})^2 + 2\lambda(x_i - \bar{x})(y_i - \bar{y}) + \lambda^2(y_i - \bar{y})^2] \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 + 2\lambda \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + \frac{\lambda^2}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \\ &= s_x^2 + 2\lambda s_{x,y} + \lambda^2 s_y^2 \end{aligned}$$

Distinguons deux cas :

- Si $s_y = 0$, alors f est une fonction affine et est positif. Il s'agit donc d'une fonction constante donc $s_{x,y} = 0$. On a donc bien $|s_{x,y}| \leq s_x s_y$ dans ce cas.
- Si $s_y \neq 0$, alors f est un trinôme et est positif. Il s'agit donc d'un trinôme de discriminant négatif ou nul. On a donc $(2s_{x,y})^2 - 4s_x^2 s_y^2 \leq 0$ soit $4s_{x,y}^2 \leq 4s_x^2 s_y^2$ on en déduit donc

$$|s_{x,y}| \leq s_x s_y.$$

On a donc bien dans tout les cas $|s_{x,y}| \leq s_x s_y$.

On a également

$$|r_{x,y}| = \left| \frac{s_{x,y}}{s_x s_y} \right| \leq \left| \frac{s_x s_y}{s_x s_y} \right| = 1.$$

□

Exemple 4.3.7. Dans notre ensemble de données, on a $s_{x,y} \simeq 0.494$ et $s_x s_y \simeq 0.501$. On a donc bien $|s_{x,y}| \leq s_x s_y$. De même, on a bien $|r_{x,y}| \simeq 0.986 \leq 1$.

Ce coefficient de régression linéaire nous permet de quantifier la position de points du nuage de points d'une série statistique double par rapport à une droite particulière que l'on va définir dans la prochaine partie.

4.4 Régression et méthode des moindres carrés

Pour un couple de variable (X, Y) où X est explicable et où Y est à expliquer l'idéal serait d'avoir une relation entre les deux sauf qu'en général cela n'est pas le cas. L'idée de cette fin de chapitre est de déterminer un moyen d'approximer un tel cas.

Définitions 4.4.1. Soit (X, Y) un couple de variable où X est explicable et où Y est à expliquer. Si il existe une fonction f tel que $Y = f(X)$ on parle alors d'un modèle de régression.

En pratique on est en général dans le cas où $Y = f(X) + \epsilon$ où f est appelé fonction de régression et ϵ est une variable aléatoire appelé erreur d'ajustement.

Le modèle qui va nous intéresser est le modèle linéaire :

Définition 4.4.2. Un modèle de régression linéaire est un modèle où f est du type $X \mapsto aX + b$.

On va à présent se demander comment construire un modèle linéaire qui représentera au mieux nos données d'une série statistique double. Autrement dit quelle

est la droite du plan qui passe le plus proche de tout les points du nuage de points.

Pour cela on va utiliser la méthode des moindres carrés :

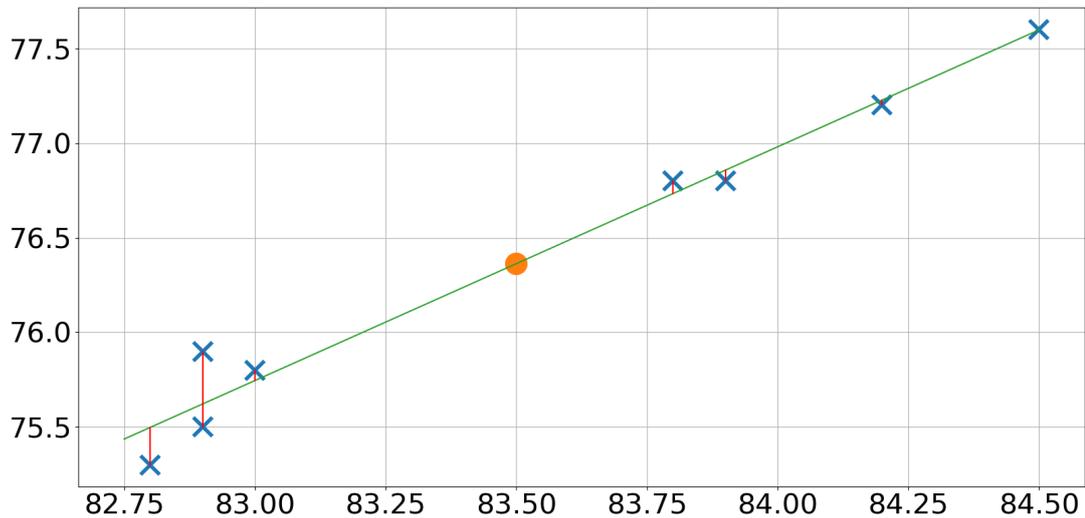
Méthode 4.4.3. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double. La méthode des moindres carrés consiste, quand le nuage de point à une forme allongé dans une direction, à déterminer une droite d'équation $y = ax + b$ passant au plus près de tout les points.

Pour cela on cherche la droite qui minimise la distance entre le point de données (x_i, y_i) et la courbe en x_i soit le point $(x_i, ax_i + b)$. On cherche donc à minimiser pour tout les points la quantité $(x_i - x_i)^2 + (y_i - (ax_i + b))^2 = (y_i - ax_i - b)^2$. On cherche donc à déterminer deux réels a et b tels que

$$\sum_{i=1}^n (y_i - ax_i - b)^2$$

soit minimale. Cette quantité s'appelle l'erreur quadratique.

Exemple 4.4.4. Sur la droite ci dessous, on recherche à minimiser la somme des longueurs des segments rouges mises au carré :



La droite recherché existe et est unique par le théorème suivant :

Théorème 4.4.5. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double. Il existe une unique droite minimisant l'erreur quadratique. Il s'agit de la droite portée par l'équation

$$y = \frac{s_{x,y}}{s_x^2}(x - \bar{x}) + \bar{y}.$$

On appelle cette droite, droite de régression linéaire de Y en X .

Remarque 4.4.6. Cela se prouve par des arguments d'analyse sur \mathbb{R}^2 , nous avons pour cela besoin d'un chapitre du deuxième semestre.

Propriété 4.4.7. La droite de régression linéaire passe par le point moyen.

Démonstration. Il suffit de vérifier que $\frac{s_{x,y}}{s_x^2}(\bar{x} - \bar{x}) + \bar{y} = \bar{y}$ ce qui est trivial. \square

Exemple 4.4.8. Dans le graphique précédent, la droite tracée est la droite de régression linéaire qui passe bien par le point moyen !

On observe que le coefficient de régression est relié au coefficient directeur de cette droite par $\frac{s_{x,y}}{s_x^2} = r_{x,y} \frac{s_y}{s_x}$ ce qui nous permet d'obtenir :

Propriété 4.4.9. Soit $((x_i, y_i))_{i \in \llbracket 1, n \rrbracket}$ une série statistique double dont le coefficient de régression linéaire $r_{x,y}$ existe. On a alors :

- $|r_{x,y}|$ vaut 1 si et seulement si les points du nuage de points sont alignés.
- Si $r_{x,y} > 0$ (resp $r_{x,y} < 0$) alors la droite de régression linéaire a une pente positive (resp négative).

Démonstration. On a $|r_{x,y}| = 1$ si et seulement si $|s_{x,y}| = s_x s_y$ ce qui par une preuve précédente correspond au cas où le trinôme $s_x^2 + 2\lambda s_{x,y} + \lambda^2 s_y^2$ est de discriminant nul, soit qu'il admet une unique racine λ_0 . On a donc $s_{x+\lambda_0 y} = 0$ or un écart type est nul si et seulement si $x + \lambda_0 y$ est une variable constante, il existe $c \in \mathbb{R}$ tel que $x + \lambda_0 y = c$ ce qui correspond à une relation linéaire entre x et y . On a donc bien que les points sont alignés si et seulement si $|r_{x,y}| = 1$.

On remarque que le coefficient directeur de la droite de régression linéaire est $\frac{s_{x,y}}{s_x^2} = r_{x,y} \frac{s_y}{s_x}$ or les écarts types s_x et s_y sont des constantes positives, on a que $\frac{s_{x,y}}{s_x^2}$ et $r_{x,y}$ sont de même signe, soit que $s_{x,y}$ et $r_{x,y}$ sont de même signe. \square

Remarque 4.4.10. De manière générale, plus $|r_{x,y}|$ est proche de 1 plus la droite de régression linéaire fournit un modèle pertinent.

Remarque 4.4.11. Toutes les études que l'on a faites dans ce chapitre ne nous amènent pas à une obligation de causalité entre les caractères étudiés. C'est un détournement classique de ces outils que de penser qu'un fort taux de corrélation implique une forte causalité comme le montre l'exemple suivant :

Exemple 4.4.12. On considère les ensembles de données suivants :

x	14.04	14.31	14.41	14.35	14.56	14.63	14.75	15.44	15.81
y	127.33	127.12	126.98	126.96	126.75	126.49	126.22	125.85	125.51

En menant les calculs on trouve que $s_{x,y} \simeq -0.973$ ce qui est très proche de 1 en valeur absolue donc on pourrait penser que les deux ensembles de données sont liés or il s'agit pour la première du prix moyen d'un kilo de fromage AOC en France entre 2013 et 2021 et de la population du Japon en millions sur les mêmes années.

Chapitre 5

Comparaisons de suites réelles

L'étude de suite est souvent complexe une fois les suites de références exclues. Nous nous proposons dans ce chapitre d'établir deux outils qui nous permettront de comparer deux suites entre elles et qui auront de nombreuses applications comme par exemple une manière d'étudier facilement la convergence ou la divergence d'une série numérique que nous verrons au prochain chapitre.

5.1 Relation de négligeabilité

5.1.1 Définition et caractérisation

Définition 5.1.1. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$. On dit que la suite $(u_n)_{n \in \mathbb{N}}$ est négligeable devant $(v_n)_{n \in \mathbb{N}}$ si il existe un entier n_0 et une suite $(\epsilon_n)_{n \in \mathbb{N}}$ convergente et de limite nulle tels que pour tout $n \geq n_0$,

$$u_n = \epsilon_n v_n.$$

On note $u_n = o_{n \rightarrow +\infty}(v_n)$ ou $u_n = o(v_n)$ si il n'y pas d'ambiguïté.

Remarques 5.1.2. On appelle "petit o " la notation $o_{n \rightarrow +\infty}(v_n)$ (inclus dans l'ensemble des notations de Landau). Il faut toujours avoir l'esprit en traitant ces objets qu'il s'agit d'une très petite quantité comparée à la suite v_n et dont on ne peut pas donner une valeur explicite en générale.

En pratique, on remarque que la suite $(\epsilon_n)_{n \in \mathbb{N}}$ correspond au quotient des deux suites. On utilise donc la caractérisation suivante pour établir une relation de négligeabilité :

Théorème 5.1.3. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles tel qu'il existe un entier n_0 où pour tout $n \geq n_0$, $v_n \neq 0$. On a alors que $u_n = o_{n \rightarrow +\infty}(v_n)$ si et seulement si

$$\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0.$$

- Exemples 5.1.4.**
1. On a $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ car pour tout $n \in \mathbb{N}^*$, $\frac{1}{n^3} = \frac{n}{n^3} = \frac{1}{n^2}$ donc comme $\lim_{n \rightarrow +\infty} \frac{1}{n^2} = 0$ on a bien $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$;
 2. On a $\frac{1}{n^2+n+1} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ car pour tout $n \in \mathbb{N}^*$, $\frac{1}{n^2+n+1} = \frac{n}{n^2+n+1} = \frac{1}{n+1+\frac{1}{n}}$ donc comme $\lim_{n \rightarrow +\infty} \frac{1}{n+1+\frac{1}{n}} = 0$ on a bien $\frac{1}{n^2+n+1} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$;
 3. On a $7n^{42} = o_{n \rightarrow +\infty}(e^n)$ car par croissance comparée on a $\lim_{n \rightarrow +\infty} \frac{7n^{42}}{e^n} = 0$.
 4. La suite $(n^2 + n + 1)_{n \in \mathbb{N}}$ n'est pas négligeable devant la suite $(n^2)_{n \in \mathbb{N}}$ car $\lim_{n \rightarrow +\infty} \frac{n^2+n+1}{n^2} = \lim_{n \rightarrow +\infty} 1 + \frac{1}{n} + \frac{1}{n^2} = 1 \neq 0$.

Remarque 5.1.5. Les deux premiers exemples précédents fournissent un contre exemple d'une erreur classique à éviter. En effet si $u_n = o_{n \rightarrow +\infty}(v_n)$ et $w_n = o_{n \rightarrow +\infty}(v_n)$, en général on a $u_n \neq w_n$. Dans les exemples précédents on a vu $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ et $\frac{1}{n^2+n+1} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ or on a bien $\frac{1}{n^3} \neq \frac{1}{n^2+n+1}$! Cela provient du fait qu'un petit o est une quantité négligeable par rapport à une autre mais deux choses négligeables par rapport à un autre objet ne sont pas identiques en général.

Un exemple classique de petit o est le cas des suites convergent vers 0 :

Propriété 5.1.6. Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle qui converge vers 0. On a alors que $u_n = o_{n \rightarrow +\infty}(1)$.

Démonstration. Par hypothèse $\lim_{n \rightarrow +\infty} \frac{u_n}{1} = \lim_{n \rightarrow +\infty} u_n = 0$ donc par la caractérisation on a bien $u_n = o_{n \rightarrow +\infty}(1)$. □

Un corollaire direct est :

Corollaire 5.1.7. Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle convergent vers un réel l . On a alors que $u_n - l = o_{n \rightarrow +\infty}(1)$.

Remarque 5.1.8. Ces résultats sont en réalités des caractérisations.

5.1.2 Règles de calculs

5.1.2.1 Opérations sur les o

La nature même d'un petit o complique les calculs et nous imposent de déterminer précisément quelques règles de calculs :

Théorème 5.1.9. Soit $(u_n)_{n \in \mathbb{N}}$, $(v_n)_{n \in \mathbb{N}}$ et $(w_n)_{n \in \mathbb{N}}$ trois suites réelles et soit $(\lambda, \mu) \in \mathbb{R}^2$. On a alors :

1. (Transitivité) Si $u_n = o_{n \rightarrow +\infty}(v_n)$ et $v_n = o_{n \rightarrow +\infty}(w_n)$ alors $u_n = o_{n \rightarrow +\infty}(w_n)$;
2. (Combinaison linéaire) Si $u_n = o_{n \rightarrow +\infty}(w_n)$ et $v_n = o_{n \rightarrow +\infty}(w_n)$ on a alors $\lambda u_n + \mu v_n = o_{n \rightarrow +\infty}(w_n)$;
3. (Multiplication par un réel non nul) Si $\lambda \neq 0$ et $u_n = o_{n \rightarrow +\infty}(v_n)$ alors $u_n = o_{n \rightarrow +\infty}(\lambda v_n)$;
4. (Produit) Si $u_n = o_{n \rightarrow +\infty}(v_n)$ alors $u_n w_n = o_{n \rightarrow +\infty}(v_n w_n)$.

Remarques 5.1.10. 1. On fera en sorte de toujours avoir qu'un seul o dans nos expressions et les deux premières propriétés imposent donc dans un premier temps de "mettre tout les petits o au même niveau" puis de les combiner entre eux.

2. La multiplication par un réel non nul permet usuellement de ne considérer que des suites "normalisées" dans notre petit o . Par exemple on peut remplacer un $o_{n \rightarrow +\infty}(2)$ par un $o_{n \rightarrow +\infty}(1)$ ou encore un $o_{n \rightarrow +\infty}\left(\frac{4n}{\sqrt{7}}\right)$ par $o_{n \rightarrow +\infty}(n)$.
3. Le produit sert usuellement à modifier un petit o pour le mettre au même niveau qu'un autre.

Exemple 5.1.11. On peut vérifier que $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right)$ et $\frac{1}{n^2} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$ donc on a aussi $\frac{1}{n^3} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$.

On peut de plus en déduire que $\frac{2}{n^3} - \frac{7}{n^2} = o_{n \rightarrow +\infty}\left(\frac{1}{n}\right)$.

En multipliant par n la dernière expression on a $\frac{2}{n^2} - \frac{7}{n} = o_{n \rightarrow +\infty}(1)$. C'est une autre manière de vérifier que la suite $\left(\frac{2}{n^2} - \frac{7}{n}\right)_{n \in \mathbb{N}^*}$ converge vers 0.

5.1.2.2 Croissances comparées

On peut réécrire le théorème de croissances comparées de première année sous le prisme des relations de négligeabilités :

Théorème 5.1.12. Soit $q \in]1, +\infty[$ et soit $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$. On a alors

- $n^\alpha = o_{n \rightarrow +\infty}(q^n)$;
- $\ln(n)^\beta = o_{n \rightarrow +\infty}(q^n)$;
- $\ln(n)^\beta = o_{n \rightarrow +\infty}(n^\alpha)$.

Exemples 5.1.13. 1. $n^7 = o_{n \rightarrow +\infty}(e^n)$;

2. $\ln(n)^4 = o_{n \rightarrow +\infty}(5^n)$;

3. $\ln(n) = o_{n \rightarrow +\infty}(\sqrt{n})$.

5.2 Relation d'équivalence

5.2.1 Définition et caractérisation

Définition 5.2.1. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$. On dit que les suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont équivalentes si il existe un entier n_0 et une suite $(\epsilon_n)_{n \in \mathbb{N}}$ convergente et de limite 1 tels que pour tout $n \geq n_0$,

$$u_n = \epsilon_n v_n.$$

On note $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ ou $u_n \sim v_n$ si il n'y pas d'ambiguïté.

Exemple 5.2.2. Montrons qu'un équivalent de $\frac{n-7}{n^2+n+1}$ est $\frac{1}{n}$. On observe que $\frac{n-7}{n^2+n+1} = \frac{n(n-7)}{n^2+n+1} \frac{1}{n}$ et on doit donc montrer que $\lim_{n \rightarrow +\infty} \frac{n(n-7)}{n^2+n+1} = 1$. On a

$$\begin{aligned} \frac{n(n-7)}{n^2+n+1} &= \frac{n^2-7n}{n^2+n+1} = \frac{n^2 \left(1 - \frac{7}{n}\right)}{n^2 \left(1 + \frac{1}{n} + \frac{1}{n^2}\right)} \\ &= \frac{1 - \frac{7}{n}}{1 + \frac{1}{n} + \frac{1}{n^2}} \xrightarrow{n \rightarrow +\infty} \frac{1-0}{1+0+0} = 1 \end{aligned}$$

On a donc bien

$$\frac{n-7}{n^2+n+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}.$$

Comme pour la relation de négligeabilité, cette définition est compliqué à utilisé en pratique et l'on utilisera en général la caractérisation suivante :

Théorème 5.2.3. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles. Elles sont équivalentes entre elles si et seulement si

$$u_n = v_n + o(v_n)$$

Si il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$ on a $v_n \neq 0$ alors $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont équivalentes entre elles si et seulement si

$$\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1.$$

Exemples 5.2.4. 1. Dans l'exemple précédent nous avons employé cette méthodologie.

2. Montrons que $(e^n - 7n)(n^2 + \sqrt{n})$ est équivalent à $n^2 e^n$. On a que pour tout $n \in \mathbb{N}^*$, $n^2 e^n$ est non nul donc pour tout $n \in \mathbb{N}^*$:

$$\frac{(e^n - 7n)(n^2 + \sqrt{n})}{n^2 e^n} = \frac{e^n - 7n}{e^n} \frac{n^2 + \sqrt{n}}{n^2} = (1 - 7ne^{-n}) \left(1 + \frac{1}{n^{\frac{3}{2}}}\right)$$

Par croissance comparée dans le premier terme du produit, on a $\lim_{n \rightarrow +\infty} \frac{(e^n - 7n)(n^2 + \sqrt{n})}{n^2 e^n} = (1 - 0)(1 + 0) = 1$ donc on a bien

$$(e^n - 7n)(n^2 + \sqrt{n}) \underset{n \rightarrow +\infty}{\sim} n^2 e^n.$$

5.2.2 Règles de calculs

Les règles de calculs avec les équivalents sont légèrement plus souples que pour les relation de négligeabilité mais il faut tout de même redoubler d'attention car de nombreuses choses ne sont pas possible avec ces dernières.

Théorème 5.2.5. Soit $(u_n)_{n \in \mathbb{N}}$, $(v_n)_{n \in \mathbb{N}}$, $(w_n)_{n \in \mathbb{N}}$ et $(z_n)_{n \in \mathbb{N}}$ quatre suites réelles. On a alors :

- (Symétrie) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ alors $v_n \underset{n \rightarrow +\infty}{\sim} u_n$;
- (Transitivité) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et $v_n \underset{n \rightarrow +\infty}{\sim} w_n$ alors $u_n \underset{n \rightarrow +\infty}{\sim} w_n$;
- (Produit "simplifié") Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ alors $u_n w_n \underset{n \rightarrow +\infty}{\sim} v_n w_n$;
- (Produit) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et $w_n \underset{n \rightarrow +\infty}{\sim} z_n$ alors $u_n w_n \underset{n \rightarrow +\infty}{\sim} v_n z_n$;
- (Inverse) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et qu'à partir d'un certain rang les deux suites sont non nulles, on a $\frac{1}{u_n} \underset{n \rightarrow +\infty}{\sim} \frac{1}{v_n}$;
- (Quotient) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, $w_n \underset{n \rightarrow +\infty}{\sim} z_n$ et qu'à partir d'un certain rang les deux suites $(w_n)_{n \in \mathbb{N}}$ et $(z_n)_{n \in \mathbb{N}}$ sont non nulles, on a $\frac{u_n}{w_n} \underset{n \rightarrow +\infty}{\sim} \frac{v_n}{z_n}$;
- (Puissance) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, alors pour tout $k \in \mathbb{N}$, on a $u_n^k \underset{n \rightarrow +\infty}{\sim} v_n^k$;
- (Valeur absolue) Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ alors $|u_n| \underset{n \rightarrow +\infty}{\sim} |v_n|$.

Remarque 5.2.6. Contrairement aux relations de négligeabilité, on ne peut pas simplifier par une constante multiplicative, ici elles ont toutes une importance capitale !

Remarque 5.2.7. Deux opérations sont impossibles. Il est INTERDIT d'additionner des équivalents ou appliquer une fonction à une équivalence !

Exemples 5.2.8. 1. On a $n + 1 \underset{n \rightarrow +\infty}{\sim} n$ et $-n \underset{n \rightarrow +\infty}{\sim} -n$ or $n + 1 - n = 1$ n'est pas équivalent à $n - n = 0$!

2. On a $n + 1 \underset{n \rightarrow +\infty}{\sim} n$ or e^{n+1} n'est pas équivalent à e^n puisque $\lim_{n \rightarrow +\infty} \frac{e^{n+1}}{e^n} = e \neq 1$.

Un résultat qui nous servira beaucoup est le suivant :

Théorème 5.2.9. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles telles que $u_n \underset{n \rightarrow +\infty}{\sim} v_n$. On a alors qu'à partir d'un certain rang les suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont de même signe.

Démonstration. Par définition, il existe une suite $(\epsilon_n)_{n \in \mathbb{N}}$ tel qu'à partir d'un certain rang n_0 , $u_n = \epsilon_n v_n$ et $\lim_{n \rightarrow +\infty} \epsilon_n = 1$. En particulier, il existe donc un rang n_1 tel que $\frac{1}{2} \leq \epsilon_n$ donc en particulier que ϵ_n est strictement positif.

On en déduit donc que pour tout $n \geq \max(n_0, n_1)$, u_n et v_n ont le même signe. \square

5.2.3 Méthodes de calculs

5.2.3.1 Équivalents usuels

Les règles de calculs précédentes combinées à des exemples usuelles nous permettrons de trouver des équivalents simples à la majorité des suites que nous étudierons, y compris dans les cas où nous n'avons pas d'intuition sur l'équivalent à déterminer.

Propriété 5.2.10. Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles telles que $u_n \underset{n \rightarrow +\infty}{\sim} v_n$. On a que $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont de même nature.

En particulier si elles sont convergentes, elles convergent vers la même limite $l \in \mathbb{R}$ et si $l \neq 0$, on a $u_n \underset{n \rightarrow +\infty}{\sim} l$ et $v_n \underset{n \rightarrow +\infty}{\sim} l$.

Exemple 5.2.11. On a déjà vu que $\frac{n-7}{n^2+n+1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$ or comme $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$ on a que $\lim_{n \rightarrow +\infty} \frac{n-7}{n^2+n+1} = 0$.

Remarques 5.2.12. • Une suite convergente de limite l non nulle est donc toujours équivalente à la suite constante égale à l .

- Dans le cas d'une suite convergente $(u_n)_{n \in \mathbb{N}}$ de limite 0, on n'a pas $u_n \underset{n \rightarrow +\infty}{\sim} 0$ sauf dans le cas où la suite est nulle à partir d'un certain rang. En effet, cette équivalence entraîne la nullité de la suite $(u_n)_{n \in \mathbb{N}}$ à partir d'un certain rang.

Exemple 5.2.13. On a pour tout $n \in \mathbb{N}^*$, $\frac{0}{n} = 0$ donc $\lim_{n \rightarrow +\infty} \frac{0}{n} = 0$ donc $(\frac{0}{n})_{n \in \mathbb{N}^*}$ n'est pas équivalente à la suite nulle alors que $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$.

Pour une suite polynômiale, la détermination d'un équivalent est simple :

Théorème 5.2.14. Soit $(a_0, \dots, a_k) \in \mathbb{R}^{k+1}$ tel que $a_k \neq 0$, on a alors

$$a_k n^k + a_{k-1} n^{k-1} + \dots + a_1 n + a_0 \underset{n \rightarrow +\infty}{\sim} a_k n^k.$$

Exemple 5.2.15. En utilisant ce résultat et la compatibilité des équivalences avec les quotients, on a

$$\frac{8n^4 + n^3 - 156n + \pi}{16n^3 + 5n + 15654156} \underset{n \rightarrow +\infty}{\sim} \frac{8n^4}{16n^3} = \frac{n}{2}.$$

Comme on l'a déjà vu précédemment, on ne peut pas appliquer de fonction à un équivalent mais nous pouvons traiter de nombreux problèmes que nous avons à l'aide des équivalents usuels suivant :

Théorème 5.2.16. Soit $(u_n)_{n \in \mathbb{N}}$ une suite convergent vers 0 et soit $\alpha \in \mathbb{R}^*$. On a alors

- $\ln(1 + u_n) \underset{n \rightarrow +\infty}{\sim} u_n$;
- $e^{u_n} - 1 \underset{n \rightarrow +\infty}{\sim} u_n$;
- $(1 + u_n)^\alpha - 1 \underset{n \rightarrow +\infty}{\sim} \alpha u_n$.

Exemples 5.2.17. Comme pour tout $n \in \mathbb{N}^*$, $\frac{1}{n} \neq 0$ et $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$, on a que

1. $\ln\left(1 + \frac{1}{n}\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$;
2. $e^{\frac{1}{n}} - 1 \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$;
3. $\sqrt{1 + \frac{1}{n}} - 1 \underset{n \rightarrow +\infty}{\sim} \frac{1}{2n}$.

5.2.3.2 Deux méthodes classiques

Méthode 5.2.18. Une manière de calculer de nombreux équivalents est d'utiliser les mêmes outils que lorsque nous levons des formes indéterminées. On peut donc par exemple factoriser par le terme "le plus puissant" ou multiplier par l'expression conjuguée.

Exemples 5.2.19. 1. Déterminons un équivalent en $+\infty$ de $\frac{e^n - n^3}{2e^{2n} - \ln(n)}$.

Par croissance comparée $-n^3 = o_{n \rightarrow +\infty}(e^n)$ donc $e^n - n^3 = e^n + o_{n \rightarrow +\infty}(e^n)$ on a $e^n \underset{n \rightarrow +\infty}{\sim} e^n - n^3$. De même, $-\ln(n) = o_{n \rightarrow +\infty}(2e^{2n})$ donc $2e^{2n} - \ln(n) = 2e^{2n} + o_{n \rightarrow +\infty}(2e^{2n})$ donc $2e^{2n} - \ln(n) \underset{n \rightarrow +\infty}{\sim} 2e^{2n}$ et comme cette dernière quantité est non nul, on a par quotient :

$$\frac{e^n - n^3}{2e^{2n} - \ln(n)} \underset{n \rightarrow +\infty}{\sim} \frac{e^n}{2e^{2n}} = \frac{1}{2e^n}.$$

2. Déterminons un équivalent en $+\infty$ de $\sqrt{n+1} - \sqrt{n}$. En multipliant par l'expression conjuguée, on a

$$\begin{aligned} \sqrt{n+1} - \sqrt{n} &= \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} \\ &= \frac{1}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n} \left(\sqrt{1 + \frac{1}{n}} + 1 \right)} \end{aligned}$$

Comme $\lim_{n \rightarrow +\infty} \sqrt{1 + \frac{1}{n}} + 1 = 2$, on a par règles de calcul sur les équivalents :

$$\sqrt{n+1} - \sqrt{n} \underset{n \rightarrow +\infty}{\sim} \frac{1}{2\sqrt{n}}.$$

Méthode 5.2.20. On peut encadrer une suite par deux autres suites équivalentes entre elle ce qui impliquera que la suite encadrée suit la même équivalence.

