

Table des matières

I	Rappels séries numériques	5
I	Généralités	5
II	Cas des séries à terme général réel positif	6
III	Séries de terme général quelconque	9
II	Familles sommables	13
I	Dénombrabilité	13
II	Familles sommables de réels positifs	15
III	Familles de complexes.	17
IV	Applications	17
III	Espaces vectoriels normés	19
I	Généralités	19
II	Suites d'un EVN	23
II.1	Séries d'un EVN	25
II.2	Application dans une algèbre normée	26
II.3	Familles sommables d'un K-ev	27
II.4	Suites extraites	27
III	Topologie	28
III.1	Ouverts	28
III.2	Adhérence	29
III.3	Relativité vaguement générale quand même	32
III.4	Compacts	32
IV	Limites de fonctions	35
IV.1	Définitions	35
IV.2	Opérations sur les limites	37
IV.3	Continuité de fonctions	38
IV.4	Propriétés globales des fonctions continues	39
IV.5	Fonctions lipschitziennes	40
IV.6	Continuité des applications linéaires	42
IV.7	Continuité des applications p-linéaires	44
V	Connexité par arcs	45
VI	Théorème d'équivalence des normes	47
IV	Réduction des endomorphismes	49
I	Généralités	49
II	Polynôme caractéristique	52
II.1	Coefficients du polynôme caractéristique	52
III	Trigonalisation	54
IV	Endomorphismes nilpotents	56
V	Polynômes d'endomorphismes	56

V Suites et séries de fonctions	61
I Suites de fonctions	61
II Séries de fonctions	62
III Continuité d'une limite uniforme	64
IV Intégration et dérivation	67
V Résultats de densité	70
VI Séries entières	73
I Généralités	73
II Calculs de rayon de convergence	74
III Opérations sur les séries entières	75
IV Continuité de la somme d'une série entière sur les complexes	76
V Développements en séries entières	77
V.1 Généralités	77
V.2 Rappels sur Taylor	78
V.3 Développements en série entière découlant de l'exponentielle sur les réels	78
V.4 Développements en série entière découlant de la série géométrique	79
V.5 Développement de l'autre là	79
VI Séries entières dans une algèbre de dimension finie	81
VII Arithmétique dans un anneau euclidien	83
I Généralités	83
II Opérations sur les idéaux	84
III PGCD et PPCM	85
IV Application à la réduction des endomorphismes	88
VIII Intégration	91
I Intégrale sur un segment (rappels)	91
II Intégration sur un segment d'une fonction vectorielle	93
III Intégration sur un intervalle quelconque	95
III.1 Intégration sur un intervalle semi-ouvert	95
III.2 Intégration sur un intervalle ouvert	97
III.3 Intégration des fonctions positives	97
III.4 Intégrales de référence de Riemann	98
III.5 Intégrabilité	100
III.6 Propriétés des intégrales impropres	101
III.7 Relation de Chasles	101
III.8 Fonction dépendant de la borne supérieure	101
III.9 Intégration par parties	102
III.10 Changement de variable	102
III.11 Intégration des relations de comparaisons	103
IV Intégrale dépendant d'un paramètre	104
IV.1 Cas d'un paramètre entier	104
IV.2 Cas d'un paramètre réel	105
IV.3 Dérivabilité d'une intégrale à paramètres	107
IX Probabilités	109
I Espace probabilisé	109
I.1 Univers, tribu	109
I.2 Probabilités	110
I.3 Univers finis, équiprobabilité, dénombrement	113

I.4	Probabilité conditionnelle, indépendance	114
II	Variables aléatoires discrètes	118
II.1	Définition d'une variable aléatoire discrète	118
II.2	Loi d'une variable aléatoire discrète	119
II.3	Indépendance et variables aléatoires discrètes	120
II.4	Espérance d'une variable aléatoire discrète réelle ou complexe	121
II.5	Variance d'une variable aléatoire discrète réelle, écart type	126
II.6	Covariance, coefficient de corrélation linéaire	128
III	Lois usuelles	130
III.1	Lois usuelles finies	130
III.2	Lois usuelles infinies	132
IV	Fonctions génératrices	134
V	Couples de variables aléatoires réelles	138
V.1	Loi conjointe, lois marginales	138
V.2	Lois conditionnelles d'un couple	139
VI	Résultats probabilistes asymptotiques	139
VI.1	Inégalités de Markov et Bienaymé-Tchebychev	139
VI.2	Loi faible des grands nombres	140
X	Espaces préhilbertiens	143
I	Préliminaires	143
II	Orthogonaux	144
III	Projections orthogonales	148
III.1	Généralités	148
III.2	Calcul pratique du projeté orthogonal	149
III.3	Exemples d'utilisation du projeté orthogonal	149
IV	Endomorphismes particuliers des espaces euclidiens	151
V	Isométries vectorielles	152
V.1	Isométries directes, indirectes	154
V.2	Orientation de l'espace	154
VI	Matrices orthogonales	155
VI.1	Généralités	155
VI.2	Classification des matrices orthogonales du plan	156
VI.3	Classification des matrices orthogonales d'un espace de dimension 2	157
VI.4	Classification des matrices orthogonales d'un espace euclidien orienté de dimension 3	158
VII	Réduction des isométries dans un espace euclidien	160
VIII	Les endomorphismes auto-adjoints	161
IX	Endomorphismes auto-adjoints et formes bilinéaires symétriques	163
X	Hors-programme	165
X.1	Racine carrées des matrices symétriques définies positives	165
X.2	Matrices de Gramm d'une famille de vecteurs	165
X.3	Familles isogonales	167
XI	Equations différentielles linéaires	169
I	Définitions	169
II	Généralités sur l'équation du premier ordre	171
III	Topo sur les exponentielles	176
III.1	Rappel sur les familles sommables	176
III.2	Exponentielle dans une algèbre de dimension finie	176

XI	Calcul différentiel	179
	.1 Conditions du chapitre	179
	.2 Rappels sur la continuité	179
I	Applications différentiables	180
	I.1 Dérivée selon un vecteur	180
	I.2 Dérivées partielles	181
	I.3 Différentielle	182
	I.4 Matrice jacobienne	185
II	Opérations sur les applications différentiables	186
	II.1 Différentielle d'une combinaison linéaire d'applications différentiables	186
	II.2 Différentielle de $B(f,g)$ où B est bilinéaire et f et g sont deux applications différentiables	186
	II.3 Différentielle d'une composée d'applications différentiables	188
	II.4 Règle de la chaîne	189
III	Applications de classe \mathcal{C}^1	190
	III.1 Définition et caractérisation	190
	III.2 Opérations algébriques sur les applications de classe \mathcal{C}^1	192
IV	Dérivées partielles d'ordre supérieur	193
	IV.1 Dérivées partielles d'ordre k	193
	IV.2 Théorème de Schwarz	193
	IV.3 Opérations sur les applications de classe \mathcal{C}^k	194
V	Cas des applications numériques	194
	V.1 Gradient	194
	V.2 Dérivée le long d'un arc	195
	V.3 Intégration le long d'un arc d'une application de classe \mathcal{C}^1	197
	V.4 Vecteurs tangents à une partie d'un espace normé de dimension finie	198
	V.5 Optimisation	199
VI	Hors programme	206
	VI.1 Fonctions k fois différentiables	206
	VI.2 Démo hors-programme : fonction constante différentiable	207
	VI.3 Différentielle d'un inverse	207
XII	Groupes et anneaux	209
I	Rappels de sup	209
II	Groupes engendrés par une partie non-vide	212

Chapitre I

Rappels séries numériques

I Généralités

Définition : Convergence série

Avec $u_n \in \mathbb{K}^{\mathbb{N}}$ on dit que la série $(\sum u_n)$ converge si la suite $(\sum_{k=1}^n u_k)_{n \in \mathbb{N}}$ converge. En cas de convergence, on définit la somme de la série $\sum_{n=0}^{\infty} u_n = \lim_{n \rightarrow +\infty} \sum_{k=1}^n u_k$ seulement si la série est convergente. $(S_n) = \left(\sum_{k=0}^n u_k\right)$ est la suite des sommes partielles de la série. En cas de convergence, $(R_n) = \left(\sum_{k=n+1}^{\infty} u_k\right)$ est la suite des restes (d'ordre n) de la série. R_n tend vers 0 par définition.

Théorème : de divergence grossière

$(u_n) \in \mathbb{K}^{\mathbb{N}}$, si $(\sum u_n)$ converge, alors $(u_n) \rightarrow 0$.

Preuve

$\forall n \in \mathbb{N}, u_n = S_n - S_{n-1}$ donc quand (u_n) tend vers une limite l , alors u_n tend vers 0.

Théorème : Séries géométriques

On prend $a \in \mathbb{C}$, la série $(\sum a^n)$ converge si, et seulement si, $|a| < 1$ et alors : $\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$

Preuve

Si $|a| \geq 1$ alors la série diverge grossièrement. Si $|a| < 1$ en particulier $a \neq 1$, donc $\forall n \in \mathbb{N}, \sum_{k=0}^n a^k = \frac{1 - a^{n+1}}{1 - a}$

Théorème : Dominos ou série télescopique

$(a_n) \in \mathbb{K}^{\mathbb{N}}$, la série $(\sum a_n - a_{n+1})$ converge si, et seulement si, la suite (a_n) converge.

Preuve

Soit $n \in \mathbb{N}$, $\sum_{k=0}^n a_k - a_{k+1} = a_0 - a_1 + a_1 - a_2 + a_2 \dots = a_0 + a_{n+1}$. Alors la suite des sommes partielles converge si, et seulement si, a_n converge. En cas de convergence, on a $\sum_{n=0}^{\infty} a_n - a_{n+1} = a_0 - \lim_{n \rightarrow +\infty} a_n$

Proposition : Opérations sur les séries convergentes

$(u_n), (v_n) \in \mathbb{K}^{\mathbb{N}}$ termes généraux de séries convergentes, alors $(\sum u_n + v_n)$ converge. Pour tout $\lambda \in \mathbb{K}$, alors $(\sum \lambda u_n)$ converge aussi.

Exemple I.1. Avec $\sum u_n$ convergente et $\sum v_n$ divergente, quel est le comportement de $\sum u_n + v_n$? Elle diverge. Prouvons-le par l'absurde. Supposons que $\sum (u_n + v_n)$ converge. On a que $\forall n \in \mathbb{N}, v_n = (u_n + v_n) - u_n$. Donc $\sum v_n$ converge. D'où la contradiction.

II Cas des séries à terme général réel positif

Définition : Série à terme général réel positif

Soit $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$, alors $(\sum_{k=0}^n u_k)$ est croissante. Elle converge si, et seulement si, $(\sum_{k=0}^n u_k)$ est majorée, et $(\sum u_n)$ diverge si, et seulement si, $(\sum_{k=0}^n u_k) \rightarrow +\infty$.

Théorème : Critères de convergence d'une série

Si $(u_n), (v_n) \in \mathbb{R}_+^{\mathbb{N}}$:

- Si $\forall n \in \mathbb{N}, v_n \leq u_n$, alors si $\sum u_n$ converge alors $\sum v_n$ converge. De même, si $\sum u_n$ diverge, alors $\sum v_n$ aussi (critère de majoration positif)
- Si $u_n = o(v_n)$, alors si $\sum v_n$ converge, alors $\sum u_n$ converge (critère de domination positif)
- Si $u_n \sim v_n$, alors $\sum v_n$ et $\sum u_n$ sont de même nature (critère d'équivalent positif)

Théorème : Séries de Riemann

Pour $\alpha \in \mathbb{R}$, alors $(\sum \frac{1}{n^\alpha})$ converge si, et seulement si, $\alpha > 1$.

Preuve

Avec $\alpha \in \mathbb{R}, \alpha \neq 1$, on a que $(n+1)^{\alpha+1} = n^{\alpha+1}(1 + \frac{1}{n})^{\alpha+1} = n^{\alpha+1}(1 + \frac{\alpha+1}{n} + o(\frac{1}{n})) = n^{\alpha+1} + (\alpha+1)n^\alpha + o(n^\alpha)$.
 $(n+1)^{\alpha+1} - n^{\alpha+1} \sim (\alpha+1)n^\alpha$, $(n^\alpha) > 0$ donc $\sum n^\alpha$ et $\sum (n+1)^{\alpha+1} - n^{\alpha+1}$ sont de même nature. Alors la suite des $\sum n^\alpha$ converge si, et seulement si, la somme des $n^{\alpha+1}$ converge, donc quand $\alpha+1 < 0$, donc $\alpha < -1$.
 Dans le cas de $\alpha = -1$...

Exemples II.1. 1. Avec $u_n = \frac{(n^2 + n + 3)^{2/3}}{n(n + \sqrt{n})^{3/2}}$. C'est un terme général positif. $(u_n) \sim \frac{n^{4/3}}{n \times n^{3/2}}$ (ce qu'on prouve en mettant en facteur le prépondérant sur le dénominateur et le numérateur).

$$u_n \sim \frac{n^{4/3}}{n \times n^{3/2}} \sim n^{-7/6} \sim \frac{1}{n^{7/6}}.$$

- $\sum \frac{1}{n^{7/6}}$ est une série de Riemann convergente, donc par critère d'équivalent positif, $\sum u_n$ converge.
2. $u_n = th(n) + \frac{1}{n}$, d'un côté $\sum th(n)$ diverge grossièrement et accessoirement $\sum \frac{1}{n}$ diverge, donc par critère de majoration positive, $\sum th(n) + \frac{1}{n}$ diverge.
3. $u_n = th(\frac{1}{n}) + \ln(1 - \frac{1}{n})$ or $th(x) = \frac{1}{2}(1 + x + o(x) - 1 - x + o(x)) = x + o(x)$ et $\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{6} + o(x^3)$ en 0. Donc en l'infini $u_n = \frac{1}{n} + o(\frac{1}{n^2}) - \frac{1}{n} + -\frac{1}{2n^2} + o(\frac{1}{n^2}) = -\frac{1}{2n^2} + o(\frac{1}{n^2})$. Donc par critère d'équivalent positif avec une série convergente, $\sum u_n$ converge.
4. $u_n = e - \left(1 + \frac{1}{n}\right)^n$, et $\left(1 + \frac{1}{n}\right)^n = \exp(n \ln(1 + \frac{1}{n})) = \exp(n(\frac{1}{n} - \frac{1}{2n^2} +))$

Théorème : Critère de d'Alembert

Soit $(u_n) \in \mathbb{R}_+^{*\mathbb{N}}$. Si $\left(\frac{u_{n+1}}{u_n}\right) \rightarrow l \in \mathbb{R}$, alors :

- Si $l < 1$ alors $\sum u_n$ converge.
- Si $l > 1$ alors $\sum u_n$ diverge grossièrement.
- Si $l = 1$ alors on ne peut rien dire.

Preuve

C'est le rapport entre un terme et son successeur, donc c'est la raison locale en comparaison à une série géométrique. Si $l < 1$, prenons $\varepsilon = \frac{1-l}{2}$. On peut fixer n_0 tel que $\forall n \geq n_0, \frac{u_{n+1}}{u_n} < l + \varepsilon \Rightarrow \frac{u_{n+1}}{u_n} < \frac{1+l}{2}$ donc pour $n \geq n_0 : u_n \leq \left(\frac{1+l}{2}\right)^{n-n_0} u_{n_0}$. $\frac{1+l}{2} < 1$ donc $\sum u_n$ converge par critère de majoration positif.

Si $l > 1$, on prend $\varepsilon = \frac{l-1}{2}$, $\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow u_n \geq \left(\frac{1+L}{2}\right)^{n-n_0} u_{n_0}$

Proposition : Comparaisons séries-intégrales

Si $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ fonction réelle positive décroissante, continue par morceaux alors : $\forall n \in \mathbb{N}^*, \int_n^{n+1} f(t)dt \leq f(n) \leq \int_{n-1}^n f(t)dt$.

Théorème : Comparaisons séries-intégrales

Si $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ fonction positive décroissante continue par morceaux, alors $(\sum f(n))$ est de même nature que la suite $(\int_0^n f(t)dt)$

Preuve

Montrer que la suite des sommes partielles converge c'est que la suite est majorée, et on doit utiliser la croissance de l'intégrale.

Remarque II.1. L'encadrement de $f(n)$ par les intégrales $\int_{n-1}^n f(t)dt, \int_n^{n+1} f(t)dt$ peut être exploité dans d'autres cadres, par exemple dans le cas où la série $\sum f(n)$ converge, on peut écrire pour $n, p \in \mathbb{N}^*, n < p$, $\int_n^{p+1} f(t)dt \leq \sum_{k=n}^p f(k) \leq \int_{n-1}^p f(t)dt$, et par passage à la limite quand $p \rightarrow +\infty$, $\lim_{p \rightarrow +\infty} \int_n^p f(t)dt \leq \sum_{k=n}^{\infty} f(k) \leq \lim_{p \rightarrow +\infty} \int_{n-1}^p f(t)dt$.

Exemple II.1. On peut retrouver par comparaisons séries-intégrales que $\sum \frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha < 1$ (ce qui n'est valable que pour $\alpha > 0$). Soit $\alpha \in \mathbb{R}, \alpha > 1, R_n = \sum_{k=n+1}^{\infty} \frac{1}{k^\alpha}$. On cherche un équivalent de R_n .

Alors : $t \mapsto \frac{1}{t^\alpha}$ est positive décroissante. Donc : $\forall n, p \in \mathbb{N}, 2 \leq n < p, \int_n^{p+1} \frac{dt}{t^\alpha} \leq \sum_{k=n}^p f(k) \leq \int_{n-1}^p \frac{dt}{t^\alpha}$. On a que $\int_a^b \frac{dt}{t^\alpha} = \left[-\frac{1}{-\alpha+1} t^{-\alpha+1} \right]_a^b$.

Par passage à la limite quand p tend vers l'infini, on a pour $n \in \mathbb{N}$: $\frac{1}{(\alpha-1)(n+1)^{\alpha-1}} \leq \sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \leq \frac{1}{(\alpha-1)n^{\alpha-1}}$

Et, les suites encadrantes étant positives et équivalentes à $\frac{1}{n^{\alpha-1}}$ donc par théorème d'encadrement des équivalents, $R_n \sim \frac{1}{(\alpha-1)n^{\alpha-1}}$

On cherche un équivalent de $\left(\sum_{k=1}^n \frac{1}{n} \right)$. Par séries-intégrales, on obtient que $\left(\sum_{k=1}^n \frac{1}{n} \right) \sim (\log(n))$

Théorème : Sommation des ordres de grandeur

Avec $(a_n), (b_n)$ deux suites réelles positives :

- Si $b_n = O(a_n)$: si $\sum a_n$ converge, alors $\sum b_n$ converge et $\sum_{k=n+1}^{+\infty} b_k = O\left(\sum_{k=n+1}^{+\infty} a_k\right)$; si $\sum a_n$ diverge, alors $\sum_{k=0}^n b_k = O\left(\sum_{k=0}^n a_k\right)$.
- Si $b_n = o(a_n)$: si $\sum a_n$ converge, alors $\sum b_n$ converge et $\sum_{k=n+1}^{+\infty} b_k = o\left(\sum_{k=n+1}^{+\infty} a_k\right)$; si $\sum a_n$ diverge, alors $\sum_{k=0}^n b_k = o\left(\sum_{k=0}^n a_k\right)$.
- Si $b_n \sim (a_n)$: si $\sum a_n$ converge, alors $\sum b_n$ converge et $\sum_{k=n+1}^{+\infty} b_k \sim \sum_{k=n+1}^{+\infty} a_k$; si $\sum a_n$ diverge, alors $\sum b_n$ diverge et $\sum_{k=0}^n b_k \sim \sum_{k=0}^n a_k$.

Preuve

Prenons $(u_n), (v_n) \in \mathbb{R}_+^{\mathbb{N}}$, avec $(u_n) = o(v_n)$. On suppose que $\sum v_n$ converge.

Alors $\sum u_n$ converge par critère de domination positif. Donc $(u_n) = (v_n \varepsilon_n)$ avec $(\varepsilon_n) \rightarrow 0$. Soit $\varepsilon > 0$, on peut fixer n_0 tel que $\forall n \geq n_0, \varepsilon_n < \varepsilon$

Donc pour $n_0 \leq n \leq p$: $\sum_{k=p}^n u_k \leq \left(\sum_{k=n}^p v_k \right) \varepsilon$.

Donc par passage à la limite quand $p \rightarrow +\infty$, $\sum_{k=n}^{\infty} u_k \leq \varepsilon \sum_{k=n}^{\infty} v_k$

Si (v_n) n'est pas la suite nulle, $\sum v_k > 0$. On a montré que : $\forall \varepsilon \in \mathbb{R}_+^*, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \frac{\sum_{k=n}^{\infty} u_k}{\sum_{k=n}^{\infty} v_k} < \varepsilon$

$(u_n), (v_n) \geq 0$, si $\sum v_n$ diverge alors : on pose aussi $(u_n) = (\varepsilon v_n)$ avec $\varepsilon_n \rightarrow 0$

Soit $n \in \mathbb{N}$, $\sum_{k=0}^n u_k = \sum_{k=0}^n v_k \varepsilon_k$

Soit $\varepsilon \in \mathbb{R}_+^*$, on fixe n_0 tel que $\forall n \geq n_0, \varepsilon_n \leq \varepsilon$.

Pour $n \geq n_0$, $\sum u_k = \sum_{k=0}^{n_0} \varepsilon_k v_k + \sum_{k=n_0}^n \varepsilon_k v_k$

On peut supposer $\sum v_k > 0$, alors : $\frac{\sum_{k=0}^n u_k}{\sum_{k=0}^n v_k} \leq \frac{\sum_{k=0}^{n_0} \varepsilon_k v_k}{\sum_{k=0}^n v_k} + \varepsilon$

$\left(\frac{\sum_{k=0}^{n_0} \varepsilon_k v_k}{\sum_{k=0}^n v_k} \right) \rightarrow 0$ donc on peut fixer $n_1 \in \mathbb{N}, \forall n \geq n_1, \frac{\sum_{k=0}^{n_0} \varepsilon_k v_k}{\sum_{k=0}^n v_k} < \varepsilon$

Alors $\forall \varepsilon \in \mathbb{R}_+^*, \exists N = \max(n_0, n_1), \forall n \geq N, \frac{\sum_{k=0}^n u_k}{\sum_{k=0}^n v_k} < 2\varepsilon$

Pour les équivalents : $u_n \sim v_n \Rightarrow (u_n) = (v_n) + o(v_n) = (v_n) + (w_n)$ avec $\frac{w_n}{v_n} \rightarrow 0$.

Dans le cas convergent : $\left(\sum_{k=n}^{\infty} u_k \right) = \left(\sum_{k=n}^{\infty} v_k \right) + \left(\sum_{k=n}^{\infty} w_k \right)$ mais comme $w_n = o(v_n)$, on a bien l'équivalence.

Pour les O : $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leq M v_n$. Il suffit de majorer les sommes.

III Séries de terme général quelconque

Ici, $(u_n) \in \mathbb{K}^{\mathbb{N}}$

Définition : Absolue convergence

Soit $(u_n) \in \mathbb{C}^{\mathbb{N}}$, on dit que $\sum u_n$ converge absolument si $\sum |u_n|$ converge.

Théorème : Conséquences de l'absolue convergence

Toute série absolument convergente est convergente

Preuve

Dans le cas réel, en utilisant que $u_n^+ \leq |u_n|$ et $u_n^- \leq |u_n|$ donc les deux convergent par critère de majoration, comme $\sum u_n = \sum u_n^+ - u_n^-$.

Dans le cas complexes, les séries des parties réelles et imaginaires sont majorées par les modules.

Remarque III.1. Pour une série absolument convergente, $\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|$. Comme on a $\forall p \in \mathbb{N}, \left| \sum_{n=0}^p u_n \right| \leq \sum_{n=0}^p |u_n|$, on passe à la limite et on a la propriété.

Exemple III.1. Quelques exemples de séries absolument convergentes sont $\sum \frac{(-1)^n}{n^2}$, $\sum \frac{\sin(n)}{n^2}$, mais il existe cependant des séries qui soient convergentes sans l'être absolument, telles que $\sum \frac{(-1)^n}{n}$ ou $\sum \frac{\sin(n)}{n}$.

Théorème : Théorème spécial des séries alternées

Soit (a_n) une suite réelle positive décroissante de limite nulle. Alors $\sum (-1)^n a_n$ converge et $\forall n \in \mathbb{N}, \left| \sum_{k=n}^{+\infty} (-1)^k a_k \right| \leq a_n$

Preuve

Pour montrer la convergence, considérons la suite (S_n) des sommes partielles : $\forall n \in \mathbb{N}, S_n = \sum_{k=0}^n (-1)^k u_k$;

Alors pour tout $n \in \mathbb{N}$ on a :

- $S_{2n+2} - S_{2n} = u_{2n+2} - u_{2n+1} \leq 0$ donc la suite (S_{2n}) est décroissante ;
- $S_{2n+3} - S_{2n+1} = -u_{2n+3} + u_{2n+2} \geq 0$ donc la suite (S_{2n+1}) est croissante ;
- $S_{2n} - S_{2n+1} = u_{2n+1} \rightarrow 0$ en l'infini

Donc les suites $(S_{2n}), (S_{2n+1})$ sont adjacentes : elles convergent vers une même limite l qui est donc la limite de (S_n) . Donc la série $\sum (-1)^n u_n$ est convergente, de somme l . Et de plus, l vérifie :

$$\forall n, m \in \mathbb{N}, S_{2m+1} \leq l \leq S_{2n}$$

Pour la majoration de $|R_n|$ on procède par disjonction de cas :

- Si n est pair : on a par propriétés des suites adjacentes que : $S_{n+1} \leq l \leq S_n$. Et donc $R_n = l - S_n \in [S_{n+1} - S_n, 0] = [-u_{n+1}, 0]$. Donc $|R_n| = -R_n \leq u_{n+1}$.
- Si n est impair : on a de même que : $S_n \leq l \leq S_{n+1}$ et donc $R_n \in [0, u_{n+1}]$. Donc $|R_n| = R_n \leq u_{n+1}$.

Exemple III.2. On a ainsi cet exemples de série semi-convergente, pour $0 < \alpha \leq 1$: $\sum \frac{(-1)^n}{n^\alpha}$. Pour $\alpha \leq 0$, on a divergence grossière, et pour $\alpha > 0$, on a $(\frac{1}{n^\alpha})$ positive décroissante tendant vers 0, donc le critère spécial des séries alternées dit que ces séries sont convergentes.

On en tire aussi cette inégalité pour $\alpha > 0$: $\forall n \in \mathbb{N}, \left| \sum_{k=n}^{+\infty} \frac{(-1)^k}{k^\alpha} \right| \leq \frac{1}{n^\alpha}$

Remarque III.2. Les critères des séries positives ne s'appliquent plus pour ces séries.

Pour le critère des équivalents : on peut prendre la série de terme général $u_n = \frac{(-1)^n}{\sqrt{n}} + \frac{1}{n}$. On a que $u_n \sim \frac{(-1)^n}{\sqrt{n}}$, dont la série associée est convergente. Mais $\sum u_n$ n'est pas convergente à cause de la somme harmonique.

Pour le critère de domination : on peut prendre la série de terme général $\frac{1}{n}$, dominée par $\frac{(-1)^n}{\sqrt{n}}$. Si le terme général qui domine est celui d'une série convergente, on sait que $\sum \frac{1}{n}$ n'est pas convergente.

On a donc un nouveau critère de domination positive : Si $(u_n), (v_n)$ termes généraux de séries, qu'on a que $v_n > 0$, que $\sum v_n$ converge et que $u_n = o(v_n)$, alors on aura bien que u_n est absolument convergente.

Le critère de d'Alembert subit une petite modification : On considère (u_n) terme général d'une série, si $\left| \frac{u_{n+1}}{u_n} \right|$ admet une limite l , on a que $\sum u_n$ est absolument convergente si $l < 1$ et diverge grossièrement si $l > 1$. Toujours pas de lois s'appliquant sur le cas où $l = 1$ cependant.

Pour le théorème de sommation des ordres de grandeur, on a pour le o et le O besoin que $v_n > 0$ pour que le théorème fonctionne, le signe de u_n n'a pas besoin d'être contrôlé. "La suite de référence doit être positive pour appliquer le théorème." Pour les équivalents, on a que les deux sont déjà du même signe.

Exemple III.3. On fait les séries de Bertrand, alternées et non.

Pour les séries de Bertrand : on finit par trouver une convergence si, et seulement si, $\alpha > 1$ ou $\alpha = 1, \beta > 1$

Pour les séries de Bertrand alternée : on étudie $\sum \frac{(-1)^n}{n^\alpha \ln(n)^\beta}$. On commence par regarder les divergences grossières : $\sum u_n$ diverge si $\alpha < 0$ ou si $\alpha = 0$ et $\beta \leq 0$.

Les cas de convergence absolue sont les cas où $\alpha > 1$ ou $\alpha = 1, \beta > 1$ d'après ce qui venait avant.

On prend le cas de $\alpha = 0$ et $\beta > 0$, alors u_n tend vers 0, et elle est décroissante en l'infini, donc par critère des séries alternées on a la convergence.

Si $0 < \alpha < 1$, alors $\exists \alpha' > 0$ tel que $|u_n| \ll \frac{1}{n^{\alpha'}}$. Ainsi, u_n est positive et décroissante à partir d'un certain rang, et tend vers 0, donc en utilisant le critère des séries alternées on a la convergence.

Si $\alpha = 1$, $\sum u_n$ est semi-convergente.

Exemple III.4. Exemple de série avec un terme en $(-1)^n$: on pose $u_n = \frac{(-1)^n}{n^\alpha + (-1)^n}$ terme général d'une série.

Les cas de divergence grossière sont si $\alpha < 0$ et on a des impossibilités si $\alpha = 0$.

Si $\alpha > 1$, alors la série converge absolument puisqu'on trouve un équivalent qui converge par Riemann à $|u_n|$

Prenons les cas $0 < \alpha \leq 1$: on va faire un "éclatement des termes" (un DL quoi) : $\frac{(-1)^n}{n^\alpha} \left(\frac{1}{1 + \frac{(-1)^n}{n^\alpha}} \right) = \frac{(-1)^n}{n^\alpha} \left(1 + \frac{(-1)^n}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right) \right) = \frac{(-1)^n}{n^\alpha} - \frac{1}{n^{2\alpha}} + o\left(\frac{1}{n^\alpha}\right)$. On pose v_n tel que $\forall n \in \mathbb{N}, u_n = \frac{(-1)^n}{n^\alpha} + v_n$, avec $(v_n) \sim \frac{1}{n^{2\alpha}}$. On a que $(-v_n)$ est positive, et que $\sum v_n$ converge si, et seulement si, $\alpha > \frac{1}{2}$. Donc on en déduit que $\sum u_n$ converge si $\alpha > \frac{1}{2}$ et diverge sinon.

Donc $\sum u_n$ converge si, et seulement si, $\alpha > \frac{1}{2}$

Chapitre II

Familles sommables

I Dénombrabilité

Définition : Dénombrable

On dit qu'un ensemble est dénombrable s'il est en bijection avec \mathbb{N} .

Définition : Au plus dénombrable

Un ensemble est au plus dénombrable s'il est fini ou dénombrable.

Exemple I.1. \mathbb{Z} est dénombrable avec $\varphi : \mathbb{N} \rightarrow \mathbb{Z}, \varphi : n \mapsto (-1)^{n+1} \left\lfloor \frac{n+1}{2} \right\rfloor$ bijective :

- *Injectivité* : supposons $\varphi(n_1) = \varphi(n_2)$, si n_1 pair, alors $n_1 \dots$
- *surjectivité* : Soit $n \in \mathbb{Z}$, trouvons un antécédent ; si n est positif on a $2n - 1$, si n est négatif $-2n$

Proposition : Parties infinies de \mathbb{N}

Toute partie infinie de \mathbb{N} est dénombrable.

Preuve

Soit X une partie infinie de \mathbb{N} . On construit $(x_n) \in X^{\mathbb{N}}$ par récurrence. $x_0 = \min(X)$ qui existe comme toute partie non-vide de \mathbb{N} possède un plus petit élément. Avec $n \in \mathbb{N}$, on suppose x_0, \dots, x_n bien définies, et on définit $x_{n+1} = \min(X \setminus \{x_0, \dots, x_n\})$.

Il faut démontrer que (x_n) est une bijection. Elle est injective parce que par construction, elle est croissante. Elle est surjective par l'absurde : s'il y avait un élément qui n'avait pas d'antécédent, considérons le plus petit élément sans antécédent, mais son prédécesseur aurait un antécédent, etc.

Proposition : Parties infinies

Toute partie infinie d'un ensemble dénombrable est dénombrable.

Preuve

Soit X dénombrable, $Y \subset X$ infinie, alors on peut composer la bijection avec laquelle on obtient les éléments de Y par les éléments de X avec la bijection entre X et \mathbb{N} , qui est donc bijective aussi.

Proposition : Produit cartésien

\mathbb{N}^2 est dénombrable.

Preuve

$\varphi : \begin{cases} \mathbb{N}^2 & \rightarrow \mathbb{N} \\ (p, q) & \mapsto 2^p(2q+1) \end{cases}$. Prouvons sa bijectivité.

Pour l'injectivité, avec $(p, q), (p', q') \in \mathbb{N}^2$, on suppose que $2^p(2q+1) = 2^{p'}(2q'+1)$. Par lemme de Gauss : $2^p | 2^{p'}(2q'+1)$, or $2^p \wedge (2q'+1) = 1$ donc $2^p | 2^{p'}$ donc $p \leq p'$, de même on a $p' \leq p$ donc $p = p'$ et $q = q'$, d'où l'injectivité.

Pour la surjectivité, soit $n \in \mathbb{N}$. Considérons $A = \{p \in \mathbb{N} | 2^p | n\}$. A est non-vidé car $2^0 = 1$ divise tout, c'est une partie de \mathbb{N} et A est majoré (2^p tend vers l'infini et n est fini). Donc A possède un plus grand élément p_0 . Donc $\frac{n}{2^{p_0}}$ est impair, alors $(p_0, \frac{1}{2}(\frac{n}{2^{p_0}} - 1))$ est antécédent de n , d'où la surjectivité.

Proposition : Produit cartésien étendu

\mathbb{N}^p est dénombrable.

Preuve

Par récurrence, avec l'application de la preuve précédente mais des ensembles $\mathbb{N}^{n-1} \times \mathbb{N} \rightarrow \mathbb{N}$ pour prouver le tout.

Proposition : Dénombrabilité de \mathbb{Q}

\mathbb{Q} est dénombrable.

Preuve

C'est une partie infinie de $\mathbb{N} \times \mathbb{N}^*$, donc elle est dénombrable.

Proposition : Dénombrabilité de \mathbb{R}

\mathbb{R} est en bijection avec $[0, 1[$, donc \mathbb{R} n'est pas dénombrable.

Preuve

Avec un nombre écrit sous la forme $0, a_1 a_2 a_3 \dots$. On peut écrire $a_1 = \lfloor 10x \rfloor$, $a_2 = \left\lfloor (x - \frac{a_1}{10}) 10 \right\rfloor, \dots$

On peut construire une bijection avec \mathbb{R} . Par diagonalisation (on indexe tous les réels entre 0 et 1 par des entiers, et ensuite on construit un nouveau réel qui n'est pas du tout présent dans la suite), on a que \mathbb{N} n'est pas en bijection avec $[0, 1[$. Ainsi, \mathbb{N} n'est vraiment pas en bijection avec \mathbb{R} .

Remarque I.1. *L'hypothèse du continu est indécidable, et dit que tout ensemble inclus dans \mathbb{R} est soit en bijection avec \mathbb{N} , soit avec \mathbb{R} , sans entre-deux. Il faut le rajouter à l'axiomatique pour ne pas créer de paradoxes en se posant la question.*

Proposition : Réunion

Toute réunion au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable. Si I est au plus dénombrable, avec $(A_i)_{i \in I}$ est une famille d'ensembles au plus dénombrable, alors $\cup_{i \in I} A_i$ est au plus dénombrables.

Preuve

E ensemble, A, B deux parties de E . Alors $A \cup B = \{x \in E | (x \in A) \vee (x \in B)\}$. Alors $(A_i)_{i \in I} \in \mathcal{P}(E)^I$, et donc $\cup_{i \in I} A_i = \{x \in E | \exists i \in I, x \in A_i\}$.

II Familles sommables de réels positifs

Définition : Somme de familles réelles positives

Soit I un ensemble, soit $(u_i)_{i \in I} \in \mathbb{R}_+^I$. On définit la somme des $(u_i)_{i \in I}$ comme : $\sum_{i \in I} u_i = \sup \left(\sum_{j \in J} u_j \right)_{J \subset I \text{ fini}}$ si ces sommes sont majorées et $+\infty$ sinon.

Définition : Familles sommables de réels positifs

Soit $(u_i)_{i \in I} \in \mathbb{R}_+^I$, on dit que $(u_i)_{i \in I}$ est sommable si $\sum_{i \in I} u_i \in \mathbb{R}_+$

Proposition : Sommabilité et dénombrabilité

Si I est non-dénombrable et $(u_i)_{i \in I} \in \mathbb{R}_+^I$, alors $(u_i)_{i \in I}$ n'est pas sommable.

Preuve

Ici, I est non-dénombrable. Pour $n \in \mathbb{N}^*$, on note $A_n = \left\{ i \in I | u_i > \frac{1}{n} \right\}$. Alors $\cup_{n \in \mathbb{N}^*} A_n = I$.

On a une réunion dénombrable d'ensembles donnant un ensemble non-dénombrable, donc il existe n_0 tel que A_{n_0} non-dénombrable. Donc A_{n_0} est infini : pour tout $p \in \mathbb{N}^*$, on peut trouver $J \subset A_{n_0}$ tel que $\text{card}(J) = p$.

Donc $\sum_{j \in J} u_j > \frac{p}{n_0}$ donc la somme n'est pas majorée. Donc $\sum_{i \in I} u_i = +\infty$

Dans la suite du cours, I est dénombrable.

Théorème :

Soit I un ensemble dénombrable, soit $\begin{matrix} \mathbb{N} & \rightarrow & I \\ n & \mapsto & i_n \end{matrix}$ une bijection de \mathbb{N} dans I . Soit $(u_i)_{i \in I} \in \mathbb{R}_+^I$. $(u_i)_{i \in I}$ est sommable si, et seulement si, la série $\sum u_{i_n}$ est convergente. En cas de convergence, $\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} u_{i_n}$.

Preuve

Dans le cadre de l'énoncé, on suppose $(u_i)_{i \in I}$ sommable.

Fixons $n \in \mathbb{N}$, $\sum_{k=0}^n u_{i_k} = \sum_{k \in \llbracket 0, n \rrbracket} u_{i_k}$ donc $\sum_{k=0}^n u_k \leq \sum_{i \in I} u_i$.

Donc $\sum u_{i_n}$ converge et $\sum_{n=0}^{+\infty} u_n \leq \sum_{i \in I} u_i$.

Soit $J \subset \mathbb{N}$ fini. $\sum_{j \in J} u_{i_j} \leq \sum_{n=0}^{\max(J)} u_{i_n} \leq \sum_{i=0}^{+\infty} u_{i_n}$

donc $\sum_{i \in I} u_i \leq \sum_{n=0}^{+\infty} u_{i_n}$

Pour $(u_i)_{i \in I} \in \mathbb{R}_+^I$, avec I dénombrable. Alors $\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} u_{\varphi(n)}$ si la série converge et $+\infty$ sinon.

Théorème : Sommation par paquets positif

Soit I dénombrable et $(J_j)_{j \in J}$ une partition de I avec J au plus dénombrable, ie $\cup_{j \in J} J_j = I, \forall j, h \in$

$I, j \neq h \Rightarrow J_j \cap J_h = \emptyset$. Soit $(u_i)_{i \in I} \in \mathbb{R}_+^I$, alors : $\sum_{i \in I} u_i = \sum_{j \in J} \left(\sum_{k \in J_j} u_k \right)$

Exemple II.1. $\sum_{n \in \mathbb{N}^*} \frac{1}{n^2} = \sum_{n \in \mathbb{N}^*} \frac{1}{(2n)^2} + \sum_{n \in \mathbb{N}^*} \frac{1}{(2n+1)^2}$ d'où $\frac{3}{4} \sum_{n \in \mathbb{N}^*} \frac{1}{(2n+1)^2}$

Mais en revanche, on n'a pas que $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = \sum_{n=1}^{+\infty} \frac{1}{2n} - \sum_{n=0}^{+\infty} \frac{1}{2n+1}$ puisque ces deux séries ne sont pas convergentes.

$\left(\frac{1}{n^2 + p^2} \right)_{(n,p) \in \mathbb{N}^{*2}}$ est-elle sommable ? D'abord, notons que $\left(\sum \frac{1}{n^2 + p^2} \right)_{np}$ est positive, et donc on a le droit d'écrire :

$\sum_{n,p \in \mathbb{N}^*} \frac{1}{n^2 + p^2} = \sum_{n \in \mathbb{N}^*} \sum_{p \in \mathbb{N}^*} \frac{1}{n^2 + p^2}$ car $\mathbb{N}^{*2} = \cup \{(n,p) | n \in \mathbb{N}^*, p \in \mathbb{N}^*\}$ et la réunion est disjointe (signe II, non homologué)

$= \sum_{n \in \mathbb{N}^*} I_n$ avec $I_n = \sum_{p \in \mathbb{N}^*} \frac{1}{n^2 + p^2}$. Soit $n \in \mathbb{N}^*$, alors $\left(\frac{1}{n^2 + p^2} \right)_p \sim \frac{1}{p^2}$, donc $\forall n \in \mathbb{N}^*, I_n < +\infty$.

Soit $n \in \mathbb{N}^*$. On définit $f : x \mapsto \frac{1}{n^2 + x^2}$ décroissante et positive. On fixe $p \in \mathbb{N}^*$. Et alors : $\int_p^{p+1} \frac{dt}{n^2 + t^2} \leq \frac{1}{n^2 + p^2}$.

Donc $\forall N \in \mathbb{N}$, $\sum_{p=1}^N \frac{1}{n^2 + p^2} \geq \int_1^{N+1} \frac{dt}{n^2 + t^2} \geq \frac{1}{n} [\arctan(t)]_{\frac{1}{n}}^{\frac{N+1}{n}} \geq \frac{1}{n} \left(\arctan\left(\frac{N+1}{n}\right) - \arctan\left(\frac{1}{n}\right) \right)$. Par passage à la limite quand N tend vers $+\infty$, on a que $I_n \geq \frac{\pi}{2n} \arctan\left(\frac{1}{n}\right)$.

Donc $\frac{\pi}{2n} - \frac{1}{n} \arctan\left(\frac{1}{n}\right) \sim \frac{\pi}{n}$. Or $\frac{\pi}{n}$ diverge, donc $\sum I_n$ diverge. Donc $\sum_{n \in \mathbb{N}^*} = +\infty$. Donc $\left(\frac{1}{n^2 + p^2} \right)_{n,p}$ n'est pas sommable.

Dans le cas plus général de $\left(\frac{1}{n^\alpha + p^\alpha} \right)_{(n,p) \in \mathbb{N}^{*2}}$

III Familles de complexes.

Définition : Sommabilité complexe

Soit $(u_i)_{i \in I} \in \mathbb{C}^I$. On dit que la famille est sommable si $(|u_i|)_{i \in I}$ est sommable, donc que $\sum_{i \in I} |u_i| < +\infty$

Proposition : Somme d'une famille sommable.

Soit I dénombrable et $(u_i)_{i \in I}$ sommable. Alors : pour toute bijection $\varphi : \begin{cases} \mathbb{N} & \rightarrow I \\ n & \mapsto \varphi(n) \end{cases}$ on a que $\sum_n u_{\varphi(n)}$ est une série absolument convergente et $\sum_{n=0}^{+\infty} u_{\varphi(n)}$ ne dépend pas de φ .

On appelle somme de $(u_i)_{i \in I} : \sum_{i \in I} = \sum_{n=0}^{\infty} u_{\varphi(n)}$ où φ est une bijection donnée. La somme d'une famille non-sommable n'est pas définie.

Preuve

(u_i) sommable. φ bijection de \mathbb{N} dans I . $\sum |u_i| = \sum |u_{\varphi(n)}| = \sum_{n=0}^{\infty} |u_{\varphi(n)}|$. $\sum u_{\varphi(n)}$ est absolument convergente. De plus, $\forall N \in \mathbb{N}$, $\sum_{k=0}^N u_{\varphi(k)} = \sum_{k=0}^N \mathcal{R}(u_{\varphi(k)})_+ - \mathcal{R}(u_{\varphi(k)})_- + i(\mathcal{I}(u_{\varphi(k)})_+ - \mathcal{I}(u_{\varphi(k)})_-)$. Tous ces restes sont majorés par $|u_{\varphi(k)}|$ donc sommables. Donc $\sum_{k=0}^{\infty} u_{\varphi(k)} \rightarrow \sum_{n \in \mathbb{N}} \mathcal{R}](u_{\varphi(n)})_+ + \sum_{n \in \mathbb{N}} \mathcal{R}](u_{\varphi(n)})_- + \dots$

Théorème : Sommation par paquets

I dénombrable, dont les J_j sont une partition. Avec $(u_i)_{i \in I} \in \mathbb{C}^I$ sommable. Alors $\sum_{i \in I} u_i = \sum_{j \in J} \sum_{k \in J_j} u_k$

IV Applications

Théorème : de Fubini

Soit $(u_{i,j})_{(i,j) \in I \times J} \in \mathbb{C}^{I \times J}$ sommable. Alors $\sum_{i,j \in I \times J} u_{i,j} = \sum_{i \in I} \sum_{j \in J} u_{i,j} = \sum_{j \in J} \sum_{i \in I} u_{i,j}$, avec le cas particulier où $u_{i,j} = a_i b_j$ où : $(a_i b_j)_{i,j \in I \times J}$ qui est sommable si, et seulement si, $(a_i)_{i \in I}$ et $(b_j)_{j \in J}$ sont sommables et dans ce cas, $\sum_{(i,j) \in I \times J} a_i b_j = \left(\sum_{i \in I} a_i \right) \left(\sum_{j \in J} b_j \right)$

Preuve

On prend $I \times J = \cup_{i \in I} \{(i,j), j \in J, i \in I\}$. Par sommation par paquets : $\sum_{(i,j) \in I \times J} |a_i b_j| = \sum_{i \in I} \left(\sum_{j \in J} |a_i| |b_j| \right) = \sum_{i \in I} |a_i| \left(\sum_{j \in J} |b_j| \right) = \left(\sum_{i \in I} |a_i| \right) \left(\sum_{j \in J} |b_j| \right)$
Donc $\sum |a_i b_j| < +\infty$ si, et seulement si, $\sum |a_i| < +\infty$ et $\sum |b_j| < +\infty$. Dans le cas de sommabilité, on peut reprendre tous ces calculs en enlevant les modules.

Théorème : Produit de Cauchy

Soit $(a_n), (b_n) \in \mathbb{C}^{\mathbb{N}}$. Si $\sum a_n, \sum b_n$ sont absolument convergentes alors $\sum \left(\sum_{k=0}^n a_k b_{n-k} \right)$ est absolument convergente et $\left(\sum_{n=0}^{+\infty} a_n \right) \left(\sum_{n=0}^{+\infty} b_n \right) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right)$

Preuve

$(a_n b_p)_{(n,p) \in \mathbb{N}^2}$ est sommable. $\mathbb{N}^2 = \cup_{s \in \mathbb{N}} \{(n, p) \in \mathbb{N}^2 | n + p = s\} = \cup_{s \in \mathbb{N}} S_s$ donc par théorème de sommation par paquets, on a que $\sum_{(n,p) \in \mathbb{N}^2} a_n b_p = \sum_{s \in \mathbb{N}} \sum_{(n,p) \in S_s} a_n b_p = \sum_{s \in \mathbb{N}} \left(\sum_{k=0}^s a_k b_{s-k} \right)$

Exemple IV.1. Avec $t \in]-1, 1[$, montrez que $\frac{1}{(1-t)^2} = \sum_{n=0}^{+\infty} (n+1)t^n$. C'est convergent par $(n+1)t^n$ est absolument convergent en remarquant que $[n+1]t^n = o(\frac{1}{n^2})$. On peut donc faire un produit de Cauchy : $\left(\sum_{n=0}^{+\infty} t^n \right) \left(\sum_{n=0}^{+\infty} t^n \right) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^n t^k t^{n-k} \right)$

Exemple IV.2. Prenons l'exponentielle complexe, définie par : pour $z \in \mathbb{C}$, on définit : $\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$.

On prouve sa convergence facilement par le critère de d'Alembert : $\left| \frac{|z|}{n+1} \right|$ qui tend vers 0. Donc par critère de d'Alembert, on a sa convergence absolue. Comme elle est absolument convergente, on peut l'écrire comme une famille sommable : $\sum_{n \in \mathbb{N}} \frac{z^n}{n!}$.

Prouvons que $\forall z, z' \in \mathbb{C}, \exp(z+z') = \exp(z) \exp(z')$: comme la famille est sommable, avec $z, z' \in \mathbb{C}$, on a donc que $\exp(z) \exp(z')$ existe en tant que limite de la famille $\sum_{p \in \mathbb{N}} \sum_{n \in \mathbb{N}} \frac{z^n}{n!} \frac{z'^{p-n}}{(p-n)!} = \sum_{p \in \mathbb{N}} \sum_{p \in \mathbb{N}} \frac{1}{p!} \binom{p}{n} z z'^{p-n} = \sum_{n \in \mathbb{N}} \frac{(z+z')^n}{n!}$

Prouvons que $\forall z \in \mathbb{C}, \exp(\bar{z}) = \overline{\exp(z)}$. Soit $z \in \mathbb{C}$, soit $n \in \mathbb{N} : \sum_{k=0}^n \frac{\bar{z}^k}{k!} = \overline{\sum_{k=0}^n \frac{z^k}{k!}}$. Par passage à la limite lorsque n tend vers l'infini, on a donc que $\exp(\bar{z}) = \overline{\exp(z)}$.

On peut alors utiliser que $\forall \theta \in \mathbb{R}, \exp(i\theta) \exp(-i\theta) = \exp(i\theta) \exp(-i\theta) = \exp(i(\theta - \theta)) = \exp(0) = 1$

La restriction de l'exponentielle à \mathbb{R} coïncide bien avec la fonction exponentielle réelle.

Chapitre III

Espaces vectoriels normés

I Généralités

On note \mathbb{K} un corps, \mathbb{R} ou \mathbb{C} . Dans tous les cours, E est un \mathbb{K} -ev.

Définition : Norme

Soit E un \mathbb{K} -ev. Soit φ une application de E dans \mathbb{R}_+ . On dit que φ est une norme si elle vérifie :

1. $\forall u \in E, \varphi(u) = 0 \Rightarrow u = 0$ (on dit que l'application est définie)
2. homogénéité : $\forall \lambda \in \mathbb{K}, \forall u \in E, \varphi(\lambda u) = |\lambda| \varphi(u)$
3. inégalité triangulaire : $\forall u, v \in E, \varphi(u + v) \leq \varphi(u) + \varphi(v)$

Définition : EVN

On dit que E est un espace vectoriel normé si on a choisi une norme dans E . On note alors $(E, \varphi), (E, N), (E, \|\cdot\|)$.

Définition : Boule unité

Avec $(E, \|\cdot\|)$, on appelle la boule unité de E l'ensemble $\mathcal{B}(\vec{0}, 1) = \{x \in \mathbb{K}^p, \|x\| \leq 1\}$. Une norme est définie par sa boule unité.

Exemple I.1. Dans \mathbb{K}^p , avec $p \in \mathbb{N}, p \geq 1$, pour $(x_1, x_2, \dots, x_p) \in \mathbb{K}^p, \|x\|_\infty = \max_{1 \leq i \leq p} |x_i|$ est une norme : prenons un vecteur $x = (x_1, \dots, x_p)$, alors si leur max est nul leur vecteur est nul, pour l'homogénéité, toutes les composantes sont multipliées fonctionne

Pour l'inégalité triangulaire, prenons aussi $y = (y_1, \dots, y_p)$. Alors $\|x+y\|_\infty = \max_{1 \leq i \leq p} |x_i + y_i| \leq \max_{1 \leq i \leq p} (|x_i| + |y_i|) \leq \max_{1 \leq i \leq p} |x_i| + \max_{1 \leq i \leq p} |y_i| \leq \|x\|_\infty + \|y\|_\infty$

Dans \mathbb{R}^2 , on note $\mathcal{B}(\vec{0}, 1)$ la boule de centre 0 de rayon 1, telle que $\mathcal{B}(\vec{0}, 1) = \{\vec{x} \in \mathbb{R}^2 | \|\vec{x}\|_\infty \leq 1\}$. Dans \mathbb{R}^2 munie de la norme infinie précédente, on obtient un carré et pas un rond, comme on aurait pu l'attendre. Les différentes normes sont donc catégorisées par leurs boules unités, ce qui découle de l'homogénéité.

On appelle la norme 1 de $x = (x_1, \dots, x_p) \in \mathbb{K}^p$ l'application : $\|x\|_1 = \sum_{i=1}^p |x_i|$, qui est bien une norme. L'inégalité triangulaire vient des propriétés du module. Sa boule unité est un carré dont les sommets sont sur les axes.

On appelle la norme 2 de $x = (x_1, \dots, x_p) \in \mathbb{K}^p$ l'application : $\|x\|_2 = \sqrt{\sum_{i=1}^p |x_i|^2}$, qui est bien une norme : la définition est assurée par le fait que si une somme de réels positifs est nulle, ils sont nuls,

l'homogénéité se fait par calcul direct (sur les modules, sur les facteurs de sommes, sur la racine carrée) et l'inégalité triangulaire est prouvée par : $\forall x, y \in \mathbb{C}^{2p}, \sqrt{\sum_{i=1}^p |x_i + y_i|^2} \leq \sqrt{\sum_{i=1}^p |x_i|^2} + \sqrt{\sum_{i=1}^p |y_i|^2}$.

En considérant \mathbb{R}^p muni du produit scalaire canonique et $x\vec{\text{tilde}} = (|x_1|, \dots, |x_p|)$, $y\vec{\text{tilde}} = (|y_1|, \dots, |y_p|)$. Alors $\|x\vec{\text{tilde}} + y\vec{\text{tilde}}\|_{\text{euclidienne}} \leq \|x\vec{\text{tilde}}\| + \|y\vec{\text{tilde}}\|$ donc on a l'égalité précédente.

Donc on a trois normes usuelles pour \mathbb{K}^p .

Exemple I.2. Si E est de dimension finie, on choisit une base $B = (e_1, \dots, e_p)$. On pourra parler de la norme infinie, de la norme 1 et de la norme 2 relativement à la base B : si x élément de E avec une décomposition $x = \sum_{i=1}^p x_i e_i$, alors on définira $\|x\|_\infty = \max_{1 \leq i \leq p} |x_i|$, $\|x\|_1 = \sum_{i=1}^p |x_i|$, $\|x\|_2 = \sqrt{\sum_{i=1}^p |x_i|^2}$

Si E est en dimension infinie, par exemple $E = \mathbb{K}[X]$ (réunion dénombrable d'espaces de dimension finie), on peut définir pour $P = \sum_{k=0}^p a_k X^k$ les mêmes normes : $\|P\|_\infty = \max_{k \in \mathbb{N}} |a_k|$, $\|P\|_1 = \sum_{k=0}^{\deg P} |a_k|$, $\|P\|_2 = \sqrt{\sum_{k=0}^{\deg P} |a_k|^2}$

Si $E = \mathcal{C}([0, 1], \mathbb{K})$, pour $f \in E$, on considère $\|f\|_\infty = \sup_{[0,1]} |f|$ qui est d'ailleurs même un max mais le sup est toujours plus adapté. C'est une norme : elle est définie car si le sup du module de la fonction est nul, elle est nulle partout ; soit $\lambda \in \mathbb{K}$, montrons l'homogénéité avec deux inégalités :

On doit majorer l'ensemble : $\forall x \in [0, 1], |\lambda f(x)| = |\lambda| |f(x)| \leq |\lambda| \sup_{[0,1]} |f(x)|$

On doit minorer l'ensemble : si $\lambda = 0$, on a $|\lambda| \sup |f| \leq \sup |\lambda f|$.

Si $\lambda \neq 0$, $\forall x \in [0, 1], |f(x)| \frac{1}{|\lambda|} |\lambda| |f(x)| \leq \frac{1}{|\lambda|} \sup_{[0,1]} |\lambda| |f|$

donc $\sup |f| \leq \frac{1}{|\lambda|} \sup |\lambda f|$. D'où l'homogénéité.

Pour l'inégalité triangulaire : $\sup_{[0,1]} |f+g| \leq \sup_{[0,1]} |f| + \sup_{[0,1]} |g| \leq \sup_{[0,1]} |f| + \sup_{[0,1]} |g|$. Soit $x \in [0, 1], |f(x)| + |g(x)| \leq \sup_{[0,1]} |f| + \sup_{[0,1]} |g|$, voilà l'inégalité prouvée.

Pour $\mathcal{C}([0, 1], \mathbb{K})$, on définit pour $f \in \mathcal{C}([0, 1], \mathbb{K})$, $\|f\|_1 = \int_0^1 |f(t)| dt$, $\|f\|_2 = \sqrt{\int_0^1 |f(t)|^2 dt}$.

Pour la norme 1 : soit $f \in E$, supposons que $\|f\|_1 = 0$, alors on a la définition par positivité (améliorée) de l'intégrale puisqu'on a une fonction continue sur tout le segment $[0, 1]$ positive d'intégrale nulle. Le reste est évident.

Pour la norme 2 : soit $f, g \in E$, montrons que $\sqrt{\int_0^1 |f+g|^2} \leq \sqrt{\int_0^1 |f|^2} + \sqrt{\int_0^1 |g|^2}$. On en déduit son produit scalaire $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$, puis on applique Cauchy-Schwarz avec les modules des deux fonctions.

Si on prend $E = \mathbb{K}^{\mathbb{N}}$, alors on peut définir pour $(u_n) \in E$, on a besoin de restreindre aux suites bornées pour que $\|u\|_\infty = \sup_{n \in \mathbb{N}} |u_n|$, ou pour que $\|u\|_1 = \sum_{n=0}^{+\infty} u_n$, il faut se restreindre à l'ev des termes généraux de séries absolument convergentes (L_1), tandis que la norme $\|u\|_2 = \sqrt{\sum_{n=0}^{+\infty} |u_n|^2}$ n'est bien définie que sur L_2 , ensemble des suites dont le module au carré est sommable. Mais il faut encore savoir si c'est un ev...

$$|u_n + v_n|^2 \leq (|u_n| + |v_n|)^2 \leq |u_n|^2 + |v_n|^2 + 2|u_n||v_n| \leq 2(|u_n|^2 + |v_n|^2)$$

Proposition : Produit fini d'EVN

Avec E, F deux EVN de normes $\|\cdot\|_E, \|\cdot\|_F$, alors $E \times F$ est un EVN qui admet comme norme l'application :
$$\begin{cases} E \times F & \rightarrow \mathbb{R}_+ \\ (e, f) & \mapsto \max(\|e\|_E, \|f\|_F) \end{cases} .$$

On peut étendre par récurrence : si E_1, \dots, E_p un nombre fini d'EVN, on peut définir une norme sur $\prod_{i=1}^p E_i$ par : $\forall (e_1, \dots, e_p) \in E_1 \times E_2 \times \dots \times E_p, \|(e_1, \dots, e_p)\| = \max_{i \in \{1, \dots, p\}} \|e_i\|$.

Définition : Distance

Soit E un EVN. On appelle distance associée à la norme sur E l'application :

$$d : \begin{cases} E & \rightarrow \mathbb{R}_+ \\ (x, y) & \mapsto \|x - y\| \end{cases}$$

Une distance définit un espace métrique.

Proposition : Propriétés de la distance

- $\forall x, y \in E, d(x, y) = 0 \Rightarrow x = y$;
- $\forall x, y, z \in E, d(x, z) \leq d(x, y) + d(y, z)$.

Proposition : Deuxième forme de l'inégalité triangulaire

Soit E un EVN, $\forall x, y \in E, \left| \|x\| - \|y\| \right| \leq \|x \pm y\| \leq \|x\| + \|y\|$

Preuve

L'inégalité de droite est déjà une des propriétés des normes. Pour l'inégalité de gauche, elle est équivalente à
$$\begin{aligned} \|x\| - \|y\| &\leq \|x - y\| \\ -\|x - y\| &\leq \|x\| - \|y\| \end{aligned} .$$

On la prouve en utilisant $x = x - y + y$ et puis $\|x\| \leq \|x - y\| + \|y\|$

Définition : Boule ouverte

Avec E un EVN, soit $x \in E, r \in \mathbb{R}_+$, on appelle la boule ouverte de centre x et de rayon r : $\mathcal{B}(x, r) = \{y \in E, \|x - y\| < r\}$

Définition : Boule fermée

Avec E un EVN, soit $x \in E, r \in \mathbb{R}_+$, on appelle la boule fermée de centre x et de rayon r : $\mathcal{B}_f(x, r) = \{y \in E, \|x - y\| \leq r\}$

Définition : Sphère

Avec E un EVN, soit $x \in E, r \in \mathbb{R}_+$, on appelle la sphère de centre x et de rayon r : $\mathcal{S}(x, r) = \{y \in E, \|x - y\| = r\}$

Définition : Partie convexe

Soit E un \mathbb{R} -ev, soit $A \subset E$, on dit que A est convexe si : $\forall x, y \in A, [x, y] \subset A$, où $[x, y] = \{tx + (1-t)y, 0 \leq t \leq 1\} = \{y + t(x-y), 0 \leq t \leq 1\}$

Proposition : Convexité des boules

Soit E un EVN. Toute boule de E est convexe.

Preuve

Prenons $x \in E, r \in \mathbb{R}_+$ et $\mathcal{B}(x, r)$ la boule ouverte qu'ils définissent. Alors, pour $y, z \in \mathcal{B}(x, r)$, on a que $\|x - z\| < r$ et $\|x - y\| < r$. Prenons $\exists t \in [0, 1]$, ce qui donne que $ty + (1-t)z \in [y, z]$.

On a que $\|ty + (1-t)z - x\| = \|ty - tx + (1-t)z - (1-t)x\| \leq |t|\|y - x\| + |1-t|\|z - x\| < (t + (1-t))r = r$ ce qui est justifié par $t \neq 0$ ou $(1-t) \neq 0$ et par les inégalités strictes de la boule ouverte.

Donc $ty + (1-t)z \in \mathcal{B}(x, r)$.

On procède de même pour les boules fermées, sauf qu'on n'a pas besoin de justifier une inégalité stricte.

Proposition : Séparabilité

Soit E un EVN. Soient $x, y \in E$ tels que $x \neq y$. Alors il existe deux boules disjointes centrées en x et y .

Preuve

Soient $x, y \in E$ différents. Alors $\|x - y\| \neq 0$, notons $\|x - y\| = d$.

Prenons les boules ouvertes $B_1 = \mathcal{B}(x, \frac{d}{4}), B_2 = \mathcal{B}(y, \frac{d}{4})$.

Soit $z \in B_1$, alors $\|z - x\| < \frac{d}{4}$.

On note que $\|z - y\| = \|z - x + x - y\|$, donc $\|z - y\| \geq \|y - x\| - \|z - x\| \geq d - \frac{d}{4}$

Donc $\|z - y\| \geq \frac{d}{4}$

Définition : Normes équivalentes

Soit E un espace vectoriel, et $\|\cdot\|_1, \|\cdot\|_2$ deux normes sur E . On dit que $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes

si : $\exists \alpha, \beta \in \mathbb{R}_+^*, \forall x \in E, \alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$, ou si $x \neq 0$: $\alpha \leq \frac{\|x\|_2}{\|x\|_1} \leq \beta$

Proposition : Bornes pour normes équivalentes

Soit E un espace vectoriel avec $\|\cdot\|_1, \|\cdot\|_2$ deux normes équivalentes. Toute partie de E bornée pour $\|\cdot\|_1$ est bornée pour $\|\cdot\|_2$

Exemple I.3. Sur \mathbb{K}^p , comparons $\|\cdot\|_\infty, \|\cdot\|_1, \|\cdot\|_2$. Soit $(x_1, \dots, x_p) \in \mathbb{K}^p$: $\|x\|_1 = \sum_{i=1}^p |x_i| \leq p\|x\|_\infty$ et $\|x\|_\infty \leq \|x\|_1$. Donc les deux normes sont équivalentes.

Comparons $\|\cdot\|_2$ et $\|\cdot\|_\infty$: $\|x\|_2 = \sqrt{\sum_{i=1}^p |x_i|^2} \leq \sqrt{p}\|x\|_\infty$ et $\|x\|_\infty \leq \|x\|_2$.

On a bien prouvé les équivalences (c'est une relation d'équivalence donc $\|\cdot\|_1$ et $\|\cdot\|_2$ aussi), mais on n'a pas forcément les constantes optimales. On a donc besoin d'établir que $\|x\|_2 \leq \|x\|_1$ (en mettant au carré)

et que $\|x\|_1 \leq \sqrt{p}\|x\|_2$ qui est immédiat en considérant $\|\cdot\|_1$ comme un produit scalaire avec un vecteur de norme infinie valant 1, et ayant sur chaque coordonnée le même signe.

On admet provisoirement le théorème : Dans un espace de dimension finie, toutes les normes sont équivalentes.

Exemple I.4. Prenons les normes qu'on a défini sur $E = \mathcal{C}([0, 1], \mathbb{R})$ (rappel : $\forall f \in E, \|f\|_\infty = \sup_{[0,1]} |f|, \|f\|_1 = \int_0^1 |f(t)| dt$).

Pour prouver que deux normes ne sont pas équivalentes, prouvons qu'un ensemble borné pour l'une ne l'est pas pour l'autre (rappel : avec $A \subset E$, A bornée signifie $\exists M \in \mathbb{R}_+, \forall x \in A, \|x\| \leq M$ et A non-bornée signifie $\forall M \in \mathbb{R}_+, \exists x \in A, \|x\| \geq M$). On peut donc prouver la non-bornaison avec l'aide d'une suite tendant vers l'infini.

Soit $f \in E$, alors $\int_0^1 |f| \leq \sup_{[0,1]} |f| (1 - 0) \leq \|f\|_\infty$. Cependant, il n'existe pas de constantes permettant une seconde inégalité, donc on peut faire des suites de fonction avec une norme 1 qui tend vers 0 et une norme infinie restant à 1. Par exemple : $\forall n \in \mathbb{N}, f_n : \begin{cases} [0, 1] & \rightarrow \mathbb{R} \\ x & \mapsto e^{-nx} \end{cases}$. Alors, $\|f_n\|_1 = \frac{1}{n}(1 - e^{-n})$ mais $\|f_n\|_\infty = 1$ donc les deux normes ne sont pas équivalentes.

II Suites d'un EVN

Définition : Convergence

Soit E un EVN sur \mathbb{K} , soit $(u_n) \in E^{\mathbb{N}}$. Soit $l \in E$. On dit que (u_n) converge vers l si $(\|u_n - l\|) \rightarrow 0$. On peut aussi écrire :

$$\forall \varepsilon \in \mathbb{R}_+, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow u_n \in \mathcal{B}(l, \varepsilon)$$

On parle de suites convergentes et de limites (notées $\lim(u_n) = l$ et $(u_n) \rightarrow l$) dans un EVN

Théorème : Unicité de la limite

La limite d'une suite convergente est unique.

Preuve

Supposons qu'une suite soit convergente avec deux limites l_1, l_2 , on prend deux boules centrées sur les deux qui soient disjointes, puis on applique le fait qu'à partir d'un certain rang, tous les termes de la suite soient dedans. On y trouve une contradiction, et donc la limite est unique.

Proposition : Convergences équivalentes

Si E est un ev et que $\|\cdot\|_1, \|\cdot\|_2$ sont deux normes de E équivalentes, alors les suites convergentes de $(E, \|\cdot\|_1)$ sont exactement les suites convergentes de $(E, \|\cdot\|_2)$

Preuve

Il existe α, β tels que : $\forall x \in E, \alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$.

Alors si $u_n \rightarrow l$ en norme 1, alors $\|u_n - l\|_1 \rightarrow 0$.

Comme $\forall n \in \mathbb{N}, \|u_n - l\|_2 \leq \beta\|u_n - l\|_1$, alors $(\|u_n - l\|_2) \rightarrow 0$

Et de même pour la réciproque.

Remarque II.1. Si $\|\cdot\|_1$ et $\|\cdot\|_2$ ne sont pas équivalentes, il existe des suites qui convergent pour l'une et pas pour l'autre.

En effet, si $\|\cdot\|_1$ et $\|\cdot\|_2$ ne sont pas équivalentes, on a soit $x \mapsto \frac{\|x\|_1}{\|x\|_2}$ non majorée, soit $x \mapsto \frac{\|x\|_2}{\|x\|_1}$.

Dans le premier cas, on peut construire $(u_n) \in E^{\mathbb{N}}$ telle que $\frac{\|u_n\|_1}{\|u_n\|_2} \rightarrow +\infty$ (existe par caractérisation séquentielle de la non-majoration). Prenons alors $v_n = \frac{1}{\|u_n\|_1} u_n$. Alors $\|v_n\|_2 = \frac{\|u_n\|_2}{\|u_n\|_1} \rightarrow 0$ et $\|v_n - 0\|_1 \rightarrow 0$ et donc v_n tend vers 0 au sens de la norme 2 alors qu'elle tend vers 1 au sens de la norme 1.

On rappelle que c'est seulement possible dans le cas de la dimension infinie, même si l'on n'a pas encore démontré qu'en dimension finie toutes les normes sont équivalentes.

Proposition : Opérations sur les suites convergentes

Avec E un EVN, (u_n) qui tend vers l_u et (v_n) qui tend vers l_v , $\lambda, \mu \in \mathbb{K}$, alors :

$$((\lambda u + \mu v)_n) \rightarrow \lambda l_u + \mu l_v.$$

Preuve

Soit $n \in \mathbb{N}$, alors $\|\lambda u_n + \mu v_n - (\lambda l_u + \mu l_v)\| \leq |\lambda| \|u_n - l_u\| + |\mu| \|v_n - l_v\|.$

Proposition : Corollaire

L'ensemble des suites convergentes est un sev de l'ensemble des suites d'un EVN.

Proposition : Extension

Si $(u_n) \in E^{\mathbb{N}}$, $(u_n) \rightarrow a$ et $(\lambda_n) \in \mathbb{K}^{\mathbb{N}}$ avec $(\lambda_n) \rightarrow \alpha$, alors $\lambda_n u_n \rightarrow \alpha a$

Preuve

Soit $n \in \mathbb{N}$:

$$\begin{aligned} \|\lambda_n u_n - \alpha a\| &= \|\lambda_n u_n - \lambda_n a + \lambda_n a - \alpha a\| \\ &\leq \|\lambda_n(u_n - a)\| + \|(\lambda_n - \alpha)a\| \\ &\leq |\lambda_n| \|u_n - a\| + |\lambda_n - \alpha| \|a\| \\ &\leq \sup |\lambda_n| \|u_n - a\| + |\lambda_n - \alpha| \|a\| \end{aligned}$$

Remarque II.2. On "rappelle" rapidement ce que sont des algèbres :

Une algèbre $(E, +, \times, \cdot)$ est un ensemble muni des lci $+$ et \times et de la lce \cdot tel que que (E, \mathbb{K}, \cdot) est un \mathbb{K} -ev et $(E, +, \times)$ est un anneau où on a pour $\lambda \in \mathbb{K}, a, b \in E, \lambda(ab) = a(\lambda b)$

$\mathbb{K}[X], (\mathcal{L}(E), +, \circ, \cdot), (\mathcal{M}_n(\mathbb{K}), +, \times, \cdot)$ sont des \mathbb{K} -algèbres. De plus, $\mathbb{K}[X]$ est commutatif intègre. $\mathcal{L}(E)$ et $\mathcal{M}_n(\mathbb{K})$ ne sont ni commutatifs ni intègres (il y a bijection entre les deux).