Exemple 5.2.21. Soit $(u_n)_{n \in \mathbb{N}}$ une suite vérifiant pour tout $n \in \mathbb{N}$, $4n - 1 \leq \frac{u_n}{3} \leq 4n + 7$.

Comme $4n - 1 \underset{n \rightarrow +\infty}{\sim} 4n$ et $4n + 7 \underset{n \rightarrow +\infty}{\sim} 4n$, on calcule $\frac{4n-1}{4n} \leq \frac{\frac{u_n}{3}}{4n} \leq \frac{4n+7}{4n}$.

Par théorème des gendarmes, on a donc $\lim_{n \rightarrow +\infty} \frac{\frac{u_n}{3}}{4n} = 1$ et donc $\frac{u_n}{3} \underset{n \rightarrow +\infty}{\sim} 4n$. On en déduit donc

$$u_n \underset{n \rightarrow +\infty}{\sim} 12n.$$

Remarque 5.2.22. Le fait d'avoir le même équivalent pour les deux suites qui encadrent u_n nous permet de nous assurer que la limite est bien de 1 et donc d'utiliser la caractérisation des équivalents.

Chapitre 6

Séries : Révisions et compléments

En première année l'étude des séries a été orienté principalement par le calcul direct mais grâce aux nouveaux outils de comparaisons de suites nous pouvons obtenir des moyens de déterminer la convergence ou la divergence d'une série numérique de manière bien plus rapide. Cela nous servira notamment en probabilité lorsque que nous devrons étudier la convergence de certains séries pour établir l'existence ou la non existence d'une espérance ou plus généralement d'un moment d'une variable aléatoire discrète.

6.1 Rappels

6.1.1 Définitions et propriétés générales

Rappelons pour commencer le vocabulaire usuel :

Définitions 6.1.1. Soit $(u_k)_{k \geq k_0}$ une suite de réels :

- on appelle alors $\sum_{k \geq k_0} u_k$ la série numérique de terme général u_k ;
- pour tout entier $n \geq k_0$, on appelle somme partielle de la série numérique $\sum_{k \geq k_0} u_k$ de rang n la somme $S_n = \sum_{k=k_0}^n u_k$.
- La série numérique $\sum_{k \geq k_0} u_k$ est dite convergente si la suite de ses sommes partielles $(S_n)_n$ est convergente. Dans le cas contraire, on dit que la série numérique est divergente.
- On dit que la série numérique $\sum_{k \geq k_0} u_k$ diverge grossièrement si $\lim_{k \rightarrow +\infty} u_k \neq 0$.

- Si la série numérique $\sum_{k \geq k_0} u_k$ est convergente, la limite $S = \lim_{n \rightarrow +\infty} S_n$ de la suite des sommes partielles est appelé somme de la série et on note

$$S = \sum_{k=k_0}^{\infty} u_k.$$

On appelle reste d'ordre n la quantité $R_n = S - S_n$

Le premier terme d'une série est important pour déterminer la valeur d'une série convergente pour l'étude de la convergence ou de la divergence, le point de départ de la série n'a que peu d'importance :

Propriété 6.1.2. Soit $(u_k)_{k \in \mathbb{N}}$ une série numérique et soit k_0 un entier positif, alors les séries numériques $\sum_{k \geq 0} u_k$ et $\sum_{k \geq k_0} u_k$ sont de même nature, c'est-à-dire qu'elles convergent ou divergent en même temps.

Remarque 6.1.3. C'est grâce à ce résultat que l'on considère dans toute la suite du chapitre que le premier terme d'une série est en grande majorité 0 ou sinon un entier arbitraire que l'on choisira implicitement.

Un lien important entre les suites et les séries est le suivant :

Propriété 6.1.4 (Lien suite-série). Soit $(u_n)_{n \in \mathbb{N}}$ une suite de réels. La suite $(u_n)_{n \in \mathbb{N}}$ est convergente si et seulement si la série numérique $\sum_{n \geq 0} u_{n+1} - u_n$ est convergente.

Pour additionner des séries entres elles, il faut prendre garde à la convergence de chaque partie :

Propriétés 6.1.5. 1. Soit $\sum_{k \geq 0} u_k$ et $\sum_{k \geq 0} v_k$ deux séries numériques convergentes alors la série $\sum_{k \geq 0} (u_k + v_k)$ converge et on a

$$\sum_{k=0}^{\infty} u_k + \sum_{k=0}^{\infty} v_k = \sum_{k=0}^{\infty} (u_k + v_k).$$

2. Soit $\sum_{k \geq 0} u_k$ une série numérique et soit λ une constante réelle non nulle, alors la série $\sum_{k \geq 0} u_k$ converge si et seulement si $\sum_{k \geq 0} \lambda u_k$ converge. Dans ce cas de convergence on a

$$\sum_{k=0}^{\infty} \lambda u_k = \lambda \sum_{k=0}^{\infty} u_k.$$

6.1.2 Séries usuelles

Commençons par la série divergente par excellence, la série harmonique :

Propriété 6.1.6 (Série harmonique). La série numérique $\sum_{k \geq 1} \frac{1}{k}$ est appelée la série harmonique et il s'agit d'une série divergente.

Pour les séries usuelles convergentes, nous en avons vu quatre en première année :

Théorème 6.1.7 (Série géométrique). On appelle série géométrique de raison $q \in \mathbb{R}$ la série numérique $\sum_{k \geq 0} q^k$ et on a :

- La série géométrique $\sum_{k \geq 0} q^k$ converge si et seulement si $|q| < 1$
- Si $|q| < 1$ alors $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$.

Théorème 6.1.8 (Série géométrique dérivée première). Soit $x \in]-1, 1[$, la série $\sum_{n \geq 1} nx^{n-1}$ est convergente et on a

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}.$$

Théorème 6.1.9 (Série géométrique dérivée seconde). Soit $x \in]-1, 1[$, la série $\sum_{n \geq 2} n(n-1)x^{n-2}$ est convergente et on a

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}.$$

Théorème 6.1.10 (Série exponentielle). Pour tout $x \in \mathbb{R}$, la série numérique $\sum_{n \geq 0} \frac{x^n}{n!}$ est convergente et on a :

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

Remarque 6.1.11. Les preuves des convergences des séries géométriques sont à savoir refaire !

6.1.3 Convergence absolue

Rappelons pour terminer cette section la notion de convergence absolue.

Définition 6.1.12. On dit que la série numérique $\sum_{k \geq 0} u_k$ est absolument convergente si la série numérique $\sum_{k \geq 0} |u_k|$ est convergente.

Propriété 6.1.13. Une série numérique absolument convergente est convergente.

Remarque 6.1.14. La réciproque est cependant fautive, il existe des séries numériques non absolument convergentes mais convergentes ! De telles séries portent un nom :

Définition 6.1.15. Une série numérique convergente mais non absolument convergente est dite semi-convergente.

Exemple 6.1.16. La série numérique $\sum_{n \geq 1} \frac{(-1)^n}{n}$ est convergente. En effet on pose pour tout $n \in \mathbb{N}^*$, $u_n = S_{2n}$ et $v_n = S_{2n+1}$ où $(S_k)_{k \in \mathbb{N}^*}$ est la somme partielle de la série numérique $\sum_{n \geq 1} \frac{(-1)^n}{n}$. On a

- $(u_n)_{n \in \mathbb{N}^*}$ est décroissante car pour tout $n \in \mathbb{N}^*$

$$\begin{aligned} u_{n+1} - u_n &= S_{2n+2} - S_{2n} = \frac{(-1)^{2n+2}}{2n+2} + \frac{(-1)^{2n+1}}{2n+1} \\ &= \frac{1}{2n+2} - \frac{1}{2n+1} = \frac{2n+1 - (2n+2)}{(2n+1)(2n+2)} \\ &= -\frac{1}{(2n+1)(2n+2)} < 0 \end{aligned}$$

- $(v_n)_{n \in \mathbb{N}^*}$ est croissante car pour tout $n \in \mathbb{N}^*$

$$\begin{aligned} v_{n+1} - v_n &= S_{2n+3} - S_{2n+1} = \frac{(-1)^{2n+3}}{2n+3} + \frac{(-1)^{2n+2}}{2n+2} \\ &= -\frac{1}{2n+3} + \frac{1}{2n+2} = \frac{-(2n+2) + (2n+3)}{(2n+2)(2n+3)} \\ &= \frac{1}{(2n+2)(2n+3)} > 0 \end{aligned}$$

- Pour tout $n \in \mathbb{N}^*$, $v_n - u_n = S_{2n+1} - S_{2n} = \frac{(-1)^{2n+1}}{2n+1} = \frac{-1}{2n+1}$ et donc on a $\lim_{n \rightarrow +\infty} v_n - u_n = 0$

Les suites $(u_n)_{n \in \mathbb{N}^*}$ et $(v_n)_{n \in \mathbb{N}^*}$ sont donc adjacentes, elles sont donc convergente de même limite. Autrement dit $(S_{2n})_{n \in \mathbb{N}^*}$ et $(S_{2n+1})_{n \in \mathbb{N}^*}$ sont convergentes de même limite donc la suite $(S_k)_{k \in \mathbb{N}^*}$ est également convergente vers cette même limite, soit que la série numérique $\sum_{k \geq 1} \frac{(-1)^k}{k}$ est convergente. (Il est possible de montrer que cette série vaut $\ln(2)$ en combinant nos arguments de séries à des arguments d'intégration.)

Cependant la série $\sum_{n \geq 1} \left| \frac{(-1)^n}{n} \right|$ correspond à la série harmonique qui est divergente.

On en conclut donc que la série numérique $\sum_{n \geq 1} \frac{(-1)^n}{n}$ est semi-convergente.

6.2 Séries à termes positifs

L'étude des séries à termes positifs est essentielle car de nombreuses séries sont de ce type là et l'étude de l'absolue convergence correspond à ce cas là et permet d'étudier de nombreuses séries dans le cas général.

6.2.1 Série de Riemann

Un exemple type de série à termes positifs dont nous connaissons parfaitement la convergence est la famille des séries de Riemann.

Définition 6.2.1. Soit $\alpha \in \mathbb{R}$, on appelle série de Riemann la série numérique $\sum_{k \geq 1} \frac{1}{k^\alpha}$.

Théorème 6.2.2 (Séries de Riemann). *La série de Riemann $\sum_{k \geq 1} \frac{1}{k^\alpha}$ converge si et seulement si $\alpha > 1$.*

Démonstration. Dans un premier temps on remarque que si $\alpha \leq 0$, alors la série numérique $\sum_{n \geq 1} \frac{1}{n^\alpha}$ diverge grossièrement. On suppose donc à présent $\alpha > 0$.

Le cas $\alpha = 1$ est bien connu car il correspond à la série harmonique. On suppose donc de plus que $\alpha \neq 1$ à présent.

On pose $f : [1, +\infty[\rightarrow \mathbb{R}_+$ et on remarque que f est une fonction continue décroissante et positive.

$$t \mapsto \frac{1}{t^\alpha}$$

Soit $k \in \mathbb{N}^*$, on remarque que, comme f est décroissante pour tout $t \in [k, k+1]$, on a $f(k+1) \leq f(t) \leq f(k)$. En intégrant cette

inégalité sur $[k, k + 1]$ par rapport à t , on obtient :

$$\frac{1}{(k+1)^\alpha} = f(k+1) = \int_k^{k+1} f(k+1) dt \leq \int_k^{k+1} f(t) dt \leq \int_k^{k+1} f(k) dt = f(k) = \frac{1}{k^\alpha}.$$

On note pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^n f(k)$ et $I_n = \int_1^n f(t) dt$. On a donc en sommant l'inégalité précédente pour tout $k \in \llbracket 1, n \rrbracket$:

$$\begin{aligned} \sum_{k=1}^n f(k+1) &\leq \sum_{k=1}^n \int_k^{k+1} f(t) dt \leq \sum_{k=1}^n f(k) \\ &\Leftrightarrow \sum_{k=2}^{n+1} f(k) \leq \int_1^{n+1} f(t) dt \leq S_n \\ &\Leftrightarrow S_{n+1} - f(1) \leq I_{n+1} \leq S_n \end{aligned}$$

On peut réécrire le premier bout d'inégalité sous la forme $S_n \leq I_n + f(1)$ et on obtient

$$I_{n+1} \leq S_n \leq I_n + f(1).$$

Comme la fonction f est positive, les suites $(I_n)_{n \in \mathbb{N}^*}$ et $(S_n)_{n \in \mathbb{N}^*}$ sont croissantes, elles convergent donc si et seulement si elles sont majorées. Cependant cette inégalité implique que l'une ne peut pas être majorée sans que la seconde ne le soit donc $(S_n)_{n \in \mathbb{N}^*}$ converge si et seulement si $(I_n)_{n \in \mathbb{N}^*}$ converge aussi. Comme la limite de $(I_n)_{n \in \mathbb{N}}$ correspond à l'intégrale de Riemann (2.2.11) qui converge lorsque $\alpha > 1$ et diverge lorsque $\alpha \leq 1$ on a :

- Si $\alpha > 1$ alors $(S_n)_{n \in \mathbb{N}^*}$ converge également et comme il s'agit de la somme partielle de la série de Riemann de paramètre α , $\sum_{k \geq 1} \frac{1}{k^\alpha}$ est convergente.
- Si $\alpha < 1$ alors $(S_n)_{n \in \mathbb{N}^*}$ diverge également et comme il s'agit de la somme partielle de la série de Riemann de paramètre α , $\sum_{k \geq 1} \frac{1}{k^\alpha}$ est divergente.

En rassemblant tout les cas, on a bien que la série de Riemann $\sum_{k \geq 1} \frac{1}{k^\alpha}$ converge si et seulement si $\alpha > 1$. □

Ce simple critère fait des séries de Riemann un exemple type simple de séries à termes positifs auxquelles comparer des séries plus complexes.

Remarque 6.2.3. La méthode utilisée ici est extrêmement classique, il s'agit de la comparaison série intégrale qui utilise le fait d'écrire la série sous la forme

$\sum_{k=1}^{\infty} f(k)$ où f est une fonction continue décroissante et positive.

C'est en particulier pour cette raison que le critère de convergence pour les séries de Riemann est le même que celui pour les intégrales de Riemann.

Exemples 6.2.4. 1. La série $\sum_{k=1}^{\infty} \frac{1}{k^2}$ est donc convergente en tant que série de

Riemann de coefficient $\alpha = 2 > 1$. Il est connu que cette série vaut $\frac{\pi^2}{6}$.

2. La série $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ est donc divergente en tant que série de Riemann de coefficient $\alpha = \frac{1}{2} < 1$.

3. La série $\sum_{k=1}^{\infty} \frac{1}{k\sqrt{k}}$ est donc convergente en tant que série de Riemann de coefficient $\alpha = \frac{3}{2} > 1$.

6.2.2 Utilisation des relations de comparaisons

En première année les résultats de comparaisons entre les suites pour les séries numériques se limitaient aux inégalités et étaient :

Propriété 6.2.5. Soit $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ deux suites réelles à termes positifs telles que pour tout $k \in \mathbb{N}$, $u_k \leq v_k$. Si la série $\sum_{k \geq 0} v_k$ est convergente alors la série

$\sum_{k \geq 0} u_k$ est également convergente et on a

$$\sum_{k=0}^{\infty} u_k \leq \sum_{k=0}^{\infty} v_k.$$

Démonstration. Soit $n \in \mathbb{N}$, on note U_n la somme partielle de rang n associé à la série numérique de terme général $(u_k)_k$ et V_n celle associé à $(v_k)_k$. Par hypothèse la suite $(V_n)_n$ converge vers la somme de la série que l'on note V . Comme pour tout $k \in \mathbb{N}$, $u_k \leq v_k$, on a pour tout n , $U_n \leq V_n$.

De plus comme les suites $(u_k)_k$ et $(v_k)_k$ sont à termes positifs, les suites $(U_n)_n$ et $(V_n)_n$ sont croissantes. On a donc pour tout n , $U_n \leq V_n \leq V$ et donc la suite $(U_n)_n$ est croissante et majorée, elle est donc convergente et par définition de la convergence d'une série numérique, la série $\sum_{k \geq 0} u_k$ est convergente. \square

Corollaire 6.2.6. Soit $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ deux suites réelles à termes positifs telles que pour tout $k \in \mathbb{N}$, $u_k \leq v_k$. Si la série $\sum_{k \geq 0} u_k$ est divergente alors la série $\sum_{k \geq 0} v_k$ est également divergente.

Démonstration. Il s'agit de la contraposée du résultat précédent. \square

Ils utilisaient le fait que, pour une série à termes positifs, la suite des sommes partielles est croissante et donc que la convergence dépend intégralement du caractère bornée des sommes partielles.

- Exemples 6.2.7.** 1. On a pour tout $k \in \mathbb{N}^*$, $0 \leq \frac{1}{k(k+1)} \leq \frac{1}{k^2}$ et comme la série de terme général $\frac{1}{k^2}$ est convergente en tant que série de Riemann de paramètre $\alpha = 2 > 1$, par théorème de comparaison sur les séries à termes positifs, on a que la série de terme général $\frac{1}{k(k+1)}$ est également convergente.
2. Par inégalité classique, on a pour tout $t \in]0, +\infty[$, on a $\ln(t) \leq t$ et donc pour tout $k \in \mathbb{N}^*$, $0 < \frac{1}{k} \leq \frac{1}{\ln(k)}$. Comme la série harmonique est divergente, on a par théorème de comparaison sur les séries à termes positifs que la série numérique $\sum_{k \geq 1} \frac{1}{\ln(k)}$ est divergente.

A présent nous avons les relations de négligeabilité et d'équivalence.

Théorème 6.2.8. Soit $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ deux suites réelles à termes positifs telles que $u_k = o_{k \rightarrow +\infty}(v_k)$. On a alors :

1. Si la série numérique $\sum_{k \geq 0} v_k$ est convergente alors la série numérique $\sum_{k \geq 0} u_k$ est également convergente.
2. Si la série numérique $\sum_{k \geq 0} u_k$ est divergente alors la série numérique $\sum_{k \geq 0} v_k$ est également divergente.

Démonstration. Par définition, il existe une suite $(\epsilon_k)_{k \in \mathbb{N}}$ de limite nulle telle qu'à partir d'un certain rang n_0 $u_k = \epsilon_k v_k$. Comme les suites $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ sont positives, au moins à partir du rang n_0 , $(\epsilon_k)_{k \in \mathbb{N}}$ l'est également.

La convergence vers 0 implique qu'il existe $n_1 \in \mathbb{N}$ tel que pour tout $k \geq n_0$, $\epsilon_k \leq 1$ et donc pour tout $k \geq \max(n_0, n_1)$, on a $u_k \leq v_k$.

Il nous suffit donc d'appliquer le précédent résultat sur les comparaisons ainsi que son corollaire pour prouver ce théorème. \square

Exemple 6.2.9. Étudions la série numérique $\sum_{k \geq 0} k e^{-k}$. On a clairement que $k e^{-k}$ est positif et par croissance comparée on a que $\lim_{k \rightarrow +\infty} k^2 (k e^{-k}) = 0$ donc $k e^{-k} = o_{k \rightarrow +\infty}\left(\frac{1}{k^2}\right)$. Par comparaison de série à terme positif avec la série de Riemann convergente de coefficient $\alpha = 2 > 1$, on a bien que la série numérique $\sum_{k \geq 0} k e^{-k}$ est convergente.

Remarque 6.2.10. La positivité est ici indispensable ! La réciproque est fautive comme le montre l'exemple suivant :

On verra en TD que la série $\sum_{k \geq 1} \frac{(-1)^k}{\sqrt{k}}$ est semi-convergente (donc convergente)

et la série harmonique est divergente or on observe que $\frac{1}{k} = o_{k \rightarrow +\infty} \left(\frac{(-1)^k}{\sqrt{k}} \right)$ car

$$\frac{\frac{1}{k}}{\frac{(-1)^k}{\sqrt{k}}} = \frac{(-1)^k}{\sqrt{k}} \rightarrow 0 \text{ à pour limite } 0.$$

La positivité dans ce théorème de comparaison est donc bien fondamentale !

Ce résultat suppose de connaître la nature de la série de terme générale d'une des deux suites et le signe des suites cependant en pratique il est parfois compliqué de connaître ces informations sur les deux séries ou bien de savoir quelle deuxième suite utiliser. Ce n'est pas le cas avec le résultat concernant les équivalents :

Théorème 6.2.11. Soit $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ deux suites réelles à termes positifs telles que $u_k \underset{k \rightarrow +\infty}{\sim} v_k$. On a alors que les séries numériques $\sum_{k \geq 0} u_k$ et $\sum_{k \geq 0} v_k$ ont la même nature.

Démonstration. Par définition, il existe une suite $(\epsilon_k)_{k \in \mathbb{N}}$ de limite 1 telle qu'à partir d'un certain rang n_0 $u_k = \epsilon_k v_k$. Comme les suites $(u_k)_{k \in \mathbb{N}}$ et $(v_k)_{k \in \mathbb{N}}$ sont positives, au moins à partir du rang n_0 , $(\epsilon_k)_{k \in \mathbb{N}}$ l'est également.

La convergence vers 1 implique qu'il existe $n_1 \in \mathbb{N}$ tel que pour tout $k \geq n_0$, $\frac{1}{2} \leq \epsilon_k \leq \frac{3}{2}$ et donc pour tout $k \geq \max(n_0, n_1)$, on a $\frac{v_k}{2} \leq u_k \leq \frac{3v_k}{2}$.

Il nous suffit donc d'appliquer le même résultat que précédemment sur les comparaisons ainsi que son corollaire pour prouver ce théorème. \square

Exemples 6.2.12. 1. Étudions la série numérique $\sum_{n \geq 0} \frac{n^4 + 3n + 7}{4n^7 + 3n^2 + 7n + \pi}$. On a que

$$\frac{n^4 + 3n + 7}{4n^7 + 3n^2 + 7n + \pi} \sim \frac{n^4}{4n^7} = \frac{1}{4n^3}$$

Les termes généraux sont donc positifs à partir d'un certain rang et comme la série numérique de terme général $\frac{1}{4n^3}$ est convergente en tant que série de Riemann de paramètre $\alpha = 3 > 1$. On a que la série numérique $\sum_{n \geq 0} \frac{n^4 + 3n + 7}{4n^7 + 3n^2 + 7n + \pi}$ est convergente.

2. Étudions la série numérique $\sum_{n \geq 1} \ln \left(1 - \frac{1}{\sqrt{n}} \right)$. Cette série est à terme négatif mais en considérant l'opposée, on se ramène à une série à terme positif. On a que

$$-\ln \left(1 - \frac{1}{\sqrt{n}} \right) \sim -\frac{-1}{\sqrt{n}} = \frac{1}{\sqrt{n}}$$

Les termes généraux sont donc positifs à partir d'un certain rang et comme la série numérique de terme général $\frac{1}{\sqrt{n}}$ est divergente en tant que série de Riemann de paramètre $\alpha = \frac{1}{2} < 1$. On a que la série numérique $\sum_{n \geq 1} -\ln\left(1 - \frac{1}{\sqrt{n}}\right)$ est divergente et donc $\sum_{n \geq 1} \ln\left(1 - \frac{1}{\sqrt{n}}\right)$ diverge également.

Remarque 6.2.13. La positivité est ici aussi indispensable ! La réciproque est fautive comme le montre l'exemple suivant : On considère de nouveau la série numérique convergente $\sum_{k \geq 1} \frac{(-1)^k}{\sqrt{k}}$ ainsi que la série numérique $\sum_{k \geq 1} \left(\frac{(-1)^k}{k} + \frac{1}{k}\right)$. Cette dernière est divergente puisque le terme général est la somme du terme général d'une série convergente et d'une série divergente. Cependant on observe que

$$\frac{\frac{(-1)^k}{k} + \frac{1}{k}}{\frac{(-1)^k}{k}} = 1 + \frac{(-1)^k}{k}$$

et $\lim_{k \rightarrow +\infty} 1 + \frac{(-1)^k}{k} = 1$ donc on a bien

$$\frac{(-1)^k}{k} \underset{k \rightarrow +\infty}{\sim} \frac{(-1)^k}{k} + \frac{1}{k}$$

La positivité de ce théorème de comparaison est donc bien fondamentale !

Chapitre 7

Applications linéaires et endomorphismes

Dans ce chapitre nous allons étendre les connaissances que nous avons obtenus sur les applications linéaires sur les espaces vectoriels \mathbb{R}^n aux applications linéaires sur les espaces vectoriels réels de dimensions finis quelconques.

Nous nous attarderons également sur le point de vue matricielle des applications linéaires et sur le lien entre les matrices d'une même application selon les bases associées aux espaces vectoriels choisies.

7.1 Définitions et règles de calculs

Définition 7.1.1. Soit E et F deux espaces vectoriels réels et soit f une application de E dans F . On dit que f est une application linéaire si elle vérifie pour tout $(x, y) \in E^2$ et pour tout $\lambda \in \mathbb{R}$,

$$\begin{cases} f(x + y) = f(x) + f(y) \\ f(\lambda x) = \lambda f(x) \end{cases}$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F .

Si $E = F$ on note cet ensemble $\mathcal{L}(E)$ et on dit que f est un endomorphisme.

En pratique on utilise la propriété suivante :

Propriété 7.1.2. Soit E et F deux espaces vectoriels sur \mathbb{R} et soit f une application de E dans F . f est une application linéaire si et seulement si f vérifie

$$\forall (x, y) \in E^2, \forall \lambda \in \mathbb{R}, f(\lambda x + y) = \lambda f(x) + f(y).$$

Remarques 7.1.3. 1. Une application linéaire est donc une application qui conserve les combinaisons linéaires. En particulier lorsque $f \in \mathcal{L}(E, F)$ on peut montrer par récurrence que pour tout $n \in \mathbb{N}$, pour tout $(x_1, \dots, x_n) \in E^n$ et pour tout $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$,

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i).$$

2. On remarquera que pour une application linéaire, on a automatiquement $f(0_E) = 0_F$ et $f(-x) = -f(x)$ mais cela ne caractérise pas les applications linéaires.

Exemples 7.1.4. 1. L'application nulle $E \rightarrow F$ et l'application identité $x \mapsto 0_F$

$$\begin{aligned} Id_E : E &\rightarrow E \\ x &\mapsto x \end{aligned} \text{ sont des applications linéaires.}$$

2. L'application $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ est une application linéaire. En effet pour tout $((x, y, z), (x', y', z')) \in (\mathbb{R}^3)^2$ et pour tout $\lambda \in \mathbb{R}$, on a

$$\begin{aligned} f(\lambda(x, y, z) + (x', y', z')) &= f((\lambda x + x', \lambda y + y', \lambda z + z')) \\ &= (\lambda x + x' + \lambda y + y', \lambda y + y' + \lambda z + z') \\ &= (\lambda(x + y) + (x' + y'), \lambda(y + z) + (y + z)) \\ &= \lambda(x + y, y + z) + (x' + y', y' + z') \\ &= \lambda f(x, y, z) + f(x', y', z') \end{aligned}$$

3. L'application cube $g : \mathbb{R} \rightarrow \mathbb{R}$ n'est pas une application linéaire car $x \mapsto x^3$

même si $g(0) = 0$ et $g(-x) = -g(x)$, on a que $g(1 + 1) \neq g(1) + g(1)$ puisque $g(1 + 1) = g(2) = 8$ et $g(1) + g(1) = 1 + 1 = 2$.

4. L'application $\varphi : \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathcal{M}_{p,n}(\mathbb{R})$ est une application linéaire. $M \mapsto {}^t M$

En effet, pour tout $(A, B) \in \mathcal{M}_{n,p}(\mathbb{R})^2$ et pour tout $\lambda \in \mathbb{R}$ on a

$$\varphi(\lambda A + B) = {}^t(\lambda A + B) = \lambda {}^t A + {}^t B = \lambda \varphi(A) + \varphi(B).$$

5. Soit $n \in \mathbb{N}^*$. L'application $D : \mathbb{R}_n[x] \rightarrow \mathbb{R}_{n-1}[x]$ est une application

$$P \mapsto P'$$

linéaire. En effet, pour tout $(P, Q) \in \mathbb{R}_n[x]^2$ et pour tout $\lambda \in \mathbb{R}$, on a

$$D(\lambda P + Q) = (\lambda P + Q)' = \lambda P' + Q' = \lambda D(P) + D(Q).$$

$$h : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathbb{R}$$

6. Soit $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + b + c + d$ est une application linéaire. En

effet, pour tout $\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \right) \in \mathcal{M}_2(\mathbb{R})^2$ et pour tout $\lambda \in \mathbb{R}$, on

a

$$\begin{aligned} h \left(\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \right) &= h \left(\begin{pmatrix} \lambda a + a' & \lambda b + b' \\ \lambda c + c' & \lambda d + d' \end{pmatrix} \right) \\ &= \lambda a + a' + \lambda b + b' + \lambda c + c' + \lambda d + d' \\ &= \lambda(a + b + c + d) + (a' + b' + c' + d') \\ &= \lambda h \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) + h \left(\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \right). \end{aligned}$$

L'ensemble des applications linéaires est un ensemble ayant les propriétés suivantes :

Propriété 7.1.5. Soit E et F deux espaces vectoriels réels alors $\mathcal{L}(E, F)$ et $\mathcal{L}(E)$ sont des espaces vectoriels réels.