Définition : Algèbre normée

Soit $(A, +, \times, \cdot)$ une algèbre, on dit qu'elle est normée si on la munit d'une norme $\|\cdot\|$ vérifiant de plus : $\forall x, y \in A, \|x \times y\| \leq \|x\| \|y\|$

Proposition : Produit de suites dans une algèbre normée

Avec \mathcal{A} algèbre normée, soient $(u_n), (v_n) \in \mathcal{A}^{\mathbb{N}}$ qui tendent vers x et y . Alors $(u_n v_n) \rightarrow xy$

Preuve

Soit $n \in \mathbb{N}$: $\|u_n v_n - xy\| = \|u_n v_n - u_n y + u_n y - xy\|$
 $\leq \|u_n(v_n - y)\| + \|(u_n - x)y\|$
 $\leq \|u_n\| \|v_n - y\| + \|u_n - x\| \|y\|$ (propriété de la norme d'algèbre)
 $\leq \sup \|u_n\| \|v_n - y\| + \|u_n - x\| \|y\|$

Théorème : Convergence dans un produit cartésien fini

Soient E_1, \dots, E_p p \mathbb{K} -ev normés. Avec $E = E_1 \times \dots \times E_p$, $(u_n) \in E^{\mathbb{N}}$ s'écrivant (u_n^1, \dots, u_n^p) .
 Alors (u_n) tend vers $l = (l_1, \dots, l_p)$ si, et seulement si, $\forall i \in \{1, \dots, p\}, u_n^i \rightarrow l_i$

Preuve

$\forall i \in \{1, \dots, p\}, \|u_n^i - l_i\| \leq \|u_n - l\|$ Donc si (u_n) tend vers l , alors $\forall i \in \{1, \dots, p\}, u_n^i \rightarrow l_i$
 Réciproquement, supposons que $\forall i \in \{1, \dots, p\}, u_n^i \rightarrow l_i$.
 Soit $\varepsilon \in \mathbb{R}_+^*$.
 Pour $i \in \{1, \dots, p\}$, comme $u_n^i \rightarrow l_i$ on peut trouver $n_i \in \mathbb{N}, \forall n \geq n_i, \|u_n^i - l_i\| < \varepsilon$
 Donc pour $n \geq \max_{i \in \{1, \dots, p\}} n_i, \|u_n - l\| < \varepsilon$

Remarque II.3. Ce théorème s'applique notamment à \mathbb{K}^n pour $n \in \mathbb{N}^*$ et donc à tout \mathbb{K} -ev de dimension finie. Et comme en dimension finie, toutes les normes sont équivalentes, on a la propriété pour les normes qui ne sont pas la norme infinie.

On a la propriété : Si E est un \mathbb{K} -ev de dimension finie p et $B = (e_1, \dots, e_p)$ base de E . Alors à toute suite de E $(u_n) = (\sum_{i=1}^p u_n^i e_i)$ on peut associer ses suites coordonnées (u_n^i) dans la base B .

Donc $(u_n) \rightarrow l = \sum_{i=1}^p l_i e_i$ si, et seulement si, $\forall i \in \{1, \dots, p\}, u_n^i \rightarrow l_i$.

Exemple II.1. $\left(\begin{pmatrix} \frac{1}{n} & \frac{2+n}{1+n} \\ \sin\left(\frac{1}{n}\right) & \exp\left(\frac{1}{n+1}\right) \end{pmatrix} \right)_{n \in \mathbb{N}} \rightarrow \begin{pmatrix} 0 & 1 \\ 0 & e \end{pmatrix}$
 $\left(\begin{pmatrix} \frac{1}{n} & \frac{2+n}{1+n} \\ \sin\left(\frac{1}{n}\right) & (-1)^n \end{pmatrix} \right)_{n \in \mathbb{N}}$ diverge.

II.1 Séries d'un EVN

Définition : Série d'un EVN

Soit E un \mathbb{K} -ev normé. Avec $(u_n) \in \mathbb{K}^{\mathbb{N}}$, on dit que $\sum u_n$ converge si $\left(\sum_{k=0}^n u_k \right)$ converge.

On a les mêmes propriétés sur le fait que la suite des termes généraux tende vers 0.

Proposition : Absolue convergence en dimension finie

Si E est de dimension finie, alors toute série absolument convergente est convergente.

Preuve

Si on fixe une base $B = (e_1, \dots, e_p)$ de E de dimension finie p .

Soit $(u_n) \in E^{\mathbb{N}}$. Pour $n \in \mathbb{N}$, $u_n = \sum_{i=1}^p u_n^i e_i$.

La série $\sum u_n$ converge si, et seulement si, $\forall i \in \{1, \dots, p\}$, $\sum u_n^i$ converge.

Si $\sum \|u_n\|$ converge, alors $\sum \|u_n\|_{\infty}$ converge (car on est en dimension finie)

Or $\forall i \in \{1, \dots, p\}$, $|u_n^i| \leq \|u_n\|_{\infty}$

Donc par critère de majoration positif, $\sum u_n^i$ converge absolument, donc converge.

Donc $\sum u_n$ converge.

II.2 Application dans une algèbre normée

Proposition : Série exponentielle

Soit \mathcal{A} une algèbre normée de dimension finie. On peut définir pour $u \in \mathcal{A}$: $\exp(u) = \sum_{n=0}^{+\infty} \frac{1}{n!} u^n$

Preuve

Prouvons la convergence de $\sum \frac{1}{n!} u^n$: $\frac{1}{n!} \|u^n\| \leq \frac{\|u\|^n}{n!}$ qui est le terme général d'une série convergente. La série exponentielle converge absolument, et comme on est en dimension finie, elle converge aussi dans \mathcal{A} .

Exemple II.2. Dans $\mathcal{M}_p(\mathbb{K})$ qui est une algèbre de dimension p^2 , on peut définir pour $A \in \mathcal{M}_p(\mathbb{K})$: $\exp(A) = \sum_{n=0}^{+\infty} \frac{1}{n!} A^n$

Proposition : Série géométrique

Avec \mathcal{A} algèbre normée, on peut définir pour $u \in \mathcal{A}$ tel que $\|u\| < 1$, la somme géométrique $\sum_{n=0}^{+\infty} u^n$.

Et si $\|u\| < 1$, alors $1 - u$ est inversible et $(1 - u)^{-1} = \sum_{n=0}^{+\infty} u^n$

Preuve

Pour $n \in \mathbb{N}$, $(1 - u) \sum_{k=0}^n u^k = 1 - u^{n+1}$ et donc par passage à la limite quand $n \rightarrow +\infty$, on a

$(1 - u) \sum_{k=0}^{+\infty} u^k = 1$. Donc $(1 - u)$ possède un inverse à droite.

De même : $\sum_{n=0}^{+\infty} (1 - u) = 1$, d'où le résultat.

Exemple II.3. Exemple d'anneau où un inverse à droite n'est pas un inverse à gauche aussi :

$$f : \begin{cases} \mathbb{R}[X] & \rightarrow & \mathbb{R}[X] \\ X^p & \mapsto & X^{p+1} \end{cases} \quad \text{et} \quad g : \begin{cases} \mathbb{R}[X] & \rightarrow & \mathbb{R}[X] \\ X^p & \mapsto & \begin{cases} X^{p-1} & \text{si } p \geq 1 \\ 0 & \text{sinon} \end{cases} \end{cases}$$

On peut alors composer les deux applications, mais selon l'ordre dans lequel on les compose, on n'aura pas nécessairement l'identité.

II.3 Familles sommables d'un K-ev

Définition : Familles sommables d'un K -ev de dimension finie

Avec E un \mathbb{K} -ev de dimension finie muni d'une base $B = (e_1, \dots, e_p)$, avec I dénombrable et avec $(u_i)_{i \in I} \in E^I$, on dit que $(u_i)_{i \in I}$ est sommable si $\|u_i\|_{i \in I}$ est sommable.

Proposition : Sommabilité des familles coordonnées

Une famille $(u_i)_{i \in I}$ s'écrivant : $\forall i \in I, u_i = \sum_{k=1}^p u_i^k e_k$ est sommable si, et seulement si, $\forall k \in \{1, \dots, p\}, (u_i^k)_{i \in I}$ est sommable.

Tous les théorèmes comme la sommations par paquets ou l'indigage en identifiant I à \mathbb{N} par une bijection sont valides pour toutes les familles sommables de E .

Exemple II.4. Si \mathcal{A} est une algèbre de dimension finie, alors avec $u, v \in \mathcal{A}, uv = vu$ (u et v commutants), on a que $\exp(u + v) = \exp(u) \exp(v)$ (la commutativité est utilisée lors du binôme de Newton)

$$\sum_{i \in \mathbb{N}} \frac{1}{i!} u^i \sum_{j \in \mathbb{N}} \frac{1}{j!} v^j = \sum_{(i,j) \in \mathbb{N}^2} \frac{1}{i!} u^i \frac{1}{j!} v^j = \sum_{s \in \mathbb{N}} \frac{1}{s!} \sum_{i+j=s} \frac{s!}{i!j!} u^i v^j = \sum_{s \in \mathbb{N}} \frac{1}{s!} (u + v)^s$$

II.4 Suites extraites

Définition : Suites extraites

Soit E un \mathbb{K} -EVN, avec $(u_n) \in E^{\mathbb{N}}$. On extrait une suite de (u_n) en se donnant une fonction $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante. $(u_{\varphi(n)})_{n \in \mathbb{N}}$ est alors une suite extraite de (u_n) .

Théorème : Convergence des suites extraites

Si une suite $(u_n) \in E^{\mathbb{N}}$ converge l alors toute suite extraite de (u_n) converge vers l .

Théorème : Union d'applications d'extraction

Si φ, ψ sont deux applications de \mathbb{N} dans \mathbb{N} strictement croissantes vérifiant $\varphi(\mathbb{N}) \cup \psi(\mathbb{N}) = \mathbb{N}$, si $(u_{\varphi(n)})$ et $(u_{\psi(n)})$ convergent vers l , alors u_n est convergente de limite l .

Théorème : Bolzano-Weierstrass

De toute suite bornée dans un \mathbb{K} -ev de dimension finie on peut extraire une suite convergente.

Preuve

Soit E un \mathbb{R} -ev de dimension finie, avec $B = (e_1, \dots, e_p)$ une base et $(u_n) \in E^{\mathbb{N}}$ bornée.

Pour $n \in \mathbb{N}, u_n = \sum_{i=1}^p u_n^i e_i$, donc $(u_n^1) \in \mathbb{R}^{\mathbb{N}}$ est bornée, donc on peut en extraire une suite convergente $(u_{\varphi_1(n)}^1)$ de (u_n^1) .

Pour (u_n^2) , on a encore une suite bornée réelle, on peut en extraire une suite convergente $(u_{\varphi_1 \circ \varphi_2(n)}^2)$. On itère ce résultat jusqu'à $(u_{\varphi_1 \circ \dots \circ \varphi_{p-1}(n)}^p)$ qui est une suite bornée de $\mathbb{R}^{\mathbb{N}}$, donc on peut en extraire la suite $(u_{\varphi_1 \circ \dots \circ \varphi_p(n)}^p)$ convergente.

On a donc que $\forall i \in \{1, \dots, p\}, (u_{\varphi_1 \circ \dots \circ \varphi_p(n)}^i)$ est convergente, donc $(u_{\varphi_1 \circ \dots \circ \varphi_p(n)})$ est convergente.

Le cas dans un \mathbb{C} -ev en découle par le fait que les \mathbb{C} -ev de dimension p soient isomorphes aux \mathbb{R} -ev de dimensions $2p$.

III Topologie

Dans toute cette section, E sera un \mathbb{K} -EVN de dimension quelconque

III.1 Ouverts

Définition : Point intérieur

Soit $A \subset E$ et $a \in A$, on dit que a est intérieur à A si $\exists \alpha \in \mathbb{R}_+^*, \mathcal{B}(a, \alpha) \subset A$

Définition : Intérieur

Soit $A \subset E$, on appelle intérieur de A l'ensemble des points intérieurs de A , noté $\overset{\circ}{A}$; on a donc $\overset{\circ}{A} \subset A$

Exemple III.1. Soient $x \in E, r \in \mathbb{R}_+^*$, alors $\mathcal{B}_f(\overset{\circ}{x}, r) = \mathcal{B}(x, r)$.

Soit $y \in \mathcal{B}(x, r)$.

Considérons $r' = r - \|y - x\|$.

Soit $z \in E$, tel que $\|z - y\| \leq r'$, alors on a que $\|z - x\| \leq \|z - y\| + \|y - x\| \leq r$

Donc $\mathcal{B}(y, r) \subset \mathcal{B}_f(x, r)$ donc $y \in \mathcal{B}_f(x, r)$

Pour l'intré inclusion : soit $y \in E$ tel que $\|y - x\| = r$. Soit $\alpha \in \mathbb{R}_+^*$.

Considérons $z = y + \frac{\alpha}{2\|y-x\|}(y-x)$

alors $z \in \mathcal{B}(y, \alpha)$

$\|z - x\| = (1 + \frac{\alpha}{2\|y-x\|})\|y - x\| = r + \frac{\alpha}{2} > r$

Donc $z \notin \mathcal{B}_f(x, r)$

donc : $\forall \alpha \in \mathbb{R}_+^*, \mathcal{B}(y, \alpha) \cap \mathcal{CB}_f(x, r) \neq \emptyset$

Donc $\mathcal{B}_f(\overset{\circ}{x}, r) = \mathcal{B}(x, r)$

Définition : Ouvert

Soit $A \subset E$, on dit que A est ouvert si tous les points de A sont intérieurs (ou si A contient un voisinage de chacun de ses points), ce qui signifie $A \subset \overset{\circ}{A}$

Proposition : Ouverture de la boule

Une boule ouverte d'un EVN est un ouvert

Preuve

Soient $x \in E, r \in \mathbb{R}_+^*$

Soit $y \in \mathcal{B}(x, r)$; on pose $r' = r - \|x - y\|$

Soit $z \in \mathcal{B}(y, r') : \|z - x\| \leq \|z - y\| + \|y - x\| < r - \|y - x\| + \|y - x\| < r$

Donc $\mathcal{B}(y, r') \subset \mathcal{B}(x, r)$

Donc $\mathcal{B}(x, r)$ est un ouvert.

Proposition : Plus grand ouvert

$\overset{\circ}{A}$ est le plus grand ouvert inclus dans A (au sens de l'inclusion)

Preuve

Montrons que $\overset{\circ}{A}$ est un ouvert :

Soit $x \in \overset{\circ}{A}$: par définition de $\overset{\circ}{A}$, on peut fixer $\alpha \in \mathbb{R}_+^*$ tel que $\mathcal{B}(x, \alpha) \subset A$

Soit $y \in \mathcal{B}(x, \alpha)$

$\exists \alpha' \in \mathbb{R}_+^*, \mathcal{B}(y, \alpha') \subset \mathcal{B}(x, \alpha) \subset A$ puisque que $\mathcal{B}(x, \alpha)$ est un ouvert

Donc $y \in \overset{\circ}{A}$

Donc $\mathcal{B}(x, \alpha) \subset \overset{\circ}{A}$

$\overset{\circ}{A}$ est donc un ouvert inclus dans A

Pour la maximalité : Soit B un ouvert, $B \subset \overset{\circ}{A}$.

Soit $b \in B$

Comme B ouvert $\exists \alpha \in \mathbb{R}_+^*, \mathcal{B}(b, \alpha) \subset B$

$b \in \overset{\circ}{A}$

Donc $B \subset \overset{\circ}{A}$

Proposition : Réunion et intersection d'ouverts

Avec E un EVN, $(O_i)_{i \in I}$ une famille d'ouverts de E , on a que :

- $\cup_{i \in I} O_i$ est un ouvert
- $\cap_{i \in I} O_i$ est un ouvert à condition que I soit fini

III.2 Adhérence

Définition : Point adhérent

Soit $A \subset E$ et $x \in E$, on dit que x est adhérent à A si : $\forall \alpha \in \mathbb{R}_+^*, \mathcal{B}(x, \alpha) \cap A \neq \emptyset$

Définition : Adhérence

Soit $A \subset E$, on appelle adhérence de A l'ensemble des points adhérents à A , noté \bar{A} . On notera que $A \subset \bar{A}$

Exemple III.2. On a que $\overset{\circ}{\mathbb{Q}} = \emptyset$ par densité des irrationnels dans les réels, et, symétriquement, $\bar{\mathbb{Q}} = \mathbb{R}$

Théorème : Caractérisation séquentielle de l'adhérence

Soient $A \subset E, x \in E$, alors x est adhérent à A si, et seulement si, $\exists (a_n) \in A^{\mathbb{N}}, (a_n) \rightarrow x$

Preuve

Procédons par double-implication.

Supposons que x soit adhérent à A : alors on peut choisir (axiome du choix) pour $n \in \mathbb{N}$, $a_n \in A \cap \mathcal{B}(x, \frac{1}{n+1})$

La suite ainsi construite converge vers x et $(a_n) \in A^{\mathbb{N}}$

Réciproquement : on suppose l'existence de $(a_n) \in A^{\mathbb{N}}, (a_n) \rightarrow x$.

soit $\alpha \in \mathbb{R}_+^*$

comme $(a_n) \rightarrow x$, $\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow a_n \in \mathcal{B}(x, \alpha)$

Donc $\mathcal{B}(x, \alpha) \cap A \neq \emptyset$

D'où l'équivalence.

Définition : Fermé

Si $A \subset E$, A est fermé si $\bar{A} = A$ (donc que $\bar{A} \subset A$), donc que A contient tous les points qui lui sont adhérents.

Théorème : Caractérisation séquentielle des fermés

Soit $A \subset E$, A est fermé si, et seulement si, pour toute suite d'éléments de A convergente vers l , $l \in A$. ie $\forall (a_n) \in A^{\mathbb{N}}, (a_n) \rightarrow l \Rightarrow l \in A$

Preuve

Découle de la caractérisation séquentielle de l'adhérence.

Exemple III.3. Soit $A = \{(x, y) \in \mathbb{R}^3 | x + 2y \geq 0\}$, montrons que A est fermé. Soit $((\alpha_n, \beta_n)) \in A^{\mathbb{N}}$.

On suppose que $((\alpha_n, \beta_n)) \rightarrow (a, b) \in \mathbb{R}^2$

Or $\forall n \in \mathbb{N}, \alpha_n + 2\beta_n \geq 0$ et on a que $(\alpha_n) \rightarrow a$ et $(\beta_n) \rightarrow b$

donc par passage à la limite : $a + 2b \geq 0$

donc $(a, b) \in A$

$\{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 \leq 1 \text{ et } 0 \leq z \leq x^2 + y^2\}$ est un fermé

$\mathcal{S}_p(\mathbb{R}) = \{M \in \mathcal{M}_p(\mathbb{R}) | M^T = M\}$, montrons que $\mathcal{S}_p(\mathbb{R})$ est un fermé de $\mathcal{M}_p(\mathbb{R})$.

Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{S}_p(\mathbb{R})$. On suppose que $(A_n) \rightarrow M \in \mathcal{M}_p(\mathbb{R})$

$(A_n^T) \rightarrow M^T$ donc que $\forall i, j \in \llbracket 1, p \rrbracket, [A_n^t]_{i,j} \rightarrow [M]_{i,j}$

et $\forall n \in \mathbb{N}, A_n^T = A_n$

Donc par passage à la limite : $M^T = M$

Donc $M \in \mathcal{S}_p(\mathbb{R})$

Donc $\mathcal{S}_p(\mathbb{R})$ est un fermé.

La sphère $S(x, r) = \{y \in E | \|y - x\| = r\}$ est un fermé.

Proposition : Fermeture de la boule

Une boule fermée est un fermé

Preuve

Soit $x \in E$

Soit $r \in \mathbb{R}_+^*$

Soit $(u_n) \in \mathcal{B}_f(x, r)^{\mathbb{N}}$ qui tend vers $y \in E$.

Or $\forall n \in \mathbb{N}, \|u_n - x\| \leq r$

et comme $(u_n) \rightarrow y$, on a $(u_n - x) \rightarrow y - x$

On a que la limite de la norme d'une suite est la norme de sa limite par la deuxième inégalité triangulaire : $|\|v_n\| - \|m\|| \leq \|v_n - m\|$

Donc $\|u_n - x\| \rightarrow \|y - x\|$

Et donc par passage à la limite : $\|y - x\| \leq r$

donc $y \in \mathcal{B}_f(x, r)$

Proposition : Produit de fermés

Tout produit cartésien fini de fermés est un fermé

Exemple III.4. Soit $(u_n) \in E^{\mathbb{N}}$. On dit que $x \in E$ est une valeur d'adhérence de u si x est limite d'une suite extraite de E . Par exemple, les valeurs d'adhérence de $(-1)^n$ sont 1 et (-1) et les valeurs d'adhérence

de $\sin(n)$ sont l'ensemble des réels entre -1 et 1 . Attention à ne pas confondre les valeurs d'adhérence d'une suite et l'adhérence d'une suite, qui est l'ensemble des points de la suite.

On peut reformuler Bolzano-Weierstrass : Toute suite bornée admet au moins une valeur d'adhérence.

Exercice : Montrer que l'ensemble des valeurs d'adhérence d'une suite est fermé.

Théorème : Complémentarité d'un ouvert

Soit E un EVN et $A \subset E$, alors A est fermé si, et seulement si, $\mathcal{C}_E A$ est ouvert.

Preuve

La proposition est équivalente à : A est ouvert si, et seulement si, $\mathcal{C}_E A$ est fermé. On montrera la contraposée des deux implications.

Soit $A \subset E$

On suppose A non-ouvert. On peut donc prendre $a \in A$ tel que $\forall \alpha \in \mathbb{R}_+^*, \mathcal{B}(a, \alpha) \cap \mathcal{C}_E A \neq \emptyset$

Donc a est adhérent à $\mathcal{C}_E A$ et $a \notin \mathcal{C}_E A$

Donc le complémentaire de A dans E n'est pas fermé.

On suppose A non-fermé. On a donc l'existence de $x \in \bar{A} \setminus A$.

Et donc $\forall \alpha \in \mathbb{R}_+^*, \mathcal{B}(x, \alpha) \cap A \neq \emptyset$

Donc $x \in \mathcal{C}_E A$ et $x \notin \mathcal{C}_E A$

Donc $\mathcal{C}_E A$ n'est pas ouvert.

"Plutôt que de me prendre l'oreille gauche avec la main droite, je prends A " - Sami Chakroun, tranquillement, 2022

Proposition : Réunion et intersection de fermés

Avec E un EVN, $(F_i)_{i \in I}$ une famille de fermés de E , on a que :

- $\cap_{i \in I} F_i$ est un fermé
- $\cup_{i \in I} F_i$ est un fermé à condition que I soit fini

Preuve

Si $(F_i)_{i \in I}$ est une famille de fermés, alors $\mathcal{C}_E \cap_{i \in I} F_i = \cup_{i \in I} \mathcal{C}_E F_i$ qui est ouvert

$\mathcal{C}_E \cup_{i \in I} F_i = \cap_{i \in I} \mathcal{C}_E F_i$

Exemple III.5. $A = \{(x, y) \in \mathbb{R}^2 | x^2 + y < 0 \text{ et } x + 3y > 0\}$ est un ouvert comme son complémentaire est une réunion de fermés.

$\cup_{n \in \mathbb{N}^*} \left[\frac{1}{n}, 1 \right] =]0, 1]$ est une réunion infinie de fermés qui est ouverte.

Définition : Densité

Soit $A \subset E$. On dit que A est dense dans E si $\bar{A} = E$

i.e. $\forall x \in E, \exists (a_n) \in A^{\mathbb{N}}, (a_n) \rightarrow x$

i.e. Tout élément de E est limite d'une suite d'éléments de A

Exemple III.6. \mathbb{Q} est dense dans \mathbb{R} , \mathbb{Q}^2 est dense dans \mathbb{R}^2 , $\mathbb{Q} \times (\mathbb{R} \setminus \mathbb{Q})$ est dense dans \mathbb{R}^2 ...

Définition : Frontière

Soit $A \subset E$, on appelle frontière de A : $F_r(A) = \bar{A} \setminus \overset{\circ}{A}$

Exemple III.7. Avec $x \in E, r \in \mathbb{R}_+^*$, et donc $F_r(B(x, r)) = S(x, r)$
On a que $F_r(\mathbb{Q}) = \mathbb{R}$

Remarque III.1. Toutes les notions topologiques évoquées dans ce chapitre sont invariantes par changement de normes équivalentes. En particulier en dimension finie, ces notions sont indépendantes de la norme.

III.3 Relativité vaguement générale quand même

Définition : Ouvert relatif

Soit E un EVN, $A \subset E$. Soit $B \subset A$. On dit que B est ou ouvert relatif de A si $\forall x \in B, \exists \alpha \in \mathbb{R}_+^*, \mathcal{B}(x, \alpha) \cap A \subset B$

Proposition : Ouvert induit

Soit E un EVN, $A \subset E$. Avec $B \subset A$, B est un ouvert relatif de A si, et seulement si, il existe un ouvert O de E tel que $B = O \cap A$

Preuve

Supposons O ouvert de E et $B = O \cap A$.

Soit $x \in B$. Comme O est ouvert, on peut fixer $\alpha \in \mathbb{R}_+^*$ tel que

$\mathcal{B}(x, \alpha) \subset O$

Donc $\mathcal{B}(x, \alpha) \cap A \subset B$

Réciproquement, si B est un ouvert relatif de A , pour $b \in B$, on peut choisir $\alpha_b \in \mathbb{R}_+^*, \mathcal{B}(b, \alpha_b) \cap A \subset B$

Donc $B = \cup_{b \in B} (\mathcal{B}(b, \alpha_b) \cap A) = (\cup_{b \in B} \mathcal{B}(b, \alpha_b)) \cap A$

Donc il existe un ouvert (réunion d'ouverts) dont l'intersection avec A donne B .

Exemple III.8. Avec $E = \mathbb{R}$ et $A =]0, 1]$. $B =]\frac{1}{2}, 1]$ n'est pas un ouvert de E car $1 \in B$ et 1 n'est pas intérieur à B . Mais c'est un ouvert relatif de A car $B =]\frac{1}{2}, 2[\cap A$.

Définition : Fermé relatif

Soit E un EVN, $A \subset E$. Soit $B \subset A$. On dit que B est un fermé relatif de A s'il vérifie l'une des propriétés équivalentes suivantes :

1. $\mathcal{C}_A B$ est un ouvert relatif de A ($\mathcal{C}_A B = A \setminus B = A \cap \mathcal{C}_E B$);
2. $\exists F$ fermé de E tel que $B = F \cap A$;
3. Caractérisation séquentielle : pour toute suite $(b_n) \in B^{\mathbb{N}}$ convergente de limite $l \in A$, alors $l \in B$.

III.4 Compacts

Définition : Compact

Soit E un EVN, soit $A \subset E$. On dit que A est compact si de toute suite de A , on peut extraire une suite convergente dans A .

Exemple III.9. Les segments sont des compacts de \mathbb{R} , les réunions de segments sont des compacts, les intervalles semi-ouverts ne sont pas des compacts.

Pour $a, b \in \mathbb{R}, a < b$, $[a, b]$ est un compact de \mathbb{R} mais $[a, b[$ n'est pas compact. $[a, b]$ est bornée, donc Bolzano-Weierstrass pour l'existence d'une suite convergente, qui a bien une limite dans le segment $[a, b]$ par passage à la limite et inégalités larges. Cependant, cette inégalité large mène à ce que b soit une limite possible, alors que b n'est pas dans l'ensemble.

Proposition : Fermeture du compact

Tout compact est fermé et borné.

Preuve

Soit K un compact.

Soit $(k_n) \in K^{\mathbb{N}}$ une suite convergent vers $l \in E$.

Comme K est compact, on peut extraire de (k_n) une suite $(k_{\varphi(n)})$ de limite dans K .

Or comme (k_n) est convergente de limite l , alors $(k_{\varphi(n)}) \rightarrow l$ donc $l \in K$.

Donc K est fermé.

Supposons K non-borné.

On peut construire une suite (k_n) de K telle que $\|k_n\| \rightarrow +\infty$ (comme K est non-borné, pour tout $n \in \mathbb{N}$, on peut choisir $k_n \in K$ tel que $\|k_n\| \geq n$)

Pour toute suite extraite de (k_n) $(k_{\varphi(n)})$, la suite $\|k_{\varphi(n)}\| \rightarrow +\infty$ donc $(k_{\varphi(n)})$ diverge.

Donc K est non-compact.

Proposition : Produit cartésien de compacts

Avec E_1, \dots, E_p des EVN, $E = E_1 \times E_2 \times \dots \times E_p$ et K_1, \dots, K_p des compacts, alors $K_1 \times K_2 \times \dots \times K_p$ est un compact.

Preuve

Soit $(x_n) \in (K_1 \times K_2 \times \dots \times K_p)^{\mathbb{N}}$ (ie $(x_n) = (x_n^1, x_n^2, \dots, x_n^p)$), et on refait la démonstration de BW. La seule différence c'est qu'au lieu d'avoir des suites scalaires, on a des suites de vecteurs.

K_1 est compact, donc on peut extraire de (x_n^1) une suite $(x_{\varphi_1(n)}^1)$ convergente dans K_1 . Ensuite $(x_{\varphi_2(n)}^2) \in K_2^{\mathbb{N}}$, qui est compact, donc on peut extraire de $(x_{\varphi_1(n)}^2)$ une suite $(x_{\varphi_1 \circ \varphi_2(n)}^2)$ convergente dans K^2 .

Au final on aura une suite extraite $(x_{\varphi_1 \circ \dots \circ \varphi_p})$ qui converge.

Proposition : Fermé d'un compact

Un fermé dans un compact est compact (ou un fermé relatif d'un compact, même si tout fermé relatif d'un compact serait un fermé)

Preuve

Soit K un compact et $B \subset K$ fermé.

Soit $(x_n) \in B^{\mathbb{N}}$, alors $(x_n) \in K^{\mathbb{N}}$.

Alors on peut extraire $(x_{\varphi(n)})$ de (x_n) convergente vers $l \in K$.

Comme B est fermé, $l \in B$.

Donc B est compact.

Théorème : Compacts

Dans un espace de dimension finie E , les compacts sont les fermés bornés.

Preuve

On a déjà vu que tout compact est fermé borné. Réciproquement :

Soit A un fermé borné. Soit $(x_n) \in A^{\mathbb{N}}$

A est bornée, donc par Bolzano-Weierstrass en dimension finie, on peut trouver une suite extraite $(x_{\varphi(n)})$ qui converge vers $l \in E$.

Comme A est fermé, $l \in A$.

Donc A est compact.

Exemple III.10. Une boule fermée en dimension finie est compacte. En dimension infinie, il est possible de démontrer qu'une boule fermée n'est jamais compacte.

Une sphère est compacte en dimension finie.

On prend E un \mathbb{K} -ev et $A \subset E$, $A \neq \emptyset$. Pour $x \in E$, on définit $d(x, A) = \inf_{a \in A} \|x - a\|$.

1) Montrer que si A est compact, $d(x, A)$ est atteinte.

2) Montrer que si E est de dimension finie et A est fermée, alors $d(x, A)$ est atteinte.

3) Pour $A, B \subset E$ on définit $d(A, B) = \inf\{\|a - b\|, a \in A, b \in B\}$.

a) Montrer que si A et B sont compacts, $d(A, B)$ est atteinte.

Si E de dim finie, A compact et B fermé, $d(A, B)$ atteinte.

c) Si A et B sont fermés, $d(A, B)$ est-elle atteinte ?

1. Dans \mathbb{R} , $\sup A \in \bar{A}$ et $\inf A \in \bar{A}$. Ici, $d(x, A) = \inf\{\|x - a\|, a \in A\}$ donc on peut trouver $(a_n) \in A^{\mathbb{N}}$ telle que $d(x, A) = \lim_{n \rightarrow +\infty} \|x - a_n\|$.

A est un compact, et comme (a_n) est à valeurs dans un compact, on peut extraire de (a_n) une suite convergente $(a_{\varphi(n)})$ de limite $\alpha \in A$.

$$(\|x - a_{\varphi(n)}\|) \rightarrow d(x, A)$$

$$\text{Et } \|x - a_{\varphi(n)}\| \rightarrow \|x - \alpha\|$$

Donc la distance est atteinte.

2. Avec E de dimension finie et A fermé, $d(x, A) = \lim_{n \rightarrow +\infty} \|x - a_n\|$ avec $(a_n) \in A^{\mathbb{N}}$ qu'on prend.

$$\forall n \in \mathbb{N}, \|a_n\| \leq \|x\| + \|x - a_n\|$$

$(\|x - a_n\|)$ est convergente donc bornée

Donc (a_n) est bornée.

Donc par Bolzano-Weierstrass, on peut extraire de (a_n) une suite $(a_{\varphi(n)})$ convergente vers $l \in E$. Comme A est fermée et que (a_n) est une suite convergente de A , $l \in A$.

$$\text{Donc } d(x, A) = \|x - l\|$$

3. — On a A et B deux compacts, donc on prend $(a_n) \in A^{\mathbb{N}}$, $(b_n) \in B^{\mathbb{N}}$ telle que $d(A, B) = \lim_{n \rightarrow +\infty} \|a_n - b_n\|$

Comme A est borné, on peut prendre $(a_{\varphi(n)})$ convergente vers $l \in A$.

De $(b_{\varphi(n)})$ bornée, on peut extraire la suite convergente $(b_{\varphi \circ \psi(n)})$ qui tend vers $l' \in B$

$$\text{Donc } (\|a_{\varphi \circ \psi(n)} - b_{\varphi \circ \psi(n)}\|) \rightarrow \|l - l'\| = d(A, B)$$

Donc la distance est atteinte.

— Avec E de dimension finie, A compact et B fermé. On a que $d(A, B) = \lim_{n \rightarrow +\infty} \|a_n - b_n\|$ en ayant pris $(a_n) \in A^{\mathbb{N}}$ et $b_n \in B^{\mathbb{N}}$.

A est compact, donc on extrait de (a_n) la suite $(a_{\varphi(n)})$ convergente vers $l \in A$.

$$\text{Donc } d(A, B) = \lim_{n \rightarrow +\infty} \|a_{\varphi(n)} - b_{\varphi(n)}\|$$

Et comme $\forall n \in \mathbb{N}, \|b_{\varphi(n)}\| \leq \|a_{\varphi(n)}\| + \|b_{\varphi(n)} - a_{\varphi(n)}\|$, alors $b_{\varphi(n)}$ est bornée.

Donc par BW (car la dimension est finie), on peut extraire de $(b_{\varphi(n)})$ une suite extraite $(b_{\varphi \circ \psi(n)})$ convergent vers l' .

Comme B est fermé, $l' \in B$. Donc $d(A, B) = \|l - l'\|$

- Prenons un contre-exemple en dimension infinie, dans l'espace vectoriel des courbes de fonctions. Avec $A = \{(x, e^x) | x \in \mathbb{R}\}$ et $B = \{(x, -e^x) | x \in \mathbb{R}\}$ on a une distance de 0, bien que les deux courbes ne se rencontrent jamais. $O_n(\mathbb{R})$ (groupe orthogonal) n'est plus du tout au programme mais permettrait d'avoir un exemple d'un compact.

Proposition : Convergence d'une suite compacte

Une suite d'un compact est convergente si, et seulement si, elle admet une unique valeur d'adhérence

Preuve

Soit A un compact.

Montrons la première implication : soit $(u_n) \in A^{\mathbb{N}}$ convergente.

Alors elle admet une unique valeur d'adhérence.

Réciproquement, montrons la contraposée : toute suite divergente dans un compact admet au moins deux valeurs d'adhérence.

Soit $(u_n) \in A^{\mathbb{N}}$ divergente.

A est compact, donc (u_n) possède une valeur d'adhérence λ .

Mais (u_n) ne tend pas vers λ , on a donc : $\exists \varepsilon \in \mathbb{R}_+^*, \forall n_0 \in \mathbb{N}, \exists n \in \mathbb{N}, n \geq n_0 \Rightarrow \|u_n - \lambda\| > \varepsilon$

Donc $\{n \in \mathbb{N} | \|u_n - \lambda\| > \varepsilon\}$ est infini.

Donc on peut extraire de (u_n) une suite $(u_{\varphi(n)})$ telle que $\forall n \in \mathbb{N}, \|u_{\varphi(n)} - \lambda\| > \varepsilon$

Mais $(u_{\varphi(n)})$ est une suite de A , donc elle possède une valeur d'adhérence $\lambda' \neq \lambda$, avec $\lambda' \in A$.

IV Limites de fonctions

IV.1 Définitions

Définition : Limite

Soient E, F deux EVN, et $A \subset E$. Soit $f \in \mathcal{F}(A, F)$, $x_0 \in \bar{A}$. Soit $l \in F$. On dit que f converge vers l en x_0 si :

$$\forall \varepsilon \in \mathbb{R}_+^*, \exists \alpha \in \mathbb{R}_+^*, \forall x \in \mathcal{B}(x_0, \alpha) \cap A, f(x) \in \mathcal{B}(l, \varepsilon)$$

On notera $f \rightarrow_{x_0} l$ ou $f(x) \rightarrow_{x \rightarrow x_0} l$ (notation plus abusive) ou $\lim f = l$ (notation plus adaptée à une conclusion) et $\lim_{x \rightarrow x_0} f(x) = l$ (le combo que Chakroun aime pas).

Définition : Vocabulaire du voisinage

Avec E un EVN et $b \in E$, on appelle voisinage de b un ensemble contenant une boule ouverte centrée en b .

Avec ce vocabulaire, $f \rightarrow_{x_0} l$ si, pour tout voisinage de l \mathcal{V}_l , il existe un voisinage relatif de x_0 \mathcal{V}_{x_0} tel que $f(\mathcal{V}_{x_0}) \subset \mathcal{V}_l$

On étend la définition de voisinage à $+\infty$:

Dans E un EVN quelconque, \mathcal{V} est un voisinage de $+\infty$ est si $\exists A \in \mathbb{R}_+^*, \{x \in E, \|x\| > A\} \subset \mathcal{V}$.

Dans \mathbb{R} on définit des voisinages de $\pm\infty$

\mathcal{V} est un voisinage de $+\infty$ si $\exists A \in \mathbb{R}_+^*,]A, +\infty[\subset \mathcal{V}$

\mathcal{V} est un voisinage de $-\infty$ si $\exists B \in \mathbb{R}_-^*,]-\infty, B[\subset \mathcal{V}$

On peut aussi étendre la notion de convergence en $+\infty$ (ou $-\infty$)

Exemple IV.1. $f \rightarrow_{+\infty} l$ donne que $\forall \varepsilon \in \mathbb{R}_+^*, \exists M \in \mathbb{R}_+^*, \forall x \in A, \|x\| \geq M \Rightarrow \|f(x) - l\| < \varepsilon$

Proposition : Caractérisation séquentielle de la limite

Avec $f \in \mathcal{F}(A, F)$ avec $A \subset E$ et E, F deux EVN. Avec $x_0 \in \bar{A}$ et $l \in F$, alors :

$$f \rightarrow_{x_0} l \Leftrightarrow \forall (u_n) \in A^{\mathbb{N}}, (u_n) \rightarrow x_0, (f(u_n)) \rightarrow l$$

Preuve

On suppose que f tend vers l en x_0 :

Soit $(u_n) \in A^{\mathbb{N}}$ tel que $(u_n) \rightarrow x_0$

Soit $\varepsilon \in \mathbb{R}_+^*$

$f \rightarrow_{x_0} l$, on peut trouver $\alpha \in \mathbb{R}_+^*, \forall x \in A \cap \mathcal{B}(x_0, \alpha), \|f(x) - l\| < \varepsilon$

Or $(u_n) \rightarrow x_0$ donc on peut trouver $n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow \|u_n - x_0\| < \alpha$

Donc pour $n \geq n_0$: $\|f(u_n) - l\| < \varepsilon$ On suppose que $\forall (u_n) \in A^{\mathbb{N}}, (u_n) \rightarrow x_0, (f(u_n)) \rightarrow l$: on prouve la contraposée :

f ne tend pas vers l en x_0 donc on peut construire une suite (u_n) tendant vers x_0 telle que $(f(u_n))$ ne tend pas vers l .

On peut donc prendre un $\varepsilon \in \mathbb{R}_+^*$ tel que $\forall \alpha \in \mathbb{R}_+^*, \exists x \in A \cap \mathcal{B}(x_0, \alpha)$ et $\|f(x) - l\| > \varepsilon$

Pour $n \in \mathbb{N}^*$, on peut donc choisir $u_n \in A \cap \mathcal{B}(x_0, \frac{1}{n})$ et $\|f(u_n) - l\| > \varepsilon$

On a donc construit $(u_n) \in A^{\mathbb{N}}$ tel que $\forall n \in \mathbb{N}, \|u_n - x_0\| < \varepsilon_n$ et $\|f(u_n) - l\| > \varepsilon$

Exemple IV.2. $f : \begin{cases} \mathbb{R}^2 \setminus \{(0,0)\} & \rightarrow \mathbb{R} \\ (x,y) & \mapsto \frac{xy}{|x|+|y|} \end{cases}$

f a-t-elle une limite en $(0,0)$?

Soient $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, alors $|f(x,y)| \leq \frac{(|x|+|y|)^2}{|x|+|y|} \leq |x| + |y|$

Donc $f \rightarrow_{(0,0)} 0$

$f : \begin{cases} \mathbb{R}^2 \setminus \{(0,0)\} & \rightarrow \mathbb{R} \\ (x,y) & \mapsto \frac{x^\alpha y^\beta}{\|(x,y)\|^\gamma} \end{cases}$

f a-t-elle une limite en $(0,0)$?

On a l'équivalence des normes en dimension finie, donc $\exists K_1, K_2 > 0, \frac{K_1 |x^\alpha y^\beta|}{\|(x,y)\|_\infty^\gamma} \leq |f(x,y)| \leq \frac{K_2 |x^\alpha y^\beta|}{\|(x,y)\|_\infty^\gamma}$

Dans un premier cas, où $\alpha + \beta > \gamma$: $\frac{|x^\alpha y^\beta|}{\|(x,y)\|_\infty^\gamma} \leq \|(x,y)\|_\infty^{\alpha+\beta-\gamma}$

et donc $f \rightarrow_{(0,0)} 0$

Dans un second cas, $\alpha + \beta = \gamma$: Soit $\lambda \in \mathbb{R}$,

considérons la suite $(u_n) = ((\frac{1}{n}, \frac{\lambda}{n}))_{n \in \mathbb{N}}$

Alors $f(u_n) = \frac{(\frac{1}{n})^{\alpha+\beta} \lambda^\beta}{|\frac{1}{n}|^\gamma \|(1, \lambda)\|^\gamma} = \frac{\lambda^\beta}{\|(1, \lambda)\|^\gamma}$, une constante qui dépend de λ en général.

Donc on n'a pas de limite en $(0, 0)$

Dans un troisième cas, $\alpha + \beta < \gamma$:

Prenons la suite, pour $n \in \mathbb{N}$: $f\left(\left(\frac{1}{n}, \frac{1}{n}\right)\right) = \frac{(\frac{1}{n})^{\alpha+\beta-\gamma}}{\|(1, 1)\|^\gamma} \rightarrow +\infty$

Donc f n'a pas de limite en $(0, 0)$.

Attention, on n'a pas nécessairement $f \rightarrow_{(0,0)} +\infty$:

prenons le cas particulier de $f : \begin{cases} \mathbb{R}^2 \setminus \{(0, 0)\} & \rightarrow \mathbb{R} \\ (x, y) & \mapsto \frac{x^2 y}{(|x| + |y|)^4} \end{cases}$

On a alors, pour $n \in \mathbb{N}^*$: $f\left(\left(\frac{1}{n}, \frac{1}{n^2}\right)\right) = \frac{\frac{1}{n^4}}{2^4 \frac{1}{n^4}} \rightarrow \frac{1}{2^4}$ qui tend bien vers 0.

Avec $f : \begin{cases} \mathbb{R}^2 \setminus D & \rightarrow \mathbb{R} \\ (x, y) & \mapsto \frac{x^{12} y^{15}}{x+y} \end{cases}$ et avec $D = \{(x, y) \in \mathbb{R}^2 | x + y = 0\}$. Est-ce que f admet une limite en $(0, 0)$?

On prend la suite $|f\left(\frac{-1}{n}, \frac{1}{n} + \frac{1}{n^{100}}\right)| \sim \frac{\frac{1}{n^{27}}}{\frac{1}{n^{100}}} \rightarrow +\infty$

IV.2 Opérations sur les limites

Proposition : Limite d'une fonction à valeurs dans un produit cartésien

Soit $f : A \rightarrow F_1 \times F_2 \times \dots \times F_p$ avec (F_i) famille d'EVN et $A \subset E$ une partie d'un EVN. On a alors $f : x \mapsto (f_1(x), f_2(x), \dots, f_p(x))$ et on peut caractériser f par ses fonctions composantes $f_i : A \rightarrow F_i$. Soit $x_0 \in A$ et $l = (l_1, \dots, l_p) \in F_1 \times \dots \times F_p$, alors $f \rightarrow_{x_0} l$ si, et seulement si, $\forall i \in \{1, \dots, p\}, f_i \rightarrow_{x_0} l_i$. En particulier, si F est un EVN de dimension finie ramené à une base $B = (u_1, u_2, \dots, u_p)$, F est isomorphe à \mathbb{K}^p par la bijection $\sum_{i=1}^p x_i u_i \mapsto (x_1, \dots, x_p)$. Si $f \in \mathcal{F}(A, F) : x \mapsto f(x) = \sum_{i=1}^p f_i(x) u_i$, on appelle les f_i les applications coordonnées et f a une limite si, et seulement si, les applications coordonnées ont des limites.

Proposition : Opérations sur les fonctions convergentes

- Avec $f, g \in \mathcal{F}(A, E)$, on a que f converge et g converge $\Rightarrow f + g$ converge et $\lim(f + g) = \lim f + \lim g$
- Avec $\lambda \in \mathcal{F}(A, \mathbb{K}), f \in \mathcal{F}(A, F)$, λ converge et f converge $\Rightarrow \lambda f$ converge et $\lim(\lambda f) = (\lim \lambda)(\lim f)$
- Si F est une algèbre munie d'une norme d'algèbre : Si $f, g \in \mathcal{F}(A, F)$, g converge et f converge $\Rightarrow fg$ converge et $\lim(f \times g) = (\lim f) \times (\lim g)$
- Si $F = \mathbb{K}$, alors avec $f \in \mathcal{F}(A, F)$, si f est convergente de limite $l \neq 0$, alors $\frac{1}{f}$ converge et $\lim \frac{1}{f} = \frac{1}{l}$
- Avec E, F, G des EVN, avec $A \subset E$ et $B \subset F$, $f \in \mathcal{A}, g \in \mathcal{F}(B, G)$, $x_0 \in \bar{A}, y_0 \in \bar{G}, f \rightarrow_{x_0} y_0, g \rightarrow_{y_0} l$, alors on peut affirmer que $g \circ f \rightarrow_{x_0} l$

Preuve

Pour la composition : soit \mathcal{V}_l un voisinage de l , comme $g \rightarrow_{y_0} l$, on peut trouver un voisinage relatif à B de y_0 \mathcal{V}_{y_0} tel que $g(\mathcal{V}_{y_0}) \subset \mathcal{V}_l$.

Comme $f \rightarrow_{x_0} y_0$, on peut trouver un voisinage relatif à A de x_0 \mathcal{V}_{x_0} tel que $f(\mathcal{V}_{x_0}) \subset \mathcal{V}_{y_0}$

Donc $(g \circ f)(\mathcal{V}_{x_0}) \subset \mathcal{V}_l$

Remarque IV.1. Cette démonstration s'étend au cas $\|x_0\| = +\infty, \|y_0\| = +\infty$

IV.3 Continuité de fonctions

Définition : Continuité

Soit $f \in \mathcal{F}(A, F)$ avec $A \subset E$ et E, F deux EVN. Soit $a \in A$, on dit que f est continue en a si $f \rightarrow_a f(a)$.

Remarque IV.2. C'est une notion locale ; deux fonctions coïncidant sur un voisinage ont les mêmes propriétés de continuité.

Proposition : Opérations

Localement, la somme de deux fonctions continues est continue, le produit par un scalaire d'une fonction continue est continue, le produit d'une fonction continue et d'une fonction scalaire continue est continu, dans un algèbre le produit de deux fonctions continues est continu, la composition de deux fonctions continues est continue.

Remarque IV.3. Toutes les notions vues sur les limites de fonctions s'étendent sur les fonctions continues en a , notamment sur les fonctions coordonnées.

Proposition :

Si E est de dimension finie associé à une base $B = (e_1, \dots, e_p)$, alors :

$$P_j : \begin{cases} E & \rightarrow \mathbb{K} \\ x = \sum_{i=1}^p x_i e_i & \mapsto x_j \end{cases}$$

Alors : $\forall x, y \in E, x = \sum x_i e_i, y = \sum y_i e_i, |p_j(x) - p_j(y)| \leq \|x - y\|_\infty$ et l'application projection est une application continue en tous points. (1-lipschitzienne donc uniformément continue)

Exemple IV.3. $f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^3 \\ (x, y) & \mapsto \left(\frac{\arctan(xy^2 \sin(x+y))}{\sqrt{x^2+y^2+1}}, \frac{\log(1+x^2 y^4)}{\cosh(x+3y)}, x+y \right) \end{cases}$

Cette fonction est continue car toutes ses coordonnées sont continues, en tant que produits, sommes et compositions de fonctions continues. En effet, x et y , en tant que projecteurs, sont nécessairement continus.

$$f : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \rightarrow \mathbb{R} \\ A & \mapsto \det(A) \end{cases}$$

est une application continue en tant que polynôme avec comme coefficients des coordonnées.

$$f : \begin{cases} E & \rightarrow \mathbb{R} \\ x & \mapsto \|x\| \end{cases}$$

est continue car par seconde inégalité triangulaire ($\forall x, y \in E, |||x| - |y|| \leq \|x - y\|$), elle est lipschitzienne.

Avec E un espace préhilbertien réel :

$$f : \begin{cases} E \times E & \rightarrow \mathbb{R} \\ (u, v) & \mapsto \langle u, v \rangle \end{cases}$$

est continue, montrons-le, même quand E n'est pas de dimension finie :

En $(0, 0)$: Montrons que $\langle h, k \rangle \rightarrow_{(0,0)} 0 : |\langle h, k \rangle| \leq \|h\| \|k\|$ par Cauchy-Schwarz donc f est continue en $(0, 0)$

En (u_0, v_0) quelconques : $|\langle u_0 + h, v_0 + k \rangle - \langle u_0, v_0 \rangle| = |\langle h, v_0 \rangle + \langle k, u_0 \rangle + \langle h, k \rangle| \leq \|u_0\| \|h\| + \|v_0\| \|k\| + \|h\| \|k\|$

Donc f est continue en (u_0, v_0) et est donc continue sur $E \times E$.

IV.4 Propriétés globales des fonctions continues

Proposition : Images réciproques

Soit $f \in \mathcal{F}(A, F)$ continue, alors l'image réciproque d'un ouvert de F par f est un ouvert relatif de A

L'image réciproque d'un fermé de F par f est un fermé relatif de A .

Preuve

Pour les fermés :

Soit B un fermé de F , on veut démontrer que $f^{-1}(B)$ est un fermé relatif de A .

Soit $(x_n) \in (f^{-1}(B))^{\mathbb{N}}$, on suppose que $(x_n) \rightarrow l \in A$.

Alors pour $n \in \mathbb{N}$, $f(x_n) \in B$ et f est continue en l donc $f(x_n) \rightarrow f(l)$

Donc $(f(x_n))$ est une suite convergente du fermé B , donc sa limite est dans B

Donc $f(l) \in B$, donc $l \in f^{-1}(B)$

Donc $f^{-1}(B)$ est bien un fermé relatif de A .

Pour les ouverts :

Soit O un ouvert de F , $\mathcal{C}_F O$ est un fermé de F , donc $f^{-1}(\mathcal{C}_F O)$ est un fermé relatif de A .

Donc $\mathcal{C}_A f^{-1}(O)$ est un fermé relatif de A

Donc $f^{-1}(O)$ est un ouvert relatif de A .

Exemple IV.4. $A = \{(x, y) \in \mathbb{R}^2 \mid \sin(x) + 2y < 0 \text{ et } x + 3y > 0 \text{ et } \arctan(\frac{x}{x^2+y^2+1}) > 0\}$ est l'image réciproque par $f : (x, y) \mapsto (\sin(x) + 2y, x + 3y, \frac{x}{x^2+y^2+1})$ de $(\mathbb{R}_-, \mathbb{R}_+, \mathbb{R}_+)$, qui est un ouvert. Or f est continue, donc A est ouvert.

Proposition : Continuité coïncidente

Deux applications continues qui coïncident sur une partie dense sont égales

Preuve

Soient $f, g \in \mathcal{C}(A, F)$ avec $A \subset E$, E, F des EVN, B une partie dense dans A , et $\forall x \in B, f(x) = g(x)$

Soit $x \in A$, alors on prend $(x_n) \in B^{\mathbb{N}}$, $(x_n) \rightarrow x$ qui existe par densité de B

$\forall n \in \mathbb{N}, f(u_n) = g(u_n)$ et par passage à la limite (comme f et g sont continues) : $f(u_n) \rightarrow f(x)$ et $g(u_n) \rightarrow g(x)$

Et donc par unicité de la limite, $f(x) = g(x)$

Exemple IV.5. On recherche la partie de $\mathcal{C}(\mathbb{R}, \mathbb{R})$ telles que : $\forall x, y \in \mathbb{R}, f(x + y) = f(x) + f(y)$.

On prouve d'abord que $f(0) = 0$ et que $\forall x \in \mathbb{R}, \forall n \in \mathbb{N}, f(x) = nf(x)$, puis on l'étend à \mathbb{Z} et enfin à \mathbb{Q} . Et ensuite on se sert du théorème pour conclure

Définition : Uniforme continuité

Soit $f \in \mathcal{F}(A, F)$ où $A \subset E$ avec E, F deux EVN. On dit que f est uniformément continue si :

$$\forall \varepsilon \in \mathbb{R}_+^*, \exists \eta \in \mathbb{R}_+^*, \forall x, y \in A, \|x - y\| < \eta \Rightarrow \|f(x) - f(y)\| < \varepsilon$$

Remarque IV.4. La continuité uniforme implique la continuité.

On peut trouver un contre-exemple à la réciproque, comme $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto e^x \end{cases}$

La négation de la continuité uniforme, c'est que : $\exists \varepsilon \in \mathbb{R}_+^*, \forall \eta \in \mathbb{R}_+^*, \exists x, y \in A, (\|x - y\| < \eta) \text{ et } (\|f(x) - f(y)\| \geq \varepsilon)$

C'est peu exploitable, donc on fait une caractérisation séquentielle de la non-continuité uniforme : $\exists \varepsilon \in \mathbb{R}_+^*, \exists (x_n), (y_n) \in A^\mathbb{N}, (\|x_n - y_n\|) \rightarrow 0 \text{ et } \|f(x_n) - f(y_n)\| \geq \varepsilon$

On note P1 la négation et P2 la caractérisation séquentielle. Prouvons leur équivalence.

Supposons P1, et prenons $\varepsilon \in \mathbb{R}_+^*$ vérifiant la propriété.

Soit $(\eta_n) = (\frac{1}{n+1})$, prenons $(x_n), (y_n) \in A^\mathbb{N}$ telles que :

$$\|x_n - y_n\| < \eta_n \text{ et } \|f(x_n) - f(y_n)\| \geq \varepsilon$$

On a ainsi défini $(x_n), (y_n)$ avec $(\|x_n - y_n\|) \rightarrow 0$ et $\|f(x_n) - f(y_n)\| \geq \varepsilon$

L'autre implication est immédiate.

On reprend f telle que définie plus haut, on note $(x_n) = (n), (y_n) = (n + \frac{1}{n})$.

On a $(x_n - y_n) \rightarrow 0$

$$\text{Alors } e^{x_n} - e^{y_n} = e^n - e^{n+\frac{1}{n}} = e^n(1 - e^{\frac{1}{n}}) = e^n(\frac{1}{n} + o(\frac{1}{n})) = \frac{e^n}{n} + o(\frac{e^n}{n})$$

Or ça tend vers $+\infty$, donc on a bien une fonction non-uniformément continue.

"L'exponentielle ça amplifie à mort, donc on s'en fout." - Chakroun, 2022

IV.5 Fonctions lipschitziennes

Définition : Fonctions lipschitziennes

Soit $f \in \mathcal{F}(A, F)$, où $A \subset E$ avec E, F EVN. Soit $k \in \mathbb{R}_+^*$. On dit que f est k -lipschitzienne si :

$$\forall x, y \in A, \|f(x) - f(y)\| \leq k\|x - y\|$$

Proposition : Uniforme continuité des fonctions lipschitziennes

Toute fonction lipschitzienne est uniformément continue

Preuve

Le η qu'on prend est le k .

Remarque IV.5. Dans \mathbb{R} , si f est dérivable sur un intervalle I , f est lipschitzienne si, et seulement si, sa dérivée est bornée.

Si f' est bornée, on conclut par inégalité des accroissements finis.

Si f est dérivable, on passe à la limite de l'expression de la dérivée.

Exemple IV.6. Donnons un exemple de fonction uniformément continue non-lipschitzienne.

$$\text{Prenons } f : \begin{cases} [0, 1] & \rightarrow \mathbb{R} \\ x & \mapsto \sqrt{x} \end{cases}$$

Alors f est uniformément continue mais non lipschitzienne (elle est dérivable de dérivée non-bornée)

Proposition : Distance lipschitzienne

Avec E un EVN, $A \subset E$, l'application $\begin{matrix} E & \rightarrow & \mathbb{R} \\ x & \mapsto & d(x, A) \end{matrix}$ est 1-lipschitzienne.

Preuve

Soient $x, y \in E$.

Soit $a \in A$

$\|y - a\| \geq \|x - a\| - \|y - x\| \geq d(x, A) - \|x - y\|$ qui est constante

donc : $d(y, A) \geq d(x, A) - \|x - y\|$

La fonction est donc lipschitzienne

Théorème : Théorème de Heine

Toute fonction continue sur un compact est uniformément continue.

Preuve

Avec E, F deux EVN, A un compact de E , $f \in \mathcal{C}(A, F)$ non-uniformément continue.

On peut donc fixer $\varepsilon \in \mathbb{R}_+^*$ et prendre deux suites $(x_n), (y_n) \in A^\mathbb{N}$, avec $(\|x_n - y_n\|) \rightarrow 0$ et $\forall n \in \mathbb{N}, \|f(x_n) - f(y_n)\| > \varepsilon$

Comme A est compact, on peut extraire de (x_n) une suite $(x_{\varphi(n)})$ convergente vers $a \in A$. Comme $(\|x_n - y_n\|) \rightarrow 0$, alors $(y_{\varphi(n)}) \rightarrow a$ et $(f(x_{\varphi(n)}) - f(y_{\varphi(n)}))$ ne tend pas vers 0.

Donc f n'est pas continue en a .

Théorème : Bornes atteintes le retour

L'image d'un compact par une application continue est un compact.

Preuve

Soit $f \in \mathcal{F}(A, F)$ avec $A \subset E$ un compact. Montrons que $f(A)$ est compact.

Soit $(y_n) \in f(A)^\mathbb{N}$

Alors pour $n \in \mathbb{N}$, on peut prendre $x_n \in A, f(x_n) = y_n$.

Donc $(x_n) \in A^\mathbb{N}$

A est compact donc on peut extraire de (x_n) une suite $(x_{\varphi(n)})$ convergente vers $a \in A$.

Comme f est continue, la suite $(f(x_{\varphi(n)})) \rightarrow f(a)$

Donc $(y_{\varphi(n)}) \rightarrow f(a)$ et par compacité de A , $f(a) \in f(A)$.

Donc $f(A)$ est compact.

Proposition : Corollaire

Le théorème des bornes atteintes : si $f \in \mathcal{C}(A, \mathbb{R})$ avec A compact, alors f est borné et "atteint ses bornes". ie f est bornée ($f(A)$ est compact donc bornée).

$\exists a \in A, \sup f = f(a)$

$\exists b \in A, \inf f = f(b)$

Exemple IV.7. Soit $f : \begin{cases} [0, 1]^2 & \rightarrow \mathbb{R} \\ (x, y) & \rightarrow e^{x^2 \sin(3y)} \arctan(x+y) \end{cases}$
Montrer que $\exists d \in \mathbb{R}_+^*, \forall (x, y) \in [0, 1]^2, f(x, y) \geq d$.

Par compacité comme $[0, 1]^2$ compact, la borne inférieure de f est atteinte et est supérieure à 0, d'où l'existence de $]0, f(x, y)]$

Proposition : Corollaire

Si $f \in \mathcal{C}(A, \mathbb{R})$, avec A compact de F , l'application $\|f\| : \begin{cases} A & \rightarrow \mathbb{R} \\ x & \mapsto \|f(x)\| \end{cases}$ atteint ses bornes.

Exemple IV.8. Soit $f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto (2x^4 + 3y^2 + 1)e^{-(x^2+3y^2)} \end{cases}$
 $\forall x, y \in \mathbb{R}^2, f(x, y) \geq 0$, donc f est minorée et admet une borne inférieure.
Avec $x \in \mathbb{R}, f(x, 0) = (2x^4 + 1)e^{-x^2}$ et alors $(f(x, 0)) \rightarrow 0$ donc $\inf f = 0$

IV.6 Continuité des applications linéaires

Théorème : Critère de continuité des applications linéaires

Soit E, F des \mathbb{K} -EVN et $f \in \mathcal{L}(E, F)$. f est continue si, et seulement si, elle vérifie l'une des propriétés équivalentes suivantes :

1. f est continue en 0 ;
2. $\exists k \in \mathbb{R}_+^*, \forall x \in E, \|f(x)\|_F \leq k\|x\|_E$;
3. f est lipschitzienne.

Preuve

f continue implique 1 de manière immédiate.
Pour $1 \Rightarrow 2$: supposons f continue en 0.
Donc on peut prendre $\alpha \in \mathbb{R}_+^*, \forall x \in E, \|x\| < \alpha \Rightarrow \|f(x)\| < 37$
Soit $x \in E, x \neq 0$, prenons $u = \frac{\alpha}{2\|x\|}x$
On a bien $\|u\| < \alpha$
Donc $\|f(u)\| < 37$
Donc par linéarité de f et homogénéité de $\|\cdot\| : \|f(x)\| \leq \frac{37}{2} < \|x\|$
 $2 \Rightarrow 3$: On suppose $\exists k \in \mathbb{R}_+^*, \forall x \in E, \|f(x)\| \leq k\|x\|$
Donc $\forall x, y \in E, \|f(x) - f(y)\| \leq \|f(x - y)\| \leq k\|x - y\|$

Exemple IV.9. Soit $E = \mathcal{C}([0, 1], \mathbb{R})$, qu'on munit de $\|\cdot\|_\infty$ et $\|\cdot\|_1$.

On cherche à connaître la continuité de l'application linéaire :

$$\varphi : \begin{cases} E & \rightarrow \mathbb{R} \\ f & \mapsto f(0) \end{cases}$$

On peut avoir la continuité si le rapport $\frac{|\varphi(f)|}{\|f\|}$ pour $f \neq 0$ est borné.

Soit $f \in E$

Alors $|f(0)| \leq \|f\|_\infty$ donc φ est continue pour la norme infinie.

Pour la norme 1, considérons $(f_n)_{n \in \mathbb{N}^*}$ telle que : $f_n : t \mapsto e^{-nt}$

$\forall n \in \mathbb{N}^*, \varphi(f_n) = 1$ et $\|f_n\| = \frac{1}{n}(1 - e^{-n})$

Et donc $\frac{|\varphi(f_n)|}{\|f_n\|} \rightarrow +\infty$ et donc si E est muni de $\|\cdot\|_1$, alors φ n'est pas continue.

Remarque IV.6. En dimension finie, toute application linéaire est continue, par continuité des projecteurs (les applications linéaires sont des polynômes de degré au plus 1 sur les coordonnées).

Si φ est linéaire et injective en dimension finie, on a que $\|\varphi\|$ est une norme, qu'on appelle la norme φ .

Proposition : Fermeture des sev

Tout sous-espace de dimension finie d'un \mathbb{K} -ev est fermé.

Preuve

Soit F un sev de E , F de dimension finie.

Si E est de dimension finie, on peut considérer G un supplémentaire de F . Alors la projection p_G sur G parallèlement à F est continue puisqu'on est en dimension finie.

Et comme $F = \ker p_G = p_G^{-1}(\{0_E\})$ donc F est fermé.

Le résultat reste vrai si E n'est pas de dimension finie :

Soit f_n une suite de F convergente vers $l \in E$.

(f_n) est convergente donc bornée, donc il existe une boule fermée B contenant l'ensemble des f_n pour $n \in \mathbb{N}$

$\forall n \in \mathbb{N}, f_n \in B \cap F$

$B \cap F$ est une boule fermée de F pour la restriction de $\|\cdot\|$ à F , et comme F est de dimension finie, $B \cap F$ est un compact.

Donc on peut extraire de (f_n) une suite convergente vers $l' \in B \cap F$

Et par unicité de la limite, $l = l'$ donc $l \in F$.

Remarque IV.7. Si F, G deux sev de E , tels que $F \oplus G = E$, notons : p_F la projection sur F parallèlement à G et p_G la projection sur G parallèlement à F .

Si p_F et p_G sont continues, alors F et G sont fermés.

Exemple IV.10. Prenons $E = \mathcal{C}([0, 1], \|\cdot\|)$ et $\varphi : f \mapsto f(0)$

Alors $F = \{f \in E, \varphi(f) = 0\}$ est un hyperplan, mais F n'est pas fermé.

Proposition : Norme triple

Soit E, F deux EVN, on appelle $\mathcal{L}_C(E, F)$ l'ensemble des applications linéaires continues de E dans F . Alors $\mathcal{L}_C(E, F)$ est un espace vectoriel normé pour la norme $\varphi \mapsto \|\varphi\| = \sup_{x \in E, x \neq 0} \frac{\|\varphi(x)\|_F}{\|x\|_E} =$

$$\sup_{x \in E, \|x\|=1} \|\varphi(x)\|$$

Preuve

soit $\varphi \in \mathcal{L}_C(E, F)$

On suppose $\|\varphi\| = 0$ alors $\forall x \in E, x \neq 0, \|\varphi(x)\| \leq 0$ donc $\varphi(x) = 0$

Et comme φ est linéaire, $\varphi(0) = 0$ donc $\varphi = 0$

L'inégalité triangulaire et l'homogénéité sont immédiates.

Remarque IV.8. Dans le cas où E et F sont de dimensions finies : $\mathcal{L}_C(E, F) = \mathcal{L}(E, F)$

et $\mathcal{L}_C(E, F)$ est isomorphe à $\mathcal{M}_{n,p}(\mathbb{K})$

Si on fixe une norme dans $\mathcal{M}_{n,1}(\mathbb{K})$ et une norme dans $\mathcal{M}_{p,1}(\mathbb{K})$, il est naturel de parler de la norme $\|\cdot\|$ associée pour $\mathcal{M}_{n,p}(\mathbb{K})$

$$A \in \mathcal{M}_{n,p}(\mathbb{K}), \|\|A\|\| = \sup_{\|X\|=1, X \in \mathcal{M}_{p,1}(\mathbb{K})} \|AX\|_n = \sup_{X \in \mathcal{M}_{n,1}(\mathbb{K}), X \neq 0} \frac{\|AX\|_n}{\|X\|_p}$$

Exemple IV.11. $\mathcal{M}_{n,1}(\mathbb{K})$ muni de la norme $\|\cdot\|_\infty$, déterminons la norme triple sur $\mathcal{M}_n(\mathbb{K})$

$$\|\|A\|\| = \sup_{X \in \mathcal{M}_{n,1}(\mathbb{K}), x \neq 0} \frac{\|AX\|_\infty}{\|X\|_\infty}$$

$$\|AX\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n A_{i,j} X_j \right|$$

$$\left| \sum_{j=1}^n A_{i,j} X_j \right| \leq \sum_{j=1}^n |A_{i,j} X_j| \leq \left(\sum_{j=1}^n |A_{i,j}| \right) \|X\|_\infty \leq \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |A_{i,j}| \right) \|X\|_\infty$$

Pour prouver qu'on a bien la meilleure majoration possible, on doit trouver un exemple de vecteur X pour que l'inégalité soit une égalité.

Prenons i_0 tel que $\sum_{j=1}^n |A_{i_0,j}| = \max_i \sum_{j=1}^n |A_{i,j}|$

Prenons $X = (e^{i\theta_1}, \dots, e^{i\theta_n})$ avec $\theta_1, \dots, \theta_n \in \mathbb{R}$ tel que $\theta_j = -\arg(A_{i_0,j})$

Alors pour tout $j \in \llbracket 1, n \rrbracket$: $A_{i_0,j} e^{i\theta_j} = |A_{i_0,j}|$

$$\text{Donc } \|AX\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |A_{i,j} X_j| \right) \|X\|_\infty$$

$$\text{Donc } \|A\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |A_{i,j}|$$

Faisons de même pour $\|\cdot\|_1$: soit $X \in \mathcal{M}_{n,1}(\mathbb{K})$

$$\|AX\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n A_{i,j} X_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |A_{i,j} X_j| \leq \sum_{j=1}^n |X_j| \sum_{i=1}^n |A_{i,j}| \leq \left(\max_{1 \leq i \leq n} \sum_{j=1}^n |A_{i,j}| \right) \sum_{k=1}^n X_k$$

$$\text{Donc } \|A\| \leq \left(\max_{1 \leq i \leq n} \sum_{j=1}^n |A_{i,j}| \right) \sum_{k=1}^n X_k$$

Pour la majoration atteinte, on prend le $j_0 \in \llbracket 1, n \rrbracket$ tel que $\sum_{i=1}^n |A_{i,j_0}| = \max_j \sum_{i=1}^n |A_{i,j}|$

Soit $X = (0_1, \dots, 0, 1, 0, \dots, 0)$ avec le 1 en position j_0

$$\|AX\|_1 = \sum_{i=1}^n |A_{i,j_0}| \text{ donc } \|A\| = \max_{1 \leq j \leq n} \sum_{i=1}^n |A_{i,j}|$$

Théorème : Sous-multiplicativité

Soit E, F, G des EVN, soit $f \in \mathcal{L}_C(E, F)$ et $g \in \mathcal{L}_C(F, G)$, alors $\|g \circ f\| \leq \|g\| \|f\|$

Preuve

Soit $x \in E$, tel que $\|x\| = 1$

Alors $\|g \circ f(x)\| = \|g(f(x))\| \leq \|g\| \|f(x)\|$ par définition de $\|\cdot\|$ dans $\mathcal{L}_C(E, F)$

Et $\|g\| \|f(x)\| \leq \|g\| \|f\| \|x\| \leq \|g\| \|f\|$

Donc $\sup_{\|x\|=1} \|g \circ f(x)\| \leq \|g\| \|f\|$

Proposition : Corollaire

$\mathcal{L}_C(E)$ est munie naturellement d'une norme d'algèbre.

IV.7 Continuité des applications p-linéaires

Définition : Application p-linéaire

Soit $f \in \mathcal{F}(E_1 \times \dots \times E_p, F)$ est dite p-linéaire si :

$\forall x_1, \dots, x_p \in E_1 \times \dots \times E_p, \forall j \in \llbracket 1, p \rrbracket, x_j \mapsto f(x_1, \dots, x_j, \dots, x_p)$ est linéaire.

Exemple IV.12. Le déterminant est une application n -linéaire de $(\mathbb{K}^n)^n$

Le produit scalaire dans une espace préhilbertien est bilinéaire.

Proposition : Critère de continuité des applications multilinéaires

Soit E_1, \dots, E_p, F des EVN, soit f p-linéaire de $E_1 \times \dots \times E_p$ dans F .

f est continue si, et seulement si, $\exists k \in \mathbb{R}_+, \forall x_1, \dots, x_p \in E_1 \times \dots \times E_p, \|f(x_1, \dots, x_p)\| \leq k\|x_1\| \dots \|x_p\| \leq k\|x\|^p$

Preuve

Supposons f continue, donc f est continue en 0.

Donc on peut prendre $\alpha \in \mathbb{R}_+^*$ tel que $\forall x \in E_1 \times \dots \times E_p, \|x\| \leq \alpha \Rightarrow \|f(x)\| \leq 1$

Soit $x = (x_1, \dots, x_p) \in E_1 \times \dots \times E_p$, on suppose pour tout $i \in \llbracket 1, p \rrbracket x_i \neq 0$

Posons $u = (\frac{\alpha}{\|x_1\|}x_1, \frac{\alpha}{\|x_2\|}x_2, \dots, \frac{\alpha}{\|x_p\|}x_p)$, alors $\|u\| \leq \alpha$

Donc $\|f(u)\| \leq 1$

Par p-linéarité de f , on a $\|f(x)\| \leq \frac{1}{\alpha^p} \|x_1\| \dots \|x_p\| \leq \frac{1}{\alpha^p} \|x\|^p$

L'inégalité reste vraie si $x_1 = 0$ ou $x_2 = 0$ ou... ou $x_p = 0$.