Remarque 7.1.6. Cela signifie en particulier que toutes combinaisons linéaires d'applications linéaires est une application linéaire.

Concernant la composition de d'applications linéaires, on a le résultat suivant :

Propriété 7.1.7. Soit E, F et G trois espaces vectoriels réels sur \mathbb{R} . Soit $f \in \mathcal{L}(E, F)$ et soit $g \in \mathcal{L}(F, G)$, alors $g \circ f \in \mathcal{L}(E, G)$.

Remarque 7.1.8. En particulier si $E = F = G$, on a $g \circ f \in \mathcal{L}(E)$, autrement dit, la composition de deux endomorphismes de E est toujours un endomorphisme de E .

Exemple 7.1.9. On considère $f : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ et

$$(x, y, z) \mapsto (x - y, 2y, 0, z)$$

$g : \mathbb{R}^4 \rightarrow \mathbb{R}^2$. On admet que f et g sont des applications linéaires. On a donc que $g \circ f$ est une application linéaire de \mathbb{R}^3 dans \mathbb{R}^2 et pour tout $(x, y, z) \in \mathbb{R}^3$, on a :

$$(x, y, z, t) \mapsto (x + y, 3z - 2t)$$

tions linéaires. On a donc que $g \circ f$ est une application linéaire de \mathbb{R}^3 dans \mathbb{R}^2 et pour tout $(x, y, z) \in \mathbb{R}^3$, on a :

$$\begin{aligned} (g \circ f)(x, y, z) &= g(f(x, y, z)) = g(x - y, 2y, 0, z) \\ &= ((x - y) + 2y, 3 \cdot 0 - 2 \cdot z) = (x + y, -2z) \end{aligned}$$

Dans le cas des endomorphismes la notation suivante est à connaître absolument :

Notation 7.1.10. Soit E un espace vectoriel réel et soit f un endomorphisme de E . Soit $n \in \mathbb{N}$, on note f^n la composée de f avec lui même n fois. On a donc

$$f^n = \underbrace{f \circ \dots \circ f}_{n \text{ fois}}$$

Exemple 7.1.11. On considère l'application linéaire $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

$$(x, y, z) \mapsto (0, x, y)$$

On a alors pour tout $(x, y, z) \in E$

- $f^2((x, y, z)) = f(f((x, y, z))) = f((0, x, y)) = (0, 0, x)$
- $f^3((x, y, z)) = f(f^2((x, y, z))) = f((0, 0, x)) = (0, 0, 0)$
- Soit k un entier valant au moins 4, on a alors $f^k((x, y, z)) = f^{k-3}(f^3((x, y, z))) = f^{k-3}((0, 0, 0)) = (0, 0, 0)$ donc pour tout $k \geq 3$, f^k est l'application nulle.

La dernière notion que nous allons considérer pour les applications linéaires est la notion d'isomorphisme.

Définitions 7.1.12. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$.

1. On dit que f est un isomorphisme de E dans F si f est bijective.
On note $\mathcal{GL}(E, F)$ l'ensemble des isomorphismes de E dans F et si $\mathcal{GL}(E, F) \neq \emptyset$ on dit alors que E est isomorphe à F .
2. Si $E = F$ et que f est un isomorphisme de E dans E , on parle alors d'automorphisme de E . On note $\mathcal{GL}(E)$ l'ensemble des automorphismes de E .

Exemples 7.1.13. Les applications linéaires

$$f : \mathbb{R}_n[x] \rightarrow \mathbb{R}^{n+1}$$

$$\sum_{i=0}^n a_i x^i \mapsto (a_0, a_1, \dots, a_n)$$

et

$$f : \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathbb{R}^{np}$$

$$(a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, p \rrbracket}} \mapsto (a_{1,1}, a_{1,2}, \dots, a_{1,p}, a_{2,1}, \dots, a_{n,p})$$

sont respectivement des éléments de $\mathcal{GL}(\mathbb{R}_n[x], \mathbb{R}^{n+1})$ et $\mathcal{GL}(\mathcal{M}_{n,p}(\mathbb{R}), \mathbb{R}^{np})$ comme vu dans la démonstration du théorème 3.1.11.

Propriété 7.1.14. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{GL}(E, F)$. On a alors que $f^{-1} \in \mathcal{GL}(F, E)$.

Remarque 7.1.15. En particulier, pour un automorphisme f de E , on a que f^{-1} est également un automorphisme de E .

Exemple 7.1.16. L'application identité est toujours un automorphisme et l'on a $Id_E^{-1} = Id_E$ qui est bien également un automorphisme.

7.2 Notion d'image et de noyau

Deux sous-espaces vectoriels que nous considérons de façons quasi-permanentes dans nos études d'applications linéaires sont les noyaux et les images associées.

7.2.1 Noyau

Définition 7.2.1. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On appelle noyau de f l'ensemble

$$\ker(f) := \{x \in E \mid f(x) = 0_F\}.$$

Remarque 7.2.2. Il s'agit d'un sous-ensemble de E et on a toujours $0_E \in \ker(f)$.

Propriété 7.2.3. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a que $\ker(f)$ est un sous-espace vectoriel de E .

Démonstration. • On a par définition que $\ker(f) \subset E$.

- On a $f(0_E) = 0_F$ car f est linéaire donc $0_E \in \ker(f)$.
- Soit $(x, y) \in \ker(f)^2$ et soit $\lambda \in \mathbb{R}$. On a

$$f(\lambda x + y) = \lambda f(x) + f(y) = \lambda 0_F + 0_F = 0_F$$

donc $\lambda x + y \in \ker(f)$.

$\ker(f)$ est donc bien un sous-espace vectoriel de E . □

Exemples 7.2.4. 1. On considère l'application linéaire $D : \mathbb{R}_n[x] \rightarrow \mathbb{R}_{n-1}[x]$

$$P \mapsto P'$$

et déterminons $\ker(D)$. On cherche donc les polynômes de degré au plus n dont la dérivée est le polynôme nulle. Il s'agit des polynômes de degré au plus 0 donc $\ker(D) = \mathbb{R}_0[x]$.

2. On considère l'application linéaire $\varphi : \mathcal{M}_{n,p}(\mathbb{R}) \rightarrow \mathbb{M}_{p,n}(\mathbb{R})$ et dé-

$$M \mapsto {}^t M$$

terminons son noyau. Soit $M = (a_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 1, p \rrbracket}} \in \ker(\varphi)$. On a alors ${}^t M = 0_{p,n}$

or $\varphi(M) = {}^t M = (a_{j,i})_{\substack{j \in \llbracket 1, p \rrbracket \\ i \in \llbracket 1, n \rrbracket}}$ donc pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, $a_{i,j} = 0$

soit $M = 0_{n,p}$. On a donc $\ker(\varphi) = \{0_{n,p}\}$.

Remarque 7.2.5. Le calcul de noyaux sera au centre d'un futur chapitre, il est impératif de maîtriser parfaitement cette notion.

Dans le cas où le noyau est réduit à $\{0_E\}$ est un cas qui nous apporte des informations théoriques sur l'application linéaire associée.

Propriété 7.2.6. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a que f est injective si et seulement si $\ker(f) = \{0_E\}$ si et seulement si $\dim(\ker(f)) = 0$.

Démonstration. Si f est injective alors l'équation $f(x) = 0_F$ admet une unique solution qui est 0_E donc $\ker(f) = \{0_E\}$.

Réciproquement supposons que $\ker(f) = \{0_E\}$. Soit $(x, y) \in E^2$ tel que $f(x) = f(y)$ donc par linéarité de f , on a $f(x - y) = 0_F$ donc $x - y \in \ker(f)$ donc $x - y = 0_E$ soit $x = y$. On a donc bien que f est injective.

L'équivalence entre $\ker(f) = \{0_E\}$ et $\dim(\ker(f)) = 0$ est immédiate. □

- Exemples 7.2.7.** 1. En considérant D de l'exemple précédent, on a $\dim(\ker(D)) = 1 \neq 0$ donc D n'est pas injectif.
2. De même, en considérant φ de l'exemple précédent, on a $\ker(\varphi) = \{0_{n,p}\}$ donc φ est une application injective.

7.2.2 Image

Définition 7.2.8. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On appelle image de f l'ensemble

$$\text{Im}(f) := \{f(x) \mid x \in E\}.$$

Remarque 7.2.9. Il s'agit d'un sous-ensemble de F et on a toujours $0_F \in \text{Im}(f)$.

Propriété 7.2.10. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a que $\text{Im}(f)$ est un sous-espace vectoriel de F .

Démonstration. • On a par définition que $\text{Im}(f) \subset F$.

- On a $f(0_E) = 0_F$ car f est linéaire donc $0_F \in \text{Im}(f)$.
- Soit $(x, y) \in \text{Im}(f)^2$ et soit $\lambda \in \mathbb{R}$. Il existe donc $(x_0, y_0) \in E^2$ tel que $f(x_0) = x$ et $f(y_0) = y$. On a $f(\lambda x_0 + y_0) = \lambda f(x_0) + f(y_0) = \lambda x + y$ donc $\lambda x + y \in \text{Im}(f)$.

$\text{Im}(f)$ est donc bien un sous-espace vectoriel de F . □

Exemples 7.2.11. 1. On considère l'application linéaire

$$f : \mathbb{R}_2[x] \rightarrow \mathbb{R}$$

$$P \mapsto P(0)$$

et montrons que $\text{Im}(f) = \mathbb{R}$.

On a déjà $\text{Im}(f) \subset \mathbb{R}$ par définition de l'image. Réciproquement, soit $c \in \mathbb{R}$. On a alors, pour c vu comme le polynôme constant, $f(c) = c$ donc $c \in \text{Im}(f)$ donc $\mathbb{R} \subset \text{Im}(f)$, soit $\text{Im}(f) = \mathbb{R}$.

2. On considère l'application linéaire

$$g : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathcal{M}_2(\mathbb{R})$$

$$M \mapsto M - {}^t M$$

On observe que pour tout $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$, on a

$$g(M) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 0 & b-c \\ c-b & 0 \end{pmatrix} = (b-c) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

On en déduit donc que $Im(g) = Vect \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$.

De manière similaire au lien entre le noyau et l'injectivité, nous avons un lien entre l'image et la surjectivité :

Propriété 7.2.12. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a que f est surjective si et seulement si $Im(f) = F$.

Exemples 7.2.13. On considère les deux applications linéaires de l'exemple précédent :

1. On a $Im(f) = \mathbb{R}$ donc f est surjective.
2. On a $Im(g) \neq \mathcal{M}_2(\mathbb{R})$ donc g n'est pas surjective.

7.3 Utilisation des dimensions et des bases

Jusqu'à présent nous n'avons pas fait intervenir la notion de dimension des espaces vectoriels que nous considérons, cependant les dimensions des espaces vectoriels que l'on considère vont nous permettre de nouveaux outils. L'utilisation des bases va nous permettre également d'obtenir de nouveaux résultats.

7.3.1 Caractérisation d'une application linéaire

Le résultat suivant nous permet de caractériser les applications linéaires à partir d'une base :

Théorème 7.3.1. Soit E et F deux espaces vectoriels et soit $f \in \mathcal{L}(E, F)$. Soit $n = \dim(E)$ et soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E . f est entièrement déterminée par la donnée des vecteurs $\{f(e_1), \dots, f(e_n)\}$. Autrement dit pour toute famille $\{u_1, \dots, u_n\} \in F^n$ il existe une unique application linéaire $f \in \mathcal{L}(E, F)$ telle que pour tout $i \in \llbracket 1, n \rrbracket$,

$$f(e_i) = u_i.$$

Remarque 7.3.2. L'image des vecteurs d'une base suffit donc à définir intégralement une application linéaire, il arrive donc qu'une application linéaire soit définie par ce moyen là dans nos exercices.

Exemple 7.3.3. Soit $\{e_1, e_2, e_3\}$ la base canonique de \mathbb{R}^3 et on considère $f \in \mathcal{L}(E)$ définie par $f(e_1) = e_1 + e_2 + e_3$, $f(e_2) = e_2 + 2e_3$ et $f(e_3) = 3e_1 - e_2$. On a alors que pour tout $(x, y, z) \in \mathbb{R}^3$:

$$\begin{aligned} f((x, y, z)) &= xf(e_1) + yf(e_2) + zf(e_3) \\ &= x(e_1 + e_2 + e_3) + y(e_2 + 2e_3) + z(3e_1 - e_2) \\ &= (x + 3z)e_1 + (x + y - z)e_2 + (x + 2y)e_3 \\ &= (x + 3z, x + y - z, x + 2y) \end{aligned}$$

On a donc que f est

$$\begin{aligned} f : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ (x, y, z) &\mapsto (x + 3z, x + y - z, x + 2y) \end{aligned}$$

Corollaire 7.3.4. Soit E et F deux espaces vectoriels et soit $(f, g) \in \mathcal{L}(E, F)^2$. Soit $\{e_1, \dots, e_n\}$ une base de E tel que pour tout $i \in \llbracket 1, n \rrbracket$, $f(e_i) = g(e_i)$ alors $f = g$.

Exemple 7.3.5. Soit E et F deux espaces vectoriels et soit $f \in \mathcal{L}(E, F)$. Soit $\{e_1, \dots, e_n\}$ une base de E tel que pour tout $i \in \llbracket 1, n \rrbracket$, $f(e_i) = 0_F$ alors f est l'application nulle.

7.3.2 Notion de rang

Définition 7.3.6. Soit E et F deux espaces vectoriels et soit $f \in \mathcal{L}(E, F)$. On appelle rang de f la dimension de $Im(f)$. On note ce dernier $rg(f)$.

Exemples 7.3.7. On considère de nouveau les applications linéaires

$$\begin{aligned} f : \mathbb{R}_2[x] &\rightarrow \mathbb{R} \\ P &\mapsto P(0) \end{aligned}$$

et

$$\begin{aligned} g : \mathcal{M}_2(\mathbb{R}) &\rightarrow \mathcal{M}_2(\mathbb{R}) \\ M &\mapsto M^{-t} M \end{aligned}$$

On avait vu que $Im(f) = \mathbb{R}_0[x]$ et $Im(g) = Vect \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$ donc on a $rg(f) = rg(g) = 1$.

En pratique pour déterminer le rang, une des deux méthodes que l'on utilise est de déterminer explicitement l'image de f à l'aide du résultat suivant :

Propriété 7.3.8. Soit E et F deux espaces vectoriels et soit $f \in \mathcal{L}(E, F)$. Soit $n = \dim(E)$ et soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E . On a alors que

$$\text{Im}(f) = \text{Vect}(f(e_1), \dots, f(e_n)).$$

Remarque 7.3.9. Tout les outils vu dans le chapitre 3 peuvent donc nous servir à déterminer le rang, soit en simplifiant l'écriture de $\text{Im}(f)$ comme sous-espace vectoriel engendrée soit en déterminant le rang de la famille de vecteurs $\{f(e_1), \dots, f(e_n)\}$.

Exemples 7.3.10. Retrouvons les images des applications linéaires précédentes à l'aide de ce résultat :

1. On considère $\{1, x, x^2\}$ la base canonique de $\mathbb{R}_2[x]$. On a donc

$$\text{Im}(f) = \text{Vect}(f(1), f(x), f(x^2)) = \text{Vect}(1, 0, 0) = \text{Vect}(1) = \mathbb{R}_0[x].$$

2. On considère $\{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$ la base canonique de $\mathcal{M}_2(\mathbb{R})$. On a donc

$$\begin{aligned} \text{Im}(g) &= \text{Vect}(g(E_{1,1}), g(E_{1,2}), g(E_{2,1}), g(E_{2,2})) \\ &= \text{Vect} \left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right) \\ &= \text{Vect} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) \\ &= \text{Vect} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) \quad \text{car} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned}$$

Remarque 7.3.11. Comme $\text{Im}(f)$ est une sous espace vectoriel de F , on a

$$\text{rg}(f) \leq \dim(F).$$

De plus comme $\text{Im}(f) = \text{Vect}(f(e_1), \dots, f(e_n))$ on a également

$$\text{rg}(f) \leq \dim(E).$$

La notion de rang nous permet également de redéfinir le critère de surjectivité :

Propriété 7.3.12. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a que f est surjective si et seulement si $\text{rg}(f) = \dim(F)$.

7.3.3 Théorème du rang

Le principal théorème concernant le rang est le suivant :

Théorème 7.3.13 (Théorème du rang). Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$. On a alors

$$\dim(\ker(f)) + \text{rg}(f) = \dim(E).$$

Ce théorème a de nombreuses conséquences :

Corollaire 7.3.14. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{L}(E, F)$.

- Si $\dim(F) < \dim(E)$ alors f n'est pas injective ;
- Si $\dim(E) < \dim(F)$ alors f n'est pas surjective ;
- Si $\dim(E) \neq \dim(F)$ alors f n'est pas un isomorphisme.

Corollaire 7.3.15. Soit E et F deux espaces vectoriels réels de même dimension et soit $f \in \mathcal{L}(E, F)$. On a alors que f est injective si et seulement si f est surjective si et seulement si f est un isomorphisme.

Remarque 7.3.16. Le cas d'un endomorphisme correspond donc à ce cas là. Ainsi pour un endomorphisme l'injectivité, la surjectivité et le fait d'être un automorphisme sont équivalents.

Exemples 7.3.17. On considère de nouveau les deux applications linéaires du dernier exemple.

1. Pour $f : \mathbb{R}_2[x] \rightarrow \mathbb{R}$, on a $\dim(\mathbb{R}_2[x]) = 3$ et $\dim(\mathbb{R}) = 1$ donc

$$P \mapsto P(0)$$

comme $1 < 3$, on a que f n'est pas injective et donc n'est pas un isomorphisme.

De plus comme $\text{rg}(f) = 1 = \dim(\mathbb{R})$, f est surjective.

Enfin par le théorème du rang $\dim(\ker(f)) = \dim(\mathbb{R}_2[x]) - \text{rg}(f) = 3 - 1 = 2$ et l'on pourrait montrer que $\ker(f) = \text{Vect}(x, x^2)$.

2. Pour $g : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathcal{M}_2(\mathbb{R})$ il s'agit d'un endomorphisme donc on

$$M \mapsto M - {}^t M$$

ne peut pas conclure quant à la non injectivité ou la non surjectivité par argument de dimension des espaces vectoriels de départ et d'arrivée.

En revanche, on a vu que $\text{rg}(g) = 1 \neq \dim(\mathcal{M}_2(\mathbb{R}))$ donc f n'est pas

surjective et l'égalité des dimensions nous donne aussi que g n'est pas injective.

On a de plus par le théorème du rang que $\dim(\ker(f)) = \dim(\mathcal{M}_2(\mathbb{R})) - \text{rg}(g) = 4 - 1 = 3$ et l'on pourrait montrer que

$$\ker(f) = \text{Vect}(E_{1,1}, E_{2,2}, E_{1,2} + E_{2,1}).$$

Corollaire 7.3.18. Soit E et F deux espaces vectoriels réels et soit $f \in \mathcal{GL}(E, F)$ alors $\dim(E) = \dim(F)$.

Remarque 7.3.19. Il ne s'agit pas d'une équivalence ! Par exemple pour l'application g utilisée précédemment, on a égalité des dimensions mais g n'est pas un isomorphisme !

7.4 Matrice associée à une application linéaire

Les applications linéaires entre espaces vectoriels de dimensions finies sont en correspondance avec les matrices. Nous allons à présent établir cette équivalence et traduire les résultats associés.

7.4.1 Généralités

Avant de pouvoir faire interagir les applications linéaires et les matrices, nous avons besoin de faire le lien entre les vecteurs et les matrices.

Définition 7.4.1. Soit E un espace vectoriel de dimension n et soit \mathcal{B} une base de E .

- Soit $u \in E$ de coordonnées $(x_1, \dots, x_n) \in \mathbb{R}^n$ dans la base \mathcal{B} . On appelle matrice de u dans la base \mathcal{B} la matrice colonne de $\mathcal{M}_{n,1}(\mathbb{R})$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

et on la note $\text{Mat}_{\mathcal{B}}(u)$.

- Soit $p \in \mathbb{N}^*$ et soit $(u_1, \dots, u_p) \in E^p$. On appelle matrice de la famille $\{u_1, \dots, u_p\}$ dans la base \mathcal{B} la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ dont la j -ème ligne est la colonne $\text{Mat}_{\mathcal{B}}(u_j)$. On note cette dernière $\text{Mat}_{\mathcal{B}}(u_1, \dots, u_p)$.

Exemples 7.4.2. 1. On considère \mathcal{B}_{can} la base canonique de \mathbb{R}^3 et on admet que $\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ est une autre base de \mathbb{R}^3 . On pose $u = (0, 6, -1)$ et $v = (2, 2, 0)$ deux vecteurs de \mathbb{R}^3 . On a

$$Mat_{\mathcal{B}_{can}}(u) = \begin{pmatrix} 0 \\ 6 \\ -1 \end{pmatrix} \quad \text{et} \quad Mat_{\mathcal{B}_{can}}(v) = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$$

On peut trouver par le calcul que

$$u = -(1, 1, 1) + 7(1, 1, 0) - 6(1, 0, 0)$$

et

$$v = 0(1, 1, 1) + 2(1, 1, 0) + 0(0, 0, 0).$$

On a alors

$$Mat_{\mathcal{B}}(u) = \begin{pmatrix} -1 \\ 7 \\ -6 \end{pmatrix} \quad \text{et} \quad Mat_{\mathcal{B}}(v) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

On a ainsi

$$Mat_{\mathcal{B}_{can}}(u, v) = \begin{pmatrix} 0 & 2 \\ 6 & 2 \\ -1 & 0 \end{pmatrix} \quad \text{et} \quad Mat_{\mathcal{B}}(u, v) = \begin{pmatrix} -1 & 0 \\ 7 & 2 \\ -6 & 0 \end{pmatrix}$$

2. On considère \mathcal{B}_{can} la base canonique de $\mathbb{R}_2[x]$ et on admet que $\mathcal{B} = \{1, (x-1), (x-1)^2\}$ est une autre base de $\mathbb{R}_2[x]$. On pose $P = 2(x-1)^2 - 3(x-1) - 4$ un vecteur de $\mathbb{R}_2[x]$. On a

$$P = 2x^2 - 4x + 2 - 3x + 3 - 4 = 2x^2 - 7x + 1.$$

On a donc

$$Mat_{\mathcal{B}_{can}}(P) = \begin{pmatrix} 1 \\ -7 \\ 2 \end{pmatrix} \quad \text{et} \quad Mat_{\mathcal{B}}(P) = \begin{pmatrix} -4 \\ -3 \\ 2 \end{pmatrix}$$

La difficulté pour définir la matrice d'une application linéaire vient du fait que l'on doit considérer deux bases, une pour l'espace vectoriel de départ et une pour l'espace vectoriel d'arrivée.

Définition 7.4.3. Soit E et F deux espaces vectoriels de dimensions respectives n et p . Soit $\mathcal{B}_E = \{e_1, \dots, e_n\}$ une base de E et soit \mathcal{B}_F une base de F . Soit $f \in \mathcal{L}(E, F)$ On appelle matrice de f dans les bases \mathcal{B}_E et \mathcal{B}_F la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$

$$Mat_{\mathcal{B}_F}(f(e_1), \dots, f(e_n)).$$

On note cette dernière $Mat_{\mathcal{B}_E, \mathcal{B}_F}(f)$.

Remarques 7.4.4. 1. Dans le cas d'un endomorphisme de E , il peut arriver que la base associée à E vu comme ensemble de départ et la base associée à E vu comme ensemble d'arrivée soit différentes, on prendra donc garde à cela. En revanche si il s'agit de la même base pour E vu comme ensemble de départ et d'arrivée, on pourra alors noter $Mat_{\mathcal{B}_E}(f)$ au lieu de $Mat_{\mathcal{B}_E, \mathcal{B}_E}(f)$.

2. Dans le cas d'un endomorphisme la matrice associée est toujours une matrice carré !

Méthode 7.4.5. On garde les mêmes notations que dans la définition et pour déterminer la matrice d'une application linéaire f on doit donc commencer par calculer $f(e_j)$ pour tout $j \in \llbracket 1, n \rrbracket$ et on exprime ces vecteurs de F dans la base \mathcal{B}_F . La matrice $Mat_{\mathcal{B}_E, \mathcal{B}_F}(f)$ est donc la matrice où la j -ème colonne est la matrice colonne des coordonnées de $f(e_j)$ trouvées précédemment.

Exemples 7.4.6. 1. On considère l'application linéaire

$$\begin{aligned} f : \mathbb{R}^3 &\rightarrow \mathbb{R}^4 \\ (x, y, z) &\mapsto (x - y, 2y, 0, z) \end{aligned}$$

et on munit les espaces vectoriels de leurs bases canoniques. On a

- $f(1, 0, 0) = (1, 0, 0, 0)$
- $f(0, 1, 0) = (-1, 2, 0, 0)$
- $f(0, 0, 1) = (0, 0, 0, 1)$

On en déduit donc que

$$\text{Mat}_{\mathbb{B}_{\text{can}_{\mathbb{R}^3}}, \mathbb{B}_{\text{can}_{\mathbb{R}^4}}}(f) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. On considère l'application linéaire

$$\begin{aligned} g : \mathbb{R}_2[x] &\rightarrow \mathbb{R}^3 \\ P &\mapsto (P(1), P'(1), P''(1)) \end{aligned}$$

et on munit les espaces vectoriels de leurs bases canoniques. On a

- $g(1) = (1, 0, 0)$
- $g(x) = (1, 1, 0)$
- $g(x^2) = (1, 2, 2)$

On en déduit donc que

$$\text{Mat}_{\mathbb{B}_{\text{can}_{\mathbb{R}_2[x]}}, \mathbb{B}_{\text{can}_{\mathbb{R}^3}}}(g) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Cette écriture entre applications linéaires selon deux bases et matrices est une caractérisation des applications linéaires :

Théorème 7.4.7. Soit E et F deux espaces vectoriels de dimensions respectives n et p . Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . L'application

$$\begin{aligned} \varphi : \mathcal{L}(E, F) &\rightarrow \mathcal{M}_{n,p}(\mathbb{R}) \\ f &\mapsto \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) \end{aligned}$$

est un isomorphisme.

Remarques 7.4.8. Pour \mathcal{B}_E une base fixé de E et \mathcal{B}_F une base fixé de F on a :

1. Une application linéaire a une unique écriture matricielle dans ces bases fixées ;
2. Réciproquement, une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ correspond à une unique application linéaire exprimé dans ces bases de $\mathcal{L}(E, F)$.

Exemple 7.4.9. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$ et considérons qu'il s'agit de la matrice

d'une application linéaire f de $\mathbb{R}_2[x]$ dans $\mathbb{R}_2[x]$ munit de la base canonique. On a donc $f(1) = 1$, $f(x) = 4x + 2$ et $f(x^2) = 6x^2 + 5x + 3$ et donc pour $P = ax^2 + bx + c \in \mathbb{R}_2[x]$ on a :

$$\begin{aligned} f(P) &= af(x^2) + bf(x) + cf(1) \\ &= a(6x^2 + 5x + 3) + b(4x + 2) + c \\ &= 6ax^2 + (5a + 4b)x + 3a + 2b + c \end{aligned}$$

Le caractère d'application linéaire de φ dans ce théorème nous donne également les propriétés suivantes :

Propriétés 7.4.10. Soit E et F deux espaces vectoriels de dimensions respectives n et p . Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . Soit $(f, g) \in \mathcal{L}(E, F)^2$ et soit $\lambda \in \mathbb{R}$. On a

$$\text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f + g) = \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) + \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(g)$$

et

$$\text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(\lambda f) = \lambda \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f).$$

Remarque 7.4.11. Cela signifie que pour toutes combinaisons linéaires d'applications linéaires a pour matrice associée la même combinaison linéaire des matrices associées à chaque applications linéaires.

Exemple 7.4.12. On considère

$$\begin{array}{ccc} f : \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (x, y) & \mapsto & (2y, 4x - y) \end{array} \quad \text{et} \quad \begin{array}{ccc} g : \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + y, x - y) \end{array} .$$

Les matrices de f et g où \mathbb{R}^2 est muni de la base canonique sont

$$\text{Mat}_{\mathcal{B}_{can}}(f) = \begin{pmatrix} 0 & 2 \\ 4 & -1 \end{pmatrix} \quad \text{et} \quad \text{Mat}_{\mathcal{B}_{can}}(g) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

On a donc

$$\text{Mat}_{\mathcal{B}_{can}}(2f - 3g) = \begin{pmatrix} -3 & 1 \\ 5 & 1 \end{pmatrix}.$$

Une dernière conséquence de ce théorème est le lemme suivant :

Lemme 7.4.13. Soit E et F deux espaces vectoriels de dimensions respectives n et p , alors $\dim(\mathcal{L}(E, F)) = np$ et $\dim(\mathcal{L}(E)) = n^2$.

Cette écriture matricielle des applications linéaires nous permet également de redéfinir le rang d'une matrice comme le rang de l'application linéaire associée :

Théorème 7.4.14. Soit E et F deux espaces vectoriels de dimensions respectives n et p et soit $f \in \mathcal{L}(E, F)$. Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . On a alors

$$\text{rg}(f) = \text{rg}(\text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f)).$$

Exemple 7.4.15. On considère de nouveau l'application linéaire $g : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathcal{M}_2(\mathbb{R})$.

$$M \mapsto M - {}^t M$$

Le rang de g est de 1 et en considérant la base canonique $\mathcal{B}_{can} = \{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$ la matrice de g dans cette base est

$$\text{Mat}_{\mathcal{B}_{can}}(g) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

et on observe que cette matrice est bien de rang 1.

7.4.2 Parallélisme dans le calcul

Il a été vu précédemment que la matrice associée à une combinaison linéaire d'applications linéaires et la même combinaison linéaire des matrices associées mais le parallélisme dans le calcul va bien plus loin. Nous avons une correspondance parfaite entre les deux

Théorème 7.4.16. Soit E et F deux espaces vectoriels de dimensions respectives n et p . Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . Soit $f \in \mathcal{L}(E, F)$ et soit $(x, y) \in E \times F$. On a alors

$$f(x) = y \Leftrightarrow \text{Mat}_{\mathcal{B}_F}(f(x)) = \text{Mat}_{\mathcal{B}_F}(y) \Leftrightarrow \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) \text{Mat}_{\mathcal{B}_E}(x) = \text{Mat}_{\mathcal{B}_F}(y).$$

Ce résultat nous permet donc de passer par l'écriture matricielle pour déterminer l'image de nos vecteurs par des applications linéaires et cela quelque soit les bases de E et F que nous considérons !

Exemples 7.4.17. 1. Soit f l'endomorphisme de $\mathbb{R}_2[x]$ dont la matrice dans la

base canonique est $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 0 & -1 & 0 \end{pmatrix}$. Calculons $f(x^2 + 1)$. Cela revient à

calculer

$$A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 0 + 1 \\ 1 + 0 + 3 \\ 0 + 0 + 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}$$

On a donc $f(x^2 + 1) = 4x + 2$.

2. Soit g l'endomorphisme de \mathbb{R}^3 dont la matrice dans la base canonique est

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & -1 & 1 \end{pmatrix}. \text{ Calculons } g(e_1 + e_2 + e_3). \text{ On a}$$

$$B \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 1 + 2 \\ 1 + 0 + 3 \\ 0 - 1 + 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}$$

donc $g(e_1 + e_2 + e_3) = 4(e_1 + e_2)$.

Corollaire 7.4.18. Soit E et F deux espaces vectoriels de dimensions respectives n et p et soit $f \in \mathcal{L}(E, F)$. Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . On a pour $x \in E$:

$$x \in \ker(f) \Leftrightarrow \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) \text{Mat}_{\mathcal{B}_E}(x) = 0_{p,1} \Leftrightarrow \text{Mat}_{\mathcal{B}_E}(x) \in \ker(\text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f)).$$

Remarque 7.4.19. On peut donc également utiliser la forme matricielle pour déterminer un noyau en utilisant la correspondance applications linéaires-matrices et cela quelque soient les bases choisis pour l'écriture matricielle.

Exemple 7.4.20. On considère g de l'exemple précédent et déterminons $\ker(g)$.

Soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. On a

$$\begin{aligned} (x, y, z) \in \ker(g) &\Leftrightarrow BX = 0_{3,1} \\ &\Leftrightarrow \begin{pmatrix} x + y + 2z \\ x + 3z \\ -y + z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\Leftrightarrow \begin{cases} x + y + 2z = 0 \\ x + 3z = 0 \\ -y + z = 0 \end{cases} \Leftrightarrow \begin{cases} 0 = 0 \\ x = -3z \\ y = z \end{cases} \\ &\Leftrightarrow X = z \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \Leftrightarrow X \in \text{Vect} \left(\begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \right) \\ &\Leftrightarrow (x, y, z) \in \text{Vect}((-3, 1, 1)) \end{aligned}$$

On en déduit donc que $\ker(g) = \text{Vect}((-3, 1, 1))$.

Il nous reste à traiter le cas des compositions des applications linéaires dans la correspondances avec les matrices, on a pour cela le théorème suivant :

Théorème 7.4.21. Soit E, F et G trois espaces vectoriels que l'on munit respectivement des bases $\mathcal{B}_E, \mathcal{B}_F$ et \mathcal{B}_G . Soit $f \in \mathcal{L}(E, F)$ et soit $g \in \mathcal{L}(F, G)$, on a alors

$$\text{Mat}_{\mathcal{B}_E, \mathcal{B}_G}(g \circ f) = \text{Mat}_{\mathcal{B}_F, \mathcal{B}_G}(g) \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(f).$$

Exemple 7.4.22. On considère de nouveau les applications linéaires

$$\begin{aligned} f : \mathbb{R}^3 &\rightarrow \mathbb{R}^4 \\ (x, y, z) &\mapsto (x - y, 2y, 0, z) \end{aligned}$$

et

$$\begin{aligned} g : \mathbb{R}^4 &\rightarrow \mathbb{R}^2 \\ (x, y, z, t) &\mapsto (x + y, 3z - 2t) \end{aligned}$$

et on a vu dans un exemple précédent 7.1.9 que pour tout $(x, y, z) \in \mathbb{R}^3$, on a $(g \circ f)(x, y, z) = (x + y, -2z)$.

Dans la base canonique, on a déjà vu que $Mat_{\mathbb{B}_{can_{\mathbb{R}^3}}, \mathbb{B}_{can_{\mathbb{R}^4}}}(f) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et

on a $Mat_{\mathbb{B}_{can_{\mathbb{R}^2}}, \mathbb{B}_{can_{\mathbb{R}^4}}}(g) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -2 \end{pmatrix}$.

On a

$$Mat_{\mathbb{B}_{can_{\mathbb{R}^3}}, \mathbb{B}_{can_{\mathbb{R}^2}}}(g \circ f) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

ce qui correspond bien au calcul fait précédemment.

On en déduit donc dans le cas d'égalité des dimensions :

Corollaire 7.4.23. Soit E et F deux espaces vectoriels de dimension commune n . Soit \mathcal{B}_E une base de E et soit \mathcal{B}_F une base de F . On a alors :

1. Soit $f \in \mathcal{L}(E)$ et soit $k \in \mathbb{N}$, on a

$$Mat_{\mathcal{B}_E}(f^k) = Mat_{\mathcal{B}_E}(f)^k.$$

2. Soit $f \in \mathcal{L}(E, F)$. f est un isomorphisme si et seulement si $Mat_{\mathcal{B}_E, \mathcal{B}_F}(f)$ est inversible. Dans ce cas on a de plus

$$Mat_{\mathcal{B}_E, \mathcal{B}_F}(f^{-1}) = Mat_{\mathcal{B}_E, \mathcal{B}_F}(f)^{-1}.$$

Exemple 7.4.24. On considère de nouveau l'application linéaire de l'exemple

7.1.11 $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. La matrice associée dans la base canonique est $A :=$

$$(x, y, z) \mapsto (0, x, y)$$

$$A := \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \text{ On a par le calcul } A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ et pour tout}$$

$k \geq 3$, $A^k = 0_3$ ce qui correspond bien aux calculs effectués sur f^k .

De plus comme $A^3 = 0_3$, on a que A n'est pas inversible donc f n'est pas un automorphisme.

Remarque 7.4.25. On a besoin pour le calcul d'une puissance d'une matrice associée à un endomorphisme E que la base associée à E soit la même pour E vu comme ensemble de départ que pour E vu comme ensemble d'arrivé.

7.4.3 Notion de changement de base

Un problème que nous rencontrerons de nombreuses fois est la nécessité de changer de base pour avoir des écritures matricielles sur lesquelles on pourrait plus facilement déduire des données sur les applications linéaires que l'on considère. Nous faisons intervenir pour cela la notion de changement de base à l'aide des matrices de passages :

Définition 7.4.26. Soit E un espace vectoriel de dimension n et soit \mathcal{B} et $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ deux bases de E . On appelle matrice de passage de \mathcal{B} à \mathcal{B}' la matrice de la famille \mathcal{B}' dans la base \mathcal{B} . On note cette dernière $P_{\mathcal{B}, \mathcal{B}'}$. Autrement dit, on a

$$P_{\mathcal{B}, \mathcal{B}'} := Mat_{\mathcal{B}}(e'_1, \dots, e'_n).$$

Remarque 7.4.27. La matrice de passage de \mathcal{B} à \mathcal{B}' est la matrice de l'identité de (E, \mathcal{B}') dans (E, \mathcal{B}) , c'est à dire $P_{\mathcal{B}, \mathcal{B}'} = Mat_{\mathcal{B}', \mathcal{B}}(Id_E)$. On prendra garde à l'ordre des bases !

Exemples 7.4.28. 1. On note \mathcal{B}_{can} la base canonique de \mathbb{R}^2 et $\mathcal{B}' = \{(1, 1), (1, -1)\}$. Les deux vecteurs n'étant pas colinéaires, ils forment une famille libre et par un argument de dimension \mathcal{B}' est une base de \mathbb{R}^2 . On a

$$P_{\mathcal{B}_{can}, \mathcal{B}'} = Mat_{\mathcal{B}_{can}}((1, 1), (1, -1)) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

2. On note \mathcal{B}_{can} la base canonique de $\mathbb{R}_2[x]$ et $\mathcal{B}' = \{1, x + 1, (x + 1)^2\}$. Il s'agit d'une famille échelonnée de polynômes, ils forment donc une famille libre et par un argument de dimension \mathcal{B}' est une base de $\mathbb{R}_2[x]$. On a

$$P_{\mathcal{B}_{can}, \mathcal{B}'} = Mat_{\mathcal{B}_{can}}(1, x + 1, (x + 1)^2) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pour faire le chemin inverse, il nous suffit d'utiliser la propriété suivante :

Propriété 7.4.29. Soit E un espace vectoriel de dimension n et soit \mathcal{B} et \mathcal{B}' deux bases de E . On a alors que $P_{\mathcal{B}, \mathcal{B}'}$ est inversible et on a

$$P_{\mathcal{B}, \mathcal{B}'}^{-1} = P_{\mathcal{B}', \mathcal{B}}.$$

Exemples 7.4.30. 1. On considère l'exemple de \mathbb{R}^2 précédent et la matrice de passage de \mathcal{B}' dans \mathcal{B}_{can} est

$$P_{\mathcal{B}', \mathcal{B}_{can}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

2. On considère l'exemple de $\mathbb{R}_2[x]$ précédent et la matrice de passage de \mathcal{B}' dans \mathcal{B}_{can} est

$$P_{\mathcal{B}', \mathcal{B}_{can}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

par un logiciel de calcul.

Nous avons donc les règles de calculs suivantes :

Propriété 7.4.31. Soit E un espace vectoriel munit de trois bases \mathcal{B} , \mathcal{B}' et \mathcal{B}'' . On a alors :

1. $P_{\mathcal{B},\mathcal{B}''} = P_{\mathcal{B},\mathcal{B}'}P_{\mathcal{B}',\mathcal{B}''}$
2. Pour tout $u \in E$, $Mat_{\mathcal{B}}(u) = P_{\mathcal{B},\mathcal{B}'}Mat_{\mathcal{B}'}(u)$.
3. Pour $f \in \mathcal{L}(E)$,

$$Mat_{\mathcal{B}'}(f) = P_{\mathcal{B}',\mathcal{B}}Mat_{\mathcal{B}}(f)P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B},\mathcal{B}'}^{-1}Mat_{\mathcal{B}}(f)P_{\mathcal{B},\mathcal{B}'}$$

Exemple 7.4.32. On considère l'application linéaire f de $\mathbb{R}_2[x]$ dont la matrice

dans la base canonique est $Mat_{\mathcal{B}_{can}}(f) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$ et déterminons la matrice

de f dans la base $\mathcal{B}' = \{1, x + 1, (x + 1)^2\}$. On a donc

$$\begin{aligned} Mat_{\mathcal{B}'}(f) &= P_{\mathcal{B}',\mathcal{B}_{can}}Mat_{\mathcal{B}_{can}}(f)P_{\mathcal{B}_{can},\mathcal{B}'} = P_{\mathcal{B}_{can},\mathcal{B}'}^{-1}Mat_{\mathcal{B}_{can}}(f)P_{\mathcal{B}_{can},\mathcal{B}'} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

Remarque 7.4.33. On observe dans le dernier exemple un type d'égalité matricielle que l'on a déjà très souvent rencontré. L'objet du prochain chapitre d'algèbre nous donnera la méthode pour trouver par nous même les coefficients diagonaux ainsi qu'une base qui nous permettra d'avoir les bonnes matrices de passages pour arriver à ce résultat.

Les matrices représentant le même endomorphisme mais dans des bases différentes sont donc liées par les matrices de passages. De telles matrices portent un nom :

Définition 7.4.34. Soit $n \in \mathbb{N}^*$. Soit $(A, B) \in \mathcal{M}_n(\mathbb{R})^2$. On dit que A et B sont des matrices semblables si il existe une matrice $P \in \mathcal{GL}_n(\mathbb{R})$ telle que

$$B = P^{-1}AP.$$

Remarque 7.4.35. Pour P une matrice de passage d'une base \mathcal{B} à une base \mathcal{B}' , cela veut donc dire que A et B représentent le même endomorphisme respectivement dans les bases \mathcal{B} et \mathcal{B}' .

Exemples 7.4.36. 1. Par l'exemple précédent, les matrices $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$

sont semblables.

2. Les matrices I_2 et $\begin{pmatrix} 4 & 8 \\ 1 & -6 \end{pmatrix}$ ne sont pas semblables. En effet si elles sont semblables alors il existe une matrice inversible P tel que

$$\begin{pmatrix} 4 & 8 \\ 1 & -6 \end{pmatrix} = P^{-1}I_2P = P^{-1}P = I_2$$

ce qui est absurde.

Chapitre 8

Révisions : Variables aléatoires discrètes

L'objectif des probabilités de l'année dernière ont été de traduire les expériences aléatoires à issues discrètes en les modélisant mathématiquement et en développant le vocabulaire adapté, principalement à travers les variables aléatoires discrètes. Nous nous proposons donc dans ce chapitre de revoir ces notions pour nous aider dans les futurs chapitres de probabilités.

En effet, cette année pour les probabilités finis et discrètes l'objectif va être de considérer plusieurs variables aléatoires discrètes en même temps, nous devons donc maîtriser les variables aléatoires discrètes individuellement parfaitement.

8.1 Généralités

Définitions 8.1.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

- Une variable aléatoire réelle est une application $X : \Omega \mapsto \mathbb{R}$ vérifiant pour tout $x \in \mathbb{R}$ que l'ensemble $[X \leq x] := \{\omega \in \Omega \mid X(\omega) \leq x\}$ est un élément de \mathcal{A} .
- On note $X(\Omega)$ l'ensemble des issues possibles que peut prendre la variable aléatoire X . On appelle ce dernier univers associée à X (ou le support de X parfois).

Remarques 8.1.2. 1. $X(\Omega)$ est donc l'image de l'application X .

2. Il n'est pas demandé de savoir prouver qu'une variable aléatoire en est une. En revanche il est fondamentale de savoir déterminer $X(\Omega)$.

Une manière de caractériser une variable aléatoire est d'utiliser la notion de fonction de répartition.

Définition 8.1.3. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et soit X une variable aléatoire. On appelle fonction de répartition associée à la variable aléatoire X la fonction

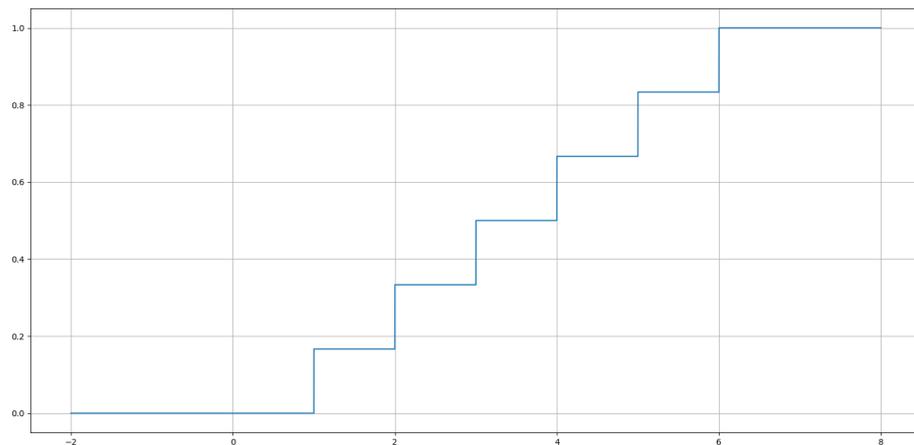
$$F_X : \mathbb{R} \rightarrow [0, 1]$$

$$x \mapsto \mathbb{P}([X \leq x])$$

Exemple 8.1.4. Soit X la variable aléatoire correspondant au résultat d'un dé à six faces équilibré. On a F_X qui s'exprime comme suit pour tout $x \in \mathbb{R}$:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ \frac{1}{6} & \text{si } 1 \leq x < 2 \\ \frac{2}{6} & \text{si } 2 \leq x < 3 \\ \frac{3}{6} & \text{si } 3 \leq x < 4 \\ \frac{4}{6} & \text{si } 4 \leq x < 5 \\ \frac{5}{6} & \text{si } 5 \leq x < 6 \\ 1 & \text{si } 6 \leq x \end{cases}$$

Graphiquement cela correspond à



Une fonction de répartition à les propriétés suivante :

Propriétés 8.1.5. En conservant les notations de la définition précédente, on a que F_X vérifie :

- F_X est croissante ;
- F_X est continue à droite en tout point de \mathbb{R} ;
- On a $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

On se concentre ici principalement sur le cas des variables aléatoires discrètes :

Définition 8.1.6. Soit X une variable aléatoire réelle. On dit que X est une variable aléatoire discrète si $X(\Omega)$ est fini ou dénombrable.

Définition 8.1.7. On appelle loi de la variable aléatoire discrète la donnée de $\mathbb{P}(X = k)$ pour tout $k \in X(\Omega)$.

Exemples 8.1.8. 1. On note X la variable aléatoire telle que $X(\Omega) = \{-1, 0, 1, 2\}$ et $\mathbb{P}(X = -1) = \frac{3}{10}$, $\mathbb{P}(X = 0) = \frac{1}{5}$, $\mathbb{P}(X = 1) = \frac{2}{5}$ et $\mathbb{P}(X = 2) = \frac{1}{10}$.
 X est une variable aléatoire discrète fini.

2. On note Y la variable aléatoire telle que $Y(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(Y = k) = \frac{1}{k(k+1)}$.
 Y est une variable aléatoire discrète infini.

Pour une variable aléatoire discrète nous avons facilement un système complet d'évènements :

Propriété 8.1.9. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et soit X une variable aléatoire sur Ω tel que $X(\Omega)$ soit discret, c'est-à-dire tel qu'il existe I une partie (fini ou dénombrable) de \mathbb{N} telle que $X(\Omega) = \{x_i \mid i \in I\}$.

On a alors que $([X = x_i])_{i \in I}$ est un système complet d'évènements que l'on appelle système complet d'évènements associée à la variable aléatoire X .

Une conséquence directe de ce résultat est :

Propriété 8.1.10. En conservant les notations précédentes, on a

$$\sum_{i \in I} \mathbb{P}(X = x_i) = 1.$$

Remarque 8.1.11. Ce dernier résultat peut être une bonne manière de se donner une idée de si nos calculs sont juste en calculant la valeur de la série.

La loi d'une variable aléatoire discrète est caractérisé par sa fonction de répartition :

Théorème 8.1.12. Soit X une variable aléatoire discrète où $X(\Omega) = \mathbb{N}$ de fonction de répartition F_X . On a alors pour tout $k \in X(\omega) = \mathbb{N}$:

$$\mathbb{P}(X = k) = F_X(k) - F_X(k - 1).$$

Remarque 8.1.13. Ce résultat est adaptable suivant les situations et l'écriture de $X(\Omega)$. On prendra donc soin de bien réécrire les évènements $X = k$ selon des évènements du type $X \leq a$ pour déterminer la relation entre $\mathbb{P}(X = k)$ et F_X pour chaque exemple.

Une manière classique de travailler sur les variables aléatoires et en particulier sur les variables aléatoires discrètes de calculer "directement" sur la variable aléatoire :

Théorème 8.1.14. Soit X une variable aléatoire discrète et soit g une fonction sur $X(\Omega)$. On a alors que $g(X)$ est une variable aléatoire discrète. On a de plus pour tout $j \in g(X)(\Omega)$,

$$\mathbb{P}(g(X) = j) = \sum_{k \in X(\Omega) \text{ tel que } g(k)=j} \mathbb{P}(X = k).$$

Exemples 8.1.15. On considère les exemples précédents :

1. X^2 est une variable aléatoire discrète d'univers associée $\{0, 1, 4\}$ et on a $\mathbb{P}(X^2 = 0) = \mathbb{P}(X = 0) = \frac{1}{5}$, $\mathbb{P}(X^2 = 1) = \mathbb{P}(X = -1) + \mathbb{P}(X = 1) = \frac{3}{10} + \frac{2}{5} = \frac{7}{10}$ et $\mathbb{P}(X^2 = 4) = \mathbb{P}(X = 2) = \frac{1}{10}$.
2. $Z = Y - 1$ est une variable aléatoire discrète d'univers associée \mathbb{N} et pour tout $k \in \mathbb{N}$, $\mathbb{P}(Z = k) = \mathbb{P}(Y = k + 1) = \frac{1}{(k+1)(k+2)}$.

8.2 Notion de moments

8.2.1 Espérance

Lorsque l'on joue à un jeu, une donnée intéressante est la gain moyen obtenue à chaque partie. Pour cela on introduit la notion d'espérance :

Définitions 8.2.1. Soit X une variable aléatoire discrète.

- Si $X(\Omega)$ est fini, X admet une espérance et l'on a

$$\mathbb{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i).$$

- Si $X(\Omega)$ est infini, X admet une espérance si la série $\sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i)$ est absolument convergente et l'on a dans ce cas

$$\mathbb{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i).$$

Remarque 8.2.2. Dans le cas d'un univers associée à la variable aléatoire fini, l'espérance existe toujours car nous sommes en présence d'une somme fini. En revanche dans le cas où l'univers associée à la variable aléatoire est infini, la question de l'espérance revient à travailler sur des séries, on doit donc commencer par montrer la convergence.

Exemples 8.2.3. On considère de nouveau les exemples précédents.

1. X admet une espérance car l'univers associé est fini et l'on a :

$$\begin{aligned} \mathbb{E}(X) &= (-1)\mathbb{P}(X = -1) + 0 \cdot \mathbb{P}(X = 0) + 1\mathbb{P}(X = 1) + 2\mathbb{P}(X = 2) \\ &= -\frac{3}{10} + 0 + \frac{2}{5} + 2\frac{1}{10} = \frac{-3 + 4 + 2}{10} = \frac{3}{10} \end{aligned}$$

2. Y admet une espérance si et seulement si la série $\sum_{k \geq 1} \frac{k}{k(k+1)} = \sum_{k \geq 1} \frac{1}{k+1}$ est absolument convergente, or on reconnaît la série harmonique qui est divergente donc Y n'admet pas d'espérance.

Propriété 8.2.4. Soit X et Y deux variables aléatoires admettant chacune une espérance et soit $(a, b) \in \mathbb{R}^2$. On a alors :

- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;
- $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$;
- Si $X(\omega) \subset \mathbb{R}_+$ (*c'est-à-dire, X est à valeurs positives*), alors $\mathbb{E}(X) \geq 0$.

Remarque 8.2.5. En particulier, on a donc $\mathbb{E}(1) = 1$.

Une manière générale de calculer des espérances est d'utiliser le théorème de transfert :

Théorème 8.2.6 (Théorème de transfert 1). Soit X une variable aléatoire discrète d'univers associée fini $X(\Omega)$ et soit g une application sur $X(\Omega)$. La variable aléatoire $g(X)$ admet une espérance et l'on a

$$\mathbb{E}(g(X)) = \sum_{x_i \in X(\Omega)} g(x_i) \mathbb{P}(X = x_i).$$

Théorème 8.2.7 (Théorème de transfert 2). Soit X une variable aléatoire discrète d'univers associée infini $X(\Omega)$ et soit g une application sur $X(\Omega)$. La variable aléatoire $g(X)$ admet une espérance si et seulement si la série

$$\sum_{x_i \in X(\Omega)} g(x_i) \mathbb{P}(X = x_i)$$

converge absolument et dans ce cas on a

$$\mathbb{E}(g(X)) = \sum_{x_i \in X(\Omega)} g(x_i) \mathbb{P}(X = x_i).$$

Exemple 8.2.8. On considère de nouveau la variable aléatoire X des exemples précédents et l'on a donc comme les sommes sont finis que l'espérance existe. Par la formule du transfert on a

$$\begin{aligned} \mathbb{E}(X^2) &= (-1)^2 \mathbb{P}(X = -1) + 0^2 \cdot \mathbb{P}(X = 0) + 1^2 \mathbb{P}(X = 1) + 2^2 \mathbb{P}(X = 2) \\ &= \frac{3}{10} + 0 + \frac{2}{5} + 4 \frac{1}{10} = \frac{3 + 4 + 4}{10} = \frac{11}{10} \end{aligned}$$

Une application classique à ce théorème est de l'utiliser avec une fonction puissance :

Définition 8.2.9. Soit X une variable aléatoire discrète et soit $r \in \mathbb{N}^*$. On dit que X admet un moment d'ordre r si la variable aléatoire X^r admet une espérance, autrement dit si la quantité $\mathbb{E}(X^r)$ existe. Dans ce cas on appelle $\mathbb{E}(X^r)$ le moment d'ordre r de X .

Propriété 8.2.10. Une variable aléatoire admettant un moment d'ordre r admet également un moment pour tout $k \in \llbracket 1, r \rrbracket$.

Remarque 8.2.11. En particulier, admettre un moment d'ordre 2 implique donc d'admettre un moment d'ordre 1 ce qui correspond exactement à admettre une espérance.

8.2.2 Variance

Nous allons nous intéresser en particulier à présent aux variables aléatoires admettant des moments d'ordres 2. Cela nous permettra de redéfinir la variance qui nous donnera des informations sur la disparité des valeurs associée à la variable aléatoire par rapport à son espérance.

Définition 8.2.12. Soit X une variable aléatoire discrète admettant un moment d'ordre 2.

1. On appelle variance de X la quantité $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$;
2. On appelle écart type la quantité $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Remarque 8.2.13. La variance est toujours positive car la variable aléatoire $(X - \mathbb{E}(X))^2$ est à valeur positive, d'où la possibilité de considérer l'écart-type.

Pour calculer efficacement une variance on utilise la formule suivante :

Propriété 8.2.14 (Formule de Koenig-Huygens). Soit X une variable aléatoire admettant un moment d'ordre 2, alors on a

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Démonstration. On a

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2\mathbb{E}(1) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned}$$

□

Remarque 8.2.15. Une variable aléatoire admettant une variance admet une espérance. On a donc que montrer que X admet un moment d'ordre 2 implique X admet une espérance et une variance.

Exemple 8.2.16. On considère de nouveau l'exemple précédent et l'on a donc

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{11}{10} - \left(\frac{3}{10}\right)^2 = \frac{110}{100} - \frac{9}{100} = \frac{101}{100}.$$

De plus on a $\sigma(X) = \sqrt{\frac{101}{100}} = \frac{\sqrt{101}}{10}$.