Réciproquement, on suppose que : $\exists k \in \mathbb{R}_+^*, \forall x_1, \dots, x_p \in E_1 \times \dots \times E_p, \|f(x_1, \dots, x_p)\| \leq k\|x_1\| \dots \|x_p\|$

Soit $x = (x_1, \dots, x_p) \in E_1 \times \dots \times E_p$

Soit $h = (h_1, \dots, h_p) \in E_1 \times \dots \times E_p$ tel que $\|h\| \leq \|x\|$

$f(x+h) - f(x) = f(x_1+h_1, \dots, x_p+h_p) - f(x_1, \dots, x_p)$

$= \sum_{y_1 \in [x_1, h_1], y_p \in [x_p, h_p]} f(y_1, \dots, y_p) - f(x_1, \dots, x_p)$

On obtient une somme de 2^{p-1} termes, chaque terme de la forme $f(y_1, \dots, y_p)$ où au moins un des y_i est égale à h_i

Pour un tel terme : $\|f(y_1, \dots, y_p)\| \leq k\|y_1\| \dots \|y_p\| \leq k\|h\|\|x\|^{p-1}$

Donc $\|f(x+h) - f(x)\| \leq K(2^p - 1)\|x\|^{p-1}\|h\|$

Donc $f(x+h) - f(x) \rightarrow_{h \rightarrow 0} 0$

Donc f est continu en x .

Attention, ça ne veut pas dire que toute fonction multilinéaire continue est lipschitzienne, elle l'est juste sur des voisinages.

Proposition : Application

Si E est un espace préhilbertien réel, alors
$$\begin{array}{ccc} E \times E & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & \langle x, y \rangle \end{array}$$
 est continue ($\forall x, y \in E, \langle x, y \rangle \leq \|x\|\|y\|$)

V Connexité par arcs**Définition : Chemin**

Soit E un EVN, $x, y \in E$, on appelle chemin de x à y une application continue φ de $[0, 1]$ dans E telle que $\varphi(0) = x$ et $\varphi(1) = y$

Remarque V.1. Dans la définition, 0 et 1 n'ont pas d'importance. Si on a $\varphi : [a, b] \rightarrow E$ continue telle que $\varphi(a) = x$ et $\varphi(b) = y$

On a $\psi : \begin{cases} [0, 1] & \rightarrow [a, b] \\ t & \mapsto a + t(b-a) \end{cases}$

Alors $\varphi \circ \psi$ est continue, $\varphi \circ \psi(0) = x$ et $\varphi \circ \psi(1) = y$

Définition : Chemin dans une partie

Si $A \subset E$, si $x, y \in A$ un chemin φ est un chemin dans A si $\forall t \in [0, 1], \varphi(t) \in A$

Proposition : Relation d'équivalence

Avec E un EVN, $A \subset E$, on définit sur A la relation $x\mathcal{R}_A y \Leftrightarrow$ il existe un chemin dans A joignant x à y .

\mathcal{R}_A est une relation d'équivalence.

Définition : Connexes

On appelle composantes connexes les classes d'équivalences par la relation précédente

Un ensemble est connexe par arc si \mathcal{R}_A a une seule classe d'équivalence.

Exemple V.1. *Les convexes sont tous connexes par arcs. La réciproque est en générale fausse, sauf sur \mathbb{R} .*

Les parties étoilées (A est étoilée s'il existe $c \in A$ tel que $\forall x \in A, [x, c] \subset A$) sont connexes par arcs.

Les connexes par arcs de \mathbb{R} sont les intervalles. $\mathbb{R} \setminus \{a\}$ n'est pas connexe. $\mathbb{R}^2 \setminus D$ avec D une droite n'est pas connexe. $\mathbb{R}^2 \setminus \{a\}$ est connexe par arcs.

Théorème : Valeurs intermédiaires

L'image d'un connexe par arcs par une application continue est connexe par arcs.

Preuve

Soient E, F EVN, $A \subset E$ et $f \in \mathcal{C}(A, F)$, et $B \subset A$ connexe par arcs.

Soit $y_1, y_2 \in f(B)$ on prend donc $x_1, x_2 \in B$ tels que $f(x_1) = y_1$ et $f(x_2) = y_2$

Comme B est connexe par arcs on a $\varphi : \begin{cases} [0, 1] & \rightarrow B \\ t & \mapsto \varphi(t) \end{cases}$ continue avec $\varphi(0) = x_1$ et $\varphi(1) = x_2$

Alors $f \circ \varphi$ est un chemin dans $f(B)$ joignant y_1 à y_2

Donc $f(B)$ est connexe par arc.

Exemple V.2. $\mathcal{GL}_n(\mathbb{R})$ n'est pas connexe par arcs : $\det(\mathcal{GL}_n(\mathbb{R})) = \mathbb{R}^*$

Il n'existe aucune application injective continue de \mathbb{R}^2 dans \mathbb{R} :

Supposons par l'absurde $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ continue et injective.

L'image de \mathbb{R}^2 par f est alors connexe par arcs, donc un intervalle I .

Notons $(x, y) \in \mathbb{R}^2$ tel que $f(x, y) \in \overset{\circ}{I}$. Alors $f(\mathbb{R}^2 \setminus \{(x, y)\})$ est privé de $f(x, y)$ par injectivité et n'est donc pas l'intervalle I , d'où la contradiction comme $\mathbb{R}^2 \setminus \{(x, y)\}$ est connexe par arc.

Montrons que $\mathcal{GL}_n(\mathbb{C})$ est connexe par arc.

Soit $M \in \mathcal{GL}_n(\mathbb{C})$, on prend $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\exp(A) = M$

Notons : $P : \begin{cases} [0, 1] & \rightarrow \mathcal{GL}_n(\mathbb{C}) \\ t & \mapsto \exp(tA) \end{cases}$

Alors $P(0) = I$ et $P(1) = M$, donc $\mathcal{GL}_n(\mathbb{C})$ est connexe par arc.

Pour $\mathcal{GL}_n(\mathbb{R})$, il y a au moins deux composantes connexes : celles avec un déterminant positif et celles avec un déterminant négatif.

VI Théorème d'équivalence des normes

Théorème : L'équivalence des normes

Toutes les normes sont équivalentes en dimension finie

Preuve

Soit E un \mathbb{K} -ev de dimension finie, avec $B = (e_1, \dots, e_p)$ une base de E .

Soit $\|\cdot\|_\infty \begin{cases} E \rightarrow \mathbb{K} \\ x = \sum_{i=1}^p x_i e_i \mapsto \max_{1 \leq i \leq p} |x_i| \end{cases}$

Soit N une norme sur E

Pour $x \in E$, en notant $x = \sum_{i=1}^p x_i e_i$:

$$N(x) \leq \sum_{i=1}^p |x_i| N(e_i) \leq \|x\|_\infty \sum_{i=1}^p N(e_i)$$

Or, par la seconde inégalité triangulaire :

$\forall x, y \in E, |N(x) - N(y)| \leq N(x - y) \leq K \|x - y\|_\infty$, où K est une constante.

Si on considère $(E, \|\cdot\|_\infty)$, on a donc que N est lipschitzienne, et donc continue.

On a que $\mathcal{S}_\infty(0, 1) = \{x \in E \mid \|x\|_\infty = 1\}$ est compact.

Donc N étant continue, elle atteint ses bornes supérieures et inférieures sur $\mathcal{S}_\infty(0, 1)$.

Donc $m = \inf_{x \in \mathcal{S}(0, 1)} N(x) > 0$

Et donc $\forall x \in E, x \neq 0 \Rightarrow N\left(\frac{1}{\|x\|_\infty} x\right) \leq m \Rightarrow N(x) \geq m \|x\|_\infty$

D'où l'équivalence entre N quelconque et $\|\cdot\|_\infty$. Par transitivité de la relation d'équivalence des normes, on a donc que toutes les normes en dimension finie sont équivalentes.

Chapitre IV

Réduction des endomorphismes

I Généralités

Définition : Valeur propre

Soit E un \mathbb{K} -ev et $f \in \mathcal{L}(E)$. Soit $\lambda \in \mathbb{K}$.

On dit que λ est une valeur propre de f si : $f - \lambda id$ n'est pas injective.

Ce qui correspond à $\ker(f - \lambda id) \neq \{0\}$

Ce qui correspond à $\exists x \in E, x \neq 0, (f - \lambda id)(x) = 0$

Ce qui correspond à $\exists x \in E, x \neq 0, f(x) - \lambda x = 0$

Remarque I.1. On étend la notion de valeurs propres à l'ensemble des matrices $\mathcal{M}_n(\mathbb{K})$: pour $A \in \mathcal{M}_n(\mathbb{K})$, λ est une valeur propre de A si λ est une valeur propre de l'endomorphisme canoniquement associé.

Ce qui correspond à $\ker(A - \lambda I_n) \neq \{0\}$

Ce qui correspond à $\exists X \in \mathcal{M}_{n,1}(\mathbb{K}), X \neq 0, (A - \lambda I_n)(X) = 0$

Ce qui correspond à $\exists X \in \mathcal{M}_{n,1}(\mathbb{K}), X \neq 0, AX = \lambda X$

Définition : Spectre d'un endomorphisme

Soit E un \mathbb{K} -ev et $f \in \mathcal{L}(E)$. On appelle spectre de f l'ensemble des valeurs propres de f , noté $S_p(f)$.

On peut étendre la notion aux matrices.

Proposition :

Deux matrices semblables ont même spectre.

Exemple I.1. Si A est triangulaire supérieure avec une diagonale $\lambda_1, \dots, \lambda_n$, alors $S_p(A) = \{\lambda_1, \dots, \lambda_n\}$. Attention, un spectre n'est pas forcément de cardinal n .

Soit $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ et $\varphi : \begin{cases} E & \rightarrow E \\ f & \mapsto f' \end{cases}$

Déterminons les valeurs propres de φ

Supposons $\lambda \in S_p(\varphi) \Leftrightarrow \exists f \in E, f \neq 0, \varphi(f) = \lambda f$

\Leftrightarrow l'équation $\varphi(f) = \lambda f$ admet au moins une solution non-nulle.

Soit $\lambda \in \mathbb{C}$, soit $f \in E$. Alors :

$\varphi(f) = \lambda f \Leftrightarrow f' = \lambda f \Leftrightarrow f \in \text{Vect}(t \mapsto e^{\lambda t})$

Donc $\lambda \in S_p(\varphi)$

Donc $S_p(\varphi) = \mathbb{C}$

Soit $E = \mathbb{R}[X]$, et $\varphi : \begin{cases} E & \rightarrow & E \\ P & \mapsto & P' \end{cases}$

Déterminons les valeurs propres de φ . Soit $\lambda \in \mathbb{K}$, soit $P \in E$:

Si $\lambda = 0$:

$$\varphi(P) = \lambda P \Leftrightarrow P' = \lambda P \Leftrightarrow P \in \text{Vect}(1)$$

Donc 0 est valeur propre.

Si $\lambda \neq 0$:

$$P' = \lambda P \Rightarrow \deg P' = \deg P \Rightarrow P = 0$$

Donc si $\lambda = 0$, alors $\lambda \notin S_p(\varphi)$

$$\text{Donc } S_p(\varphi) = \{0\}$$

Définition : Espace propre

Soit E un \mathbb{K} -ev, $f \in \mathcal{L}(E)$ et $\lambda \in S_p(f)$. On appelle espace propre associé à λ l'ensemble $\ker(f - \lambda \text{id})$, qu'on note $E_\lambda(f)$

C'est un sev de E non-réduit à 0.

On étend la notion aux matrices de $\mathcal{M}_n(\mathbb{K})$. Pour $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in S_p(A)$, alors $E_\lambda(A) = \ker(A - \lambda I_n)$ est un sev de $\mathcal{M}_{n,1}(\mathbb{K})$.

Remarque I.2. Attention, deux matrices semblables n'ont pas les mêmes espaces propres. Ils sont simplement isomorphes : $E_\lambda(A') = P(E_\lambda(A))$ avec P la matrice de passage entre A et A' .

Définition : Vecteur propre

Soit E un \mathbb{K} -ev, $f \in \mathcal{L}(E)$ et $\lambda \in S_p(f)$. Soit $x \in E$, on dit que x est un vecteur propre de f relativement à λ si x est un vecteur non-nul de $E_\lambda(f)$.

Théorème :

Une somme finie de sous-espaces propres d'un endomorphisme associés à des valeurs propres distinctes est directe.

Ce qui correspond à : si $\lambda_1, \dots, \lambda_p$ sont des valeurs propres distinctes de $f \in \mathcal{L}(E)$, alors $E_1 + \dots + E_p$ est une somme directe.

Preuve

Par récurrence sur le nombre d'espaces propres : Soit, pour $p \in \mathbb{N}^*$, H_p : Pour tout \mathbb{K} -ev E et toute application $f \in \mathcal{L}(E)$, la somme de p espaces propres de f associés à des valeurs propres distinctes est directe.

H_1 est vérifiée.

Soit $p \in \mathbb{N}^*$, on suppose H_p . Soit E un \mathbb{K} -ev, soit $f \in \mathcal{L}(E)$ et $\lambda_1, \dots, \lambda_{p+1}$ $p+1$ valeurs propres de f .

Soit $(u_1, \dots, u_{p+1}) \in E_{\lambda_1} \times \dots \times E_{\lambda_{p+1}}$, et on suppose $u_1 + \dots + u_{p+1} = 0$ (1).

On a donc $f(u_1) + \dots + f(u_{p+1}) = 0$ ie $\lambda_1 u_1 + \dots + \lambda_{p+1} u_{p+1} = 0$ (2)

On fait $\lambda_1(1) - (2)$:

$$(\lambda_1 - \lambda_1)u_2 + \dots + (\lambda_{p+1} - \lambda_1)u_{p+1} = 0$$

Or par H_p , on a que $E_{\lambda_2} + \dots + E_{\lambda_{p+1}}$ est directe, donc $\forall i \in \llbracket 2, p+1 \rrbracket, (\lambda_i - \lambda_1)u_i = 0$ donc $u_i = 0$

Donc d'après (1), on a aussi $u_1 = 0$

Donc $E_{\lambda_1}(f) + \dots + E_{\lambda_{p+1}}(f)$ est directe.

D'où la récurrence.

Proposition : Corollaire

Si $\dim E$ est finie : $f \in \mathcal{L}(E)$, $n = \dim E$:

- $S_p(f)$ est fini et $\text{card} S_p(f) \leq n$ ($\lambda \in S_p(f) \Rightarrow \dim E_\lambda(f) \geq 1$) ;
- f est diagonalisable (ie il existe une base de E dans laquelle $\text{Mat}_B(f)$ est diagonale) si, et seulement si, $\sum_{\lambda \in S_p(f)} \dim E_\lambda(f) = \dim E$
- Si f admet n valeurs propres distinctes, alors f est diagonalisable.

Preuve

Si $\sum_{\lambda \in S_p(f)} \dim E_\lambda(f) = \dim E$:

$S_p(f) = \{\lambda_1, \dots, \lambda_p\}$, on prend B_1, \dots, B_p bases respectivement de $E_{\lambda_1}(f), \dots, E_{\lambda_p}(f)$

On considère $B = (B_1, \dots, B_p)$

B est une famille libre (la somme des espaces propres est directe) et $\dim E_{\lambda_1}(f) + \dots + \dim E_{\lambda_p}(f) = \dim E$ vecteurs de E

Donc B est une base. Et alors la matrice de f dans cette base est diagonale de coefficients les éléments du spectre selon leurs dimensions.

On suppose qu'il existe $B = (e_1, \dots, e_n)$ telle qu'il existe $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ telle que la matrice de f dans B soit diagonale de coefficients les λ_i .

On a donc $\forall i \in \llbracket 1, n \rrbracket, e_i \in E_{\lambda_i}(f)$

Donc $E \subset E_{\lambda_1}(f) + \dots + E_{\lambda_n}(f) \subset \bigoplus_{\lambda \in S_p(f)} E_\lambda(f)$

Donc $\dim E \leq \sum_{\lambda \in S_p(f)} \dim E_\lambda(f)$, l'égalité étant assurée par le fait que les espaces propres sont des sev de E .

Exemple I.2. Soit $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ et $(e_\lambda)_{\lambda \in \mathbb{C}}$ telle que : $e_\lambda : \begin{cases} \mathbb{R} & \rightarrow \mathbb{C} \\ t & \mapsto e^{\lambda t} \end{cases}$

Pour montrer que (e_λ) est libre, on considère l'application φ qui à tout élément de E associe sa dérivée.

Alors $S_p(\varphi) = \mathbb{C}$

Prenons $\lambda_1, \dots, \lambda_p \in \mathbb{C}$ deux-à-deux distincts : alors $(e_{\lambda_1}, \dots, e_{\lambda_p})$ est une famille de p vecteurs non-nuls des espaces propres de φ $E_{\lambda_1}(\varphi), \dots, E_{\lambda_p}(\varphi)$

Donc $(e_{\lambda_1}, \dots, e_{\lambda_p})$ est libre.

Proposition :

Soit E un \mathbb{K} -ev. Soient $u, v \in \mathcal{L}(E)$ tels que $u \circ v = v \circ u$. Alors :

- $\text{Im} v$ et $\ker v$ sont stables par u
- $\forall \lambda \in S_p(v)$, $E_\lambda(v)$ est stable par u

Preuve

Soit $y \in \text{Im} v$

Donc on peut prendre $x \in E, y = v(x)$

donc $u(y) = u(v(x)) = v(u(x))$ donc $u(x) \in \text{Im} v$

Soit $\lambda \in S_p(f)$

Soit $x \in E_\lambda(v)$

alors $v(u(x)) = u(v(x)) = u(\lambda x) = \lambda u(x)$

Et donc $u(x) \in \text{Ker}(v - \lambda \text{id})$

Donc $E_\lambda(v)$ est stable.

II Polynôme caractéristique

Dans la suite du chapitre, E est de dimension finie. Avec $f \in \mathcal{L}(E)$, $\lambda \in \mathbb{K}$ on a :

$\lambda \in S_p(f) \Leftrightarrow (f - \lambda id)$ est non injective

$\Leftrightarrow (f - \lambda id)$ est non inversible dans $\mathcal{L}(E)$ en dim finie

$\Leftrightarrow \det(f - \lambda id) = 0$

Proposition : Polynôme caractéristique

Avec $f \in \mathcal{L}(E)$, on a que $\begin{matrix} \mathbb{K} & \rightarrow & \mathbb{K} \\ \lambda & \mapsto & \det(f - \lambda id) \end{matrix}$ est polynomiale. Le polynôme associé est de degré n unitaire.

On appelle ce polynôme le polynôme caractéristique de f , noté χ_f

Preuve

Soit A la matrice de f dans une base quelconque fixée.

$$\det(\lambda id - f) = \det(\lambda I - A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^n (\lambda I - A)_{\sigma(k),k}$$

Cette fonction est polynomiale, associée au polynôme :

$$\det(XI - A) = \sum_{\sigma \in \mathfrak{S}(n)} \varepsilon(\sigma) \prod_{k=1}^n [XI - A]_{\sigma(k),k}$$

Or $[XI - A]_{\sigma(k),k}$ est un polynôme de degré inférieur ou égal à 1 (qui vaut 1 si, et seulement si, $\sigma(k) = k$)

Donc $\det(XI - A)$ est une somme de polynômes de degré $\leq n$, donc un polynôme de degré $\leq n$

On remarque que $\prod_{k=1}^n [XI - A]_{\sigma(k),k}$ est de degré n si, et seulement si, $\sigma = id$

$$\text{donc } \det(XI - A) = \prod_{i=1}^n (X - A_{i,i}) + Q \text{ avec } Q \in \mathbb{K}_{n-1}[X]$$

Donc $\deg(\det(XI - A)) = n$ est le coefficient le degré n de ce polynôme et celui de $\sum_{i=1}^n (X - A_{i,i})$ c'est-à-dire 1

Proposition : Valeurs propres

Soit E un \mathbb{K} -ev de dimension finie.

λ est une valeur propre de f si, et seulement si, λ est une racine de χ_f

Remarque II.1. On retrouve le fait qu'un endomorphisme en dimension finie n possède au plus n valeurs propres réelles.

II.1 Coefficients du polynôme caractéristique

Avec $f \in \mathcal{L}(E)$, $A = \text{Mat}_B(f)$ dans une base B fixée.

$$\chi_f = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^n [XI - A]_{\sigma(k),k}$$

Pour $j \in \{1, n\}$, calculons le coefficient devant X^{n-j} de χ_f : α_{n-j}

Pour le coefficient de degré $n - 1$: soit $\sigma \in \mathfrak{S}_n$, si $\sigma \neq id$, alors σ a au plus $n - 2$ points fixes

$$\text{Donc } \forall \sigma \in \mathfrak{S}_n, \sigma \neq id, \deg\left(\prod_{k=1}^n [XI - A]_{\sigma(k),k}\right) \leq n - 2$$

$$\text{Donc } \chi_f = \prod_{i=1}^n [X - A_{i,i}] + Q \text{ avec } \deg Q \leq n - 2$$

$$\text{Donc le coefficient de degré } n - 1 \text{ de } \chi_f \text{ est : } -\sum_{i=1}^n A_{i,i} = -\text{tr}(A)$$

Le coefficient constant de χ_f est la valeur en 0 du polynôme caractéristique, donc c'est $\chi_f(0) = \det(-A) = (-1)^n \det(A)$

$$\text{Donc } \chi_f = X^n - \text{tr}(f)X^{n-1} + \dots + (-1)^n \det(A)$$

Exemple II.1. Application : si χ_f est scindé, alors $\chi_f = \prod_{k=1}^n (X - \lambda_i)$ où les λ_i sont les racines de χ_f , non nécessairement distinctes.

$$\text{Alors } \begin{cases} \sum_{i=1}^n \lambda_i = \text{tr}(f) \\ \prod_{i=1}^n \lambda_i = \det(f) \end{cases}$$

Définition : Ordre d'une valeur propre

Soit E un \mathbb{K} -ev de dimension finie. Soit $f \in \mathcal{L}(E)$. Soit $\lambda \in \mathbb{K}$.

On dit que λ est une valeur propre d'ordre α si λ est une racine de χ_f d'ordre α

Ce qui correspond à $(X - \lambda)^\alpha | \chi_f$ et $(X - \lambda)^{\alpha+1}$ ne divise pas χ_f

Ce qui correspond à $\chi_f = (X - \lambda)^\alpha Q$ pour $Q \in \mathbb{K}[X]$, $Q(\lambda) \neq 0$

Théorème :

Soit E un \mathbb{K} -ev de dimension finie. Soit $f \in \mathcal{L}(E)$. Soit λ une valeur propre de f d'ordre α

Alors $\dim E_\lambda(f) \leq \alpha$

Preuve

Dans le cadre de l'énoncé, on note $p = \dim E_\lambda(f)$.

On peut prendre (e_1, \dots, e_p) une base de $E_\lambda(f)$

On complète cette famille libre de e par e_{p+1}, \dots, e_n en une base B de E .

Alors $\text{Mat}_B(A) = \begin{pmatrix} \lambda I_p & C \\ 0 & D \end{pmatrix}$ avec $C \in \mathcal{M}_{p, n-p}(\mathbb{K})$ et $D \in \mathcal{M}_{n-p, n-p}(\mathbb{K})$

Donc $\chi_f = \det \begin{pmatrix} (X - \lambda)I_p & -C \\ 0 & XI - D \end{pmatrix} = (X - \lambda)^p \det(XI - D)$

Donc $(X - \lambda)^p$ divise χ_f

Donc $p \leq \alpha$

Proposition : Corollaire

Si E de dimension finie, $f \in \mathcal{L}(E)$, alors f est diagonalisable si, et seulement si :

$$\begin{cases} \chi_f \text{ est scindé} \\ \forall \lambda \in S_p(f), \dim E_\lambda(f) = \alpha_\lambda \end{cases}$$

Preuve

On note α_λ l'ordre de λ pour $\lambda \in S_p(f)$

Par contraposée : si χ_f est non-scindé.

On a alors $\sum_{\lambda \in S_p(f)} \alpha_\lambda$

On sait que $\dim \bigoplus_{\lambda \in S_p(f)} E_\lambda(f) = \sum_{\lambda \in S_p(f)} \dim E_\lambda(f)$

Donc $\dim \bigoplus_{\lambda \in S_p(f)} E_\lambda(f) \leq \sum_{\lambda \in S_p(f)} \alpha_\lambda < n$

Donc $\bigoplus E_\lambda(f) \neq E$

f n'est pas diagonalisable.

On suppose maintenant χ_f scindé.

Pour $\lambda \in S_p(f)$, on note n_λ la dimension de $E_\lambda(f)$

f est diagonalisable si, et seulement si $\bigoplus_{\lambda \in S_p(f)} E_\lambda(f) = E$

$$\Leftrightarrow \sum_{\lambda \in S_p(f)} n_\lambda = n$$

$$\Leftrightarrow \sum_{\lambda \in S_p(f)} (\alpha_\lambda - n_\lambda) = 0$$

$$\forall \lambda \in S_p(f), \alpha_\lambda = n_\lambda$$

Exemple II.2. Avec $A = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}$

On a immédiatement $\text{rg}(A - 3I_3) = 1$ et donc $\dim \ker(A - 3I_3) = 2$

Donc 3 est une valeur propre de A d'ordre au moins 2.

Donc les valeurs propres de A comptées avec ordre de multiplicité sont 3, 3, X telles que $3+3+X = \text{tr}(A)$

Donc $X = 6$ et donc $\chi_A = (X - 3)^2(X - 6)$

III Trigonalisation

Définition : Trigonalisation

Soit E un \mathbb{K} -ev de dimension finie, on dit que $f \in \mathcal{L}(E)$ est trigonalisable si il existe une base B telle que $\text{Mat}_B(f)$ est triangulaire.

Remarque III.1. Si $\text{Mat}_B(f) \in T_n^+(\mathbb{K}) : f(e_1) \in \text{Vect}(e_1), f(e_2) \in \text{Vect}(e_1, e_2), \dots, f(e_n) \in \text{Vect}(e_1, \dots, e_n)$ et donc $\forall i \in \llbracket 1, n \rrbracket, f(F_i) \subset F_i$.

Posons $B'(e_n, \dots, e_1)$. Alors : $f(e_n) \in \text{Vect}(e_n, \dots, e_1), f(e_{n-1}) \in \text{Vect}(e_{n-1}, \dots, e_1), \dots, f(e_1) \in \text{Vect}(e_1)$

Et alors $\text{Mat}_{B'}(f) \in T_n^+(\mathbb{K})$

Donc quand une matrice est trigonalisable, elle l'est de manière inférieure et supérieure.

Une matrice est dite trigonalisable si son endomorphisme canoniquement associé est trigonalisable. ie elle est semblable à une matrice triangulaire.

Théorème :

Un endomorphisme d'un \mathbb{K} -ev E de dimension finie est trigonalisable si, et seulement si, son polynôme caractéristique est scindé.

Preuve

On suppose f trigonalisable. On a donc B une base de E telle que : $Mat_B(f) = T$ est triangulaire, de coefficients diagonaux $\alpha_1, \dots, \alpha_n$.

Et donc $\chi_f = \chi_T = (X - \alpha_1) \dots (X - \alpha_n)$

Donc χ_f est scindé.

Réciproquement, raisonnons par récurrence sur la dimension de l'espace.

Montrons que $\forall n \in \mathbb{N}^*, H_n$: pour tout \mathbb{K} -ev E de dimension n , pour tout endomorphisme f de E de polynôme caractéristique scindé, f est trigonalisable.

H_1 : En dimension 1, les endomorphismes sont uniquement $x \mapsto \lambda x$, on a donc une matrice à un seul coefficient λ .

Soit $n \in \mathbb{N}$, on suppose H_n . Soit E un \mathbb{K} -ev de dimension $(n + 1)$.

Soit $f \in \mathcal{L}(E)$ tel que χ_f est scindé.

Donc χ_f a au moins une racine dans \mathbb{K} , notée λ . Donc $\lambda \in S_p(f)$.

On considère $e_1 \in E$ un vecteur propre (non nul) associé à λ et on complète e_1 par (e_2, \dots, e_{n+1}) en une base B de E .

Alors $Mat_B(f) = \begin{pmatrix} \lambda & L \\ 0 & A \end{pmatrix}$ avec $A \in \mathcal{M}_n(\mathbb{K}), L \in \mathcal{M}_{1,n}(\mathbb{K})$

On note $F = Vect(e_2, \dots, e_{n+1})$. On note p la projection sur F parallèlement à $Vect(e_1)$.

On appelle g l'application $g : \begin{cases} F & \rightarrow & F \\ x & \mapsto & p \circ f(x) \end{cases}$.

Et alors $A = Mat_{(e_2, \dots, e_{n+1})}(g)$

On a que $\dim F = n$ et donc $g \in \mathcal{L}(F)$

On a que $\chi_f = (X - \lambda)\chi_A = (X - \lambda)\chi_g$ or χ_f est scindé, donc χ_g est scindé.

Donc par H_n , on peut trouver une base B'_f de F dans laquelle $Mat_{B'_f}(g) = T \in T_n^+(\mathbb{K})$

Donc $B' = (e_1, e'_2, \dots, e'_n)$ est une base de E

Donc $Mat_{B'}(f) = \begin{pmatrix} \lambda & L' \\ 0 & T \end{pmatrix}$ qui est triangulaire supérieure.

Ce qui conclut la récurrence.

Exemple III.1. Soit E un \mathbb{C} -ev de dimension n et $f \in \mathcal{L}(E)$

Notons $\chi_f = \prod_{i=1}^n (X - \lambda_i)$

Déterminer χ_{f^p} pour $p \in \mathbb{N}$ et χ_{f^p} pour $p \in \mathbb{Z}$ si $f \in \mathcal{GL}_n(\mathbb{C})$

Comme $\chi_f \in \mathbb{C}[X]$, il est scindé : on fixe B une base de E dans laquelle $Mat_B(f)$ est triangulaire supérieure de coefficients diagonaux $\alpha_1, \dots, \alpha_n$

Comme un produit de matrices triangulaires supérieures est triangulaire supérieur avec les coefficients diagonaux qui sont les produits des éléments des deux diagonales, et donc $Mat_B(f^2)$ est une matrice triangulaire supérieure avec comme coefficients diagonaux les $\lambda_1^2, \dots, \lambda_n^2$

Donc $\chi_{f^2} = \prod_{i=1}^n (X - \lambda_i^2)$

Par récurrence immédiate, on a $\chi_{f^p} = \prod_{i=1}^n (X - \lambda_i^p)$

Si $f \in \mathcal{GL}_n(\mathbb{C})$, alors tous ses coefficients diagonaux dans B sont non-nuls et son inverse a comme coefficients diagonaux les $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$. On peut alors étendre $\chi_{f^p}(X - \lambda_i^p)$ pour $n \in \mathbb{Z}$

Exemple III.2. Est-ce qu'il y a un lien entre f diagonalisable et f^2 diagonalisable ?

Si f est diagonalisable, alors f^2 l'est pour les mêmes matrices de passage. On peut trouver $B = (e_1, \dots, e_n)$ une base de vecteurs propres de f tels que $f(e_i) = \lambda_i e_i$

On a alors $f^2(e_i) = f(\lambda_i e_i) = \lambda_i^2 e_i$ et par récurrence immédiate $f^p(e_i) = \lambda_i^p e_i$

On a cependant des exemples de réciproque fausse, avec la valeur propre 0. Par exemple, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} =$

$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ n'est pas diagonalisable mais son carré est la matrice nulle, qui est diagonale.

Remarque III.2. Avec la trigonalisation, on retrouve le fait que :

$\forall A \in \mathcal{GL}_n(\mathbb{K}), \text{tr}(A)$ est la somme des vp de A comptées avec leur ordre de multiplicité.

$\det(A)$ est le produit des vp de A comptées avec leur ordre de multiplicité.

IV Endomorphismes nilpotents

Définition : Nilpotence

Soit E un \mathbb{K} -ev de dimension finie et f un endomorphisme de E . f est nilpotent si il existe $k \in \mathbb{N}^*$ tel que $f^k = 0$. On appelle ordre de nilpotence de f le plus petit entier p tel que $f^p = 0$

Théorème :

Soit E un \mathbb{K} -ev de dimension finie n , $f \in \mathcal{L}(E)$. Alors f est nilpotente si, et seulement si, $\chi_f = X^n$ et $\chi_f = X^n$ si, et seulement si, il existe une base B de E dans laquelle $\text{Mat}_B(f)$ est triangulaire avec des 0 sur la diagonale.

Preuve

Supposons $\chi_f = X^n$. Alors χ_f est scindé, et f est trigonalisable.

Mais comme ses valeurs propres sont $0, \dots, 0$, il existe donc une base B de E telle que $\text{Mat}_B(f)$ est triangulaire avec des zéros sur la diagonale.

On suppose qu'il existe B base de E telle que $\text{Mat}_B(f)$ est triangulaire avec une diagonale nulle.

On note pour $i \in \llbracket 1, n \rrbracket : F_i = \text{Vect}(e_1, \dots, e_i)$.

Pour $i \in \llbracket 1, n \rrbracket, f(F_i) \subset F_{i-1}$ avec $F_0 = \{0\}$

Et donc par récurrence : $\forall p \in \mathbb{N}^*, p \leq i, f^p(F_i) \subset F_{i-p}$

Et en particulier, $f^n(F_n) \subset F_0$

Et donc f est nilpotente.

Supposons f nilpotente. On considère une base B de E et $A = \text{Mat}_f(B)$

Donc $A \in \mathcal{M}_n(\mathbb{K})$.

Donc on sait que $\chi_A = \chi_f$

Montrons que la seule valeur propre de A dans \mathbb{C} est 0, donc on a besoin d'une complexification de l'espace.

Donc on prend $\lambda \in S_{p, \mathbb{C}}(A)$

On peut donc prendre $X \in \mathcal{M}_{n,1}(\mathbb{C}), X \neq 0$ tel que $AX = \lambda X$

Par récurrence, $\forall k \in \mathbb{N}, A^k X = \lambda^k X$

f est nilpotente d'ordre $\leq n$, donc A aussi.

Donc $\lambda^n X = 0$

Mais $X \neq 0$ donc $\lambda^n = 0$ donc $\lambda = 0$

Et donc 0 est valeur propre d'ordre n

Donc $\chi_f = \chi_A = X^n$

V Polynômes d'endomorphismes

Un projecteur, défini polynomialement, est un endomorphisme vérifiant $p^2 - p = 0$. Une symétrie est un endomorphisme qui vérifie $s^2 - id = 0$.

On parle dans les deux cas de polynômes annulateurs. Dans les deux cas, les endomorphismes sont diagonalisables.

Dans toute cette partie, E est un \mathbb{K} -ev de dimension finie.

Définition : Polynôme d'endomorphismes

Soit $f \in \mathcal{L}(E)$. On note $\mathbb{K}[f]$ la sous-algèbre de $\mathcal{L}(E)$ engendrée par f .

Donc $\mathbb{K}[f] = Vect((f^n)_{n \in \mathbb{N}}) = \left\{ \sum_{k=0}^n a_k f^k \mid n \in \mathbb{N}, (a_0, \dots, a_n) \in \mathbb{K}^{n+1} \right\}$

On peut donc considérer l'application
$$\begin{cases} \mathbb{K}[X] & \rightarrow \mathbb{K}[f] \\ P = \sum_{k=0}^n a_k X^k & \mapsto \sum_{k=0}^n a_k f^k = P(f) \end{cases}$$

Proposition : Opérations sur les polynômes d'endomorphismes

Pour $P, Q \in \mathbb{K}[X]$, $\lambda \in \mathbb{K}$ et $f \in \mathcal{L}(E)$, on a :

- $(P + Q)(f) = P(f) + Q(f)$
- $(\lambda \cdot P)(f) = \lambda P(f)$
- $(P \times Q)(f) = P(f) \circ Q(f) = Q(f) \circ P(f)$
- $(P \circ Q)(f) = P(Q(f))$

Preuve

Les deux premières sont immédiates.