Le carré dans la définition de la variance n'est pas linéaire contrairement à l'espérance, en revanche on a un résultat dans le cas où les variables aléatoires sont indépendantes :

Définition 8.2.17. Soit X et Y deux variables aléatoires discrètes. On dit que X et Y sont indépendantes lorsque pour tout $x \in X(\Omega)$ et pour tout $y \in Y(\Omega)$, on a

$$\mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Propriété 8.2.18. Soit X et Y deux variables aléatoires indépendantes, alors

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

Remarques 8.2.19. • L'indépendance est indispensable, sinon un nouveau coefficient appelé covariance intervient et le calcul est plus complexe. Nous verrons cela au prochain chapitre.

- Ce résultat peut s'étendre à une somme de variables aléatoires deux à deux indépendantes.

Propriété 8.2.20. Soit X une variable aléatoire admettant un moment d'ordre 2 et soit $(a, b) \in \mathbb{R}^2$. On a alors

$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X).$$

Il est parfois utile de "normaliser" les variables aléatoires :

Définition 8.2.21. Soit X une variable aléatoire discrète :

- Si X admet une espérance et que $\mathbb{E}(X) = 0$, on dit que X est centrée ;
- Si X admet une espérance, la variable aléatoire $X - \mathbb{E}(X)$ est la variable centrée associée à X ;
- Si X admet une variance et que $\mathbb{V}(X) = 1$, on dit que X est réduite ;
- Si X admet une variance non nulle, la variable aléatoire $\frac{X - \mathbb{E}(X)}{\sigma(X)}$ est la variable aléatoire centrée réduite associée à X . On la note en général X^* .

Exemple 8.2.22. Si on considère une dernière fois la variable aléatoire X défini dans les exemples précédents $X^* = \frac{X - \frac{3}{10}}{\frac{\sqrt{101}}{10}} = \frac{10X - 3}{\sqrt{101}}$ est centrée réduite.

En effet on a

$$\mathbb{E}(X^*) = \frac{10\mathbb{E}(X) - 3}{\sqrt{101}} = \frac{10 \cdot \frac{3}{10} - 3}{\sqrt{101}} = \frac{0}{\sqrt{101}} = 0$$

et

$$\mathbb{V}(X^*) = \left(\frac{10}{\sqrt{101}} \right)^2 \mathbb{V}(X) = \frac{100}{101} \frac{101}{100} = 1.$$

8.3 Lois usuelles

On utilisera régulièrement des lois usuelles et toutes les propriétés que l'on a revu dans ce chapitre pour étudier nos problèmes. La connaissance des lois usuelles suivantes ainsi que les manières de retrouver les résultats associées est donc indispensable.

8.3.1 Lois discrètes finis

8.3.1.1 Lois certaine et uniforme

Les lois associées aux situations d'équiprobabilités sont parmi les lois classiques que nous avons déjà étudié plusieurs fois sans les nommer. Formalisons cela à présent.

Commençant par le cas où une expérience à une unique issue possible.

Définition 8.3.1. Soit X une variable aléatoire et soit $a \in \mathbb{R}$ où $X(\Omega) = \{a\}$. On a donc $\mathbb{P}(X = a) = 1$. On dit que X suit une loi certaine de paramètre a .

Propriété 8.3.2. Soit X une variable aléatoire suivant une loi certaine de paramètre a . X admet une espérance et une variance et l'on a

$$\mathbb{E}(X) = a \text{ et } \mathbb{V}(X) = 0.$$

Démonstration. Comme $X(\Omega)$ est fini, X admet une espérance et une variance. On a par définition $\mathbb{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i) = a \mathbb{P}(X = a) = a$ et $\mathbb{V}(X) = \mathbb{E}((X - a)^2) = \mathbb{E}(0) = 0$. □

Le fait que la variance soit nulle est une caractérisation des lois certaines :

Propriété 8.3.3. Soit X une variable aléatoire discrète. X suit une loi certaine si et seulement si $\mathbb{V}(X) = 0$.

Remarque 8.3.4. Cette caractérisation nous implique donc que dès que $X(\Omega)$ n'est pas singleton, la variance et l'écart type (sous réserve d'existence) sont non nuls, donc en particulier on peut déterminer la variable aléatoire centrée réduite associée.

En général une variable aléatoire ne prend pas qu'une seule valeur et dans le cas d'équiprobabilité on a la loi suivante :

Définition 8.3.5. Soit X une variable aléatoire discrète et soit $(a, b) \in \mathbb{Z}^2$ (avec $a \leq b$). On dit que X suit une loi uniforme sur $[[a, b]]$ si pour tout $k \in [[a, b]]$,

$$\mathbb{P}(X = k) = \frac{1}{b - a + 1}.$$

On note alors $X \hookrightarrow \mathcal{U}([a, b])$.

Remarque 8.3.6. Lorsque $a = b$, on retombe sur le cas d'une loi certaine de paramètre a .

Exemple 8.3.7. On utilise couramment le cas où $a = 1$ et $b = n$ pour $n \in \mathbb{N}^*$. On a donc dans ce cas pour tout $k \in \llbracket 1, n \rrbracket$, $\mathbb{P}(X = k) = \frac{1}{n}$.

Propriété 8.3.8. Soit X une variable aléatoire suivant une uniforme sur $\llbracket a, b \rrbracket$. X admet une espérance et une variance et l'on a

$$\mathbb{E}(X) = \frac{a+b}{2} \text{ et } \mathbb{V}(X) = \frac{(b-a+1)^2-1}{12}.$$

Dans le cas particulier où $\llbracket a, b \rrbracket = \llbracket 1, n \rrbracket$, on a :

Propriété 8.3.9. Soit X une variable aléatoire suivant une uniforme sur $\llbracket 1, n \rrbracket$. X admet une espérance et une variance et l'on a

$$\mathbb{E}(X) = \frac{n+1}{2} \text{ et } \mathbb{V}(X) = \frac{n^2-1}{12}.$$

Démonstration. Considérons le cas où $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$. Comme $X(\Omega)$ est fini, X admet une espérance et une variance. On a donc :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^n k \mathbb{P}(X = k) \\ &= \sum_{k=1}^n \frac{k}{n} = \frac{1}{n} \sum_{k=1}^n k \\ &= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2} \end{aligned}$$

De même, on a :

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=1}^n k^2 \mathbb{P}(X = k) \\ &= \sum_{k=1}^n \frac{k^2}{n} = \frac{1}{n} \sum_{k=1}^n k^2 \\ &= \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6} \end{aligned}$$

Donc par la formule de Koenig-Huygens, on a

$$\mathbb{V}(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$\begin{aligned}
&= (n+1) \left(\frac{2n+1}{6} - \frac{n+1}{4} \right) = (n+1) \left(\frac{4n+2}{12} - \frac{3n+3}{12} \right) \\
&= \frac{n+1}{12} (4n+2-3n-3) = \frac{n+1}{6} (n-1) = \frac{n^2-1}{12}
\end{aligned}$$

□

Remarque 8.3.10. Pour démontrer le cas général, on peut utiliser les propriétés sur l'espérance et la variance en posant $Y = X + a - 1$ où $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$ et où $n = b - a + 1$. On a ainsi que $Y \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)$. On a ainsi :

$$\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}(X) + a - 1 = \frac{n+1}{2} + a - 1 \\
&= \frac{b-a+1+1+2a-2}{2} = \frac{b+a}{2}
\end{aligned}$$

et

$$\mathbb{V}(Y) = 1^2 \mathbb{V}(X) = \frac{n^2-1}{12} = \frac{(b-a+1)^2-1}{12}$$

8.3.1.2 Lois de Bernoulli et Binomial

Un type d'expérience que nous étudions régulièrement sont les expériences dont l'issue est un succès ou un échec ou des répétitions de plusieurs fois la même épreuve aboutissant à un succès ou à un échec.

Définition 8.3.11. Soit X une variable aléatoire discrète tel que $X(\Omega) = \{0, 1\}$ et soit $p \in]0, 1[$. On dit que X suit une loi de Bernoulli de paramètre p lorsque $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$. On note $X \hookrightarrow \mathcal{B}(p)$.

Remarques 8.3.12. 1. En général l'évènement $X = 1$ correspond au succès à l'épreuve et $X = 0$ à l'échec.

2. Si $p = 0$ ou $p = 1$, X correspond à une loi certaine.

Exemples 8.3.13. 1. Le lancer d'une pièce truquée ayant une probabilité p de tomber sur pile correspond à une loi de Bernoulli de paramètre p lorsque l'on considère un seul lancer et qu'un succès revient à avoir pile.

2. Le lancer d'un dé équilibré à 6 faces où l'on réussit l'épreuve lorsque l'on obtient 5 ou 6 correspond à une épreuve de Bernoulli avec $p = \frac{1}{3}$.

Propriété 8.3.14. Soit X une loi de Bernoulli de paramètre $p \in]0, 1[$. X admet une espérance et une variance. On a $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p(1-p)$.

Démonstration. Comme $X(\Omega)$ est fini X admet une espérance et une variance. On a donc :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=0}^1 k\mathbb{P}(X = k) \\ &= 0\mathbb{P}(X = 0) + 1\mathbb{P}(X = 1) \\ &= 0 \cdot (1 - p) + p = p\end{aligned}$$

De plus, par la formule du transfert :

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=0}^1 k^2\mathbb{P}(X = k) \\ &= 0^2\mathbb{P}(X = 0) + 1^2\mathbb{P}(X = 1) \\ &= 0 \cdot (1 - p) + p = p\end{aligned}$$

Par la formule de Koenig-Hyugens, on a

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1 - p).$$

□

En général on ne considère pas qu'une seule épreuve mais une succession d'épreuves indépendantes de Bernoulli et l'on compte le nombre de succès. Cela nous mène à la loi binomial.

Définition 8.3.15. Soit X une variable aléatoire discrète, $n \in \mathbb{N}$ et $p \in]0, 1[$. On dit que X suit une loi binomiale de paramètre n, p lorsque $X(\Omega) = \llbracket 0, n \rrbracket$ et que pour tout $k \in \llbracket 0, n \rrbracket$,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

On note alors $X \hookrightarrow \mathcal{B}(n, p)$.

Remarques 8.3.16.

1. Le cas $n = 1$ correspond donc à une loi de Bernoulli.
2. Une variable aléatoire suivant une loi binomiale de paramètre n, p correspond à la somme de n variables aléatoires indépendantes suivant chacune une loi de Bernoulli de paramètre p .

Exemple 8.3.17. Le nombre de 6 sur 10 lancers successif du même dé équilibré à 6 faces correspond à une variable aléatoire suivant une loi binomial $\mathcal{B}(10, \frac{1}{6})$.

Propriété 8.3.18. Soit X une variable aléatoire suivant une loi binomiale de paramètre n, p . X admet donc une espérance et une variance et l'on a

$$\mathbb{E}(X) = np \text{ et } \mathbb{V}(X) = np(1 - p).$$

Démonstration. Comme $X(\Omega)$ est fini, X admet une espérance et une variance. On pose pour tout $i \in \llbracket 1, n \rrbracket$ Y_i une variable aléatoire discrète suivant une loi de Bernoulli de paramètre p . Elles sont automatiquement toutes indépendantes deux à deux. On a $X = \sum_{i=1}^n Y_i$ donc par linéarité de l'espérance, on a :

$$\mathbb{E}(X) = \sum_{i=1}^n \mathbb{E}(Y_i) = \sum_{i=1}^n p = p(n - 1 + 1) = np.$$

De même, comme les variables aléatoires $(Y_i)_{i \in \llbracket 1, n \rrbracket}$ sont deux à deux indépendantes, on a

$$\mathbb{V}(X) = \sum_{i=1}^n \mathbb{V}(Y_i) = \sum_{i=1}^n p(1 - p) = p(1 - p)(n - 1 + 1) = np(1 - p).$$

□

8.3.2 Lois discrètes infinis

8.3.2.1 Loi géométrique

Une expérience classique est le cas d'une première réussite d'une expérience aléatoire qui se répète comme par exemple dans le cas d'un lancer de pièce jusqu'à avoir un pile ou du lancer de dé jusqu'à avoir un 6.

Cela correspond à la loi géométrique :

Définition 8.3.19. Soit X une variable aléatoire réelle et soit $p \in]0, 1[$. On dit que X suit une loi géométrique de paramètre p si X correspond au rang du premier succès d'une expérience de Bernoulli sans mémoire.

On note une telle loi $X \hookrightarrow \mathcal{G}(p)$.

On utilise en général la forme suivante pour une loi géométrique :

Théorème 8.3.20. Soit X une variable aléatoire réelle et soit $p \in]0, 1[$. On a que $X \hookrightarrow \mathcal{G}(p)$ si et seulement si $X(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$,

$$\mathbb{P}(X = k) = p(1 - p)^{k-1}.$$

- Exemples 8.3.21.** 1. Deux joueurs jouent à pile ou face avec une pièce équilibré chacun leur tour et le premier à avoir un pile remporte le jeu. On note X le nombre de tour nécessaire pour mettre fin à la partie. X représentant donc le premier pile sur une succession de lancer d'une même pièce de manière indépendante, soit le rang du premier succès à l'épreuve de Bernoulli "Avoir un pile" répété sans mémoire et indépendamment des épreuves précédentes. X suit donc une loi géométrique de paramètre $\frac{1}{2}$, c'est-à-dire $X \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$. On a pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = \frac{1}{2} \frac{1}{2^{k-1}} = \frac{1}{2^k}$.
2. Soit Y le nombre de lancers nécessaires pour obtenir un 6 sur un dé équilibré. Le dé étant équilibré et les lancers étant indépendant, Y suit une loi géométrique de paramètre $\frac{1}{6}$, soit $Y \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$. On a pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(Y = k) = \frac{1}{6} \frac{5^{k-1}}{6^{k-1}} = \frac{5^{k-1}}{6^k}$.

Pour les moments, la loi géométrique possèdent les résultats suivants :

Théorème 8.3.22. Soit $p \in]0, 1[$ et soit X une variable aléatoire suivant une loi géométrique de paramètre p , c'est-à-dire $X \hookrightarrow \mathcal{G}(p)$. X admet une espérance et une variance et l'on a

$$E(X) = \frac{1}{p} \text{ et } \mathbb{V}(X) = \frac{1-p}{p^2}.$$

Démonstration. Comme $X(\Omega) = \mathbb{N}^*$ et que la série $\sum_{k \geq 1} k^2 p(1-p)^{k-1}$ est convergente en tant que combinaison linéaire de séries numériques convergentes (puisque $1-p \in]0, 1[$) car

$$k^2 p(1-p)^{k-1} = p(1-p)k(k-1)(1-p)^{k-2} + pk(1-p)^{k-1}.$$

X admet donc un moment d'ordre 2 donc X admet une espérance et une variance. On a donc :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} \\ &= \frac{p}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p} \end{aligned}$$

De plus, on a par la formule de transfert

$$\mathbb{E}(X^2) = \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{\infty} k(k-1)p(1-p)^{k-1} + \sum_{k=1}^{\infty} kp(1-p)^{k-1}$$

$$\begin{aligned}
&= p(1-p) \sum_{k=1}^{\infty} (k(k-1)(1-p)^{k-2}) + \frac{1}{p} \\
&= p(1-p) \sum_{k=2}^{\infty} (k(k-1)(1-p)^{k-2}) + \frac{1}{p} \text{ par nullité du premier terme} \\
&= p(1-p) \frac{2}{(1-(1-p))^3} + \frac{1}{p} \\
&= \frac{2p(1-p)}{p^3} + \frac{1}{p} = \frac{2-2p}{p^2} + \frac{p}{p^2} = \frac{2-p}{p^2}
\end{aligned}$$

On a donc par la formule de Koenig-Huygens

$$\mathbb{V}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

□

8.3.2.2 Loi de Poisson

Il ne nous reste que la loi de Poisson à redéfinir dans les lois usuelles :

Définition 8.3.23. Soit X une variable aléatoire réelle et soit $\lambda \in \mathbb{R}_+^*$. On dit que X suit une loi de Poisson de paramètre λ si $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}.$$

On note une telle loi $X \hookrightarrow \mathcal{P}(\lambda)$.

Remarque 8.3.24. La modélisation d'une loi de Poisson est plus complexe, elle sera donc en principe toujours annoncé dans les exercices.

Au niveau des moments, la loi de Poisson possèdent les résultats suivants :

Théorème 8.3.25. Soit $\lambda \in \mathbb{R}_+^*$ et soit X une variable aléatoire suivant une loi de Poisson de paramètre λ , c'est-à-dire $X \hookrightarrow \mathcal{P}(\lambda)$. X admet une espérance et une variance et l'on a

$$E(X) = \lambda \text{ et } \mathbb{V}(X) = \lambda.$$

Démonstration. Comme $X(\Omega) = \mathbb{N}$ et que la série $\sum_{k \geq 0} k^2 e^{-\lambda} \frac{\lambda^k}{k!}$ est convergente en tant que combinaison linéaire de séries numériques convergentes car pour tout $k \geq 2$,

$$k^2 e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} k(k-1) \frac{\lambda^k}{k!} + e^{-\lambda} k \frac{\lambda^k}{k!}$$

$$= \lambda^2 e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!}.$$

X admet donc un moment d'ordre 2 donc X admet une espérance et une variance. On a donc :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{\infty} e^{-\lambda} k \frac{\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} e^{-\lambda} k \frac{\lambda^k}{k!} \text{ par nullité du premier terme} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

De plus, on a par la formule de transfert

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=0}^{\infty} e^{-\lambda} k^2 \frac{\lambda^k}{k!} \\ &= \sum_{l=0}^{\infty} e^{-\lambda} k(k-1) \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} e^{-\lambda} k \frac{\lambda^k}{k!} \\ &= \sum_{k=2}^{\infty} \left(e^{-\lambda} k(k-1) \frac{\lambda^k}{k!} \right) + \lambda \text{ par nullité des premiers termes} \\ &= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \left(\frac{\lambda^{k-2}}{(k-2)!} \right) + \lambda \\ &= \lambda^2 e^{-\lambda} \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{k!} \right) + \lambda \\ &= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda \end{aligned}$$

On a donc par la formule de Koenig-Huygens

$$\mathbb{V}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

□

Remarque 8.3.26. Nous avons déjà vu en TD cette année que la loi "limite" de lois binomiales $X_n \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$ au sens où l'on peut montrer que pour tout $k \in \mathbb{N}$, $\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ est une loi de Poisson $X \hookrightarrow \mathcal{P}(\lambda)$. Cette notion de limite de lois fera l'objet d'un chapitre du prochain semestre.

Chapitre 9

Couples de variables aléatoires

L'étude d'une expérience aléatoire ne se traduit pas toujours, mathématiquement parlant, sous la forme d'une unique variable aléatoire mais plusieurs phénomènes aléatoires peuvent se produire en même temps et nécessite donc l'usage de plusieurs variables aléatoires.

On se propose dans ce chapitre d'étudier la cas de deux variables aléatoires discrètes et de voir comment deux telles variables peuvent dépendre et varier l'une par rapport à l'autre.

9.1 Loi d'un couple de variables aléatoires discrètes

9.1.1 Définitions

Définition 9.1.1. Soit X et Y deux variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. L'application

$$\begin{aligned} (X, Y) &: \Omega \rightarrow \mathbb{R}^2 \\ \omega &\mapsto (X(\omega), Y(\omega)) \end{aligned}$$

est appelée couple de variables aléatoires discrètes.

Définition 9.1.2. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle loi du couple (X, Y) ou loi conjointe de X et Y la donnée de

$$\mathbb{P}([X = x] \cap [Y = y])$$

pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

On pourra parfois noter $\mathbb{P}(X = x, Y = y)$ au lieu de $\mathbb{P}([X = x] \cap [Y = y])$.

Exemples 9.1.3. 1. On pose X et Y deux variables aléatoires représentant respectivement le résultat d'un dé équilibré à 4 faces et le résultat d'un dé équilibré à 6 faces. On a donc $X \hookrightarrow \mathcal{U}(\llbracket 1, 4 \rrbracket)$ et $Y \hookrightarrow \mathcal{U}(\llbracket 1, 6 \rrbracket)$. On a pour tout $(i, j) \in \llbracket 1, 4 \rrbracket \times \llbracket 1, 6 \rrbracket$,

$$\mathbb{P}([X = i] \cap [Y = j]) = \frac{1}{24}.$$

2. On considère à présent que X représente le rang du premier pile d'une série de lancers d'une pièce équilibrée et que Y représente le rang du premier pile d'une série de lancers d'une pièce truquée qui à une probabilité $\frac{1}{3}$ d'obtenir pile. On a donc $X \hookrightarrow \mathcal{G}(\frac{1}{2})$ et $Y \hookrightarrow \mathcal{G}(\frac{1}{3})$. On a ainsi pour tout $(i, j) \in (\mathbb{N}^*)^2$,

$$\mathbb{P}([X = i] \cap [Y = j]) = \frac{1}{2^i} \frac{1}{3} \left(\frac{2}{3}\right)^{j-1} = \frac{1}{2^{i-j+1} 3^j}.$$

9.1.2 Aparté sur les séries doubles et un système complet d'évènements

En considérant une variable aléatoire discrète on est confronté à calculer des sommes et des séries. Avec des couples de variables aléatoires on se retrouve confronté à des sommes sur deux indices finis ou infinis. On ne considérera pas les problèmes théoriques intrinsèques des sommations doubles en se contentant d'étendre naïvement les résultats des séries et en considérant les points suivants :

Soit I et J deux ensembles de \mathbb{N} , une famille de réels $(a_{i,j})_{(i,j) \in I \times J}$ et considérons $\sum_{(i,j) \in I \times J} a_{i,j}$.

- Dans le cas où I et J sont finis, $\sum_{(i,j) \in I \times J} a_{i,j}$ correspond à une somme finie et on a

$$\sum_{(i,j) \in I \times J} a_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} a_{i,j} \right) = \sum_{j \in J} \left(\sum_{i \in I} a_{i,j} \right).$$

- Dans le cas où I ou J est infini, définissons la notion de convergence absolue de la série double $\sum_{(i,j) \in I \times J} a_{i,j}$.

Si la série $\sum_{j \in J} a_{i,j}$ est absolument convergente pour tout $i \in I$ et si la série

$\sum_{i \in I} \left(\sum_{j \in J} a_{i,j} \right)$ est absolument convergente on dit alors que la série double

$\sum_{(i,j) \in I \times J} a_{i,j}$ est absolument convergente et on a

$$\sum_{(i,j) \in I \times J} a_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} a_{i,j} \right) = \sum_{j \in J} \left(\sum_{i \in I} a_{i,j} \right).$$

Le réel $\sum_{i \in I} \left(\sum_{j \in J} a_{i,j} \right)$ est appelé somme double de la série double.

Remarque 9.1.4. Il est bien évidemment possible de montrer que la série $\sum_{i \in I} a_{i,j}$ est absolument convergente pour tout $j \in J$ et que la série $\sum_{j \in J} \left(\sum_{i \in I} a_{i,j} \right)$ est absolument convergente pour établir la convergence absolue de la série double.

Remarque 9.1.5. Ici aussi, la convergence absolue implique la convergence de la série double. Comme nous considérerons des séries doubles de probabilités, il n'y aura pas de différence entre les deux dans nos études.

Exemples 9.1.6. 1. Calculons $\sum_{(i,j) \in \llbracket 0, n \rrbracket \times \llbracket 0, m \rrbracket} ij$. Comme il s'agit d'une somme fini, nous n'avons aucun problème de convergence à vérifier et l'on a

$$\begin{aligned} \sum_{(i,j) \in \llbracket 0, n \rrbracket \times \llbracket 0, m \rrbracket} ij &= \sum_{i=0}^n \left(\sum_{j=0}^m ij \right) = \sum_{i=0}^n \left(i \sum_{j=0}^m j \right) \\ &= \sum_{i=0}^n \left(i \frac{m(m+1)}{2} \right) = \frac{m(m+1)}{2} \sum_{i=0}^n i \\ &= \frac{m(m+1)}{2} \frac{n(n+1)}{2} = \frac{nm(n+1)(m+1)}{4}. \end{aligned}$$

2. Considérons $\sum_{(i,j) \in \mathbb{N}^2} \frac{1}{2^{i+j}}$. Pour tout $i \in \mathbb{N}$, on a que $\sum_{j \geq 0} \frac{1}{2^{i+j}}$ est absolument convergente en tant que série à terme positif et série géométrique de raison $\frac{1}{2} \in]-1, 1[$. On a

$$\sum_{j=0}^{\infty} \frac{1}{2^{i+j}} = \frac{1}{2^i} \sum_{j=0}^{\infty} \frac{1}{2^j} = \frac{1}{2^i} \frac{1}{1 - \frac{1}{2}} = 2 \frac{1}{2^i}.$$

La série $\sum_{i \geq 0} 2 \frac{1}{2^i}$ est absolument convergente pour les mêmes raisons que précédemment donc la série double $\sum_{(i,j) \in \mathbb{N}^2} \frac{1}{2^{i+j}}$ est absolument convergente

et l'on a

$$\sum_{(i,j) \in \mathbb{N}^2} \frac{1}{2^{i+j}} = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} \frac{1}{2^{i+j}} \right) = \sum_{i=0}^{\infty} 2 \frac{1}{2^i} = 2 \frac{1}{1 - \frac{1}{2}} = 4.$$

On utilisera cette théorie des séries doubles principalement dans le contexte de ce système complet d'évènements :

Propriété 9.1.7. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. La famille

$$([X = x] \cap [Y = y])_{(x,y) \in X(\Omega) \times Y(\Omega)}$$

est un système complet d'évènements.

Remarque 9.1.8. Dans le cas où $X(\Omega)$ et $Y(\Omega)$ sont finis il n'y a aucun problème de convergence, sinon on a que la série double

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \mathbb{P}([X = x] \cap [Y = y])$$

est convergente. Dans tout les cas la somme vaut 1.

Exemple 9.1.9. Considérons de nouveau l'exemple où $X \hookrightarrow \mathcal{G}(\frac{1}{2})$ et $Y \hookrightarrow \mathcal{G}(\frac{1}{3})$ et on a

$$\begin{aligned} \sum_{(i,j) \in (\mathbb{N}^*)^2} \mathbb{P}([X = i] \cap [Y = j]) &= \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} \frac{1}{2^{i-j+1} 3^j} \right) = \frac{1}{2} \sum_{i=1}^{\infty} \left(\frac{1}{2^i} \sum_{j=1}^{\infty} \frac{2^j}{3^j} \right) \\ &= \frac{1}{2} \sum_{i=1}^{\infty} \left(\frac{1}{2^i} \left(\frac{1}{1 - \frac{2}{3}} - 1 \right) \right) \\ &= \frac{1}{2} \sum_{i=1}^{\infty} \frac{1}{2^i} 2 = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{1 - \frac{1}{2}} - 1 = 2 - 1 = 1. \end{aligned}$$

9.1.3 Lois marginales et loi conditionnelles

Rappelons la définition des probabilités conditionnelles par le prisme des couples de variables aléatoires :

Définition 9.1.10. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

- Soit $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, on appelle loi conditionnelle de X sachant $[Y = y]$ la donnée pour tout $x \in X(\Omega)$ du réel

$$\mathbb{P}_{[Y=y]}(X = x) = \frac{\mathbb{P}([X = x] \cap [Y = y])}{\mathbb{P}(Y = y)};$$

- Soit $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, on appelle loi conditionnelle de Y sachant $[X = x]$ la donnée pour tout $y \in Y(\Omega)$ du réel

$$\mathbb{P}_{[X=x]}(Y = y) = \frac{\mathbb{P}([X = x] \cap [Y = y])}{\mathbb{P}(X = x)}.$$

- On dit que X et Y sont indépendantes pour la probabilité \mathbb{P} si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a

$$\mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Exemples 9.1.11. Tout les exemples précédents ont représentés des cas où X et Y sont indépendant.

On a besoin pour utiliser cette définition de connaître les valeurs de $\mathbb{P}(X = x)$ et de $\mathbb{P}(Y = y)$. Il s'agit des lois marginales :

Définition 9.1.12. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1. On appelle première loi marginale du couple (X, Y) la loi de X ;
2. On appelle deuxième loi marginale du couple (X, Y) la loi de Y .

Pour calculer ces lois, en général on utilise la formule des probabilités totales :

Propriété 9.1.13. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1. Soit $x \in X(\Omega)$, on a

$$\mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}([X = x] \cap [Y = y])$$

Si pour tout $y \in Y(\Omega)$, $\mathbb{P}(Y = y) \neq 0$, on a de plus

$$\mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(Y = y)\mathbb{P}_{[Y=y]}(X = x).$$

2. Soit $y \in Y(\Omega)$, on a

$$\mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [Y = y])$$

Si pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x) \neq 0$, on a de plus

$$\mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \mathbb{P}_{[X=x]}(Y = y).$$

Remarque 9.1.14. Dans le cas où $X(\Omega)$ ou $Y(\Omega)$ sont infinis, cette propriété impose la convergence des séries que l'on considère.