Pour le produit, si $P = \sum_{k=0}^n a_k X^k$, $Q = \sum_{k=0}^n b_k X^k$ alors :

$$PQ = \sum_{0 \leq i \leq n, 0 \leq j \leq n} a_i b_j X^{i+j}$$

$$\text{Et donc } (PQ)(f) = \sum_{0 \leq i \leq n, 0 \leq j \leq n} a_i b_j f^{i+j} = \left(\sum_{i=0}^n a_i f^i \right) \circ \left(\sum_{j=0}^n b_j f^j \right)$$

Pour la composée, on montre par récurrence que $(X^n \circ Q)(f) = Q^n(f) = Q(f)^n$

Et ensuite on a que pour $P = \sum_{k=0}^n a_k X^k$: $P \circ Q(f) = \sum_{k=0}^n a_k Q^k(f) = \sum_{k=0}^n a_k Q(f)^k = P(Q(f))$

Remarque V.1. Si E est de dimension finie, alors $\mathbb{K}[f]$ est un sev de $\mathcal{L}(E)$, donc $\mathbb{K}[f]$ est de dimension finie. Donc l'application $\varphi : \begin{cases} \mathbb{K}[X] & \rightarrow \mathbb{K}[f] \\ P & \mapsto P(f) \end{cases}$ n'est pas injective.

$\ker \varphi$ n'est pas réduit à $\{0\}$

Définition : Polynôme annulateur

Soit $P \in \mathbb{K}[X]$ et $f \in \mathcal{L}(E)$. On dit que P est un polynôme annulateur de f si $P(f) = 0$

Définition : Polynôme minimal

Soit $f \in \mathcal{L}(E)$. On appelle polynôme minimal de f un polynôme non-nul unitaire annulateur de f de degré minimal, qu'on note π_f ou μ_f (cette notation existe mais elle est très rare).

Proposition :

Avec E un \mathbb{K} -ev de dimension finie et $f \in \mathcal{L}(E)$:

1. $\mathbb{K}[f]$ est de dimension $\deg(\pi_f)$ et $(id, f, \dots, f^{\deg(\pi_f)-1})$ est une base de $\mathbb{K}[f]$
2. P est annulateur de f si, et seulement si, $\pi_f | P$
3. Si P est annulateur de f , alors toute valeur propre de f est une racine de P .
4. Les valeurs propres de f sont les racines de π_f

Preuve

Montrons le 1 : On note $p = \deg \pi_f$. Soit $(a_0, \dots, a_{p-1}) \in \mathbb{K}^p$.

On suppose $a_0 id + a_1 f + \dots + a_{p-1} f^{p-1} = 0$

Par l'absurde, si $(a_0, \dots, a_{p-1}) \neq (0, \dots, 0)$, alors le polynôme $Q = \sum_{k=0}^{p-1} a_k X^k$ est non-nul, annulateur de f et de degré strictement inférieur à $\deg \pi_f$. D'où la contradiction.

Donc la famille (id, f, \dots, f^{p-1}) est libre.

Soit $g \in \mathbb{K}[X]$. on a donc $P \in \mathbb{K}[X], g = P(f)$.

Par division euclidienne, $P = \pi_f Q + R$ avec $\deg R < \deg \pi_f$.

Alors $P(f) = (\pi_f Q + R)(f) = \pi_f(f) \circ Q(f) + R(f) = R(f)$

Donc $g \in Vect(id, f, \dots, f^{p-1})$

D'où le résultat.

Montrons le 2 : Tout multiple de π_f est annulateur (immédiat)

Réciproquement, soit P un polynôme annulateur de f . Par division euclidienne, soit $P = \pi_f Q + R$ avec $\deg R < \deg \pi_f$

Alors $P(f) = 0 = R(f)$

Donc R est annulateur de degré strictement inférieur à $\deg \pi_f$, donc $R = 0$.

De là, on peut assurer l'unicité du polynôme minimal : s'il y en avait deux, ils se diviseraient l'un l'autre et seraient associés. Mais puisqu'ils sont unitaires, ils sont égaux.

Montrons le 3 : Soit $\lambda \in S_p(f)$

On peut donc prendre $x \in E$ non-nul tel que $f(x) = \lambda(x)$

Par récurrence, $\forall i \in \mathbb{N}, f^i(x) = \lambda^i x$

Et donc, $\forall P \in \mathbb{K}[X], P(f)(x) = P(\lambda)x$

Et si P est annulateur de f , $P(\lambda)x = 0$, et comme $x \neq 0$, $P(\lambda) = 0$

Donc λ est une racine de P .

Montrons le 4 : π_f est annulateur de f , donc toute valeur propre de f est racine de π_f .

Réciproquement, soit λ une racine de π_f

On a donc $Q \in \mathbb{K}[X]$ tel que $\pi_f = (X - \lambda)Q$

$\pi_f(f) = 0$, ce qui s'écrit $(f - \lambda id) \circ Q(f) = 0$

Par l'absurde, si $\lambda \notin S_p(f)$, alors $(f - \lambda id)$ est bijective.

Donc $(f - \lambda id)^{-1} \circ (f - \lambda) \circ Q(f) = Q(f) = 0$

Donc Q est annulateur de f , mais $Q \neq 0$ alors que $\deg Q < \deg \pi_f$, d'où la contradiction.

Donc $\lambda \in S_p(f)$

Remarque V.2. On étend toutes ces notions aux racines de $\mathcal{M}_n(\mathbb{K})$ par passage à l'endomorphisme canoniquement associés. On note avec les notations $\mathbb{K}[A]$ et π_f pour $A \in \mathcal{M}_n(\mathbb{K})$ et au passage si A et B sont semblables, alors $\pi_A = \pi_B$.

Exemple V.1. À refaire pendant la toussaint.

Théorème :

Avec E un \mathbb{K} -ev de dimension finie, $f \in \mathcal{L}(E)$. f est diagonalisable si, et seulement si, π_f est scindé à racines simples

Ce qui correspond à ce qu'il existe un polynôme annulateur non-nul scindé à racines simples de f .

Preuve

Supposons f diagonalisable.

$S_p(f) = \{\lambda_1, \dots, \lambda_p\}$ et donc $\chi_f = (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_p)^{\alpha_p}$

Donc il existe une base dans laquelle la matrice de l'endomorphisme est une diagonale avec α_1 fois λ_1 , α_2 fois λ_2 ...

Et alors $\pi_f = \pi_A = (X - \lambda_1) \dots (X - \lambda_p)$

Réciproquement, on suppose qu'il existe $\lambda_1, \dots, \lambda_p$ distincts tels que $\pi_f = (X - \lambda_1) \dots (X - \lambda_p)$. On

considère L_i les polynômes de Lagrange associés à $\lambda_1, \dots, \lambda_p$ (on rappelle que $L_i = \prod_{j=1, j \neq i}^p \frac{X - \lambda_j}{\lambda_i - \lambda_j}$ et que $\forall i, j \in \llbracket 1, p \rrbracket, L_i(\lambda_j) = \delta_{i,j}$)

Donc (L_1, \dots, L_p) est une base de $\mathbb{K}_{p-1}[X]$ où on peut écrire $P = \sum_{k=1}^p P(\lambda_k) L_k$. En particulier, le polynôme constant de valeur 1 est $1 = \sum_{k=1}^p L_k$.

Alors $1|(f) = \sum_{i=1}^p L_i(f)$ ce qui correspond à $id = \sum_{i=1}^p L_i(f)$

Donc $\forall x \in E, x = \sum_{i=1}^p L_i(f)(x)$ avec $L_i(f)(x) \in \ker(f - \lambda_i id)$

Donc E est la somme des espaces propres

Donc f est diagonalisable.

En effet, $(f - \lambda_i id)(L_i(f)(x)) = ((f - \lambda_i id) \circ L_i(f))(x) = ((X - \lambda_i)L_i)(f)(x) = (\mu \pi_f)(f)(x)$ avec

$$\mu = \prod_{j=1, j \neq i}^p \frac{1}{\lambda_i - \lambda_j}$$

Donc $E \subset \bigoplus_{i=1}^p E_{\lambda_i}(f)$

Théorème : Cayley-Hamilton

Avec E un \mathbb{K} -ev de dimension finie, $f \in \mathcal{L}(E)$, alors χ_f est un polynôme annulateur de f .

Preuve

On se place d'abord dans le cas où $\mathbb{K} = \mathbb{C}$.

On sait que les valeurs propres sont les racines de π_f et de χ_f .

$$\pi_f = \prod_{\lambda \in S_p(f)} (X - \lambda)^{\alpha_\lambda} \text{ et } \chi_f = \prod_{\lambda \in S_p(f)} (X - \lambda)^{\beta_\lambda}$$

On va montrer que $\forall \lambda \in S_p(f), \alpha_\lambda \leq \beta_\lambda$

Soit $\lambda \in S_p(f)$. On considère $F = \text{Ker}(f - \lambda \text{id})^{\alpha_\lambda}$

F est stable par F , donc on peut considérer f_λ l'endomorphisme induit par f sur F

Donc $(X - \lambda)^{\alpha_\lambda}$ est annulateur de f_λ

Donc $\pi_{f_\lambda} = (X - \lambda)^\gamma$ où $\gamma \leq \alpha_\lambda$

Supposons par l'absurde $\gamma < \alpha_\lambda$:

Rappel : $\pi_f = (X - \lambda)^{\alpha_\lambda} Q$ avec $Q(\lambda) \neq 0$

$$\forall x \in E, ((X - \lambda)^\gamma Q)(f)(x) = (f - \lambda \text{id})^\gamma(Q(f)(x))$$

$$\text{Or } (f - \lambda \text{id})^{\alpha_\lambda} Q(f)(x) = 0$$

$$\text{Donc } Q(f)(x) \in F$$

$$\text{Donc } (f - \lambda \text{id})^\gamma(Q(f)(x)) = (f_\lambda - \lambda \text{id})^\gamma(Q(f)(x)) = 0$$

Donc $(X - \lambda)^\gamma Q$ est un polynôme annulateur de f de degré strictement inférieur à $\deg \pi_f$

D'où la contradiction. Donc $\gamma = \alpha_\lambda$

Donc $(f_\lambda - \lambda \text{id})$ est donc nilpotent d'ordre α_λ , donc $\dim F \geq \alpha_\lambda$

On prend une base de F (e_1, \dots, e_p) qu'on complète en une base de E . La matrice de f dans cette

base est une matrice par blocs triangulaire supérieure. $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$

$$\text{Donc } \chi_f = \chi_A \chi_B = (X - \lambda)^p \chi_B$$

$$\text{Donc } \alpha_\lambda \leq p \leq \beta_\lambda$$

Donc le polynôme minimal divise le polynôme caractéristique.

Avec $\mathbb{K} = \mathbb{R}$, avec B une base de E . On considère $M = \text{Mat}_B(f) \in \mathcal{M}_n(\mathbb{R})$

On a que $\pi_f = \pi_M$ et $\chi_f = \chi_M$

$$M \in \mathcal{M}_n(\mathbb{R}) \text{ et } \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$$

$$\text{Donc } \pi_{M, \mathbb{C}} | \chi_{M, \mathbb{C}}$$

$$\chi_M = \det(XI - M) \text{ ne dépend pas du corps considéré. Donc } \chi_{M, \mathbb{C}} = \chi_{M, \mathbb{R}}$$

$$\pi_{M, \mathbb{C}} | \pi_{M, \mathbb{R}}$$

$$\text{Rappelons que pour } P \in \mathbb{C}[X] : P(M) = 0 \Rightarrow \bar{P}(\bar{M})$$

Mais si $M \in \mathcal{M}_n(\mathbb{R})$, alors $\bar{P}(M) = 0$ aussi.

$\bar{\pi}_{M, \mathbb{C}}$ est annulateur de M et de même degré que $\pi_{M, \mathbb{C}}$, donc $\pi_{M, \mathbb{C}} = \bar{\pi}_{M, \mathbb{C}}$

$$\text{Donc } \pi_{M, \mathbb{C}} \in \mathbb{R}[X]$$

$$\text{Donc } \pi_{M, \mathbb{R}} | \pi_{M, \mathbb{C}}$$

Chapitre V

Suites et séries de fonctions

I Suites de fonctions

Définition : Convergence simple

Soit E, F , 2 \mathbb{K} -ev de dimension finie. Soit $A \subset E$. Soit $(f_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$.
On dit que (f_n) converge simplement vers $g \in \mathcal{F}(A, F)$ si $\forall t \in A, (f_n(t)) \rightarrow g(t)$

Définition : Convergence uniforme

Soit E, F , 2 \mathbb{K} -ev de dimension finie. Soit $A \subset E$. Soit $(f_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$.
On dit que (f_n) converge uniformément vers $g \in \mathcal{F}(A, F)$ si $\sup_A \|f_n - g\| \rightarrow_{n \rightarrow \infty} 0$

Remarque I.1. (f_n) converge simplement (CVS) vers g si, et seulement si, $\forall t \in A, \forall \varepsilon \in \mathbb{R}_+^*, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow \|f_n(t) - g(t)\| \leq \varepsilon$

(f_n) converge uniformément (CVU) vers g si, et seulement si, $\forall \varepsilon \in \mathbb{R}_+^*, \exists n_0 \in \mathbb{N}, \forall t \in A, \forall n \in \mathbb{N}, n \geq n_0 \Rightarrow \|f_n(t) - g(t)\| \leq \varepsilon$

Exemple I.1. $f_n : \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} 0 & \text{si } x \leq n \\ 1 & \text{si } x > n \end{cases} \end{cases}$ converge simplement vers 0, mais ne converge pas uniformément vers 0 : $\sup_A \|f_n(t) - g(t)\| = 1$

Exemple I.2. $f_n : \begin{cases}]0, 1[\rightarrow \mathbb{R} \\ x \mapsto \begin{cases} 0 & \text{si } x \leq 1 - \frac{2}{n} \\ n & \text{si } 1 - \frac{2}{n} < x < \frac{1}{n} - \frac{2}{n} \\ 0 & \text{sinon} \end{cases} \end{cases}$ converge simplement vers 0, mais ne converge pas uniformément vers 0 : $\sup_A \|f_n(t) - g(t)\| \rightarrow_{n \rightarrow +\infty} +\infty$

Remarque I.2. — Si f_n converge uniformément vers g alors f_n converge simplement vers g .
— La notion de convergence simple n'est pas une notion de convergence dans un EVN (suites non-bornées convergentes)
— Sur $\mathcal{B}(A, F)$ (fonctions bornées) la notion de convergence uniforme est celle de convergence pour la norme infinie ($\sup_A \|f\|$)

Exemple I.3. Prenons un exemple de fonction qui converge uniformément sans notion de norme infinie :
 $f_n : \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto x + \frac{1}{n} \end{cases}$ qui converge uniformément vers $x \mapsto x$ car, $\forall x \in \mathbb{R}, |f_n(x) - x| \leq \frac{1}{n}$ et donc $\sup |f_n - id| \leq \frac{1}{n}$

Exemple I.4. Etudier la convergence simple et uniforme de $f_n : \begin{cases} \mathbb{R}_+^* \rightarrow \mathbb{R} \\ x \mapsto \frac{n^\alpha}{1+nx} \end{cases}$

À $x \in \mathbb{R}_+^*$ fixé : $f_n \sim \frac{n^\alpha}{nx} \sim \frac{n^{\alpha-1}}{x}$ qui converge simplement si $\alpha \leq 1$.

— si $\alpha < 1$, $f_n(x)$ tend vers 0.

— si $\alpha = 1$, $f_n(x)$ tend vers $\frac{1}{x}$

— si $\alpha > 1$, $f_n(x)$ diverge.

Donc :

— si $\alpha < 1$, (f_n) converge simplement vers 0.

— si $\alpha = 1$, (f_n) converge simplement vers $\frac{1}{x}$

— si $\alpha > 1$, (f_n) ne converge pas simplement.

Si $\alpha = 1$, pour la convergence uniforme, il y a deux candidats, $x \mapsto \frac{1}{x}$ et $x \mapsto 0$.

Pour $x \in \mathbb{R}_+^* : f_n(x) - \frac{1}{x} = \frac{n}{1+nx} - \frac{1}{x} = \frac{-1}{x(1+nx)} = g_n(x)$

Alors $\forall n \in \mathbb{N}^*, |g(\frac{1}{n})| = \frac{n}{2}$ et donc (f_n) ne converge pas uniformément vers $x \mapsto \frac{1}{x}$

Si $\alpha < 1$, étudions f_n pour $n \in \mathbb{N}^*$. f_n est dérivable et $\forall x \in \mathbb{R}_+^*, f'_n(x) = -n^\alpha n \frac{1}{(1+nx)^2} = \frac{-n^{\alpha+1}}{(1+nx)^2}$ et donc f_n est décroissante positive donc $\sup |f_n| = n^\alpha$

Et donc (f_n) converge uniformément vers la fonction nulle si, et seulement si, $\alpha < 0$

II Séries de fonctions

Dans la suite du chapitre, on aura E, F deux \mathbb{K} -ev de dimension finie, et $A \subset E$

Définition : Convergences

Soit $(u_n) \in \mathcal{F}(A, F)^\mathbb{N}$, on dit que la série de fonction $(\sum u_n)$ converge simplement si la suite des sommes partielles $(\sum_{k=0}^n u_k)$ converge simplement.

On dit que $(\sum u_n)$ converge uniformément si la suite des sommes partielles $(\sum_{k=0}^n u_k)$ converge uniformément.

Remarque II.1. Si $(\sum u_n)$ converge simplement sur A , on peut définir la fonction "somme" de $\sum u_n$ sur

$$A \rightarrow F$$

$A : c'est l'application \quad t \mapsto \sum_{n=0}^{+\infty} u_n(t)$

Exemple II.1. $(\sum (z \mapsto z^n))$ (on notera le plus souvent $(u_n) \in \mathcal{C}, \mathcal{C}^\mathbb{N}$ définie par $u_n(z) = z^n$ pour $z \in \mathbb{C}$). Pour quelle valeur de z $(\sum u_n)$ est-elle définie ? $(\sum u_n)$ converge simplement sur $\mathbb{D}(0, 1)$ (disque ouvert de centre 0 de rayon 1) et $\forall z \in \mathbb{D}(0, 1); \sum_{n=0}^{+\infty} u_n = \frac{1}{1-z}$

$(\sum u_n)$ converge simplement vers $z \mapsto \frac{1}{1-z}$

Pour la convergence uniforme : soit $z \in \mathbb{D}(0, 1)$, $n \in \mathbb{N}$

$$|\sum_{k=0}^n z^k - \frac{1}{1-z}| = |\sum_{k=n+1}^{+\infty} z^k| = \frac{|z|^{n+1}}{1-|z|}$$

On trouve une suite de $z \in \mathbb{D}(0, 1)$ telle que le reste diverge : $|R_n(1 - \frac{1}{n})| = (n+1)(1 + \frac{1}{n+1})^{n+1} \sim ne^{-1}$

Et donc il n'y a pas convergence uniforme sur $\mathbb{D}(0, 1)$.

Par contre, avec $R \in [0, 1[: \forall n \in \mathbb{N}, \forall z \in \mathbb{D}(0, R), |R_n(z)| \leq \frac{R^{n+1}}{1-R}$ (on minore le dénominateur et on majore le numérateur)

$$\text{Donc } \sup_{z \in \mathbb{D}(0, R)} |R_n(z)| \leq \frac{R^{n+1}}{1-R}$$

Donc $\sum z \mapsto z^n$ converge uniformément sur $\mathbb{D}(0, R)$ pour tout $R \in [0, 1[$

Remarque II.2. $\cup_{R \in [0,1[} \mathbb{D}(0, R) = \mathbb{D}(0, 1)$. Il y a convergence uniforme sur $\mathbb{D}(0, R)$ pour tout R mais pas convergence uniforme sur $\mathbb{D}(0, 1)$. Cependant, si on avait une réunion finie, il y aurait convergence uniforme (en prenant le maximum des deux majorants).

La convergence uniforme n'est pas une notion locale.

Proposition :

Soit $(u_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$, alors $(\sum u_n)$ converge uniformément si, et seulement si, $\sum u_n$ converge simplement et $(R_n) = \sum_{k=n+1}^{+\infty} u_k$ converge uniformément vers 0.

Exemple II.2. $f :]-1, 1[\rightarrow \mathbb{R}$
 $x \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^n x^n}{n}$

Montrer que f est bien définie. La série de fonctions $\sum (-1)^n \frac{x^n}{n}$ converge-t-elle uniformément vers f ? Trouver des domaines sur lesquels $(\sum (-1)^n \frac{x^n}{n})$ converge uniformément.

Soit $n \in \mathbb{N}^*$, alors $|\frac{(-1)^n x^n}{n}| \leq |\frac{x^n}{n}| \leq |x|^n$

Or $\sum |x|^n$ converge pour $|x| < 1$

Donc $\sum \frac{(-1)^n x^n}{n}$ converge absolument.

D'autre part, $(\frac{1}{n})$ est positive, décroissante et de limite nulle, donc $\sum \frac{(-1)^n}{n}$ converge pour $x \in]-1, 1[$

Et donc on a bien la convergence simple.

Soit $x \in]-1, 1[$, alors $|R_n(x)| = |\sum_{k=n+1}^{+\infty} (-1)^k \frac{x^k}{k}|$

$|R_n(-1 + \frac{1}{n+1})| = |\sum_{k=n+1}^{+\infty} (-1)^k \frac{(-1 + \frac{1}{n+1})^k}{k}| = |(-1 + \frac{1}{n+1})^{n+1}| |\sum_{k=0}^{+\infty} \frac{(-1 + \frac{1}{n+1})^k}{n+k+1}|$

$\forall k \in \mathbb{N}^*$, $(-1 + \frac{1}{n+1})^k$ est du signe de $(-1)^k$

Donc $|B_n| = |\sum_{k=0}^{+\infty} \frac{(-1)^{n+1+k} (-1 + \frac{1}{n+1})^k}{n+k+1}| = \sum_{k=0}^{+\infty} \frac{(1 - \frac{1}{n+1})^k}{n+k+1}$

$\geq \sum_{k=0}^{+\infty} \frac{(1 - \frac{1}{n+1})^k}{n} \geq \sum_{k=0}^N \frac{(1 - \frac{1}{n+1})^k}{n+k+1}$ pour $N \in \mathbb{N}$ quelconque

$\geq \sum_{k=0}^N \frac{(\frac{n}{n+1})^k}{n+N+1} \geq \frac{1 - (\frac{n}{n+1})^{N+1}}{(1 - \frac{n}{n+1})(n+N+1)} \geq (n+1) \frac{(1 - (\frac{n}{n+1})^{N+1})}{n+N+1}$

ceci est vrai en particulier pour $N = n$ et alors $|B_n| \geq \frac{(n+1)(1 - (\frac{n}{n+1})^{n+1})}{2n+1} \sim \frac{1}{2}e^{-1}$

Donc $|R_n(1 - \frac{1}{n+1})| \geq (1 - \frac{1}{n+1})^{n+1} |B_n|$ qui ne tend pas vers 0 : la série de fonctions ne converge pas uniformément sur $] -1, 1[$.

Soit $a \in]0, 1[$, $\forall x \in [-a, a]$, $\forall n \in \mathbb{N}$, $|\sum_{k=n+1}^{+\infty} (-1)^k \frac{x^k}{k}| \leq \sum_{k=n+1}^{+\infty} \frac{|x|^k}{k} \leq \sum_{k=n+1}^{+\infty} \frac{a^k}{k}$ qui tend vers 0 et est indépendant de x , d'où la convergence uniforme. Donc $\sup_{x \in [-a, a]} |R_n(x)| \rightarrow_{n \rightarrow +\infty} 0$ donc la série converge uniformément sur $[-a, a]$ pour $a \in]0, 1[$

Soit $x \in [0, 1]$, soit $n \in \mathbb{N}$

$(\frac{x^k}{k})$ est positive, décroissante et de limite nulle.

Donc $|\sum_{k=n+1}^{+\infty} (-1)^k \frac{x^k}{k}| \leq \frac{x^{n+1}}{n+1} \frac{1}{n+1}$

Donc $\sup_{x \in [0, 1]} R_n(x) \leq \frac{1}{n+1}$ et tend donc vers 0

Donc la série converge uniformément sur $[0, 1]$

Donc $(\sum \frac{(-1)^n x^n}{n})$ converge uniformément sur $[-a, 1]$ pour tout $a \in]0, 1[$

Définition : Convergence normale

Soit $(u_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$. On dit que $\sum u_n$ converge normalement si $\sum \sup_A \|u_n\|$ converge.

Exemple II.3. $\sum (-1)^n \frac{x^n}{n}$ converge normalement sur $[-\frac{3}{4}, \frac{3}{4}]$: soit $x \in [-\frac{3}{4}, \frac{3}{4}]$, $\forall n \in \mathbb{N}$, $|(-1)^n \frac{x^n}{n}| \leq \frac{1}{n} \left(\frac{3}{4}\right)^n$ et $\sum \frac{(\frac{3}{4})^n}{n}$ converge

Donc $\sup_{x \in [-\frac{3}{4}, \frac{3}{4}]} |(-1)^n \frac{x^n}{n}| \leq \frac{1}{n} \left(\frac{3}{4}\right)^n$ donc $\sum \sup \left| \frac{(-1)^n x^n}{n} \right|$ converge par critère de majoration positif.

Proposition :

Tout série de fonction qui converge normalement converge uniformément

Preuve

Si $\sum u_n$ converge normalement :

$\forall x \in A$, $\sum u_n(x)$ converge absolument

Par critère de majoration positif et $\forall x \in A$, $\|R_n(x)\| = \left\| \sum_{k=n+1}^{+\infty} u_k(x) \right\| \leq \sum_{k=n+1}^{+\infty} \|u_k(x)\| \leq \sum_{k=n+1}^{+\infty} \sup_A \|u_k\|$.

Donc $\sup_A \|R_n(x)\| \leq \sum_{k=n+1}^{+\infty} \sup_A \|u_k\|$

Donc $\sup_A \|R_n(x)\| \rightarrow 0$

Remarque II.3. Résumé :

- Pour les suites de fonctions, on a deux notions, et la convergence uniforme entraîne la convergence simple.
- Pour les séries de fonctions, on a trois notions : la convergence normale, la convergence uniforme, la convergence simple et on peut rajouter la convergence absolue en tout point. On a que la convergence normale entraîne la convergence uniforme. La convergence uniforme entraîne la convergence simple. La convergence absolue en tout point entraîne la convergence simple. Par définition, la convergence normale entraîne la la convergence absolue en tous points. Par transitivité, la convergence normale entraîne la convergence simple.

III Continuité d'une limite uniforme**Théorème : Continuité uniforme**

Soit $(f_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$ une suite de fonctions.

Si $\forall n \in \mathbb{N}$, f_n est continue sur A et que (f_n) converge uniformément vers g , alors g est continue.

Ce qui correspond à : une limite uniforme de fonctions continues est continue.

Preuve

Soit $x \in A$

Soit $\varepsilon \in \mathbb{R}_+^*$

f_n converge uniformément vers g , on fixe donc $n_0 \in \mathbb{N}$ tel que $\sup_{x \in A} \|g(x) - f_{n_0}(x)\| < \frac{\varepsilon}{3}$

f_{n_0} est continue en x , on peut donc prendre $\alpha \in \mathbb{R}_+^*$ tel que :

$\forall h \in E, \|h\| < \alpha$ et $x + h \in A \Rightarrow \|f_{n_0}(x + h) - f_{n_0}(x)\| < \frac{\varepsilon}{3}$

Pour $h \in E$ tel que $\|h\| < \alpha$ et $x + h \in A$, on a :

$\|g(x + h) - g(x)\| \leq \|g(x + h) - f_{n_0}(x + h)\| + \|f_{n_0}(x + h) - f_{n_0}(x)\| + \|f_{n_0}(x) - g(x)\| \leq \varepsilon$

Et donc $g(x + h) \rightarrow_{h \rightarrow 0} g(x)$

Donc g est continue sur A

Théorème : Application aux séries de fonctions

Soit $(u_n) \in \mathcal{C}(A, F)^{\mathbb{N}}$, si $\sum u_n$ converge uniformément sur A , alors $x \mapsto \sum_{n=0}^{+\infty} u_n(x)$ est continue.

Preuve

On a que la suite des sommes partielles est une suite de sommes finies de fonctions continues, donc de fonctions continues. Et donc par le théorème précédent, on a que $\sum_{n=0}^{+\infty} u_n$ est continue.

Exemple III.1. Notons $f : \begin{cases} \mathbb{R}_+ \rightarrow \mathbb{R} \\ x \mapsto \sum_{n=1}^{+\infty} \frac{1}{n\sqrt{n+x}} \end{cases}$

Montrer que f est bien définie et continue.

Soit $x \in \mathbb{R}_+$, on a donc que $\frac{1}{n\sqrt{n+x}} > 0$ et que $\frac{1}{n\sqrt{n+x}} \sim \frac{1}{n^{3/2}}$. Or $\sum \frac{1}{n^{3/2}}$ est convergente

Donc $\sum \frac{1}{n\sqrt{n+x}}$ converge simplement vers f par critère d'équivalent positif.

Pour $n \in \mathbb{N}^*$, on a que $x \mapsto \frac{1}{n\sqrt{n+x}}$ est continue.

Montrons la convergence uniforme.

Soient $n \in \mathbb{N}^*, x \in \mathbb{R}_+$.

Alors on a que $|\frac{1}{n\sqrt{n+x}}| \leq \frac{1}{n^{3/2}}$

Or $\sum \frac{1}{n^{3/2}}$ converge donc $(\sum x \mapsto \frac{1}{n\sqrt{n+x}})$ converge normalement sur \mathbb{R}_+ donc converge uniformément sur \mathbb{R}_+

Donc f est continue par théorème de continuité uniforme

Exemple III.2. Notons $f : \begin{cases} \mathbb{R}_+ \rightarrow \mathbb{R} \\ x \mapsto \sum_{n=1}^{+\infty} \frac{\sqrt{x^2+n}}{n(n+2)} \end{cases}$

Montrer que f est bien définie et continue.

Soit $x \in \mathbb{R}_+$, on a donc que $\frac{\sqrt{x^2+n}}{n(n+2)} > 0$ et que $\frac{\sqrt{x^2+n}}{n(n+2)} \sim \frac{1}{n^{3/2}}$. Or $\sum \frac{1}{n^{3/2}}$ est convergente

Donc $\sum \frac{\sqrt{x^2+n}}{n(n+2)}$ converge simplement vers f par critère d'équivalent positif.

Pour $n \in \mathbb{N}^*$, on a que $x \mapsto \frac{\sqrt{x^2+n}}{n(n+2)}$ est continue.

Montrons la convergence uniforme.

On a que $u_n(n) \sim \frac{1}{2n}$, et donc $\sup_{\mathbb{R}_+} \|u_n\| > \frac{1}{2n}$, donc $\sum \frac{\sqrt{x^2+n}}{n(n+2)}$ ne converge pas normalement sur \mathbb{R}_+ .

Soit $A \in \mathbb{R}_+$, soit $x \in [0, A]$

Soit $n \in \mathbb{N}^*$, alors $|\frac{\sqrt{x^2+n}}{n(n+2)}| \leq |\frac{\sqrt{A^2+n}}{n^2}| \sim \frac{1}{n^{3/2}}$

Par critère d'équivalent positif, on a que $\sum u_n$ converge normalement sur $[0, A]$ pour tout $A \in \mathbb{R}_+$

Donc pour tout $A \in \mathbb{R}_+$, $(\sum u_n)$ converge uniformément sur $[0, A]$

Par théorème de continuité uniforme : Pour $A \in \mathbb{R}_+$, $(\sum_{n=1}^{+\infty} u_n)$ est continue sur $[0, A]$.

Donc $\sum_{n=1}^{+\infty} u_n$ est continue sur \mathbb{R}_+ : en effet, la continuité est une notion locale, l'intersection infinie conserve la propriété.

Exemple III.3. Notons $f : \begin{cases} \mathbb{R}_+^* \rightarrow \mathbb{R} \\ x \mapsto \sum_{n=1}^{+\infty} \frac{1}{1+n^2x} \end{cases}$

Montrer que f est bien définie et continue.

Soit $x \in \mathbb{R}_+$, on a donc que $\frac{1}{1+n^2x} > 0$ et que $\frac{1}{1+n^2x} \sim \frac{1}{n^2x}$. Or $\sum \frac{1}{n^2x}$ est convergente

Donc $\frac{1}{1+n^2x}$ converge simplement vers f par critère d'équivalent positif.

Pour $n \in \mathbb{N}^*$, on a que $x \mapsto \frac{1}{1+n^2x}$ est continue.

Montrons la convergence uniforme.

Soit $A \in \mathbb{R}_+$, soit $x \in [0, A]$

Soit $n \in \mathbb{N}^*$, alors $|\frac{1}{1+n^2x}| \leq |\frac{1}{1+n^2A}| \sim \frac{1}{n^2A}$

Par critère d'équivalent positif, on a que $\sum u_n$ converge normalement sur $[0, A]$ pour tout $A \in \mathbb{R}_+$

Donc pour tout $A \in \mathbb{R}_+$, $(\sum u_n)$ converge uniformément sur $[0, A]$

Par théorème de continuité uniforme : Pour $A \in \mathbb{R}_+$, $(\sum_{n=1}^{+\infty} u_n)$ est continue sur $[0, A]$.

Donc $\sum_{n=1}^{+\infty} u_n$ est continue sur \mathbb{R}_+ .

Exemple III.4. Notons $\exp : \begin{cases} \mathcal{M}_n(\mathbb{K}) \rightarrow \mathcal{M}_n(\mathbb{K}) \\ A \mapsto \exp(A) \end{cases}$

Montrer que f est bien définie et continue.

Soit N une norme d'algèbre sur $\mathcal{M}_n(\mathbb{K})$

Par récurrence on prouve que $N(A^n) \leq N(A)^n$

Et donc $\frac{1}{n!}N(A^n) \leq \frac{1}{n!}N(A)^n$ et donc $\sum \frac{1}{n!}A^n$ converge simplement dans $\mathcal{M}_n(\mathbb{K})$

$\forall n \in \mathbb{N}$, $\frac{1}{n!}A^n$ est continue

Montrons la convergence normale :

Si $R \in \mathbb{R}_+^*$, $A \in \mathcal{B}(0, R)$ alors $\frac{1}{n!}N(A^n) \leq \frac{R^n}{n!}$ et comme $\sum \frac{R^n}{n!}$ converge, alors on a que \exp converge normalement (et donc uniformément) dans $\mathcal{B}(0, R)$.

Donc par théorème de continuité uniforme, $A \mapsto \exp(A)$ est continue sur $\mathcal{B}(0, R)$ pour tout $R \in \mathbb{R}_+$

Et donc $A \mapsto \exp(A)$ est continue sur $\mathcal{M}_n(\mathbb{K})$

Théorème : Extension de limite uniforme

Soit $(f_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$, soit $a \in \bar{A}$.

Si $\forall n \in \mathbb{N}$, $f_n(x) \rightarrow_{x \rightarrow a} l_n$ et f_n converge uniformément vers g

Alors (l_n) est convergente et $g(x) \rightarrow_{x \rightarrow a} \lim_{n \rightarrow +\infty} l_n$

ie : g a une limite finie en a et $\lim_{x \rightarrow a} \lim_{n \rightarrow +\infty} l_n = \lim_{n \rightarrow +\infty} \lim_{x \rightarrow a} f_n$

Ce théorème s'étend lorsque $E = \mathbb{R}$ et que $a = \pm\infty$

Preuve

Sous les hypothèses du théorème : si on admet que (l_n) est convergente, on a une démonstration identique à celle de la continuité.

Pour $x \in A$, on a : $\|g(x) - \lim l_n\| \leq \|g(x) - f_n(x)\| + \|f_n(x) - l_n\| + \|l_n - \lim l_n\|$

Ces trois termes sont plus petit qu'un ε donné pour n suffisamment grand pour le premier et troisième terme et pour $\|x - a\| < \alpha$ pour le deuxième terme.

Théorème : échange de limites de séries

Avec $(u_n) \in \mathcal{F}(A, F)^{\mathbb{N}}$ et $a \in \bar{A}$

Si $\forall n \in \mathbb{N}, u_n(x) \rightarrow_{x \rightarrow a} v_n$ et que $(\sum u_n)$ converge uniformément, alors $(\sum_{n=0}^{+\infty} u_n)$ a une limite en a

$$\text{et } \lim_{x \rightarrow a} \sum_{n=0}^{+\infty} u_n(x) = \sum_{n=0}^{+\infty} v_n$$

Exemple III.5. Reprenons $f : \begin{cases} \mathbb{R}_+ \rightarrow \mathbb{R} \\ x \mapsto \sum_{n=1}^{+\infty} \frac{1}{n\sqrt{n+x}} \end{cases}$

On rappelle que f converge normalement sur \mathbb{R}_+ puisque $\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}_+, |u_n(x)| \leq \frac{1}{n^{3/2}}$, terme général d'une série convergente.

Pour déterminer la limite de f en l'infini, on note que $\forall n \in \mathbb{N}, u_n(x) \rightarrow_{x \rightarrow +\infty} 0$

$\sum u_n$ converge normalement donc uniformément sur \mathbb{R}_+ , donc $\lim_{x \rightarrow +\infty} f_x$ existe

$$\text{et alors } f(x) \rightarrow \sum_{n=0}^{+\infty} \lim_{x \rightarrow +\infty} u_n(x) = 0$$

Exemple III.6. Soit $\zeta : x \mapsto \sum_{n=1}^{+\infty} \frac{1}{n^x}$ sur \mathbb{R} . Donner son domaine de définition, sa continuité et sa limite en $+\infty$.

Soit $x \in \mathbb{R}$, alors $\sum \frac{1}{n^x}$ converge si, et seulement si, $x > 1$

Donc le domaine de définition de ζ est $D =]1, +\infty[$

Soit $a \in D, n \in \mathbb{N}^*, x \in [a, +\infty[$, on a $0 < \frac{1}{x} \leq \frac{1}{n^a}$

$\sum \frac{1}{n^a}$ converge et donc $\sum (x \mapsto \frac{1}{n^x})$ converge normalement sur $[a, +\infty[$

De plus, pour tout $n \in \mathbb{N}^*$, on a que $x \mapsto \frac{1}{n^x}$ est uniforme

Donc par théorème de continuité uniforme, ζ est continue sur $[a, +\infty[$ pour tout $a > 1$ donc sur $]1, +\infty[$

Pour la limite en l'infini : $\forall n \in \mathbb{N}^*, \frac{1}{n^x} \rightarrow_{x \rightarrow \infty} \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}$

$\sum x \mapsto \frac{1}{n^x}$ converge normalement donc uniformément sur $[2022, +\infty[$

Donc par théorème d'échange de limites de séries, $\zeta(x) \rightarrow_{x \rightarrow +\infty} 1$

Pour la "limite" en 1 de ζ . Soit $A > 0$

$\sum \frac{1}{n}$ diverge comme $\frac{1}{n} > 0$ et $(\sum_{k=1}^n \frac{1}{k}) \rightarrow +\infty$

On peut donc trouver N tel que $\sum_{k=1}^N \frac{1}{k} > A + 1$

$$\sum_{k=1}^N \frac{1}{k^x} \rightarrow_{x \rightarrow 1} \sum_{k=1}^N \frac{1}{k}$$

On peut donc fixer α tel que $1 < x < \alpha \Rightarrow \left| \sum_{k=1}^N \frac{1}{k^x} - \sum_{k=1}^N \frac{1}{k} \right| < 1$

Pour $x \in]1, \alpha[$, alors $\zeta(x) \geq \sum_{k=1}^N \frac{1}{k^x} \geq \left| - \sum_{k=1}^N \frac{1}{k^x} - \sum_{k=1}^N \frac{1}{k} \right| + \sum_{k=1}^N \frac{1}{k} \geq A$

Donc $\lim_{x \rightarrow 1} \zeta(x) = +\infty$

IV Intégration et dérivation

On s'intéresse aux fonctions de $I \subset \mathbb{R}$ à valeurs dans F où I est un intervalle.

Théorème : Intégration uniforme ou théorème d'échange limite-intégrale uniforme

Soit a, b un segment, $(f_n) \in \mathcal{C}_0([a, b], F)^{\mathbb{N}}$
 si (f_n) converge uniformément vers g sur $[a, b]$, alors $\int_a^b f_n \rightarrow \int_a^b g$
 Ce qui correspond à $\lim_{n \rightarrow +\infty} \int_a^b f_n = \int_a^b \lim_{n \rightarrow +\infty} f_n$

Preuve

$$\left\| \int_{[a,b]} f_n - \int_{[a,b]} g \right\| \leq \int_{[a,b]} \|f_n - g\| \leq |b - a| \|f_n - g\|_{\infty}$$

Comme f_n converge uniformément vers g , alors $\|f_n - g\|_{\infty} \rightarrow 0$
 Donc $\int_{[a,b]} f_n \rightarrow \int_{[a,b]} g$

Exemple IV.1. $\int_0^1 n^2 t^n (1 - t) dt = n^2 \left(\frac{1}{n+1} - \frac{1}{n+2} \right) = \frac{n^2}{(n+1)(n+2)}$

$$\text{Et } \frac{n^2}{(n+1)(n+2)} \xrightarrow{n \rightarrow +\infty} 0$$

La fonction $f_n : \begin{cases} [0, 1] & \rightarrow \mathbb{R} \\ t & \mapsto n^2 t^n (1 - t) \end{cases}$ converge simplement vers 0.

Donc si elle converge uniformément, ça sera vers 0, cependant son intégrale ne tend pas vers 0, donc il n'y a pas de convergence uniforme.

Remarque IV.1. Pour une série de fonctions, le théorème devient : si $(u_n) \in \mathcal{C}_0([a, b], F)^{\mathbb{N}}$ et si $\sum u_n$ converge uniformément sur $[a, b]$ alors $\int_a^b \sum_{n=0}^{+\infty} u_n(t) dt = \sum_{n=0}^{+\infty} \int_a^b u_n(t) dt$

Théorème :

Soit (f_n) une suite de fonctions continues d'un intervalle I de \mathbb{R} à valeurs dans F convergeant uniformément vers g sur tout segment de I

Soit $a \in I$, on a alors : $F_n : \begin{matrix} I & \rightarrow & F \\ x & \mapsto & \int_a^b f_n(t) dt \end{matrix}$ et $G : \begin{matrix} I & \rightarrow & F \\ x & \mapsto & \int_a^b g(t) dt \end{matrix}$

Alors F_n converge uniformément vers G sur tout segment de I .

Preuve

La définition de G est correcte car f_n converge uniformément vers g sur tout segment de I donc g est continue

À x fixé, comme f_n converge uniformément vers g sur le segment $[a, x]$, d'après le théorème précédent

$\int_{[a,x]} f_n \rightarrow \int_{[a,x]} g$ donc F_n converge simplement vers G sur I

Soit S un segment inclu dans I , soit b un point de S , alors $\forall x \in S, \|F_n(x) - G(x)\| = \left\| \int_b^x f_n(t) - g(t) dt \right\|$
 $\leq \int_b^x \|f_n(t) - g(t)\| dt \leq |x - b| \sup_{t \in [b,x]} \|f_n(t) - g(t)\| \leq \text{diam}(S) \sup_S \|f_n - g\|$

Qui tend vers 0 quand n tend vers l'infini

Ce qui conclut

Théorème : Dérivation uniforme des suites de fonctions

Avec I un intervalle, si :

- $(f_n) \in \mathcal{C}^1(I, F)^{\mathbb{N}}$
- (f_n) converge simplement vers g_0
- (f'_n) converge uniformément vers g_1 sur tout segment de I

alors g_0 est \mathcal{C}^1 tel que $g'_0 = g_1$ et (f_n) converge uniformément sur tout segment de I .

Ce qui correspond à $g'_0 = g_1 \Rightarrow (\lim f_n) = \lim(f'_n)$

Preuve

Pour $n \in \mathbb{N}$, (f_n) est \mathcal{C}^1 , donc (f'_n) est continue.

Donc par théorème de continuité uniforme, g_1 est continue.

D'autre part, si on fixe $a \in I, \forall x \in I, f_n(x) = f_n(a) + \int_a^x f'_n(t)dt$

Donc par théorème d'intégration uniforme, $g_0 = \lim_{n \rightarrow +\infty} f_n = g_0(a) + \int_a^x g_1(t)dt$

Et donc g_0 est \mathcal{C}^1 et $g'_0 = g_1$

La suite (f_n) converge uniformément sur tout segment vers g_0 .

Théorème : Théorème de dérivation des suites à l'ordre k

Avec I un intervalle, si :

- $(f_n) \in \mathcal{C}^k(I, F)^{\mathbb{N}}$
- $\forall j \in \llbracket 0, k-1 \rrbracket, (f_n^{(j)})$ converge simplement vers g_j
- $(f_n^{(k)})$ converge uniformément vers g_k sur tout segment de I

alors g_0 est \mathcal{C}^k tel que $\forall j \in \llbracket 0, k \rrbracket, g_0^{(j)} = g_j$ et $(f_n^{(j)})$ converge uniformément sur tout segment de I .

Théorème : Théorème de dérivation terme à terme à l'ordre k

Avec I un intervalle, si :

- $(u_n) \in \mathcal{C}^1(I, F)^{\mathbb{N}}$
- $(\sum u_n)$ converge simplement
- $(\sum u_n)$ converge uniformément sur tout segment de I

alors $\sum_{n=0}^{+\infty} u_n$ est \mathcal{C}^1 tel que $\forall j \in \llbracket 0, k \rrbracket, \left(\sum_{n=0}^{+\infty} u_n\right)' = \sum_{n=0}^{+\infty} u'_n$ et la somme converge uniformément sur tout segment de I .

Exemple IV.2. $f : \begin{cases} \mathbb{R}_+^* \rightarrow \mathbb{R} \\ x \mapsto \sum_{n=0}^{+\infty} e^{-n^2 x} \end{cases}$ en notant $u_n = e^{-n^2 x}$

Pour $n \in \mathbb{N}$, u_n est \mathcal{C}^∞ , $u'_n : x \mapsto -n^2 e^{-n^2 x}$ et $u''_n : x \mapsto n^4 e^{-n^2 x}$

Soit $x \in \mathbb{R}_+^*$, alors $|u_n(x)| = o(\frac{1}{n^2})$, $|u'_n(x)| = o(\frac{1}{n^2})$

$\sum u_n$ et $\sum u'_n$ convergent simplement sur \mathbb{R}_+^*

Soit $a \in \mathbb{R}_+^*$, soit $n \in \mathbb{N}$, soit $x \in [a, +\infty[$

$$|n^4 e^{-n^2 x}| \leq n^4 e^{-an^2}$$

Or $\sum n^4 e^{-an^2}$ est convergente

Donc la série $\sum x \mapsto n^4 e^{-n^2 x}$ converge normalement donc uniformément sur $[a, +\infty[$ pour tout a

Donc par théorème de dérivation terme à terme d'ordre 2, f est \mathcal{C}^2 et $f' : x \mapsto \sum_{n=0}^{+\infty} -n^2 e^{-n^2 x}$ et

$$f'' : x \mapsto \sum_{n=0}^{+\infty} n^4 e^{-n^2 x}$$

V Résultats de densité

Théorème : Stone-Weierstrass (admis)

Toute fonction continue sur un segment à valeurs dans \mathbb{K} est limite uniforme d'une suite de fonctions polynomiales.

Ce qui correspond à : l'ensemble des fonctions polynomiales est dense dans $(\mathcal{C}([a, b], \mathbb{K}), \|\cdot\|_\infty)$

Exemple V.1. Soit $E = \mathcal{C}([a, b], \mathbb{R})$ un espace préhilbertien réel pour le produit scalaire $\langle f, g \rangle = \int_{[a, b]} fg$

On a que $\mathbb{R}[X] \subset E$, déterminer $\mathbb{R}[X]^\perp$

Soit $f \in \mathbb{R}[X]^\perp$. Alors f vérifie : $\forall P \in \mathbb{R}[X], \langle f, P \rangle = 0$

Par théorème de Stone-Weierstrass, on peut prendre $(P_n) \in \mathbb{R}[X]^\mathbb{N}$ qui converge uniformément vers f

Donc $P_n f$ converge uniformément vers f^2 (comme P_n borné sur $[a, b]$, avec $x \in [a, b], \|P_n(x)g(x) - f(x)g(x)\|_\infty = \|g(x)\|_\infty \|P_n(x) - f(x)\|_\infty \leq \|g(x)\| \sup \|f - P_n\|$)

Donc $\left(\int_a^b P_n(t)f(t)dt\right) \rightarrow \int_a^b f^2(t)dt$

Or $\forall n \in \mathbb{N}, \langle P_n, f \rangle = 0$

Donc $\langle f, f \rangle = 0$ donc $f = 0$

Donc $\mathbb{R}[X]^\perp = \{0\}$

Remarque V.1. Dans le théorème de Stone-Weierstrass, l'hypothèse du segment est fondamentale.

Exemple V.2. Il n'existe aucune suite de polynôme réels qui converge uniformément vers \exp sur \mathbb{R} .

Supposons par l'absurde que $(P_n) \in \mathbb{R}[X]^\mathbb{N}$ converge uniformément vers \exp sur \mathbb{R}_+

On a que $\frac{1}{\exp}$ est bornée sur \mathbb{R}_+ , donc $t \mapsto P_n(t)e^{-t}$ converge uniformément vers la fonction constante de valeur 1.

Pour $n \in \mathbb{N}$ fixé, $P_n(t)e^{-t} \rightarrow_{t \rightarrow +\infty} 0$ par croissance comparée. Comme $t \mapsto P_n(t)e^{-t}$ converge uniformément sur \mathbb{R}_+ , on peut appliquer le théorème d'échange des limites

Donc $\lim_{t \rightarrow +\infty} \lim_{n \rightarrow +\infty} P_n(t)e^{-t} = \lim_{n \rightarrow +\infty} \lim_{t \rightarrow +\infty} P_n(t)e^{-t}$

Et donc $1 = 0$, ce qui est absurde. D'où le résultat.

Théorème : Approximation uniforme par des fonctions en escalier

Toute fonction continue sur un segment à valeurs dans F est limite uniforme d'une suite de fonctions en escaliers

Ce théorème est encore valable pour les fonctions continues par morceaux sur un segment.

Preuve

Soit $f \in \mathcal{C}([a, b], F)$. Par théorème de Heine, f est uniformément continue.

Soit $n \in \mathbb{N}^*$, par continuité uniforme de f on peut trouver $\alpha \in \mathbb{R}_+^*$ tel que $\forall x, y \in [a, b], |x - y| < \alpha \Rightarrow \|f(x) - f(y)\| < \frac{1}{n}$

On fixe $p \in \mathbb{N}^*$ tel que $\frac{b-a}{p} < \alpha$ (qui existe comme \mathbb{R} est archimédien)

On considère la subdivision $(x_i)_{0 \leq i \leq p} = \left(a + i \frac{b-a}{p}\right)_{0 \leq i \leq p}$

On définit la fonction en escalier φ_n par $\forall t \in [a, b], (\exists i \in \llbracket 0, p \rrbracket \text{ tel que } t = x_i) \Rightarrow \varphi(t) = f(t)$ et $\forall i \in \llbracket 0, n-1 \rrbracket, \forall t \in]x_i, x_{i+1}[, \varphi(t) = f\left(\frac{x_i + x_{i+1}}{2}\right)$

Alors $\forall t \in [a, b], \|\varphi_n(t) - f(t)\| \leq \frac{1}{n}$ car $\forall t \in [a, b], \exists i \in \llbracket 0, p-1 \rrbracket, t \in]x_i, x_{i+1}[, |t - \frac{x_i + x_{i+1}}{2}| \leq |x_{i+1} - x_i| \leq \frac{b-a}{p} < \alpha$

Donc $\|f(t) - \varphi\left(\frac{x_i + x_{i+1}}{2}\right)\| < \frac{1}{n}$

Donc $\sup_{t \in [a, b]} \|\varphi_n(t) - f(t)\| \leq \frac{1}{n}$

Remarque V.2. *Ce résultat s'étend sans difficulté aux fonctions continues par morceaux. Il suffit de prendre une subdivision plus fine qu'une subdivision adaptée sur chaque intervalle continu.*

Chapitre VI

Séries entières

I Généralités

Définition :

Soit $(a_n) \in \mathbb{C}^{\mathbb{N}}$. On appelle série entière associée à (a_n) (de variable complexe) la série de fonctions $(\sum(z \mapsto a_n z^n))$ qu'on notera en général $(\sum a_n z^n)$
 $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ est la somme de cette série entière.

On va étudier en général le domaine de définition, les paramètres de continuité, de dérivabilité de la restriction à \mathbb{R} de ces fonctions.

Exemple I.1. $\exp z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$ est définie sur \mathbb{C}

$\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$ est définie pour $z \in \mathbb{C}, |z| < 1$

Proposition : Lemme d'Abel

Soit $(\sum a_n z^n)$ une série entière. Soit $z_0 \in \mathbb{C}^*$. Si $(a_n z_0^n)$ est bornée, alors la série $\sum a_n z^n$ converge absolument pour $z \in \mathbb{C}, |z| < |z_0|$

Preuve

Dans les conditions de l'énoncé :

soit $n \in \mathbb{N}$, soit $z \in \mathbb{C}$ tel que $|z| < |z_0|$

Alors $|a_n z^n| = |a_n z_0^n \frac{z^n}{z_0^n}| \leq |a_n z_0^n| \left| \frac{z}{z_0} \right|^n \leq \sup_{n \in \mathbb{N}} |a_n z_0^n| \left| \frac{z}{z_0} \right|^n$

$\left| \frac{z}{z_0} \right| < 1$ donc $\sum \left(\frac{z}{z_0} \right)^n$ est une série géométrique convergente

Donc par critère de majoration positif, $\sum a_n z^n$ converge absolument.

Définition : Rayon de convergence

Soit $\sum a_n z^n$ une série entière. On appelle rayon de cette série $R = \sup\{|z| \mid z \in \mathbb{C}, (a_n z^n) \text{ est bornée}\}$
 Par convention, $R = +\infty$ si cet ensemble n'est pas majoré.

Proposition :

Soit $\sum a_n z^n$ une série entière de rayon $R \in \bar{\mathbb{R}}$

- $\forall z \in \mathcal{D}(0, R), \sum a_n z^n$ converge absolument
- $\forall z \in \mathbb{C}, |z| > R, a_n z^n$ est non-bornée.

Preuve

Soit $z \in \mathbb{C}, |z| < R$

Par caractérisation de la borne supérieure, on peut trouver $z_0 \in \mathbb{C}$ tel que $(a_n z_0^n)$ est bornée et $|z| < |z_0| \leq R$

Par lemme d'Abel, $\sum a_n z^n$ converge.

Soit $z \in \mathbb{C}, |z| > R$, alors $(a_n z^n)$ n'est pas bornée par définition de la borne supérieure, R étant un majorant de l'ensemble des z tels que $(a_n z^n)$ bornée.

Exemple I.2. Pour $\sum \frac{z^n}{n!}$, on a $R = +\infty$

Pour $\sum z^n$ on a $R = 1$ et son domaine de définition n'inclut pas le bord, donc $\mathcal{D}(0, 1)$

Pour $\sum_{n \geq 1} \frac{z^n}{n^2}$, on a $R = 1$ aussi, parce que l'ordre de grandeur de n^2 est négligeable face à une exponentielle. Le domaine de convergence inclut le bord du disque, et est donc $\mathcal{D}_f(0, 1)$

$\sum_{n \geq 1} \frac{z^n}{n}$ a un rayon $R = 1$. La série converge pour $z = -1$ et diverge pour $z = 1$, les deux appartenant au bord du disque.

II Calculs de rayon de convergence

Proposition :

Le rayon de $\sum a_n z^n$ est le rayon de $\sum |a_n z^n|$

Remarque II.1. Les déterminations de rayon d'une série entière ne font intervenir que des critères de séries numériques positives.

Proposition : Relations de comparaison

Soient $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières de rayons R_a et R_b :

- Si $a_n = o(b_n)$ alors $R_a \geq R_b$
- Si $a_n = \mathcal{O}(b_n)$ alors $R_a \geq R_b$
- Si $a_n \sim b_n$ alors $R_a = R_b$

Preuve

Soit $z \in \mathcal{D}(0, R_b)$, on a donc que $\sum |b_n z^n|$ converge, or $|a_n z^n| = \mathcal{O}(|b_n z^n|)$

Donc par critère de domination positive, $\sum |a_n z^n|$ converge

Donc $z \in \mathcal{D}_f(0, R_a)$

Donc $\mathcal{D}(0, R_b) \subset \mathcal{D}_f(0, R_a)$

Donc $R_a \geq R_b$

On a donc prouvé la propriété pour les \mathcal{O} , et elle en découle pour les o . Comme un équivalent est aussi un \mathcal{O} symétrique, on a l'égalité.

Théorème : Critère de D'Alembert

Si (a_n) ne s'annule pas à partir d'un certain rang :

si $\left| \frac{a_{n+1}}{a_n} \right| \rightarrow l$ avec $l \in \mathbb{R}$, alors $R_a = \frac{1}{l}$

On prend la convention de $\frac{1}{+\infty} = 0$ et que $\frac{1}{0} = +\infty$

Preuve

On suppose que $\left| \frac{a_{n+1}}{a_n} \right|$ tend vers $l \in \mathbb{R}$

Soit $z \in \mathbb{C}^*$, alors $\left| \frac{a_{n+1}z^{n+1}}{a_n z^n} \right| \rightarrow lz$

Par critère de D'Alembert appliqué aux séries numériques, si $|lz| < 1$ alors $\sum a_n z^n$ converge et si $|lz| > 1$ alors $\sum a_n z^n$ diverge.

Donc $R_a = \frac{1}{l}$.

Remarque II.2. Il faut faire attention aux séries lacunaires pour utiliser d'Alembert.

Exemple II.1. Déterminer le rayon de $\sum \frac{(-1)^n(e^n+n)}{3^n+n^2} z^{2n}$

La suite associée à la série entière est équivalente à $\left(\frac{e}{3}\right)^n$, qu'on note $u_n(z)$

Soit $z \in \mathbb{C}^*$:

$$\frac{\left| \left(\frac{e}{3}\right)^{n+1} z^{2n+2} \right|}{\left| \left(\frac{e}{3}\right)^n z^{2n} \right|} = \left| \frac{e}{3} z^2 \right| \rightarrow \frac{e}{3} z^2$$

$$\text{Donc } R = \sqrt{\frac{3}{e}}$$

Donc si $|z| < \sqrt{\frac{3}{e}}$, $\sum |u_n(z)|$ converge et diverge sinon.

Exemple II.2. Pour la série entière associée à la suite $a_n = \begin{cases} a_{n^2} = n! \\ a_p = 0 \text{ si } p \text{ n'est pas un carré} \end{cases}$

Exemple II.3. Pour la série entière $\sum \sin(n) z^n$

III Opérations sur les séries entières

Proposition : Combinaisons linéaires

Si $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons respectifs R_a et R_b , alors :

$\sum (a_n + b_n) z^n$ converge de rayon R avec $R \geq \min(R_a, R_b)$

Si $R_a \neq R_b$, on peut affirmer que $R = \min(R_a, R_b)$

De plus, pour $z \in \mathcal{D}(0, R)$, on a $\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$

Preuve

Avec les notations de l'énoncé :

Supposons que $R_a = R_b$:

Alors

Supposons que $R_a \neq R_b$, supposons sans perte de généralité que $R_a < R_b$:

Si $z \in \mathcal{D}(0, R_a)$, alors $\sum (a_n + b_n) z^n$ est bien définie, sinon non.

Théorème : Produit de Cauchy

Pour $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons respectifs R_a et R_b

On pose pour $n \in \mathbb{N}$, $c_n = \sum_{k=0}^n a_k b_{n-k}$

La série entière $\sum c_n z^n$ est de rayon de convergence $R \geq \min(R_a, R_b)$ et $\forall z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$, $\sum_{n=0}^{+\infty} c_n z^n = \sum_{n=0}^{+\infty} a_n z^n \sum_{n=0}^{+\infty} b_n z^n$

Preuve

Par application du produit de Cauchy à deux séries absolument convergentes sur un domaine.

IV Continuité de la somme d'une série entière sur les complexes

Proposition :

Soit $(\sum a_n z^n)$ une série entière de rayon $r > 0$, alors $(\sum a_n z^n)$ converge normalement sur tout disque fermé $\mathcal{B}(0, A)$ pour $A < R$

Preuve**Proposition : Continuité**

La somme d'une série entière est continue sur son disque ouvert de convergence

Preuve**Proposition : Somme de série entière sur \mathbb{R}**

Ici, considérons la restriction à \mathbb{R} de la somme des séries entières considérées.

Alors le domaine de définition dans \mathbb{R} , noté \mathcal{D} est l'un des intervalles : $[-R, R],]-R, R[, [-R, R[,]-R, R[$

Exemple IV.1.**Proposition :**

Soit $(a_n) \in \mathbb{C}^{\mathbb{N}}$, alors les séries entières $\sum a_n z^n$ et $\sum n a_n z^n$ ont même rayon.

Preuve

Proposition : Corollaire

Soit $(a_n) \in \mathbb{C}^{\mathbb{N}}$ ($\sum a_n z^n$) est de rayon $R > 0$

Alors $\varphi : \begin{cases}]-R, R[\rightarrow \mathbb{C} \\ x \mapsto \sum_{n=0}^{+\infty} a_n x^n \end{cases}$ est \mathcal{C}^1 et :

$\varphi' : \begin{cases}]-R, R[\rightarrow \mathbb{C} \\ x \mapsto \sum_{n=0}^{+\infty} n a_n x^{n-1} \end{cases}$

Preuve**Proposition : Corollaire**

La somme d'une série entière est \mathcal{C}^∞ sur son intervalle ouvert de convergence
Les dérivées s'obtiennent par dérivation terme à terme

Preuve

Par récurrence

Remarque IV.1. Sur la valeur de dérivée :

Proposition : Corollaire

Deux séries entières dont les sommes coïncident sur un voisinage de 0 sont identiques

i.e. si $\exists \alpha \in \mathbb{R}_+^*, \forall x \in]-\alpha, \alpha[, \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} b_n x^n$, alors $\forall n \in \mathbb{N}, a_n = b_n$

Théorème : Convergence radiale

Soit $(a_n) \in \mathbb{C}^{\mathbb{N}}$, ($\sum a_n z^n$) de rayon $R > 0$

Si $\sum a_n R^n$ converge alors : $\sum_{n=0}^{+\infty} a_n x^n \rightarrow_{x \rightarrow R} \sum_{n=0}^{+\infty} a_n R^n$ pour $x \in]-R, R[$

Plus précisément, si f définie en R en tant que somme de série entière, alors f continue sur \mathcal{D}_f

Remarque IV.2. Le théorème reste valable pour $-R$: par considération de $\sum a_n (-1)^n R^n$ qui est de même rayon.

V Développement en séries entières

V.1 Généralités

Définition : Développement en série entière

Soit I un intervalle tel que $0 \in \overset{\circ}{I}$. Soit $f \in \mathcal{F}(I, \mathbb{C})$.

On dit que f est développable en série entière au voisinage de 0 (ou en 0) si :

$\exists \alpha \in \mathbb{R}_+^*, \exists (a_n) \in \mathbb{C}^{\mathbb{N}}, \forall x \in]-\alpha, \alpha[, f(x) = \sum_{n=0}^{+\infty} a_n x^n$

Proposition :

- Si f est développable en série entière (DSE) en 0, alors f est \mathcal{C}^∞ sur un voisinage de 0.
- Si f est développable en série entière, son développement est unique, et donc par unicité du DL on a :

$$\exists \alpha \in \mathbb{R}_+^*, \forall x \in]-\alpha, \alpha[, f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$$
- Si f est \mathcal{C}^∞ sur I : on peut écrire que $\forall x \in I, \forall n \in \mathbb{N}, f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + R_n(x)$
- f est développable en série entière en 0 s'il existe $\alpha \in \mathbb{R}_+^*$ tel que R_n converge simplement vers 0 sur $] -\alpha, \alpha[$

V.2 Rappels sur Taylor**Théorème : Taylor-Reste-Intégral**

Si f est \mathcal{C}^{n+1} sur $[a, b]$, alors :

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + R_n(b)$$

Où $R_n(b) = \int_a^b \frac{(t-a)^n}{n!} f^{(n+1)}(t) dt$

Théorème : Taylor-Lagrange

En posant $t = a + (b-a)u$, on transforme l'expression du reste en $\int_0^1 \frac{(b-a)^{n+1}(1-u)^n}{n!} f^{(n+1)}(a + (b-a)u) du$

On a donc : $R_n(b) = \frac{(b-a)^{n+1}}{n!} \int_0^1 (1-u)^n f^{(n+1)}(a + (b-a)u) du$

De là on déduit la majoration de Lagrange, en posant $M_{n+1} = \sup_{[a,b]} |f^{(n+1)}|$

$$|R_n(b)| \leq \frac{|b-a|^{n+1}}{n!} \int_0^1 |1-u|^n |f^{(n+1)}(a + (b-a)u)| du \leq \frac{|b-a|^{n+1}}{n!} M_{n+1} \int_0^1 (1-u)^n du \leq \frac{|b-a|^{n+1}}{(n+1)!} M_{n+1}$$

V.3 Développements en série entière découlant de l'exponentielle sur les réels

$$\forall x \in \mathbb{R}, \exp(x) = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

Preuve

Pour $x \in \mathbb{R}$, on a que $\exp(x) = \sum_{k=0}^n \frac{x^k}{k!} + R_n(x)$ par Taylor-Reste-Intégral.

$$\text{Alors } |R_n(x)| \leq \frac{|x|^{n+1}}{n!} e^{|x|}$$

$$\text{Or } \left(\frac{|x|^{n+1}}{n!} e^{|x|} \right) \rightarrow 0$$

$$\text{Et donc } R_n(x) \rightarrow 0$$

Et donc, pour tout $x \in \mathbb{R}$, on a les développements suivants :

$$\text{— } \cosh(x) = \frac{1}{2}(e^x + e^{-x}) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$$

$$\text{— } \sinh(x) = \frac{1}{2}(e^x - e^{-x}) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$$

$$\text{— } \cos(x) = \Re(e^{ix}) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

$$— \sin(x) = \Im(e^{ix}) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

V.4 Développements en série entière découlant de la série géométrique

On a, $\forall x \in]-1, 1[$:

$$— \frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$$

$$— \frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n$$

— $-\ln(1-x) = \sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1}$ par primitivation d'un DL. Cette égalité peut être prolongée en -1 par théorème de convergence radiale.

— $\ln(1+x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{n+1}}{n+1} = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$. Cette égalité peut être prolongée en 1 par théorème de convergence radiale.

Cette année est aussi au programme le développement d'Arctan :

$$\forall x \in]-1, 1[, \arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n x^{2n}$$

$$\text{Et donc : } \arctan(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

V.5 Développement de l'autre là

En notant $f_\alpha : x \mapsto (1+x)^\alpha$, on a :

$$f_\alpha(x) = 1 + \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} x^k$$

Preuve

f_α est \mathcal{C}^∞ sur $] -1, 1[$. De plus, $\forall x \in] -1, 1[, f'_\alpha(x) = \alpha(1+x)^{\alpha-1}$

Donc f_α est solution de l'équation différentielle $(1+x)y' = \alpha y : (E)$

C'est une équation différentielle linéaire d'ordre 1 résoluble, il y a donc unicité de la solution de cette équation vérifiant une condition initiale.

Et donc f_α est l'unique solution de cette équation vérifiant $f_\alpha(0) = 1$

Déterminons la suite de sa décomposition en série entière de f_α par analyse-synthèse.

Analyse : Soit $(a_n) \in \mathbb{R}^\mathbb{N}$, et on suppose que $\sum a_n x^n$ est une série entière de rayon $R > 0$. On pose

pour $x \in] -R, R[$, on pose $f(x) = \sum_{n=0}^{+\infty} a_n x^n$

On a alors que f est \mathcal{C}^∞ sur $] -R, R[$ et que $\forall x \in] -R, R[, f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$

Soit $x \in] -R, R[$:

$$\begin{aligned} (1-x)f'(x) - \alpha f(x) &= (1-x) \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n - \alpha \sum_{n=0}^{+\infty} a_n x^n \\ &= \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^{n+1} + \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n - \alpha \sum_{n=0}^{+\infty} a_n x^n \\ &= \sum_{n=1}^{+\infty} n a_n x^n + \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n - \alpha \sum_{n=0}^{+\infty} a_n x^n \quad (\text{on peut mettre 0 comme indice de départ du premier terme puisque } n a_n = 0 \text{ si } n = 0) \\ &= \sum_{n=0}^{+\infty} [(n+1) a_{n+1} + (n-\alpha) a_n] x^n \end{aligned}$$

Et donc f est solution de (E) si, et seulement si :

$$\forall x \in] -R, R[, \sum_{n=0}^{+\infty} [(n+1) a_{n+1} + (n-\alpha) a_n] x^n = 0$$

Et par unicité du développement, comme 0 a comme développement la suite nulle, on déduit que f est solution de (E) si, et seulement si :

$$\forall n \in \mathbb{N}, (n+1) a_{n+1} + (n-\alpha) a_n = 0$$

Synthèse : considérons (a_n) définie par $\begin{cases} a_0 = 1 \\ a_{n+1} = \frac{\alpha-n}{n+1} a_n \end{cases}$

Si $\alpha \in \mathbb{N}$: alors $\forall n \in \mathbb{N}, n \geq \alpha \Rightarrow a_n = 0$ et donc $\sum a_n x^n$ est de rayon infini

Si $\alpha \notin \mathbb{N}, \forall n \in \mathbb{N}, a_n \neq 0$ et $\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\alpha-n}{n+1} \right| \rightarrow 1$

Et dans tous les cas, on a bien que pour $x \in] -1, 1[, f(x) = \sum_{n=0}^{+\infty} a_n x^n$

Les calculs précédents sont valides sur $] -1, 1[, f$ est solution de (E) sur $] -1, 1[$ et $f(0) = 1$

Donc par unicité de la solution au problème de Cauchy :

$$\forall x \in] -1, 1[, f_\alpha(x) = \sum_{n=0}^{+\infty} a_n x^n$$

$$\text{Et par récurrence, } \begin{cases} a_n = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} \\ a_0 = 1 \end{cases}$$

Il y a certains cas particuliers à savoir refaire facilement, même si pas à savoir par coeur.

Pour $(1-x)^{\frac{1}{2}}$, on a :

$$\begin{aligned} (1-x)^{\frac{1}{2}} &= 1 + \sum_{n=1}^{+\infty} \frac{-\frac{1}{2}(-\frac{1}{2}-1)\dots(-\frac{1}{2}-n+1)}{n!} (-x)^n \\ &= 1 + \sum_{n=1}^{+\infty} (-1)^n \frac{\frac{1}{2}(\frac{1}{2}+1)\dots(\frac{1}{2}+n-1)}{n!} (-1)^n x^n \\ &= 1 + \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2^n (n!)} x^n \\ &= 1 + \sum_{n=1}^{+\infty} \frac{(2n)!}{2^{2n} (n!)^2} x^n = 1 + \sum_{n=1}^{+\infty} \frac{1}{4^n} \binom{2n}{n} x^n \end{aligned}$$

$$= \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} x^n$$

Et on en déduit, pour $x \in]-1, 1[$:

$$\text{--- } \arcsin'(x) = \frac{1}{1-x^2} = \sum_{n=1}^{+\infty} \frac{1}{4^n} \binom{2n}{n} x^{2n}$$

$$\text{--- } \arcsin(x) = \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} \frac{x^{2n+1}}{2n+1}$$

$$\text{--- } \arccos(x) = \frac{\pi}{2} - \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} \frac{x^{2n+1}}{2n+1}$$

VI Séries entières dans une algèbre de dimension finie

Définition :

Soit E une \mathbb{K} -algèbre de dimension finie munie d'une norme d'algèbre $\|\cdot\|$

Soit $(a_n) \in \mathbb{K}^{\mathbb{N}}$, $\sum a_n z^n$ est de rayon $R > 0$

Alors on peut définir $\sum_{n=0}^{+\infty} a_n u^n$ pour $u \in E$ si $\|u\| < R$.

Preuve

Soit $u \in E$

Pour $n \in \mathbb{N}$, $\|a_n u^n\| \leq |a_n| \|u\|^n$ (comme $\|\cdot\|$ est une norme d'algèbre)

Pour $u \in \mathcal{B}(0, R)$, $\sum |a_n| \|u\|^n$ converge

Donc $\sum a_n u^n$ converge absolument donc converge car E est de dimension finie.