Exemples 9.1.15. 1. On lance deux dés équilibrés à 4 faces et on note X le plus petit résultat et Y le plus grand. La loi de probabilités est donnée par le tableau suivant :

$x \in X(\Omega)$ \ $y \in Y(\Omega)$	1	2	3	4
1	$\frac{1}{16}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{16}$	0	0
3	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	0
4	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$

On en déduit les lois marginales :

k	1	2	3	4
$\mathbb{P}(X = k)$	$\frac{7}{16}$	$\frac{5}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
$\mathbb{P}(Y = k)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$

On remarque de plus que

$$\mathbb{P}(X = 2)\mathbb{P}(Y = 1) = \frac{15}{16^2} \neq 0 = \mathbb{P}([X = 2] \cap [Y = 1])$$

donc X et Y ne sont pas indépendants.

2. Soit (X, Y) un couple de variables aléatoires discrètes où $X(\Omega) = \mathbb{N}^*$ et $Y(\Omega) = \mathbb{N}$. On a pour tout $(i, j) \in X(\Omega) \times Y(\Omega)$

$$\mathbb{P}([X = i] \cap [Y = j]) = \frac{e^{-i} i^j}{2^i j!}.$$

- Déterminons la première loi marginale du couple. On a pour $i \in \mathbb{N}^*$ fixé, par utilisation des probabilités totales avec le système complet d'évènements $(Y = j)_{j \in \mathbb{N}}$:

$$\begin{aligned} \mathbb{P}(X = i) &= \sum_{j=0}^{\infty} \mathbb{P}([X = i] \cap [Y = j]) = \sum_{j=0}^{\infty} \frac{e^{-i} i^j}{2^i j!} \\ &= \frac{e^{-i}}{2^i} \sum_{j=0}^{\infty} \frac{i^j}{j!} = \frac{e^{-i}}{2^i} e^i = \frac{1}{2^i} \end{aligned}$$

On en déduit donc que X suit une loi géométrique de paramètre $\frac{1}{2}$.

- Calculons la loi conditionnelle de Y sachant $X = i$. Soit $j \in \mathbb{N}$, on a

$$\mathbb{P}_{X=i}(Y = j) = \frac{\mathbb{P}([X = i] \cap [Y = j])}{\mathbb{P}(X = i)} = \frac{e^{-i} i^j}{j!}$$

On reconnaît donc que sachant $X = i$, Y suit une loi de Poisson de paramètre i .

9.2 Variable aléatoire du type $Z = g(X, Y)$

9.2.1 Définition

Il est courant qu'une variable aléatoire soit en réalité définie par deux autres variables aléatoires :

"Définition" 9.2.1. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit g une fonction de $X(\Omega) \times Y(\Omega) \rightarrow \mathbb{R}$. On note $Z = g(X, Y)$ l'application

$$\begin{aligned} Z : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto g(X(\omega), Y(\omega)) \end{aligned}$$

est une variable aléatoire.

Remarques 9.2.2. On a

- $Z(\Omega) = \{g(X(\omega), Y(\omega)) \mid \omega \in \Omega\} \subset \{g(x, y) \mid (x, y) \in X(\Omega) \times Y(\Omega)\}$. Cette inclusion n'est pas toujours une égalité car l'expérience aléatoire peut empêcher certain cas de se produire.
- La loi de Z est pour tout $z \in Z(\Omega)$

$$\mathbb{P}(Z = z) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega) \mid z=g(x,y)} \mathbb{P}([X = x] \cap [Y = y]).$$

- Dans la loi précédente, si $X(\Omega)$ ou $Y(\Omega)$ est infini, le calcul correspond à une série, éventuellement double, dont la convergence est assurée par le fait que Z est une variable aléatoire.

Exemples 9.2.3. • Soit X et Y deux variables aléatoires indépendantes suivant la même loi uniforme $\mathcal{U}([1, 3])$ et soit $Z = X - Y$. On a alors $Z(\Omega) = \{-2, -1, 0, 1, 2\}$. On a ainsi

$$\begin{aligned} \mathbb{P}(Z = 0) &= \mathbb{P}([X = 1] \cap [Y = 1]) + \mathbb{P}([X = 2] \cap [Y = 2]) + \mathbb{P}([X = 3] \cap [Y = 3]) \\ &= \frac{1}{3} \frac{1}{3} + \frac{1}{3} \frac{1}{3} + \frac{1}{3} \frac{1}{3} = \frac{1}{3} \\ \mathbb{P}(Z = 1) &= \mathbb{P}([X = 2] \cap [Y = 1]) + \mathbb{P}([X = 3] \cap [Y = 2]) \\ &= \frac{1}{3} \frac{1}{3} + \frac{1}{3} \frac{1}{3} = \frac{2}{9} \\ \mathbb{P}(Z = 2) &= \mathbb{P}([X = 3] \cap [Y = 1]) = \frac{1}{3} \frac{1}{3} = \frac{1}{9} \end{aligned}$$

Par des calculs similaires on trouve également $\mathbb{P}(Z = -2) = \frac{1}{9}$ et $\mathbb{P}(Z = -1) = \frac{2}{9}$

- Soit (X, Y) un couple de variables aléatoires discrètes où $X(\Omega) = Y(\Omega) = \mathbb{N}$ dont la loi conjointe est pour tout $(i, j) \in \mathbb{N}^2$,

$$\mathbb{P}([X = i] \cap [Y = j]) = \frac{i+j}{e^{2i+j} i! j!}.$$

On pose $Z = 2^{X+Y}$ et on a par exemple

$$\begin{aligned} \mathbb{P}(Z = 4) &= \mathbb{P}([X = 2] \cap [Y = 0]) + \mathbb{P}([X = 1] \cap [Y = 1]) + \mathbb{P}([X = 0] \cap [Y = 2]) \\ &= \frac{2}{e^{2^2 2! 0!}} + \frac{2}{e^{2^2 1! 1!}} + \frac{2}{e^{2^2 0! 2!}} \\ &= \frac{1}{4e} + \frac{1}{2e} + \frac{1}{4e} = \frac{1}{e} \end{aligned}$$

9.2.2 Espérance et théorème du transfert

Pour une loi du type $Z = g(X, Y)$, il existe une version du théorème de transfert nous permettant de calculer l'espérance (sous réserve d'existence) :

Théorème 9.2.4 (Théorème du transfert). *Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit g une fonction de $X(\Omega) \times Y(\Omega) \rightarrow \mathbb{R}$. On note $Z = g(X, Y)$. Si la série*

$$\sum_{x \in X(\Omega), y \in Y(\omega)} g(x, y) \mathbb{P}([X = x] \cap [Y = y])$$

est absolument convergente alors Z possède une espérance et l'on a

$$\mathbb{E}(Z) = \sum_{x \in X(\Omega), y \in Y(\omega)} g(x, y) \mathbb{P}([X = x] \cap [Y = y]).$$

Remarque 9.2.5. Dans le cas où $X(\Omega)$ et $Y(\Omega)$ sont finis, la convergence absolue est automatique car on ne considère que des sommes finis. Dans le cas où au moins l'un des deux est infini nous sommes dans le cas d'une série (ou de séries doubles si les deux univers associées sont infinis).

Exemples 9.2.6. Considérons les deux exemples précédents :

1. Pour $Z = X - Y$ où $X \hookrightarrow \mathcal{U}(\llbracket 1, 3 \rrbracket)$ et $Y \hookrightarrow \mathcal{U}(\llbracket 1, 3 \rrbracket)$, on a que $Z(\Omega)$ est fini donc Z admet bien une espérance et par formule du transfert on a :

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{i=1}^3 \sum_{j=1}^3 (i - j) \mathbb{P}([X = i] \cap [Y = j]) \\ &= \sum_{i=1}^3 \sum_{j=1}^3 (i - j) \frac{1}{9} \\ &= \frac{1}{9} \left(\sum_{i=1}^3 \sum_{j=1}^3 i - \sum_{i=1}^3 \sum_{j=1}^3 j \right) \\ &= \frac{1}{9} \left(\sum_{i=1}^3 3i - \sum_{j=1}^3 3j \right) \\ &= \frac{1}{9} (3 \cdot 6 - 3 \cdot 6) = \frac{18 - 18}{9} = 0. \end{aligned}$$

2. Pour $Z = 2^{X+Y}$ où la loi conjointe du couple (X, Y) est $\mathbb{P}([X = i] \cap [Y = j]) = \frac{i+j}{e^{2^{i+j}} i! j!}$. Z admet une espérance si la série double

$$\sum_{(i,j) \in \mathbb{N}^2} 2^{i+j} \mathbb{P}([X = i] \cap [Y = j]) = \sum_{(i,j) \in \mathbb{N}^2} \frac{i+j}{e^{i!} j!}$$

est absolument convergente.

Soit $i \in \mathbb{N}$ fixé. Soit $N \in \mathbb{N}$ fixé et l'on a

$$\begin{aligned} \sum_{j=0}^N \frac{i+j}{ei!j!} &= \frac{1}{ei!} \sum_{j=0}^N \left(\frac{i}{j!} + \frac{j}{j!} \right) \\ &= \frac{1}{ei!} \left(i \sum_{j=0}^N \frac{1}{j!} + \sum_{j=1}^N \frac{j}{j!} \right) \\ &= \frac{1}{ei!} \left(i \sum_{j=0}^N \frac{1}{j!} + \sum_{j=0}^{N-1} \frac{1}{j!} \right) \end{aligned}$$

On reconnaît deux sommes partielles de séries exponentielles de termes positifs donc la série $\sum_{j \geq 0} \frac{i+j}{ei!j!}$ est absolument convergente et l'on a

$$\sum_{j=0}^{\infty} \frac{i+j}{ei!j!} = \frac{1}{ei!} (ie + e) = \frac{i+1}{i!}.$$

On a de plus

$$\begin{aligned} \sum_{i=0}^N \frac{i+1}{i!} &= \sum_{i=0}^N \left(\frac{1}{i!} + \frac{i}{i!} \right) \\ &= \sum_{i=0}^N \frac{1}{i!} + \sum_{i=1}^N \frac{i}{i!} \\ &= \sum_{i=0}^N \frac{1}{i!} + \sum_{i=0}^{N-1} \frac{1}{i!} \end{aligned}$$

On reconnaît deux sommes partielles de séries exponentielles de termes positifs donc la série $\sum_{i \geq 0} \frac{i+1}{i!}$ est absolument convergente donc la série double est absolument convergente donc par le théorème de transfert Z admet une espérance et l'on a

$$\mathbb{E}(Z) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} 2^{i+j} \mathbb{P}([X=i] \cap [Y=j]) = \sum_{i=0}^{\infty} \frac{i+1}{i!} = 2e.$$

9.2.3 Lois classiques

Considérons à présent quelques variables aléatoires classiques construites à partir d'un couple de variables aléatoires (X, Y) .

9.2.3.1 Somme et sommes de lois usuelles

Définition 9.2.7. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. $X + Y$ est une variable aléatoire dont la loi est donnée par pour tout $z \in (X + Y)(\Omega)$:

$$\mathbb{P}(X + Y = z) = \sum_{x \in X(\Omega) \mid z-x \in Y(\Omega)} \mathbb{P}([X = x] \cap [Y = z - x]).$$

Remarques 9.2.8. 1. La définition est symétrique donc on a aussi

$$\mathbb{P}(X + Y = z) = \sum_{y \in Y(\Omega) \mid z-y \in X(\Omega)} \mathbb{P}([X = z - y] \cap [Y = y]).$$

2. Il fait savoir être capable de retrouver la définition sur chaque cas à l'aide de la formule des probabilités totales !

Exemple 9.2.9. On considère la somme Z de deux dés équilibrés à 4 faces dont on note le résultat du premier X et le résultat du second Y . On a donc $X \hookrightarrow \mathcal{U}(\llbracket 1, 4 \rrbracket)$, $Y \hookrightarrow \mathcal{U}(\llbracket 1, 4 \rrbracket)$ et $Z = X + Y$.

On observe que $Z(\Omega) = \llbracket 2, 8 \rrbracket$ et l'on a pour tout $k \in Z(\Omega)$, par utilisation du système complet d'évènements $(Y = j)_{j \in \llbracket 1, 4 \rrbracket}$:

$$\begin{aligned} \mathbb{P}(Z = k) &= \sum_{j=1}^4 \mathbb{P}([Z = k] \cap [Y = j]) = \sum_{j=1}^4 \mathbb{P}([X + Y = k] \cap [Y = j]) \\ &= \sum_{j=1}^4 \mathbb{P}([X + j = k] \cap [Y = j]) = \sum_{j=1}^4 \mathbb{P}([X = k - j] \cap [Y = j]) \\ &= \mathbb{P}([X = k - 1] \cap [Y = 1]) + \mathbb{P}([X = k - 2] \cap [Y = 2]) \\ &\quad + \mathbb{P}([X = k - 3] \cap [Y = 3]) + \mathbb{P}([X = k - 4] \cap [Y = 4]) \end{aligned}$$

On en déduit la loi suivante de Z :

k	2	3	4	5	6	7	8
$\mathbb{P}(Z = k)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

Remarque 9.2.10. La formule du transfert nous permet également de retrouver, sous réserve d'existence la linéarité de l'espérance. On a également une nouvelle manière de calculer $\mathbb{E}(X + Y)$ qui peut parfois être plus simple que de calculer $\mathbb{E}(X)$ et $\mathbb{E}(Y)$.

Exemple 9.2.11. Si on calcule $\mathbb{E}(Z)$ de l'exemple précédent à partir de la loi déterminée, on obtient

$$\begin{aligned}\mathbb{E}(Z) &= \frac{2.1 + 3.2 + 4.3 + 5.4 + 6.3 + 7.2 + 8.1}{16} \\ &= \frac{2 + 6 + 12 + 20 + 18 + 14 + 8}{16} \\ &= \frac{80}{16} = 5.\end{aligned}$$

On a également $\mathbb{E}(X) = \mathbb{E}(Y) = \frac{1+4}{2} = 2,5$ et donc on retrouve bien la linéarité par

$$\mathbb{E}(Z) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Certains lois usuelles sont stables par additions :

Propriété 9.2.12. Soit $p \in]0, 1[$ et soit $(n, m) \in (\mathbb{N}^*)^2$. Soit $X \hookrightarrow \mathcal{B}(n, p)$ et soit $Y \hookrightarrow \mathcal{B}(m, p)$ telles que X et Y soient indépendantes. On a alors que

$$X + Y \hookrightarrow \mathcal{B}(n + m, p).$$

Propriété 9.2.13. Soit $(\lambda, \mu) \in (\mathbb{R}_+^*)^2$. Soit $X \hookrightarrow \mathcal{P}(\lambda)$ et soit $Y \hookrightarrow \mathcal{P}(\mu)$ telles que X et Y soient indépendantes. On a alors que

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu).$$

Remarque 9.2.14. Dans les deux cas, si on considère les espérances, on retrouve bien la linéarité de cette dernière.

Remarque 9.2.15. Il est possible d'itérer ce processus pour considérer la somme de n variables aléatoires mutuellement indépendantes qui suivent une loi binomiale ou une loi de Poisson. On s'attardera sur ce point dans un prochain chapitre.

9.2.3.2 Loi du produit

Définition 9.2.16. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. XY est une variable aléatoire dont la loi est donnée par pour tout $z \in (XY)(\Omega)$:

$$\mathbb{P}(XY = z) = \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [xY = z]).$$

Remarques 9.2.17. 1. La définition est symétrique donc on a aussi

$$\mathbb{P}(XY = z) = \sum_{y \in Y(\Omega)} \mathbb{P}([yX = z] \cap [Y = y]).$$

2. Il fait savoir être capable de retrouver la définition sur chaque cas à l'aide de la formule des probabilités totales !

Exemple 9.2.18. On considère la produit Z de deux dés équilibrés à 4 faces dont on note le résultat du premier X et le résultat du second Y . On a donc $X \hookrightarrow \mathcal{U}(\llbracket 1, 4 \rrbracket), Y \hookrightarrow \mathcal{U}(\llbracket 1, 4 \rrbracket)$ et $Z = XY$.

On observe que $Z(\Omega) = \{1, 2, 3, 4, 6, 8, 9, 12, 16\}$ et l'on a pour tout $k \in Z(\Omega)$, par utilisation du système complet d'évènements $(Y = j)_{j \in \llbracket 1, 4 \rrbracket}$:

$$\begin{aligned} \mathbb{P}(Z = k) &= \sum_{j=1}^4 \mathbb{P}([Z = k] \cap [Y = j]) = \sum_{j=1}^4 \mathbb{P}([XY = k] \cap [Y = j]) \\ &= \sum_{j=1}^4 \mathbb{P}([Xj = k] \cap [Y = j]) = \sum_{j=1}^4 \mathbb{P}\left(\left[X = \frac{k}{j}\right] \cap [Y = j]\right) \\ &= \mathbb{P}([X = k] \cap [Y = 1]) + \mathbb{P}\left(\left[X = \frac{k}{2}\right] \cap [Y = 2]\right) \\ &\quad + \mathbb{P}\left(\left[X = \frac{k}{3}\right] \cap [Y = 3]\right) + \mathbb{P}\left(\left[X = \frac{k}{4}\right] \cap [Y = 4]\right) \end{aligned}$$

On en déduit la loi suivante de Z :

k	1	2	3	4	6	8	9	12	16
$\mathbb{P}(Z = k)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

Dans ce cas précis, si X et Y sont indépendantes, les espérances de XY , X et Y sont liées :

Propriété 9.2.19. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que X et Y soit indépendantes et qui admettent chacune une espérance. On a alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Exemple 9.2.20. Dans notre exemple précédent nos variables aléatoires X et Y sont indépendantes et donc $\mathbb{E}(Z) = \mathbb{E}(X)\mathbb{E}(Y) = \frac{25}{4}$. Si on fait le calcul à partir de la définition, on a

$$\begin{aligned}\mathbb{E}(Z) &= \frac{1.1 + 2.2 + 3.2 + 4.3 + 6.2 + 8.2 + 9.1 + 12.2 + 16.1}{16} \\ &= \frac{1 + 4 + 6 + 12 + 12 + 16 + 9 + 24 + 16}{16} \\ &= \frac{100}{16} = \frac{25}{4}\end{aligned}$$

9.2.3.3 Loi du minimum, loi du maximum

Les deux dernières lois classiques que nous considérons sont les loi du maximum et du minimum :

Définition 9.2.21. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que X et Y sont indépendantes. On appelle loi du minimum $\min(X, Y)$ et loi du maximum $\max(X, Y)$.

Il y a pas de formule global pour la loi de telles lois mais la méthode est sensiblement toujours la même :

Méthode 9.2.22. Pour déterminer la loi de $Z = \max(X, Y)$, on utilise en général la fonction de répartition en suivant les points suivants :

1. Pour tout $z \in Z(\Omega)$, on justifie que $[Z \leq z] = [X \leq z] \cap [Y \leq z]$
2. L'indépendance de X et Y nous permet de calculer F_Z .
3. On utilise $\mathbb{P}(Z = k) = F_Z(k) - F_Z(k - 1)$.

Exemple 9.2.23. Soit $n \in \mathbb{N}^*$. Soit X et Y deux variables aléatoires indépendantes et suivant chacune une loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$. On pose $Z = \max(X, Y)$. On a $Z(\Omega) = \llbracket 1, n \rrbracket$ et pour tout $k \in Z(\Omega)$, or $[Z \leq k] = [\max(X, Y) \leq k]$ correspond à avoir $[X \leq k]$ et $[Y \leq k]$ donc

$$[Z \leq k] = [X \leq k] \cap [Y \leq k].$$

On a ainsi :

$$\begin{aligned}F_Z(k) &= \mathbb{P}(Z \leq k) = \mathbb{P}([X \leq k] \cap [Y \leq k]) \\ &= \mathbb{P}(X \leq k)\mathbb{P}(Y \leq k) \text{ par indépendance} \\ &= \frac{k}{n} \frac{k}{n} = \frac{k^2}{n^2}\end{aligned}$$

On en déduit

$$\mathbb{P}(Z = k) = F_Z(k) - F_Z(k-1) = \frac{k^2}{n^2} - \frac{(k-1)^2}{n^2} = \frac{k^2 - k^2 + 2k - 1}{n^2} = \frac{2k - 1}{n^2}.$$

Pour la loi du minimum, la méthode est presque la même :

Méthode 9.2.24. Pour déterminer la loi de $Z = \min(X, Y)$, on utilise en général la fonction de répartition en suivant les points suivants :

1. Pour tout $z \in Z(\Omega)$, on justifie que $[Z > z] = [X > z] \cap [Y > z]$
2. L'indépendance de X et Y nous permet de calculer $1 - F_Z$ et donc F_Z .
3. On utilise $\mathbb{P}(Z = k) = F_Z(k) - F_Z(k-1)$.

Exemple 9.2.25. Soit $n \in \mathbb{N}^*$. Soit X et Y deux variables aléatoires indépendantes et suivant chacune une loi uniforme $\mathcal{U}(\llbracket 1, n \rrbracket)$. On pose $Z = \min(X, Y)$. On a $Z(\Omega) = \llbracket 1, n \rrbracket$ et pour tout $k \in Z(\Omega)$, or $[Z > k] = [\min(X, Y) > k]$ correspond à avoir $[X > k]$ et $[Y > k]$ donc

$$[Z > k] = [X > k] \cap [Y > k].$$

On a ainsi :

$$\begin{aligned} 1 - F_Z(k) &= 1 - \mathbb{P}(Z \leq k) \\ &= \mathbb{P}(Z > k) \\ &= \mathbb{P}([X > k] \cap [Y > k]) \\ &= \mathbb{P}(X > k)\mathbb{P}(Y > k) \text{ par indépendance} \\ &= \frac{n-k}{n} \frac{n-k}{n} = \frac{(n-k)^2}{n^2} \end{aligned}$$

donc

$$F_Z(k) = 1 - \frac{(n-k)^2}{n^2}.$$

On en déduit

$$\begin{aligned} \mathbb{P}(Z = k) &= F_Z(k) - F_Z(k-1) = 1 - \frac{(n-k)^2}{n^2} - 1 + \frac{(n-k+1)^2}{n^2} \\ &= \frac{-(n^2 - 2nk + k^2) + n^2 - 2n(k-1) + (k-1)^2}{n^2} \\ &= \frac{-n^2 + 2nk - k^2 + n^2 - 2nk + 2n + k^2 - 2k + 1}{n^2} \\ &= \frac{2n - 2k + 1}{n^2} \end{aligned}$$

Remarque 9.2.26. Il est possible de relier ces lois à celle de la somme en remarquant que $X + Y = \min(X, Y) + \max(X, Y)$.

9.3 Notion de covariance

Jusqu'à présent deux variables aléatoires étaient indépendantes ou non indépendantes, nous allons voir à présent comment quantifier le degré de "non-indépendance" ainsi que comment elles varient l'une par rapport à l'autre à l'aide de la covariance.

9.3.1 Définition et propriétés fondamentales

Définition 9.3.1. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune un moment d'ordre 2. On appelle alors covariance la quantité

$$\text{Cov}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))).$$

En pratique on utilise la formule de Koenig-Huygens :

Théorème 9.3.2 (Koenig-Huygens). Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune un moment d'ordre 2. On a alors

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Démonstration. On observe que

$$(X - \mathbb{E}(X))(Y - \mathbb{E}(Y)) = XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)$$

et par linéarité de l'espérance on a

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}(XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \end{aligned}$$

□

Exemple 9.3.3. Soit $X \hookrightarrow \mathcal{U}(\llbracket 0, 1 \rrbracket)$ et soit Y une variable aléatoire définie par les lois conditionnelles :

- Sachant $[X = 0]$, Y suit une loi géométrique de paramètre $\frac{1}{2}$ (et en posant que $\mathbb{P}_{X=0}(Y = 0) = 0$);
- Sachant $[X = 1]$, Y suit une loi de Poisson de paramètre 3.

On a que X et Y admettent chacun un moment d'ordre 2 et calculons la covariance de X et Y .

- On a par la formule de transfert que XY admet une espérance si

$$\begin{aligned} \sum_{(i,j) \in \llbracket 0,1 \rrbracket \times \mathbb{N}} ij \mathbb{P}([X = i] \cap [Y = j]) &= \sum_{(i,j) \in \llbracket 0,1 \rrbracket \times \mathbb{N}} ij \mathbb{P}(X = i) \mathbb{P}_{X=i}(Y = j) \\ &= \sum_{j \geq 0} \frac{j}{2} \mathbb{P}_{X=1}(Y = j) \\ &= \frac{1}{2} \sum_{j \geq 0} \frac{j}{2} e^{-3} \frac{3^j}{j!} \end{aligned}$$

est absolument convergente. On reconnaît la moitié de la série donnant la variance d'une loi de Poisson qui est donc bien convergente et on a

$$\mathbb{E}(XY) = \frac{3}{2}.$$

- Par des arguments classiques on trouve que $\mathbb{E}(Y) = \frac{1}{2} \frac{1}{2} + \frac{3}{2} = \frac{5}{2}$.

On en déduit donc par la formule de Koenig-Huygens que

$$\text{Cov}(X, Y) = \frac{3}{2} - \frac{1}{2} \frac{5}{2} = \frac{6}{4} - \frac{5}{4} = \frac{1}{4}.$$

Remarque 9.3.4. Aussi bien dans la définition que dans le théorème de Koenig-Huygens, pour $X = Y$, on retrouve bien les équivalents pour la variance à savoir $\text{Cov}(X, X) = \mathbb{V}(X)$.

Lorsque X et Y sont indépendantes on a :

Propriété 9.3.5. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune un moment d'ordre 2 et sont indépendantes. On a alors

$$\text{Cov}(X, Y) = 0.$$

Remarque 9.3.6. Il ne s'agit pas d'une équivalence, la nullité de la covariance n'implique **PAS** que X et Y sont indépendantes !

Exemple 9.3.7. Soit $X \hookrightarrow \mathcal{U}(\llbracket -1, 1 \rrbracket)$ et soit $Y = X^2$.

- X et Y ne sont pas indépendantes car $[X = 0] \cap [Y = 1]$ est un évènement impossible donc $\mathbb{P}([X = 0] \cap [Y = 1]) = 0$ or $\mathbb{P}(X = 0)\mathbb{P}(Y = 1) = \frac{2}{9}$ donc

$$\mathbb{P}([X = 0] \cap [Y = 1]) \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 1).$$

- On a, en utilisant que $\mathbb{E}(X) = \frac{-1+1}{2} = 0$ et par la formule de Koenig Hyugens :

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) \\ &= -1.\mathbb{P}(XY = -1) + 0.\mathbb{P}(XY = 0) + 1.\mathbb{P}(XY = 1) \\ &= -\mathbb{P}(X = -1) + \mathbb{P}(X = 1) = -\frac{1}{3} + \frac{1}{3} \\ &= 0. \end{aligned}$$

On est donc bien dans un cas où la covariance est nulle alors que les variables aléatoires ne sont pas indépendantes !

La covariance a les propriétés suivantes :

Propriété 9.3.8. Soit X, Y et Z trois variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X, Y et Z admettent chacune un moment d'ordre 2. Soit $(\lambda, \mu) \in \mathbb{R}^2$. On a alors

1. Symétrie : $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
2. Linéarité à gauche : $\text{Cov}(\lambda X + \mu Y, Z) = \lambda \text{Cov}(X, Z) + \mu \text{Cov}(Y, Z)$;
3. Linéarité à droite : $\text{Cov}(X, \lambda Y + \mu Z) = \lambda \text{Cov}(X, Y) + \mu \text{Cov}(X, Z)$.

Exemple 9.3.9. Soit (X, Y) un couple de variables aléatoires discrète où chaque variables admet un moment d'ordre 2. On pose $U = X + Y$ et $V = X - Y$. On a alors

$$\begin{aligned} \text{Cov}(U, V) &= \text{Cov}(X + Y, X - Y) = \text{Cov}(X, X - Y) + \text{Cov}(Y, X - Y) \\ &= \text{Cov}(X, X) - \text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Cov}(Y, Y) \\ &= \mathbb{V}(X) - \text{Cov}(X, Y) + \text{Cov}(X, Y) - \mathbb{V}(Y) \\ &= \mathbb{V}(X) - \mathbb{V}(Y). \end{aligned}$$

De manière similaire à la variance qui est nulle seulement lorsque la variable aléatoire est constante, si une des deux variables aléatoires est constante, alors la covariance est nulle :

Propriété 9.3.10. Soit X une variable aléatoire admettant un moment d'ordre 2 et soit $a \in \mathbb{R}$, on a alors

$$\text{Cov}(X, a) = \text{Cov}(a, X) = 0.$$

Relions à présent la variance à la covariance :

Propriété 9.3.11. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune un moment d'ordre 2. On a alors

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X, Y).$$

Remarque 9.3.12. On retrouve ici la formule dans le cas où X et Y sont indépendantes où $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$ car $Cov(X, Y) = 0$ dans ce cas.

Exemple 9.3.13. On considère de nouveau les variables aléatoires $X \hookrightarrow \mathcal{U}(\llbracket 0, 1 \rrbracket)$ et Y défini par les lois conditionnelles :

- Sachant $[X = 0]$ Y suit une loi géométrique de paramètre $\frac{1}{2}$ (et en posant que $\mathbb{P}_{X=0}(Y = 0) = 0$);
- Sachant $[X = 1]$ Y suit une loi de Poisson de paramètre 3.

On a $\mathbb{V}(X) = \frac{1}{4}$ et par des arguments classiques on peut trouver $\mathbb{V}(Y) = \frac{11}{4}$. On a donc

$$\mathbb{V}(X + Y) = \frac{1}{4} + \frac{11}{4} + 2\frac{1}{4} = \frac{1 + 11 + 2}{4} = \frac{14}{4} = \frac{7}{2}.$$

9.3.1.1 Corrélation linéaire entre deux variables aléatoires discrètes

Définition 9.3.14. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune une variance non nulle. On appelle coefficient de corrélation linéaire de X et Y la quantité

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}.$$

Propriété 9.3.15. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent chacune une variance non nulle. On a alors

1. $|\rho(X, Y)| \leq 1$;
2. $\rho(X, Y) = 1$ si et seulement si il existe deux réels a et b tel que $a > 0$ et $\mathbb{P}(Y = aX + b) = 1$;
3. $\rho(X, Y) = -1$ si et seulement si il existe deux réels a et b tel que $a < 0$ et $\mathbb{P}(Y = aX + b) = 1$.

Chapitre 10

Réduction de matrices carrées

Chapitre 11

Comparaisons de fonctions réelles et développements limités

Dans un précédent chapitre nous avons vu comment comparer efficacement des suites entre elles au voisinage de l'infini à l'aide des relations de négligeabilité et d'équivalence. De nombreuses applications à ces outils se sont déjà présentées pour l'étude de limites, de séries numériques ou encore en probabilités. Il existe les mêmes relations pour les fonctions que nous allons définir dans ce chapitre et nous verrons de premières applications dans ce chapitre et l'application à l'intégration impropre dans le prochain.

11.1 Relations de comparaisons

11.1.1 Notion de voisinage

Contrairement aux études sur les suites que nous avons fait où toutes les limites sont obligatoirement en $+\infty$, ce n'est pas le cas pour les fonctions. On a besoin de définir un endroit où l'on pourra travailler sans risque. C'est pour cela qu'on introduit la notion de voisinage.

Notation 11.1.1. On utilisera plusieurs fois dans cette section la notation $\mathbb{R} \cup \{\pm\infty\}$ et lorsque l'on considèrera un élément $x_0 \in \mathbb{R} \cup \{\pm\infty\}$ cela signifie que $x_0 \in \mathbb{R}$ ou x_0 correspond à $+\infty$ ou encore x_0 correspond à $-\infty$.

Définition 11.1.2. Soit $x_0 \in \mathbb{R} \cup \{\pm\infty\}$. On dit qu'un intervalle I est un voisinage en x_0 si :

- Si $x_0 \in \mathbb{R}$: I contient le segment $[x_0 - r, x_0 + r]$ pour un $r \in \mathbb{R}_+^*$;
- Si $x_0 = +\infty$: I contient l'intervalle $[r, +\infty[$ où $r \in \mathbb{R}$. Si il n'y a pas d'ambiguïté on peut parler d'un voisinage au voisinage de l'infini au lieu de plus l'infini ;

- Si $x_0 = -\infty$: I contient l'intervalle $] -\infty, r]$ où $r \in \mathbb{R}$;

Définition 11.1.3. On dit qu'une propriété est vraie au voisinage de $a \in \mathbb{R} \cup \{\pm\infty\}$ si il existe un voisinage V de a tel que la propriété soit vraie pour tout $x \in V$.

Exemples 11.1.4. 1. Au voisinage de l'infini, la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$
 $x \mapsto -x^2 + 4x - 3$
est strictement négative.

En effet on peut facilement trouver que ce trinôme admet comme racines 1 et 3 donc f est strictement négative sur $] -\infty, 1[\cup]3, +\infty[$ donc $[42, +\infty[$ est un voisinage répondant à notre propriété.

2. Au voisinage de 0, la fonction f ne s'annule jamais. Cela est vrai sur $\mathbb{R} \setminus \{1; 3\}$ donc en particulier aussi sur $[-1, \frac{1}{2}]$ qui est un voisinage de 0.

Remarque 11.1.5. Le voisinage est l'équivalent pour les fonctions du "à partir d'un certain rang" des suites.

11.1.2 Relation de négligeabilité

11.1.2.1 Définition et caractérisation

Définition 11.1.6. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). On dit que f est négligeable devant g au voisinage de a si il existe un voisinage V de a et une fonction $\epsilon : V \mapsto \mathbb{R}$ tels que pour tout $x \in V \cap I$, $f(x) = \epsilon(x)g(x)$ et $\lim_{x \rightarrow a} \epsilon(x) = 0$.

On note alors $f(x) = o_{x \rightarrow a}(g(x))$.

Remarque 11.1.7. Contrairement aux suites, le " $x \rightarrow a$ " ne peut pas se retirer en cas de non ambiguïté!

Comme pour les suites, une caractérisation va nous être fort utile :

Théorème 11.1.8. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). Si g ne s'annule pas sur un voisinage de a (sauf éventuellement en a) alors on a que f est négligeable devant g au voisinage de a si et seulement si $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$.

Exemples 11.1.9. 1. On a $x = o_{x \rightarrow +\infty}(e^x)$ car $\lim_{x \rightarrow +\infty} \frac{x}{e^x} = 0$ par croissance comparée;

2. On a $\ln(x) = o_{x \rightarrow +\infty}(e^x)$ car $\lim_{x \rightarrow +\infty} \frac{\ln(x)}{e^x} = 0$ par croissance comparée ;
3. On a $x^2 = o_{x \rightarrow 0}(x)$ car $\lim_{x \rightarrow 0} \frac{x^2}{x} = 0$.
4. On a $x = o_{x \rightarrow -\infty}(x^2)$ car $\lim_{x \rightarrow -\infty} \frac{x}{x^2} = 0$.

Remarques 11.1.10. 1. Deux fonctions négligeables devant la même quantité ne sont pas égales. Dans les exemples précédents, on a $x = o_{x \rightarrow +\infty}(e^x)$ et $\ln(x) = o_{x \rightarrow +\infty}(e^x)$ or $x \neq \ln(x)$.

2. L'endroit où l'on considère notre relation de négligabilité est fondamentale, il faut toujours le préciser sous peine d'écrire d'énormes erreurs ! Cela est bien illustré par les exemples précédents où l'on a vu que $x^2 = o_{x \rightarrow 0}(x)$ et $x = o_{x \rightarrow 0}(x^2)$

11.1.2.2 Règles de calculs

De la même manière que pour les suites, on a certaines règles de calculs et la possibilité de réécrire les croissances comparées en fonction de petit o .

Théorème 11.1.11. Soit f, g et h trois fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). Soit $(\lambda, \mu) \in \mathbb{R}^2$. On a alors :

1. (Transitivité) Si $f(x) = o_{x \rightarrow a}(g(x))$ et $g(x) = o_{x \rightarrow a}(h(x))$ alors $f(x) = o_{x \rightarrow a}(h(x))$;
2. (Combinaison linéaire) Si $f(x) = o_{x \rightarrow a}(h(x))$ et $g(x) = o_{x \rightarrow a}(h(x))$ on a alors $\lambda f(x) + \mu g(x) = o_{x \rightarrow a}(g(x))$;
3. (Multiplication par un réel non nul) Si $\lambda \neq 0$ et $f(x) = o_{x \rightarrow a}(g(x))$ alors $f(x) = o_{x \rightarrow a}(\lambda g(x))$
4. (Produit) Si $f(x) = o_{x \rightarrow a}(g(x))$ alors $f(x)h(x) = o_{x \rightarrow a}(g(x)h(x))$

Remarques 11.1.12. 1. On fera en sorte de toujours avoir qu'un seul o dans nos expressions et les deux premières propriétés imposent donc dans un premier temps de "mettre tout les petits o au même niveau" puis de les combiner entres eux.

2. La multiplication par un réel non nul permet usuellement de ne considérer que des fonctions "normalisées" dans notre petit o . Par exemple on peut remplacer un $o_{x \rightarrow a}(2)$ par un $o_{x \rightarrow a}(1)$ ou encore un $o_{x \rightarrow a}\left(\frac{4x}{\sqrt{7}}\right)$ par $o_{x \rightarrow a}(x)$.

3. Le produit sert usuellement à modifier un petit o pour le mettre au même niveau qu'un autre.

Exemple 11.1.13. Simplifions $(1+x)(1+x+\frac{x^2}{2}+o_{x \rightarrow 0}(x^2))$. On a

$$\begin{aligned} (1+x)(1+x+\frac{x^2}{2}+o_{x \rightarrow 0}(x^2)) &= 1+x+\frac{x^2}{2}+o_{x \rightarrow 0}(x^2)+x+x^2+\frac{x^3}{2}+x o_{x \rightarrow 0}(x^2) \\ &= 1+2x+\frac{3x^2}{2}+\frac{x^3}{3}+o_{x \rightarrow 0}(x^2)+o_{x \rightarrow 0}(x^3) \\ &= 1+2x+\frac{3x^2}{2}+o_{x \rightarrow 0}(x^2) \end{aligned}$$

car $\frac{x^3}{2}+o_{x \rightarrow 0}(x^3)=o_{x \rightarrow 0}(x^2)$.

Pour les croissances comparées nous avons :

Théorème 11.1.14 (Croissances comparées en $+\infty$). Soit $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$. On a alors

- $x^\alpha = o_{x \rightarrow +\infty}(e^{\beta x})$;
- $\ln(x)^\beta = o_{x \rightarrow +\infty}(x^\alpha)$;
- $\ln(x)^\alpha = o_{x \rightarrow +\infty}(e^{\beta x})$;
- $\frac{1}{x^\alpha} = o_{x \rightarrow +\infty}\left(\frac{1}{\ln(x)^\beta}\right)$.

De plus si $\alpha < \beta$, on a de plus :

- $e^{\alpha x} = o_{x \rightarrow +\infty}(e^{\beta x})$;
- $x^\alpha = o_{x \rightarrow +\infty}(x^\beta)$;
- $e^{-\beta x} = o_{x \rightarrow +\infty}(e^{-\alpha x})$;
- $\frac{1}{x^\beta} = o_{x \rightarrow +\infty}\left(\frac{1}{x^\alpha}\right)$.

Théorème 11.1.15 (Croissances comparées en 0^+). Soit $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$. On a alors

- $x^\alpha = o_{x \rightarrow 0^+}\left(\frac{1}{\ln(x)^\beta}\right)$;
- $\ln(x)^\alpha = o_{x \rightarrow 0^+}\left(\frac{1}{x^\beta}\right)$.

De plus si $\alpha < \beta$, on a de plus :

- $x^\beta = o_{x \rightarrow 0^+}(x^\alpha)$;
- $\frac{1}{x^\alpha} = o_{x \rightarrow 0^+}\left(\frac{1}{x^\beta}\right)$.

Exemples 11.1.16. • $x^4 = o_{x \rightarrow +\infty}(e^{\frac{x}{2}})$;

- $x^7 = o_{x \rightarrow +\infty}(x^{42})$;
- $\sqrt{\ln(x)} = o_{x \rightarrow 0^+}\left(\frac{1}{x^{100}}\right)$;
- $\frac{1}{x} = o_{x \rightarrow 0^+}\left(\frac{1}{x^2}\right)$.

11.1.3 Relation d'équivalence

11.1.3.1 Définition et caractérisation

Définition 11.1.17. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). On dit que f est équivalent à g au voisinage de a si il existe un voisinage V de a et une fonction $\epsilon : V \mapsto \mathbb{R}$ tels que pour tout $x \in V \cap I$, $f(x) = \epsilon(x)g(x)$ et $\lim_{x \rightarrow a} \epsilon(x) = 1$.

On note alors $f(x) \underset{x \rightarrow a}{\sim} g(x)$.

Les relations d'équivalences se caractérisent de la manière suivante :

Théorème 11.1.18. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). Si g ne s'annule pas sur un voisinage de a (sauf éventuellement en a) alors on a que f est équivalent à g au voisinage de a si et seulement si

$$f(x) = g(x) + o_{x \rightarrow a}(g(x))$$

si et seulement si

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1.$$

Exemples 11.1.19. • On a $x + 2 \underset{x \rightarrow +\infty}{\sim} x$ et $x - 1 \underset{x \rightarrow +\infty}{\sim} x$ car $\lim_{x \rightarrow +\infty} \frac{x+2}{x} = \lim_{x \rightarrow +\infty} \frac{x-1}{x} = 1$.

• On a $\frac{x+1}{x-2} \underset{x \rightarrow 2}{\sim} \frac{3}{x-2}$ car $\lim_{x \rightarrow 2} \frac{\frac{x+1}{x-2}}{\frac{3}{x-2}} = \lim_{x \rightarrow 2} \frac{x+1}{3} = \frac{3}{3} = 1$.

Remarque 11.1.20. Deux fonctions équivalentes à une même troisième fonction ne sont pas égales entre elles comme on peut le voir avec l'exemple précédent. On a que $x + 2 \underset{x \rightarrow +\infty}{\sim} x$ et $x - 1 \underset{x \rightarrow +\infty}{\sim} x$ or $x + 2 \neq x - 1$.

11.1.3.2 Règles de calculs et équivalents usuelles

Comme pour les suites les règles de calculs pour les équivalents sont légèrement plus souples mais il faut tout de même redoubler d'attention :

Propriété 11.1.21. Soit f , g , u et v quatre fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). On a alors :

• (Symétrie) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ alors $g(x) \underset{x \rightarrow a}{\sim} f(x)$;

- (Transitivité) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ et $g(x) \underset{x \rightarrow a}{\sim} u(x)$ alors $f(x) \underset{x \rightarrow a}{\sim} u(x)$;
- (Produit "simplifié") Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ alors $f(x)u(x) \underset{x \rightarrow a}{\sim} g(x)u(x)$;
- (Produit) Si $f(x) \underset{x \rightarrow a}{\sim} h(x)$ et $u(x) \underset{x \rightarrow a}{\sim} v(x)$ alors $f(x)u(x) \underset{x \rightarrow a}{\sim} g(x)v(x)$;
- (Inverse) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ et que sur un voisinage de a (sauf éventuellement en a) les fonctions ne s'annulent pas, on a $\frac{1}{f(x)} \underset{x \rightarrow a}{\sim} \frac{1}{g(x)}$;
- (Quotient) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$, $u(x) \underset{x \rightarrow a}{\sim} v(x)$ et u et v ne s'annulent pas sur un voisinage de a (sauf éventuellement en a), on a $\frac{f(x)}{u(x)} \underset{x \rightarrow a}{\sim} \frac{g(x)}{v(x)}$;
- (Puissance) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$, alors pour tout $k \in \mathbb{N}$, on a $f(x)^k \underset{x \rightarrow a}{\sim} g(x)^k$,
- (Valeur absolue) Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ alors $|f(x)| \underset{x \rightarrow a}{\sim} |g(x)|$.

Exemple 11.1.22. Calculons un équivalent au voisinage de l'infini de $\frac{(x^3-1)^3}{2(x+3)}$. On montre que $x^3 - 1 \underset{x \rightarrow +\infty}{\sim} x^3$ donc on a $(x^3 - 1)^3 \underset{x \rightarrow +\infty}{\sim} x^9$. De plus on a $x + 3 \underset{x \rightarrow +\infty}{\sim} x$ donc par quotient on a

$$\frac{(x^3 - 1)^3}{2(x + 3)} \underset{x \rightarrow +\infty}{\sim} \frac{x^9}{2x} = \frac{x^8}{2}.$$

Remarque 11.1.23. On n'additionne ou soustrait JAMAIS des équivalents. En effet, on a vu précédemment que $x+2 \underset{x \rightarrow +\infty}{\sim} x$ et $x-1 \underset{x \rightarrow +\infty}{\sim} x$ or $x+2 - (x-1) = 3$ et $x - x = 0$ et 3 n'est pas équivalent au voisinage de l'infini à 0!

De même on ne peut pas appliquer une fonction à des équivalents. En effet $\lim_{x \rightarrow +\infty} \frac{e^{x+2}}{e^x} = e^2 \neq 1$ donc e^{x+2} n'est pas équivalent au voisinage de l'infini à e^x alors que $x+2$ l'est de x .

Théorème 11.1.24. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$). On a alors :

1. Si $f(x) \underset{x \rightarrow a}{\sim} g(x)$ et si $\lim_{x \rightarrow a} f(x) = l \in \mathbb{R} \cup \{\pm\infty\}$ alors $\lim_{x \rightarrow a} g(x) = l$.
2. Si $\lim_{x \rightarrow a} f(x) = l$ où $l \in \mathbb{R}^*$ alors $f(x) \underset{x \rightarrow a}{\sim} l$.

Exemple 11.1.25. On a vu précédemment que $\frac{(x^3-1)^3}{2(x+3)} \underset{x \rightarrow +\infty}{\sim} \frac{x^9}{2x} = \frac{x^8}{2}$ or $\lim_{x \rightarrow +\infty} \frac{x^8}{2} = +\infty$ donc $\lim_{x \rightarrow +\infty} \frac{(x^3-1)^3}{2(x+3)} = +\infty$.

Remarque 11.1.26. Si la limite de f est nulle en a , alors on n'a pas d'équivalent car sinon cela impose que f soit nulle sur tout un voisinage de a . De manière générale, on n'écrit jamais d'équivalent à 0 comme on n'écrit pas de $o_{x \rightarrow a}(0)$.

Concluons cette section par des équivalents usuelles à maîtriser parfaitement :

Théorème 11.1.27. On a les équivalents usuelles suivants :

- $\ln(1+x) \underset{x \rightarrow 0}{\sim} x$;
- $e^x - 1 \underset{x \rightarrow 0}{\sim} x$;
- Pour $\alpha \in \mathbb{R}_+^*$, $(1+x)^\alpha - 1 \underset{x \rightarrow 0}{\sim} \alpha x$;

Pour les polynômes, on a le résultat suivant :

Théorème 11.1.28. Soit $(a_p, \dots, a_n) \in \mathbb{R}^{n-p+1}$ où $p \leq n$, $a_n \neq 0$ et $a_p \neq 0$. On a alors :

- $a_p x^p + \dots + a_n x^n \underset{x \rightarrow +\infty}{\sim} a_n x^n$;
- $a_p x^p + \dots + a_n x^n \underset{x \rightarrow -\infty}{\sim} a_n x^n$;
- $a_p x^p + \dots + a_n x^n \underset{x \rightarrow 0}{\sim} a_p x^p$;

Exemple 11.1.29. On pose $f : \mathbb{R} \rightarrow \mathbb{R}$. On a alors

$$x \mapsto \frac{7x^4 + 3x^3 - 4x}{x^2 + 9}$$

- $f(x) \underset{x \rightarrow +\infty}{\sim} \frac{7x^4}{x^2} = 7x^2$;
- $f(x) \underset{x \rightarrow 0}{\sim} \frac{-4x}{9}$.

Un résultat qui nous servira beaucoup est le suivant :

Théorème 11.1.30. Soit f et g deux fonctions réelles définies sur un intervalle I . Soit a un élément ou une borne de I (donc en particulier $a \in \mathbb{R} \cup \{\pm\infty\}$) tel que $f(x) \underset{x \rightarrow a}{\sim} g(x)$. On a alors que sur un voisinage de a , f et g sont de même signe.

Démonstration. Par définition, il existe une fonction ϵ tel que sur un voisinage de a , $f(x) = \epsilon(x)g(x)$ et $\lim_{x \rightarrow a} \epsilon(x) = 1$. En particulier, il existe donc un voisinage de a tel que $\frac{1}{2} \leq \epsilon(x)$ donc en particulier sur lequel $\epsilon(x)$ est strictement positif. On en déduit donc que sur ce voisinage, f et g ont le même signe. \square

11.2 Développements limités

Les fonctions usuelles sous réserves de bonnes régularités, peuvent s'écrire localement comme un polynôme à un petit o près. C'est ce qu'on appelle les développements limités.

11.2.1 Ordre 0

Définition 11.2.1. Soit f une fonction réelle définie sur un intervalle I . Soit $x_0 \in I$. On dit que f admet un développement limité en x_0 d'ordre 0 si il existe un réel a_0 tel que

$$f(x) = a_0 + o_{x \rightarrow x_0}(1).$$

Remarque 11.2.2. Un développement limité se fait toujours en un réel, il n'y a pas de développement limité en $+\infty$ ou $-\infty$.

Remarque 11.2.3. L'existence d'un développement limité d'ordre 0 en x_0 est équivalente à la continuité de la fonction en x_0 . Cela implique également que le développement limité est unique.

Exemple 11.2.4. On considère la fonction $f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$. On observe

$$x \mapsto \frac{1}{1-x}$$

qu'au voisinage de 0 f admet le développement limité d'ordre 0,

$$f(x) = 1 + o_{x \rightarrow 0}(1).$$

11.2.2 Ordre 1

Définition 11.2.5. Soit f une fonction réelle définie sur un intervalle I . Soit $x_0 \in I$. On dit que f admet un développement limité en x_0 d'ordre 1 si il existe deux réels a_0 et a_1 tels que

$$f(x) = a_0 + a_1(x - x_0) + o_{x \rightarrow x_0}((x - x_0)).$$

On appelle $a_0 + a_1(x - x_0)$ la partie régulière du développement limité et $o_{x \rightarrow x_0}((x - x_0))$ le reste.

Remarque 11.2.6. La partie régulière est unique grâce au résultat suivant :

Théorème 11.2.7 (Formule de Taylor Young - ordre 1). *Soit f une fonction définie sur un intervalle I , $x_0 \in I$ et de classe \mathcal{C}^1 au voisinage de x_0 , alors f admet un développement limité d'ordre 1 en x_0 donné par*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \underset{x \rightarrow x_0}{o}((x - x_0)).$$

Exemple 11.2.8. On considère de nouveau la fonction $f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$.

$$x \mapsto \frac{1}{1-x}$$

f est bien de classe \mathcal{C}^1 en 0 donc f admet un développement limité d'ordre 1 en 0. On a de plus $f(0) = \frac{1}{1-0} = 1$ et pour tout x dans un voisinage de 0, $f'(x) = -\frac{-1}{(1-x)^2} = \frac{1}{(1-x)^2}$ donc $f'(0) = \frac{1}{(1-0)^2} = 1$.
Le développement limité de f en 0 d'ordre 1 est donc

$$f(x) = 1 + x + \underset{x \rightarrow 0}{o}(x).$$

11.2.3 Ordre 2

Définition 11.2.9. Soit f une fonction réelle définie sur un intervalle I . Soit $x_0 \in I$. On dit que f admet un développement limité en x_0 d'ordre 2 si il existe trois réels a_0 , a_1 et a_2 tels que

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \underset{x \rightarrow x_0}{o}((x - x_0)^2).$$

On appelle $a_0 + a_1(x - x_0) + a_2(x - x_0)^2$ la partie régulière du développement limité et $\underset{x \rightarrow x_0}{o}((x - x_0)^2)$ le reste.

Remarque 11.2.10. La partie régulière est unique. Sous de bonne condition on l'obtient grâce au théorème suivant :

Théorème 11.2.11 (Formule de Taylor Young - ordre 2). *Soit f une fonction définie sur un intervalle I , $x_0 \in I$ et de classe \mathcal{C}^2 au voisinage de x_0 , alors f admet un développement limité d'ordre 2 en x_0 donné par*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \underset{x \rightarrow x_0}{o}((x - x_0)^2).$$

Remarque 11.2.12. Contrairement au cas d'ordre 1, ici le fait d'être de classe \mathcal{C}^2 est plus fort que le fait d'admettre un développement limité d'ordre 2. Cependant les fonctions que nous utiliserons ne nous permettront pas d'illustrer ce problème.

Exemple 11.2.13. On considère de nouveau la fonction $f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$.

$$x \mapsto \frac{1}{1-x}$$

f est bien de classe \mathcal{C}^2 en 0 donc f admet un développement limité d'ordre 2 en 0 par la formule de Taylor Young. On a déjà vu que $f(0) = \frac{1}{1-0} = 1$ et $f'(0) = \frac{1}{(1-0)^2} = 1$.

On a au voisinage de 0, $f''(x) = -\frac{-2}{(1-x)^3} = \frac{2}{(1-x)^3}$ et donc $f''(0) = \frac{2}{(1-0)^3} = 2$. Le développement limité de f en 0 d'ordre 1 est donc

$$f(x) = 1 + x + x^2 + o_{x \rightarrow 0}(x^2).$$

Remarque 11.2.14. Admettre un développement limité d'ordre 2 implique d'admettre un développement limité d'ordre 1, qui lui même implique d'admettre un développement limité d'ordre 0.

De manière général un développement limité fourni un équivalent assez directement. La fonction est équivalente au premier terme non nul du développement limité.

Propriété 11.2.15. Soit f une fonction réelle définie sur un intervalle I . Soit $x_0 \in I$ et on suppose que f admet un développement limité en x_0 d'ordre 2 $f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + o_{x \rightarrow x_0}((x - x_0)^2)$. On a alors

- Si $a_0 \neq 0$, $f(x) \underset{x \rightarrow x_0}{\sim} a_0$;
- Si $a_0 = 0$ et $a_1 \neq 0$, $f(x) \underset{x \rightarrow x_0}{\sim} a_1 x$;
- Si $a_0 = a_1 = 0$ et $a_2 \neq 0$, $f(x) \underset{x \rightarrow x_0}{\sim} a_2 x^2$.

Remarques 11.2.16. 1. Si $a_0 = a_1 = a_2 = 0$, on ne peut rien dire.

2. Cette propriété s'adapte au cas d'un développement limité d'ordre 0 ou 1.

Exemple 11.2.17. On considère une nouvelle fois la fonction $f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$.

$$x \mapsto \frac{1}{1-x}$$

On a donc

$$f(x) \underset{x \rightarrow 0}{\sim} 1.$$

11.2.4 Développements limités usuelles

Certains développements limités sont à connaître par cœur. De même il faut absolument savoir les calculer rapidement.

Nous donnons ces trois développements limités en 0 dans le tableau suivant et la preuve sera faite en TD.

	Ordre 0	Ordre 1	Ordre 2
e^x	$1 + o_{x \rightarrow 0}(1)$	$1 + x + o_{x \rightarrow 0}(x)$	$1 + x + \frac{x^2}{2} + o_{x \rightarrow 0}(x^2)$
$\ln(1+x)$	$o_{x \rightarrow 0}(1)$	$x + o_{x \rightarrow 0}(x)$	$x - \frac{x^2}{2} + o_{x \rightarrow 0}(x^2)$
$(1+x)^\alpha, \alpha \in \mathbb{R}$	$1 + o_{x \rightarrow 0}(1)$	$1 + \alpha x + o_{x \rightarrow 0}(x)$	$1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + o_{x \rightarrow 0}(x^2)$

Remarques 11.2.18. • Avoir le même développement limité ne signifie que les fonctions sont égales, on peut illustrer cela par l'exponentielle et la fonction puissance à l'ordre 0 dans le tableau précédent.

- $(1+x)^\alpha$ permet notamment de calculer les développements limités de $x \mapsto \sqrt{1+x}$ et $x \mapsto \frac{1}{1+x}$.
- Il pourra parfois être utile de savoir que $o_{x \rightarrow 0}(x^n) o_{x \rightarrow 0}(x^m) = o_{x \rightarrow 0}(x^{n+m})$ dans les calculs de produits de développements limités. Cependant cette règle est spécifique aux développements limités en 0 et à ces fonctions puissances. Elle est de plus à la limite du programme.

11.3 Applications des développements limités

11.3.1 Aspect graphique

Un développement limité d'ordre 2 contient aussi l'information nécessaire pour donner l'allure de la courbe en le point où le développement limité est effectué.

On a le résultat suivant qui synthétise tout les cas possibles :

Théorème 11.3.1. Soit f une fonction définie sur un intervalle I et soit $x_0 \in I$. On suppose que f admet un développement limité d'ordre 2 en x_0 qui est

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + o_{x \rightarrow x_0}((x - x_0)^2).$$

On a alors :

- f est dérivable en x_0 et l'on a $f(x_0) = a_0$ et $f'(x_0) = a_1$;

- L'équation de la tangente de f en x_0 est $y = a_0 + a_1(x - x_0)$;
- Si $a_2 > 0$, alors la courbe de f est localement au dessus de la tangente en x_0 ;
- Si $a_2 < 0$, alors la courbe de f est localement au dessous de la tangente en x_0 ;

Remarque 11.3.2. Le cas où $a_2 = 0$ ne nous apportent aucune information quant à la position de la tangente.

Remarque 11.3.3. Dans le cas où $a_1 = 0$ implique que x_0 est un extremum local de f et donc dans ce cas si $a_2 > 0$, il s'agit d'un minimum local et si $a_2 < 0$ il s'agit d'un maximum local.

Exemple 11.3.4. On considère la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ $x \mapsto \frac{1}{x^2+2x+5}$. On peut vérifier aisément que cette fonction est bien de classe \mathcal{C}^∞ sur \mathbb{R} et l'on va s'intéresser à son comportement autour de -1 et de 0 .

Pour tout $x \in \mathbb{R}$, on a

$$f'(x) = -\frac{2x+2}{(x^2+2x+5)^2}$$

et

$$\begin{aligned} f''(x) &= -\frac{2(x^2+2x+5)^2 - (2x+2)2(2x+2)(x^2+2x+5)}{(x^2+2x+5)^4} \\ &= -2\frac{x^2+2x+5 - (2x+2)^2}{(x^2+2x+5)^3} = -2\frac{x^2+2x+5 - (4x^2+8x+4)}{(x^2+2x+5)^3} \\ &= -2\frac{-3x^2-6x+1}{(x^2+2x+5)^3} \end{aligned}$$

Comme f est de classe \mathcal{C}^2 en -1 et en 0 , par la formule de Taylor Young f admet des deux développements limités en ces deux points. On a :

- | | |
|-----------------------------|--|
| • $f(0) = \frac{1}{5}$ | • $f(-1) = \frac{1}{4}$ |
| • $f'(0) = \frac{-2}{25}$ | • $f'(-1) = 0$ |
| • $f''(0) = \frac{-2}{125}$ | • $f''(-1) = \frac{-8}{64} = \frac{-1}{8}$ |

On en déduit donc que le développement limité de f en 0 à l'ordre 2 est

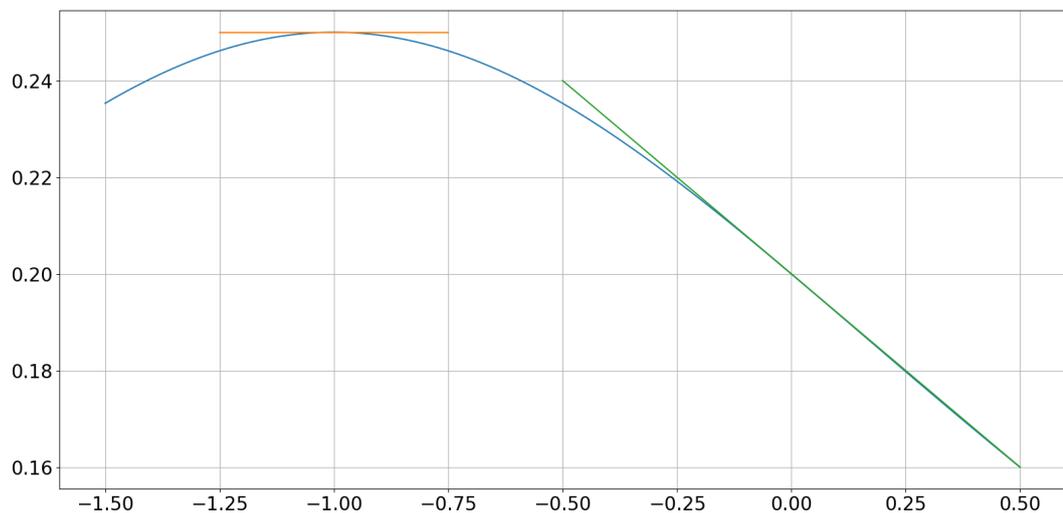
$$f(x) = \frac{1}{5} - \frac{2x}{25} - \frac{x^2}{125} + o_{x \rightarrow 0}(x^2)$$

et le développement limité de f en -1 à l'ordre 2 est

$$f(x) = \frac{1}{4} - \frac{(x+1)^2}{16} + o_{x \rightarrow -1}((x+1)^2)$$

On déduit de ces développements limités que la tangente en 0 est portée par $y = \frac{1}{5} - \frac{2x}{25}$ et que la courbe est au dessous de la tangente au voisinage de 0. La tangente en -1 est portée par $y = \frac{1}{4}$, que la courbe est au dessous de la tangente au voisinage de -1 et que -1 est un maximum local de f .

On observe tout cela sur le graphique python ci dessous :



11.3.2 Application aux suites

Une application classique des développements limités est de déduire l'allure d'une suite comme dans l'exemple suivant :

Exemple 11.3.5. Montrons que la série $\sum_{n \geq 1} u_n$ est convergente où pour tout $n \in$

$$\mathbb{N}^*, u_n = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right).$$

On sait qu'au voisinage de 0, $\ln(1+x) = x - \frac{x^2}{2} + o_{x \rightarrow 0}(x^2)$ donc comme $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$, on a

$$\ln\left(1 + \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right).$$

On a donc

$$\begin{aligned}u_n &= \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) \\&= \frac{1}{n} - \left(\frac{1}{n} - \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right)\right) \\&= \frac{1}{2n^2} + o_{n \rightarrow +\infty}\left(\frac{1}{n^2}\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{2n^2}\end{aligned}$$

Par théorème de comparaison sur les séries à termes positifs, comme la série de Riemann de terme général $\frac{1}{2n^2}$ est convergente car de coefficient $\alpha = 2 > 1$ on a que la série de terme général u_n est convergente.

Chapitre 12

Intégrales impropres de fonctions positives

L'intégration est un pan très large de l'analyse et, nous allons le voir, des probabilités. Avec les derniers outils d'analyse que nous avons construit on a de nouvelles manières de prouver la convergence d'intégrale impropre de fonctions positives.

On rappelle la définition d'une intégrale impropre :

Définition 2.2.1. Soit f une fonction continue sur un intervalle de la forme $[a, +\infty[$ où $a \in \mathbb{R}$.

- Si la limite $\lim_{x \rightarrow +\infty} \int_a^x f(t) dt$ existe et est finie alors on dit que l'intégrale impropre en $+\infty$ $\int_a^{+\infty} f(t) dt$ est convergente et sa valeur est la valeur de la limite trouvée.
- Si la limite n'existe pas ou est divergente on dit alors que l'intégrale impropre $\int_a^{+\infty} f(t) dt$ est divergente.

12.1 Généralités

Dans le cas d'une fonction positive, la convergence de l'intégrale impropre revient au caractère bornée de l'intégrale "partielle".

Propriété 12.1.1. Soit f une fonction continue et positive sur $[a, +\infty[$. L'intégrale $\int_a^{+\infty} f(t) dt$ converge si et seulement si $F : x \mapsto \int_a^x f(t) dt$ est majorée sur $[a, +\infty[$.

Démonstration. On remarque que F est une fonction croissante. En effet pour tout $(\alpha, \beta) \in [a, +\infty[^2$ tel que $\alpha \leq \beta$. On a

$$\begin{aligned} F(\beta) - F(\alpha) &= \int_a^\beta f(t) dt - \int_a^\alpha f(t) dt \\ &= \int_a^\alpha f(t) dt + \int_\alpha^\beta f(t) dt - \int_a^\alpha f(t) dt \\ &= \int_\alpha^\beta f(t) dt \geq 0 \end{aligned}$$

car la fonction f est positive sur $[\alpha, \beta]$. On a donc bien que F est croissante. Par le théorème de la limite monotone on a donc que $\lim_{x \rightarrow +\infty} F(x)$ existe bien si et seulement si $F(x)$ est majorée, autrement dit $\int_a^{+\infty} f(t) dt$ converge si et seulement si F est majorée sur $[a, +\infty[$. \square

On a de même :

Propriété 12.1.2. Soit f une fonction continue et positive sur $] -\infty, b]$. L'intégrale $\int_{-\infty}^b f(t) dt$ converge si et seulement si $F : x \mapsto \int_x^b f(t) dt$ est majorée sur $] -\infty, b]$.

Exemple 12.1.3. Considérons l'intégrale $\int_1^\infty \frac{e^{-t^2}}{t^2} dt$.

On observe que $f : [1, +\infty[\rightarrow \mathbb{R}$
 $t \mapsto \frac{e^{-t^2}}{t^2}$ est continue et positive. Pour montrer que l'intégrale est convergente il nous suffit donc de montrer que pour tout $x \in [1, +\infty[$, $\int_1^x f(t) dt$ est majorée indépendamment de x .

On a pour tout $t \in [1, +\infty[$, $e^{-t^2} \leq 1$ donc on a pour tout $x \in [1, +\infty[$:

$$\int_1^x f(t) dt \leq \int_1^x \frac{1}{t^2} dt = \left[\frac{-1}{t} \right]_1^x = \frac{-1}{x} + 1 \leq 1.$$

On en déduit donc que $\int_1^x f(t) dt$ est bien majorée donc l'intégrale $\int_1^\infty \frac{e^{-t^2}}{t^2} dt$ est convergente.

Comme pour les séries, l'équivalence entre la convergence et le caractère majorée nous permet de construire de nouveaux critères de convergences.

Remarque 12.1.4. Les intégrales de Riemann ainsi que l'intégrale de l'exponentielle sont des intégrales de fonctions positives que l'on pourra utiliser dans nos comparaisons !

12.2 Inégalités

Théorème 12.2.1. Soit f et g deux fonctions réelles continues sur $[a, +\infty[$, positives au voisinage de $+\infty$ et telles que au voisinage de $+\infty$, $f(x) \leq g(x)$. On a alors :

- Si $\int_a^{+\infty} g(t) dt$ converge alors $\int_a^{+\infty} f(t) dt$ converge également.
- Si $\int_a^{+\infty} f(t) dt$ diverge alors $\int_a^{+\infty} g(t) dt$ diverge également.

Exemple 12.2.2. On reprend l'exemple précédent, et l'on a montré que pour tout $t \in [1, +\infty[$, $0 \leq \frac{e^{-t^2}}{t^2} \leq \frac{1}{t^2}$ or $\int_1^{\infty} \frac{1}{t^2} dt$ est convergente en tant qu'intégrale de Riemann de coefficient $\alpha = 2 > 1$. On a donc par théorème de comparaison que l'intégrale $\int_1^{\infty} \frac{e^{-t^2}}{t^2} dt$ est bien convergente.

De même en $-\infty$, on a :

Théorème 12.2.3. Soit f et g deux fonctions réelles continues sur $] -\infty, b]$, positives au voisinage de $-\infty$ et telles que au voisinage de $-\infty$, $f(x) \leq g(x)$. On a alors :

- Si $\int_{-\infty}^b g(t) dt$ converge alors $\int_{-\infty}^b f(t) dt$ converge également.
- Si $\int_{-\infty}^b f(t) dt$ diverge alors $\int_{-\infty}^b g(t) dt$ diverge également.

Remarque 12.2.4. On prendra garde au fait que ces résultats ne montrent que la convergence ou la divergence de l'intégrale. Ils ne fournissent pas une inégalité sur la valeur des intégrales si ils sont convergents. En effet les hypothèses ne concernant que les fonctions au voisinage de $+\infty$, ce qui se passe avant à trop d'influence sur l'intégrale pour donner une règle général.

12.3 Négligeabilité

Théorème 12.3.1. Soit f et g deux fonctions réelles continues sur $[a, +\infty[$, positives au voisinage de $+\infty$ et telles que $f(x) = o_{x \rightarrow +\infty}(g(x))$. On a alors :

- Si $\int_a^{+\infty} g(t) dt$ converge alors $\int_a^{+\infty} f(t) dt$ converge également.
- Si $\int_a^{+\infty} f(t) dt$ diverge alors $\int_a^{+\infty} g(t) dt$ diverge également.

Exemples 12.3.2. 1. On remonte plus simplement que l'intégrale $\int_1^{\infty} \frac{e^{-t^2}}{t^2} dt$ est convergente. On observe que $\lim_{t \rightarrow +\infty} t^2 \frac{e^{-t^2}}{t^2} = 0$ par croissance comparée donc $\frac{e^{-t^2}}{t^2} = o_{t \rightarrow +\infty}\left(\frac{1}{t^2}\right)$ or $t \mapsto \frac{1}{t^2}$ est intégrale en tant qu'intégrale de

Riemann de coefficient $\alpha = 2 > 1$ donc par le critère d'intégrabilité des fonctions continues positives, on a l'intégrale que $\int_1^\infty \frac{e^{-t^2}}{t^2} dt$ est convergente.

2. Montrons que l'intégrale $\int_1^\infty \frac{1}{\ln(t)} dt$ diverge. On a par croissance comparée que $\lim_{t \rightarrow +\infty} \frac{\ln(t)}{t} = 0$ donc $\frac{1}{t} = o_{t \rightarrow +\infty} \left(\frac{1}{\ln(t)} \right)$ donc comme l'intégrale de Riemann de coefficient $\alpha = 1 \geq 1$ $\int_1^\infty \frac{dt}{t}$ est divergente par critère d'intégrabilité des fonctions continues positives au voisinage de l'infini, on a que l'intégrale $\int_1^\infty \frac{dt}{\ln(t)}$ est divergente.

De même en $-\infty$, on a :

Théorème 12.3.3. Soit f et g deux fonctions réelles continues sur $] -\infty, b]$, positives au voisinage de $-\infty$ et telles que $f(x) = o_{x \rightarrow +\infty}(g(x))$. On a alors :

- Si $\int_{-\infty}^b g(t) dt$ converge alors $\int_{-\infty}^b f(t) dt$ converge également.
- Si $\int_{-\infty}^b f(t) dt$ diverge alors $\int_{-\infty}^b g(t) dt$ diverge également.

12.4 Équivalence

Théorème 12.4.1. Soit f et g deux fonctions réelles continues sur $[a, +\infty[$, positives au voisinage de $+\infty$ et telles que $f(x) \sim_{x \rightarrow +\infty} g(x)$. On a alors que $\int_a^{+\infty} g(t) dt$ converge si et seulement si $\int_a^{+\infty} f(t) dt$ converge.

Exemple 12.4.2. Étudions l'intégrabilité de $f : \mathbb{R} \rightarrow \mathbb{R}$ sur $[-3, +\infty[$.

$t \mapsto \frac{7t-4}{t^2+1}$
On a que f est continue sur $[-3, +\infty[$ et $f(t) \sim_{t \rightarrow +\infty} \frac{7t}{t^2} = \frac{7}{t}$. Comme $t \mapsto \frac{7}{t}$ est positive au voisinage de l'infini, f l'est également. De plus, on reconnaît une intégrale de Riemann divergente car de coefficient $\alpha = 1 \leq 1$ donc par critère de comparaison l'intégrale $\int_{-3}^\infty \frac{7t-4}{t^2+1}$ est divergente.

De même en $-\infty$, on a :

Théorème 12.4.3. Soit f et g deux fonctions réelles continues sur $] -\infty, b]$, positives au voisinage de $-\infty$ et telles que $f(x) \sim_{x \rightarrow -\infty} g(x)$. On a alors que $\int_{-\infty}^b g(t) dt$ converge si et seulement si $\int_{-\infty}^b f(t) dt$ converge.

12.5 Convergence absolue

Le cas des fonctions positives nous donne donc des outils supplémentaires pour établir des convergences d'intégrales et l'on peut relier la convergence de

l'intégrale impropre d'une fonction de signe quelconque à la convergence de l'intégrale impropre d'une fonction positive dans la majorité des cas par la notion de convergence absolue.

Définition 12.5.1. Soit f une fonction continue sur $[a, +\infty[$. Si l'intégrale impropre $\int_a^{+\infty} |f(t)| dt$ est convergente on dit alors que l'intégrale $\int_a^{+\infty} f(t) dt$ est absolument convergente.

Cette notion trouve son utilité grâce au résultat suivant :

Théorème 12.5.2. *Une intégrale impropre absolument convergente est convergente.*

Remarque 12.5.3. La réciproque est fautive ! Il existe des fonctions dont l'intégrale impropre est convergente mais non absolument convergente.

Remarque 12.5.4. Cette notion s'adapte bien évidemment aux intégrales impropres en $-\infty$ et aux intégrales doublement impropres avec le même lien que l'absolue convergence implique la convergence.

On a de plus une inégalité entre ces deux intégrales :

Théorème 12.5.5. • Soit f une fonction continue sur $[a, +\infty[$ telle que $\int_a^{+\infty} f(t) dt$ soit absolument convergente. On a alors

$$\left| \int_a^{+\infty} f(t) dt \right| \leq \int_a^{+\infty} |f(t)| dt;$$

• Soit f une fonction continue sur $] -\infty, b]$ telle que $\int_{-\infty}^b f(t) dt$ soit absolument convergente. On a alors

$$\left| \int_{-\infty}^b f(t) dt \right| \leq \int_{-\infty}^b |f(t)| dt;$$

• Soit f une fonction continue sur \mathbb{R} telle que $\int_{-\infty}^{+\infty} f(t) dt$ soit absolument convergente. On a alors

$$\left| \int_{-\infty}^{+\infty} f(t) dt \right| \leq \int_{-\infty}^{+\infty} |f(t)| dt;$$

Chapitre 13

Systemes différentielles linéaires

Chapitre 14

Suites de variables aléatoires discrètes

Nous avons vu dans un précédent chapitre comment deux variables aléatoires discrètes interagissent l'une par rapport à l'autre et comment elles peuvent créer une nouvelle variable aléatoire. Nous allons nous intéresser dans ce chapitre au cas où il y a plus de deux variables aléatoires à considérer.

14.1 Généralités

14.1.1 Définitions

Définition 14.1.1. Soit $n \in \mathbb{N}^*$ et soit X_1, \dots, X_n n variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle le n -uplet $X = (X_1, \dots, X_n)$ vecteur aléatoire discret sur $(\Omega, \mathbb{A}, \mathbb{P})$.

Remarque 14.1.2. Un vecteur aléatoire est à valeur dans \mathbb{R}^n .

Exemple 14.1.3. On lance trois dés à 6 faces équilibrées, un bleu, un rouge et un jaune dont le résultat est matérialisé par respectivement les variables aléatoires B, R et J . On a donc que (B, R, J) est un vecteur aléatoire et par exemple

$$\mathbb{P}((B, R, J) = (1, 4, 3)) = \frac{1}{216}.$$

De manière similaire aux couples de variables aléatoires, un vecteur aléatoire permet de définir de nouvelles variables aléatoires :

Théorème 14.1.4. Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit g une fonction de $\prod_{i=1}^n X_i(\Omega) \rightarrow \mathbb{R}$.

On note $Z = g(X_1, \dots, X_n)$ l'application

$$\begin{aligned} Z &: \Omega \rightarrow \mathbb{R} \\ \omega &\mapsto g(X_1(\omega), \dots, X_n(\omega)) \end{aligned}$$

est une variable aléatoire.

On a les exemples classiques suivants :

Corollaire 14.1.5. Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On a alors que

$$\sum_{i=1}^n X_i, \prod_{i=1}^n X_i, \max(X_1, \dots, X_n) \text{ et } \min(X_1, \dots, X_n)$$

sont des variables aléatoires.

14.1.2 Indépendances

La notion d'indépendance est plus subtile lorsque l'on considère au moins trois variables aléatoires :

Définition 14.1.6. Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que les variables aléatoires sont mutuellement indépendantes si l'on a pour tout $(x_1, \dots, x_n) \in \prod_{i=1}^n X_i(\Omega)$:

$$\mathbb{P} \left(\bigcap_{i=1}^n [X_i = x_i] \right) = \prod_{i=1}^n \mathbb{P}(X_i = x_i).$$

Remarque 14.1.7. Cette définition implique par plusieurs utilisations de la formule des probabilités que dans le cas où les variables aléatoires sont mutuellement indépendantes, on a aussi l'égalité pour seulement certaines variables aléatoires ou également pour des évènements à la place des évènements élémentaires.

Définition 14.1.8. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit qu'il s'agit d'une suite de variables aléatoires mutuellement indépendantes si et seulement si pour tout ensemble fini I de \mathbb{N} , la famille $\{X_i\}_{i \in I}$ est une famille de variables aléatoires mutuellement indépendantes.

Exemple 14.1.9. En général, cette notion est induite par l'énoncé des exercices. Si on considère l'exemple précédent, l'indépendance mutuelle est automatique et c'est ce qui permet de pouvoir affirmer que pour tout $(i, j, k) \in \llbracket 1, 6 \rrbracket^3$,

$$\mathbb{P}((B, R, J) = (i, j, k)) = \frac{1}{6} \frac{1}{6} \frac{1}{6} = \frac{1}{216}.$$

Remarque 14.1.10. Être indépendant deux à deux est différent d'être mutuellement indépendant, on prendra garde à cela ! On a cependant que d'être mutuellement indépendant implique l'indépendance deux à deux.

Exemple 14.1.11. Soit X et Y deux variables aléatoires indépendantes suivant la même loi uniforme $\mathcal{U}(\llbracket 0, 1 \rrbracket)$ et soit Z la variable aléatoire valant 1 lorsque $X = Y$ et 0 lorsque $X \neq Y$. Après calculs, on obtient

- $\mathbb{P}(Z = 0) = \mathbb{P}(Z = 1) = \frac{1}{2}$;
- $\mathbb{P}([X = 0] \cap [Z = 0]) = \mathbb{P}([X = 0] \cap [Z = 1]) = \frac{1}{4}$;
- $\mathbb{P}([X = 1] \cap [Z = 0]) = \mathbb{P}([X = 1] \cap [Z = 1]) = \frac{1}{4}$;
- $\mathbb{P}([Y = 0] \cap [Z = 0]) = \mathbb{P}([Y = 0] \cap [Z = 1]) = \frac{1}{4}$;
- $\mathbb{P}([Y = 1] \cap [Z = 0]) = \mathbb{P}([Y = 1] \cap [Z = 1]) = \frac{1}{4}$;
- $\mathbb{P}([X = 0] \cap [Y = 0] \cap [Z = 0]) = \frac{1}{4}$.

On remarque X et Y sont indépendants par hypothèses et on a par analyse des calculs précédents que X et Z sont indépendants et de même Y et Z sont indépendants. X , Y et Z sont donc deux à deux indépendants.

Cependant on que

$$\mathbb{P}([X = 0] \cap [Y = 0] \cap [Z = 0]) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(X = 0)\mathbb{P}(Y = 0)\mathbb{P}(Z = 0).$$

On a donc X , Y et Z ne sont pas mutuellement indépendants.

Le dernier résultat qu'il nous reste à considérer pour la notion d'indépendance est le lemme suivant :

Lemme 14.1.12 (Lemme des coalitions). *Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ où les variables aléatoires sont mutuellement indépendantes. Toute variable aléatoire fonction de X_1, \dots, X_p est indépendante de toute variable aléatoire fonction de X_{p+1}, \dots, X_n .*

Exemple 14.1.13. Soit X_1, X_2, X_3 et X_4 quatre variables mutuellement indépendantes. On a alors que $e^{X_1 - X_4}$ est indépendante de $X_2 \ln(X_3^2 + 1)$.

14.2 Lois usuelles

Théorème 14.2.1. Soit $p \in]0, 1[$. (X_1, \dots, X_n) un vecteur aléatoire mutuellement indépendant où pour tout $i \in \llbracket 1, n \rrbracket$ X_i est une loi de Bernoulli de paramètre p . On a alors que $X_1 + \dots + X_n$ suit une loi binomial de paramètre n et p .

Démonstration. On pose $X = X_1 + \dots + X_n$. On a $X(\Omega) = \llbracket 0, n \rrbracket$ et pour tout $k \in X(\Omega)$,

$$\begin{aligned}
 \mathbb{P}(X = k) &= \sum_{\substack{(i_1, \dots, i_n) \in \llbracket 0, 1 \rrbracket^n \\ \sum_{j=1}^n i_j = k}} \prod_{j=1}^n \mathbb{P}(X_j = i_j) \\
 &= \sum_{\substack{(i_1, \dots, i_n) \in \llbracket 0, 1 \rrbracket^n \\ \sum_{j=1}^n i_j = k}} p^k (1-p)^{n-k} \\
 &= p^k (1-p)^{n-k} \sum_{\substack{(i_1, \dots, i_n) \in \llbracket 0, 1 \rrbracket^n \\ \sum_{j=1}^n i_j = k}} 1 \\
 &= p^k (1-p)^{n-k} \text{Card} \left((i_1, \dots, i_n) \in \llbracket 0, 1 \rrbracket^n \mid \sum_{j=1}^n i_j = k \right) \\
 &= \binom{n}{k} p^k (1-p)^{n-k}
 \end{aligned}$$

car $\text{Card} \left((i_1, \dots, i_n) \in \llbracket 0, 1 \rrbracket^n \mid \sum_{j=1}^n i_j = k \right)$ correspond au nombre de manière de choisir k fois le nombre 1 parmi les n possibilités. \square

Remarque 14.2.2. On retrouve à l'aide de ce théorème le résultat 9.2.12 car sensiblement sommer n lois de Bernoulli avec m lois de Bernoulli mutuellement indépendantes revient à en sommer $n+m$ et donne donc une loi binomial $\mathcal{B}(n+m, p)$.

On peut étendre ce résultat à :

Propriété 14.2.3. Soit $p \in]0, 1[$. Soit (X_1, \dots, X_m) des variables aléatoires mutuellement indépendantes où pour tout $i \in \llbracket 1, m \rrbracket$, $X_i \hookrightarrow \mathcal{B}(n_i, p)$. On a alors que

$$X_1 + \dots + X_m \hookrightarrow \mathcal{B} \left(\sum_{i=1}^m n_i, p \right).$$

Remarque 14.2.4. Ce résultat se démontre par application du cas $m = 2$ par récurrence.

Exemple 14.2.5. Quatre personnes jouent à un jeu qui consiste à avoir le plus de "pile" possible en commun à l'aide d'une pièce équilibrée. La première personne lance la pièce une fois, la deuxième deux fois, la troisième trois fois et la quatrième quatre fois.

En modélisant le nombre de pile obtenu par la personne numéro i par la variable aléatoire X_i , on a que X_i suit une loi binomiale $\mathcal{B}(i, \frac{1}{2})$.

On note S la variable aléatoire correspondant au nombre de "pile" total et on observe que $S = X_1 + X_2 + X_3 + X_4$.

Par la méthodologie employée X_1, X_2, X_3 et X_4 sont mutuellement indépendantes donc on a $S \hookrightarrow \mathcal{B}(1 + 2 + 3 + 4, \frac{1}{2})$ soit

$$S \hookrightarrow \mathcal{B}\left(10, \frac{1}{2}\right).$$

On a le même prolongement pour loi de Poisson du résultat 9.2.13

Propriété 14.2.6. Soit (X_1, \dots, X_m) des variables aléatoires mutuellement indépendantes où pour tout $i \in \llbracket 1, m \rrbracket$, $X_i \hookrightarrow \mathcal{P}(\lambda_i)$. On a alors que

$$X_1 + \dots + X_m \hookrightarrow \mathcal{P}\left(\sum_{i=1}^m \lambda_i\right).$$

Exemple 14.2.7. Une étude montre que trois guichets permettent de traiter respectivement X, Y et Z demandes de clients en une heure où $X \hookrightarrow \mathcal{P}(8)$, $Y \hookrightarrow \mathcal{P}(11)$ et $Z \hookrightarrow \mathcal{P}(7)$. Le nombre total de demandes de clients traitées en une heure par le service est donc représenté par la variable aléatoire $S \hookrightarrow \mathcal{P}(8 + 11 + 7)$ soit

$$S \hookrightarrow \mathcal{P}(26).$$

14.3 Moments

Pour les moments, on a besoin de connaître les résultats suivants :

Propriété 14.3.1. Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que pour tout $i \in \llbracket 1, n \rrbracket$ X_i admet une espérance. On a alors

$$\mathbb{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}(X_i).$$

Pour la variance l'indépendance mutuelle est fondamentale :

Propriété 14.3.2. Soit $n \in \mathbb{N}^*$. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret mutuellement indépendant sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que pour tout $i \in \llbracket 1, n \rrbracket$ X_i admet un moment d'ordre 2. On a alors

$$\mathbb{V} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \mathbb{V}(X_i).$$

Exemple 14.3.3. Soit (X_1, \dots, X_n) un vecteur aléatoire de loi géométrique de paramètre p mutuellement indépendant. On a alors

$$\mathbb{E} \left(\sum_{i=1}^n X_i \right) = \frac{n}{p} \quad \text{et} \quad \mathbb{V} \left(\sum_{i=1}^n X_i \right) = \frac{n(1-p)}{p^2}.$$

Remarque 14.3.4. Dans le cas où les variables aléatoires ne sont pas mutuellement indépendantes toutes les covariances possibles interviennent dans le cas calcul. Par exemple pour $n = 3$, on a

$$\mathbb{V}(X+Y+Z) = \mathbb{V}(X)+\mathbb{V}(Y)+\mathbb{V}(Z)+2Cov(X, Y)+2Cov(X, Z)+2Cov(Y, Z).$$

Deuxième partie
Quatrième semestre