Proposition :

$$\text{Et } \begin{cases} \mathcal{B}(0, R) \rightarrow E \\ u \mapsto \sum_{n=0}^{+\infty} a_n u^n \end{cases} \text{ est continue}$$

Preuve

Soit R' tel que $0 < R' < R$

$\forall n \in \mathbb{N}, \forall u \in \mathcal{B}(0, R'), \|a_n u^n\| \leq |a_n| R'^n$

Or $\sum |a_n| R'^n$ converge normalement sur $\mathcal{B}_f(0, R')$

Soit $n \in \mathbb{N}$, $u \mapsto u^n$ est continue sur

(car $E^n \mapsto E, (u_1, \dots, u_n) \mapsto u_1 \dots u_n$ est n -linéaire en dimension finie)

Donc d'après le théorème de continuité par convergence uniforme (valable en dimension finie), $u \mapsto$

$\sum_{n=0}^{+\infty} a_n u^n$ est continue sur $\mathcal{B}(0, R')$ pour tout $R' < R$ donc sur $\mathcal{B}(0, R)$

Proposition :

$$\forall u \in E, \exists P_u \in \mathbb{K}[X], \sum_{n=0}^{+\infty} a_n u^n = P_u(u)$$

Preuve

Pour $u \in E$, $\mathbb{K}[u]$ est une sous-algèbre de dimension finie donc fermée, or $\sum_{n=0}^{+\infty} a_n u^n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n a_k u^k$

Donc $\sum_{n=0}^{+\infty} a_n u^n$ est la limite d'une suite de $\mathbb{K}[u]$

Donc $\sum_{n=0}^{+\infty} a_n u^n \in \mathbb{K}[u]$

Remarque VI.1. En particulier, $\left\{ \begin{array}{l} \mathcal{M}_n(\mathbb{K}) \rightarrow \mathcal{GL}_n(\mathbb{K}) \\ A \mapsto \exp(A) = \sum_{n=0}^{+\infty} \frac{1}{n!} A^n \end{array} \right.$ est continue, bien définie, de rayon $+\infty$ et est continue.

Exemple VI.1. $\left\{ \begin{array}{l} \mathcal{B}(0, 1) \rightarrow \mathcal{GL}_n(\mathbb{K}) \\ A \mapsto \sum_{n=0}^{+\infty} \frac{1}{\sqrt{n^2+n+1}} A^n \end{array} \right.$ est définie par exemple sur $\mathcal{B}(0, 1)$ pour une norme d'algèbre, et est continue

Exemple VI.2. Pour $A \in \mathcal{M}_n(\mathbb{K})$, on définit $\varphi_A : \left\{ \begin{array}{l} \mathbb{R} \rightarrow \mathcal{M}_n(\mathbb{K}) \\ t \mapsto \exp(tA) = \sum_{n=0}^{+\infty} \frac{1}{n!} A^n \end{array} \right.$

Montrer que φ_A est \mathcal{C}^1 et exprimer φ'_A

$\sum t \mapsto \frac{t^n}{n!} A^n$ converge simplement sur \mathbb{R}

Pour $n \in \mathbb{N}$, $t \mapsto \frac{t^n}{n!} A^n$ est \mathcal{C}^1 de dérivée $t \mapsto \frac{t^{n-1}}{(n-1)!} A^n$ si $n > 0$ et 0 sinon

Soit $\alpha \in \mathbb{R}_+^*$, $t \in [-\alpha, \alpha]$

Soit $n \in \mathbb{N}^*$, alors : $\left\| \frac{t^{n-1}}{(n-1)!} A^n \right\| \leq \frac{\alpha^{n-1}}{(n-1)!} \|A\|^n$

Et $\sum \frac{\alpha^{n-1}}{(n-1)!} \|A\|^n$ converge

Donc la série $\sum t \mapsto \frac{t^{n-1}}{(n-1)!} A^n$ converge normalement donc uniformément sur $[-\alpha, \alpha]$

Donc φ_A est \mathcal{C}^1 sur $] -\alpha, \alpha[$ pour $\alpha \in \mathbb{R}_+^*$ donc sur \mathbb{R}

Et $\forall t \in \mathbb{R}$, $\varphi'_A(t) = \sum_{n=1}^{+\infty} \frac{t^{n-1}}{(n-1)!} A^n = A \left(\sum_{n=0}^{+\infty} \frac{t^n}{n!} A^n \right) = A \exp(tA)$

Application : Pour $A \in \mathcal{M}_n(\mathbb{K})$, trouver des solutions de $Y' - AY = 0$ dans $\mathcal{F}(\mathbb{R}, \mathcal{M}_{n,1}(\mathbb{K}))$

C'est une équation différentielle linéaire d'ordre 1 à coefficients constants, qui correspond à un système :

$$\begin{cases} y'_1 &= a_{1,1}y_1 + a_{1,2}y_2 + \dots + a_{1,n}y_n \\ y'_2 &= a_{2,1}y_1 + a_{2,2}y_2 + \dots + a_{2,n}y_n \\ &\dots \\ y'_n &= a_{n,1}y_1 + \dots + a_{n,n}y_n \end{cases}$$

On a que $\varphi : t \mapsto e^{tA} Y_0$ est une solution de l'équation vérifiant $\varphi(0) = Y_0$

On démontrera plus tard que c'est l'unique solution.

Chapitre VII

Arithmétique dans un anneau euclidien

I Généralités

Définition : Anneau euclidien

A est un anneau euclidien si :

- A est un anneau commutatif intègre ($AB = 0 \Rightarrow A = 0$ ou $B = 0$)
- A est muni d'une division euclidienne : il existe $\varphi \in \mathcal{F}(A \setminus \{0\}, \mathbb{N})$ telle que $\forall b \in A \setminus \{0\}, \forall a \in A, \exists q, r \in A, a = bq + r$ et ($r = 0$ ou $\varphi(r) < \varphi(b)$)

\mathbb{Z} et $\mathbb{K}[X]$ sont des anneaux euclidiens

Dans \mathbb{Z} , $\forall a \in \mathbb{Z}, \forall b \in \mathbb{Z}^*, \exists!(q, r) \in \mathbb{Z}^2, a = bq + r, 0 \leq r < |b|$

On peut prendre $\varphi : a \mapsto |a|$

Dans $\mathbb{K}[X]$, $\forall A \in \mathbb{K}[X], \forall B \in \mathbb{K}[X] \setminus \{0\}, \exists!(Q, R) \in \mathbb{K}[X]^2, A = BQ + R$ et $\deg R < \deg B$

On peut prendre $\varphi : P \mapsto \deg P$

Définition : Divisibilité

Soit A un anneau euclidien (ou intègre). Soit $a, b \in A$.

On dit que b divise a (noté $b|a$) ou que a est un multiple de b si $\exists c \in A, a = bc$

Exemple I.1. *Tout élément de A divise 0*

Tout inversible de A est un diviseur universel : si u est inversible, $\forall a \in A, a = uu^{-1}a$

Dans \mathbb{Z} , $\{-1, 1\}$ sont les diviseurs universels

Dans $\mathbb{K}[X]$, tout $a \in \mathbb{K}^$ est un diviseur universel*

Proposition : Association

Soit A un anneau euclidien

Soit $a, b \in A$, a divise b et b divise a si, et seulement si, il existe un inversible tel que $b = au$

Preuve

Le sens indirect est immédiat.

Pour le sens direct, on suppose que $a|b$ et $b|a$

On a donc $c \in A$ tel que $b = ca$

On a donc $d \in A$ tel que $a = db$

Donc $b = cdb$

Si $b = 0$: $a = 0$ et la propriété est vérifiée

Si $b \neq 0$: on peut simplifier par b , comme A est intègre : $cd = 1$ donc c et d sont inversibles et inverses l'un de l'autre, ce qui conclut.

Remarque I.1. La relation divise est presque une relation d'ordre : elle est réflexive, transitive et "presque" antisymétrique.

Définition : Idéal

Soit A un anneau euclidien. Soit $I \subset A$.

On dit que I est un idéal si :

- I est un sous-groupe de $(A, +)$
- $\forall i \in I, \forall a \in A, ia \in I$

Exemple I.2. Soit $a \in A$, alors $aA = \{ax, x \in A\}$ est l'ensemble des multiples de a et est un idéal. C'est même le plus petit idéal qui contient a , l'intersection de tous les idéaux contenant a .

II Opérations sur les idéaux

Proposition :

Avec J un ensemble quelconque, $(I_j)_{j \in J}$ une famille d'idéaux de A , alors $\cap_{j \in J} I_j$ est un idéal

Preuve

Faisons la stabilité par somme :

si $x, y \in \cap_{j \in J} I_j$, alors $\forall j \in J, x, y \in I_j$ et par nature d'idéal, $\forall j \in J, x + y \in I_j$

Donc $x + y \in \cap_{j \in J} I_j$

Faisons la stabilité par produit :

Si $x \in \cap_{j \in J} I_j$, si $a \in A$

Alors $\forall j \in J, x \in I_j$, et par nature d'idéal, $\forall j \in J, ax \in I_j$

Donc $ax \in \cap_{j \in J} I_j$

Remarque II.1. Une union d'idéaux non-inclus dans les autres n'est pas un idéal, de même que pour des groupes.

En effet, avec I, J deux groupes non-inclus l'un dans l'autre, on prend $x \in I, x \notin J, y \in J, y \notin I$

Cependant, par propriétés d'un groupe, $x + y \in I \cup J$, supposons sans perte de généralité que $x + y \in J$

Et donc $(x + y) - y \in J$

Donc $x \in J$, d'où la contradiction

Proposition :

Si I_1 et I_2 sont des idéaux, alors $I_1 + I_2$ est un idéal.

C'est le plus petit idéal contenant $I_1 \cup I_2$.

Preuve

Montrons que $I_1 + I_2$ est un idéal :

Soient $x, y \in I_1 + I_2$, on note $x = x_1 + x_2, y = y_1 + y_2$ avec $x_1, y_1 \in I_1, x_2, y_2 \in I_2$

Alors $x + y = (x_1 + y_1) + (x_2 + y_2) \in I_1 + I_2$

Donc $I_1 + I_2$ est stable par l'addition.

Soit $x = x_1 + x_2 \in I_1 + I_2, a \in A$. Alors $ax = ax_1 + ax_2 \in I_1 + I_2$

De plus, tout idéal contenant I_1 et I_2 contient $I_1 + I_2$

Théorème :

Si A est un anneau euclidien, alors tout idéal de A est principal.

Cela signifie que si I est un idéal de A , alors $\exists a \in A, I = aA$

Preuve

On dispose de $\varphi \in \mathcal{F}(A \setminus \{0\}, \mathbb{N})$ tel que $\forall a \in A, \forall b \in A \setminus \{0\}, \exists (q, r) \in A^2, a = bq + r$ et ($r = 0$ ou $\varphi(r) < \varphi(b)$)

Si $I = \{0\}$, c'est l'idéal $0A$

Si I est un idéal non-réduit à 0 :

On peut considérer $n_0 = \min\{\varphi(i) \mid i \in I \setminus \{0\}\}$, qui existe puisqu'on considère une partie non-vide de \mathbb{N}

On note i_0 tel que $\varphi(i_0) = n_0$

Soit $i \in I$, on peut alors prendre $q, r \in A$ telq que $i = qi_0 + r$ et ($r = 0$ ou $\varphi(r) < \varphi(b)$)

$i - qi_0 \in I$, et donc $r \in I$. $r = 0$, si ce n'était pas le cas alors ça contredirait la définition de n_0 .

Donc $i \in i_0A$

Donc $I \subset i_0A$

$i_0A \subset I$ est immédiat puisque $i_0 \in I$

Remarque II.2. Deux générateurs d'un même idéal sont égaux à produit par un inversible près. Pour rendre le générateur unique, on fait un choix :

- Dans \mathbb{Z} , on prend en général le générateur positif.
- Dans $\mathbb{K}[X]$, on prend en général le générateur unitaire.

Exemple II.1. Avec E un \mathbb{K} -ev de dimension finie, $f \in \mathcal{L}(E), x \in E$:

$I_f = \{P \in \mathbb{K}[X], P(f) = 0\}$ est un idéal de $\mathbb{K}[X]$ (ce qui se prouve avec les formules sur $(P + Q)(f)$ et $(PQ)(f)$)

π_f est le générateur unitaire de cet idéal.

$I_{f,x} = \{P \in \mathbb{K}[X], P(f)(x) = 0\}$ est un idéal de $\mathbb{K}[X]$

On note son générateur $\pi_{f,x}$

On a que $\pi_f \in I_{f,x}$, donc $\pi_{f,x} = \pi_f$

III PGCD et PPCM

Définition : PGCD

Avec A un anneau euclidien, soit $a, b \in A$

On appelle $\text{pgcd}(a, b)$ un générateur de $aA + bA$

Dans \mathbb{Z} ou $\mathbb{K}[X]$, on choisit un générateur spécifique, noté $a \wedge b$, et on l'appelle le pgcd.

$aA + bA = (a \wedge b)A$

Proposition :

Soit $(a, b) \in A^2$, soit $d \in A$. d est un diviseur de a et b si, et seulement si, d divise $a \wedge b$

Preuve

$$\begin{cases} d|a \\ d|b \end{cases} \Leftrightarrow \begin{cases} aA \subset dA \\ bA \subset dA \end{cases} \Leftrightarrow aA + bA \subset dA \Leftrightarrow (a \wedge b)A \subset dA \Leftrightarrow d|a \wedge b$$

Définition :

Soit $a, b \in A$, on dit que a et b sont premiers entre eux si $a \wedge b = 1$

Théorème : Bezout

$a, b \in A$ sont premiers entre eux si, et seulement si, $\exists(u, v) \in A^2, au + bv = 1$

Preuve

S'il existe $(u, v) \in A^2, au + bv = 1$

Alors $1 \in aA + bA$

Donc $aA + bA = A$

Donc $a \wedge b = 1$

Si $a \wedge b = 1$:

Alors $1 \in aA + bA$

Donc $\exists(u, v) \in A^2, au + bv = 1$

Théorème : Bezout étendu

$$\forall a, b \in A, \forall x \in A : \exists u, v \in A, x = au + bv \Leftrightarrow (a \wedge b)|x$$

Définition :

Soit A un anneau euclidien. Soit $a_1, \dots, a_n \in A$

On appelle *pgcd* de a_1, \dots, a_n le générateur de $a_1A + a_2A + \dots + a_nA$

On note que $(a_1 \wedge a_2 \wedge \dots \wedge a_n)A = a_1A + a_2A + \dots + a_nA$

Proposition :

$$\forall a_1, \dots, a_n \in A, \forall d \in A, [\forall i \in \llbracket 1, n \rrbracket, d|a_i] \Leftrightarrow d|(a_1 \wedge a_2 \wedge \dots \wedge a_n)$$

Proposition :

$$\forall a, b, c \in A, (a \wedge b) \wedge c = a \wedge (b \wedge c)$$

Définition :

Avec $(a_1, a_2, \dots, a_n) \in A^n$, on dit que (a_1, a_2, \dots, a_n) sont premiers entre eux si $(a_1 \wedge a_2 \wedge \dots \wedge a_n) = 1$

Remarque III.1. Attention à ne pas confondre premiers deux-à-deux et premiers entre eux dans leur ensemble. Par exemple, $(6, 15, 35)$ sont premiers dans leur ensemble mais pas premiers deux-à-deux.

Théorème : Bézout généralisé

$(a_1, a_2, \dots, a_n) \in A^n$ sont premiers entre eux si, et seulement si, $\exists (u_1, \dots, u_n) \in A^n, 1 = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$

Théorème :

Tout élément de \mathbb{Z} et de $\mathbb{K}[X]$ peut se décomposer en un produit unique d'éléments irréductibles (éléments qui, à un inversible près, ont un seul diviseur) et d'un inversible.

Preuve

Par récurrence forte sur la valeur absolue ou le degré pour l'existence, et par Lemme de Gauss pour l'unicité. Revoir le cours de sup.

Définition : PPCM

Soit A un anneau euclidien, soit $a, b \in A$. On appelle ppcm de a et b un générateur de $aA \cap bA$, noté $a \vee b$

Dans \mathbb{Z} ou $\mathbb{K}[X]$, on peut parler du ppcm pour l'entier positif ou polynôme unitaire ppcm.

Et donc on a : $aA \cap bA = (a \vee b)A$

Proposition :

Soit $a, b \in A$, soit $p \in A$

p est un multiple de a et de b si, et seulement si, p est un multiple du ppcm.

Preuve

Dans le cadre de l'énoncé :

$$\begin{cases} a|p \\ b|p \end{cases} \Leftrightarrow \begin{cases} pA \subset aA \\ pA \subset bA \end{cases} \Leftrightarrow pA \subset aA \cap bA \Leftrightarrow pA \subset (a \vee b)A \Leftrightarrow a \vee b | p$$

Proposition : ppcm de n éléments

Avec $a_1, \dots, a_n \in A$, alors $a_1 \vee a_2 \vee \dots \vee a_n A$ est bien un générateur de $a_1 A \cap a_2 A \cap \dots \cap a_n A$

Proposition :

$$\forall a, b, c \in A, (a \vee b) \vee c = a \vee (b \vee c)$$

Proposition :

$$\text{Si } a, b, d \in A : \begin{cases} a = da' \\ b = db' \end{cases} \text{ et } a' \wedge b' = 1 \Leftrightarrow d \equiv a \wedge b$$

Preuve

Sens direct : on suppose que $d|a$ et $d|b$.

a' et b' sont premiers entre eux, on peut donc prendre $u, v \in A$ tels que $a'u + b'v = 1$

Donc $da'u + db'v = au + bv = d$

Donc $d \in aA + bA$

Donc $a \wedge b | d$

Et donc $d \equiv a \wedge b$

Pour l'autre implication : on peut écrire $a = da'$ et $b = db'$. Par l'absurde, si a' et b' ne sont pas premiers, alors d ne serait pas congru au pgcd.

IV Application à la réduction des endomorphismes

Théorème : Lemme des noyaux

Soit E un \mathbb{K} -ev, $f \in \mathcal{L}(E)$.

Soit $P \in \mathbb{K}[X]$ tel que $P = P_1 \dots P_n$ où P_1, \dots, P_n sont premiers entre eux deux à deux.

$$\ker P(f) = \bigoplus_{i=1}^n \ker(P_i(f))$$

Preuve

Dans le cadre de l'énoncé :

On pose pour $i \in \llbracket 1, n \rrbracket$: $Q_i = \prod_{1 \leq j \leq n, j \neq i} P_j = \frac{P}{P_i}$

Montrons que (Q_1, \dots, Q_n) sont premiers entre eux dans leur ensemble.

Supposons par l'absurde que Q_1, \dots, Q_n ne sont pas premiers entre eux.

Donc $Q_1 \wedge \dots \wedge Q_n \neq 1$

Donc $Q_1 \wedge \dots \wedge Q_n$ possèdent un diviseur irréductible Q

En particulier, $Q|Q_1$

Donc $Q|P_2 P_3 \dots P_n$

Comme Q est irréductible, $\exists i_0 \in \llbracket 2, n \rrbracket Q|P_{i_0}$

Mais $Q|Q_{i_0}$, donc $Q|\prod_{j \neq i_0} P_j$

Donc $Q|P_{j_0}$ avec $j_0 \neq i_0$

Mais $P_{j_0} \wedge P_{i_0}$

D'où la contradiction.

Par Bezout généralisé, on peut trouver $U_1, \dots, U_n \in \mathbb{K}[X]$ tels que $\sum_{i=1}^n Q_i U_i = 1$

On a que $\forall i \in \llbracket 1, n \rrbracket, \ker P_i(f) \subset \ker P(f)$ donc $\ker P_1(f) + \ker P_2(f) + \dots + \ker P_n(f) \subset \ker P(f)$

Analyse : soit $x \in \ker P(f)$, on suppose qu'il existe $(x_1, \dots, x_n) \in \ker P_1(f) \times \dots \times \ker P_n(f)$ tel que $x = x_1 + \dots + x_n$

Soit $i \in \llbracket 1, n \rrbracket$, alors $Q_i(f)(x) = Q_i(f)(x_i)$

$U_i(f) \circ Q_i(f)(x) = (U_i(f) \circ Q_i(f))(x_i)$

Remarque : $\left(\sum_{j=1}^n U_j(f) \circ Q_j(f) \right) (x_i) = (U_i(f) \circ Q_i(f))(x_i)$

Or $\sum_{j=1}^n U_j Q_j = 1$

Donc $\sum_{j=1}^n U_j(f) Q_j(f) = id$

Donc $x_i = Q_i(f) \circ U_i(f)(x_i)$

Donc $x_i = (U_i Q_i)(f)(x)$

Donc $\ker P_1(f), \dots, \ker P_n(f)$ sont en somme directe.

Synthèse : soit $x \in \ker P(f)$

Alors $x = id(x) = \sum_{i=1}^n (Q_i U_i)(f)(x)$

Et $\forall i \in \llbracket 1, n \rrbracket, P_i(f)((Q_i U_i)(f)(x)) = (P_i Q_i U_i)(f)(x) = (U_i P)(f)(x) = U_i(f) \circ P(f)(x) = 0$

Exemple IV.1. E un \mathbb{K} -ev, $f \in \mathcal{L}(E)$ telle que $f^3 - id = 0$

Si $\mathbb{K} = \mathbb{C}$: $E = \ker(f - id) \oplus \ker(f - jid) \oplus \ker(f - j^2 id)$ puisque $X^3 - 1 = (X - 1)(X - j)(X - j^2)$

Si $\mathbb{K} = R$, on peut écrire par le lemme des noyaux : $E = \ker(f - id) \oplus \ker(f^2 + f - id)$

Définition : Espaces caractéristiques

Soit E un \mathbb{K} -ev de dimension finie, soit $f \in \mathcal{L}(E)$ telle que χ_f est scindé.

$$\chi_f = \prod_{i=1}^p (x - \lambda_i)^{\alpha_i} \text{ avec } S_p(f) = \{\lambda_1, \dots, \lambda_p\} \text{ et } (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^{*p}$$

Alors $E = \bigoplus_{1 \leq i \leq p} \ker((f - \lambda_i \text{id})^{\alpha_i})$ par lemme des noyaux.

Les sev de E , $F_i = \ker((f - \lambda_i)^{\alpha_i})$ sont appelés sous-espaces caractéristiques de f .

- les sous-espaces caractéristiques sont stables par f ;
- f est entièrement caractérisée par f_1, f_2, \dots, f_p les endomorphismes induits par f sur F_1, \dots, F_p ;
- $\forall i \in \llbracket 1, p \rrbracket, (f_i - \lambda_i \text{id})^{\alpha_i} = 0$, ou $f_i - \lambda_i \text{id}$ est nilpotent, donc f_i a pour unique valeur propre λ_i ;
- les f_i sont trigonalisables (leur polynôme caractéristique est scindé) ;
- dans une base $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_p)$ où $\forall i \in \llbracket 1, p \rrbracket, \mathcal{B}_i$ est une base de diagonalisation de f_i , alors $\text{Mat}_{\mathcal{B}}(f)$ est diagonale par blocs, chaque bloc étant triangulaire de taille $\alpha_i \times \alpha_i$;
- pour $i \in \llbracket 1, p \rrbracket, \dim F_i = \alpha_i$.

Chapitre VIII

Intégration

I Intégrale sur un segment (rappels)

Définition :

Soient c, d deux réels avec $c < d$

Soit f une fonction définie sur $[c, d]$ et à valeurs dans \mathbb{K} . La fonction f est dite continue par morceaux s'il existe une subdivision $\sigma = (x_i)_{i \in \llbracket 0, n \rrbracket}$ telle que pour tout $i \in \llbracket 0, n-1 \rrbracket$, la fonction $f|_{]x_i, x_{i+1}[}$ se prolonge en une fonction continue sur $[x_i, x_{i+1}]$

Une telle subdivision est dite adaptée à f

Remarque I.1. Une fonction continue par morceaux sur un segment est bornée

Définition :

Une fonction f définie sur un intervalle I et à valeurs dans \mathbb{K} est dite continue par morceaux sur l'intervalle I si sa restriction à tout segment inclus dans I est continue par morceaux.

On note $\mathcal{CM}(I, \mathbb{K})$ l'ensemble des fonctions continues par morceaux de I dans \mathbb{K}

Proposition :

Pour f, g continues par morceaux sur $[a, b]$:

— Linéarité : pour $\lambda \in \mathbb{K}$

$$\int_{[a,b]} f + \lambda g = \int_{[a,b]} f + \lambda \int_{[a,b]} g$$

— Positivité : pour f réelle positive

$$\int_{[a,b]} f \geq 0$$

— Croissante : pour f, g réelles

$$f \leq g \Rightarrow \int_{[a,b]} f \leq \int_{[a,b]} g$$

— Inégalité triangulaire :

$$\left| \int_{[a,b]} f \right| \leq \int_{[a,b]} |f|$$

— Positivité améliorée : pour f continue et positive

$$\int_{[a,b]} f = 0 \Rightarrow f = 0$$

Proposition : Chasles

Soit I un intervalle, f une fonction continue par morceaux sur I , $a, b, c \in I$.

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$$

Proposition :

Soit I un intervalle. Soit $a \in I$ et f une fonction continue par morceaux sur I à valeur dans \mathbb{K} , on définit la fonction F_a intégrale dépendant de la borne supérieure s'annulant en a par :

$$F_a : \begin{cases} I & \rightarrow \mathbb{K} \\ x & \mapsto \int_a^x f(t)dt \end{cases}$$

- F_a est continue sur I
- Si f est continue, F_a est \mathcal{C}^1 de dérivée f

Remarque I.2. Cette proposition permet d'énoncer le théorème fondamental de l'analyse : toute fonction continue sur un intervalle possède une primitive.

Proposition : Intégration par parties

Soit $[a, b]$ un segment, f, g des fonctions \mathcal{C}^1 sur $[a, b]$

$$\int_a^b f'(t)g(t)dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt$$

Proposition : Changement de variable

Soit $[a, b]$ un segment, $\varphi \in \mathcal{C}^1$ sur $[a, b]$ à valeurs réelles et f une fonction continue sur $\varphi([a, b])$

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(u)du$$

On dispose du résultat suivant, conséquence du théorème de Heine :

Proposition : Sommes de Riemann régulières

Soit $[a, b]$ un segment et f continue par morceaux sur $[a, b]$

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \xrightarrow{n \rightarrow +\infty} \int_a^b f(t)dt$$

Preuve

Dans le cas continu :

Soit $\varepsilon \in \mathbb{R}_+^*$

f est continue sur $[a, b]$ donc f est uniformément continue

On peut donc fixer $\alpha > 0, \forall t_1, t_2 \in [a, b], |t_1 - t_2| < \alpha \Rightarrow |f(t_1) - f(t_2)| < \varepsilon$

Pour $k \in \llbracket 0, n-1 \rrbracket$, on note $x_k = a + k \frac{b-a}{n}$

On choisit $n_0 \in \mathbb{N}$ tel que $\frac{b-a}{n_0} < \alpha$

Donc $\forall n \geq n_0, \frac{b-a}{n} < \alpha$

Pour $n \geq n_0$,

$$\left| \int_a^b f(t)dt - \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_k) \right| = \left| \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(t)dt - f(x_k) \right| \leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f(t) - f(x_k)|dt \leq \varepsilon(b-a)$$

Dans le cas continu par morceaux, on se sert du fait qu'on soit sur un segment pour avoir une subdivision adaptée finie et ensuite appliquer Chasles.g

Exemple I.1. Donner un équivalent de $(S_n) = \left(\sum_{k=1}^n \frac{1}{n^2+k^2} \right)$

$$S_n = \frac{1}{n^2} \sum_{k=1}^n \frac{1}{1+\frac{k^2}{n^2}}$$

$t \mapsto \frac{1}{1+t^2}$ est continue sur $[0, 1]$

$$\text{Alors } \frac{1}{n} \sum_{k=1}^n \frac{1}{1+\frac{k^2}{n^2}} \rightarrow \int_0^1 \frac{dt}{1+t^2}$$

$$\text{Et donc } S_n \sim \frac{1}{n} \int_0^1 \frac{dt}{1+t^2}$$

$$\text{Donc } S_n \sim \frac{\pi}{4n}$$

Proposition : formule de Taylor avec reste intégral et majoration de Lagrange

Soit $[a, b]$ un segment et f une fonction \mathcal{C}^{n+1} sur $[a, b]$ à valeurs dans \mathbb{K} . Alors :

$$\int_a^b f(t)dt = \sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!} + R_n(f)$$

$$\text{Où } R_n(f) = \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t)dt = \frac{(b-a)^{n+1}}{n!} \int_0^1 (1-u)^n f^{(n+1)}(a+u(b-a))du$$

En majorant la deuxième forme du reste intégrale :

$$|R_n(f)| \leq \frac{|b-a|^{n+1}}{(n+1)!} \sup_{[a,b]} |f^{(n+1)}|$$

II Intégration sur un segment d'une fonction vectorielle

La définition de l'intégrale sur un segment s'étend sans difficultés aux fonctions "vectorielles" continues par morceaux sur un segment, c'est à dire aux fonctions continues par morceaux sur un segment, à valoir dans un \mathbb{K} -espace vectoriel de **dimension finie** E

Il suffit de fixer une base \mathcal{B}

Pour une fonction vectorielle f , de fonctions coordonnées f_i dans \mathcal{B}

$$\int_{[a,b]} f$$

est le vecteur de E de coordonnées dans \mathcal{B}

$$\left(\int_{[a,b]} f_i \right)$$

On prouve la majorité des propriétés équivalentes à celles dans \mathbb{K} en le faisant coordonnées par coordonnées, et on perd toutes les propriétés liées à la relation d'ordre, sauf l'inégalité triangulaire.

Dans cette partie, E désigne un \mathbb{K} -espace vectoriel de dimension finie.

Proposition :

Pour f, g continues par morceaux de $[a, b]$ dans E :

— Linéarité : pour $\lambda \in \mathbb{K}$

$$\int_{[a,b]} f + \lambda g = \int_{[a,b]} f + \lambda \int_{[a,b]} g$$

— Composition linéaire : si L est une application linéaire de E dans un \mathbb{K} -ev de dimension finie F

$$\int_{[a,b]} L(f) = L \left(\int_{[a,b]} f \right)$$

— Inégalité triangulaire : pour une norme quelconque $\|\cdot\|$

$$\left\| \int_{[a,b]} f \right\| \leq \int_{[a,b]} \|f\|$$

Preuve

Preuve 8 : Pour l'inégalité triangulaire

Soit $\|\cdot\|$ une norme sur E

Soit f continue par morceaux sur $[a, b]$, $a < b$

Soit $n \in \mathbb{N}$, alors $\left\| \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right\| \leq \frac{b-a}{n} \sum_{k=0}^{n-1} \left\| f\left(a + k \frac{b-a}{n}\right) \right\|$

On passe à la limite, par continuité de la norme et par sommes de Riemann.

Alors $\left\| \int_{[a,b]} f \right\| \leq \int_{[a,b]} \|f\|$

Proposition : Chasles

Soit I un intervalle, f une fonction continue par morceaux sur I , $a, b, c \in I$.

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$$

Proposition :

Soit I un intervalle. Soit $a \in I$ et f une fonction continue par morceaux sur I à valeur dans \mathbb{K} , on définit la fonction F_a intégrale dépendant de la borne supérieur s'annulant en a par :

$$F_a : \begin{cases} I & \rightarrow \mathbb{K} \\ x & \mapsto \int_a^x f(t)dt \end{cases}$$

- F_a est continue sur I
- Si f est continue, F_a est \mathcal{C}^1 de dérivée f

Remarque II.1. *Le théorème fondamental de l'analyse reste valable pour une fonction vectorielle : toute fonction continue sur un intervalle possède une primitive.*

Proposition : Intégration par parties

Soit $[a, b]$ un segment, λ une fonction à valeur dans \mathbb{K} , \mathcal{C}^1 sur $[a, b]$, f une fonction \mathcal{C}^1 sur $[a, b]$ à valeurs dans E

$$\int_a^b \lambda'(t)f(t)dt = [\lambda(t)f(t)]_a^b - \int_a^b \lambda(t)f'(t)dt$$

Proposition : Changement de variable

Soit $[a, b]$ un segment, φ une fonction \mathcal{C}^1 sur $[a, b]$ à valeurs réelles et f une fonction continue sur $\varphi([a, b])$ à valeurs dans E

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(u)du$$

Proposition : Sommes de Riemann régulières

Soit $[a, b]$ un segment et f continue par morceaux sur $[a, b]$

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \xrightarrow{n \rightarrow +\infty} \int_a^b f(t)dt$$

Proposition : formule de Taylor avec reste intégral et majoration de Lagrange

Soit $[a, b]$ un segment et f une fonction \mathcal{C}^{n+1} sur $[a, b]$ à valeurs dans \mathbb{K} . Alors :

$$\int_a^b f(t)dt = \sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!} + R_n(f)$$

Où $R_n(f) = \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t)dt = \frac{(b-a)^{n+1}}{(n+1)!} \int_0^1 (1-u)^n f^{(n+1)}(a+u(b-a))du$

En majorant la deuxième forme du reste intégrale :

$$\|R_n(f)\| \leq \frac{|b-a|^{n+1}}{(n+1)!} \sup_{[a,b]} \|f^{(n+1)}\|$$

Remarque II.2. La majoration de Lagrange à l'ordre 0 constitue ce qu'on appelle l'inégalité des accroissements finis : Soit $[a, b]$ un segment, et f une fonction \mathcal{C}^1 sur $[a, b]$ à valeur dans E

$$\|f(b) - f(a)\| \leq |b - a| \sup_{[a,b]} \|f'\|$$

III Intégration sur un intervalle quelconque

Dans tout ce chapitre, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} et les fonctions sont à valeurs dans \mathbb{K}

III.1 Intégration sur un intervalle semi-ouvert

Définition :

Soit $f : [a, b[\rightarrow \mathbb{K}$ une fonction continue par morceaux avec $b \in \mathbb{R} \cup \{+\infty\}$

Notons F la fonction :

$$F : \begin{cases} [a, b[& \rightarrow \mathbb{K} \\ x & \mapsto \int_a^x f \end{cases}$$

On dit que $\int_a^b f$ est convergente si $F(x)$ a une limite finie quand x tend vers b .

Dans ce cas, on note :

$$\int_a^b f = \lim_{x \rightarrow b} \int_a^x f$$

Dans le cas contraire, on dit que $\int_a^b f$ est divergente.

Etudier la nature de $\int_a^b f$, c'est étudier si l'intégrale est convergente ou divergente.

Exemple III.1. Pour $\int_1^{+\infty} \frac{dt}{t^2}$

$t \mapsto \frac{1}{t^2}$ est continue sur $[1, +\infty[$, donc continue par morceaux.

Pour $x \geq 1$: $\int_1^x \frac{dt}{t^2} = \left[\frac{-1}{t} \right]_1^x = 1 - \frac{1}{x}$

Et $1 - \frac{1}{x} \xrightarrow{x \rightarrow +\infty} 1$

Donc $\int_1^{+\infty} \frac{dt}{t^2}$ converge et $\int_1^{+\infty} \frac{dt}{t^2} = 1$

Pour $\int_1^{+\infty} \frac{dt}{\sqrt{t^4 + t^{3/2} + 1}}$, on a aussi la convergence par des critères qu'on verra plus loin.

Pour calculer $\int_0^1 -\ln(1-t)dt$, on commence par établir que la valeur de $-\ln(1-t)$ en 1 est infinie, et continue en $[0, 1[$

Pour $x \in [0, 1[$:

$\int_0^x -\ln(1-t) = \int_1^{1-x} \ln(u)du$ en faisant un changement de variable $u = (1-t)$

$$\begin{aligned}
&= -\int_{1-x}^1 \ln(u) du = [t \ln(t) - t]_{1-x}^1 = \ln(1) - 1 - ((1-x)\ln(1-x) - (1-x)) \\
&= x + (1-x)\ln(1-x) = F(x) \\
&\text{Et donc } F(x) \rightarrow_{x \rightarrow 1} 1 \\
&\text{Donc } \int_0^1 -\ln(1-t) dt \text{ converge et } \int_0^1 -\ln(1-t) dt = 1 \\
&\int_1^{+\infty} \frac{dt}{t} \text{ diverge.} \\
&\int_0^{+\infty} \cos(t) dt \text{ diverge.}
\end{aligned}$$

Remarque III.1. La notation $\int_a^b f$ désigne à la fois l'intégrale impropre qui peut être convergente ou divergente, et en cas de convergence la "valeur" de l'intégrale qui est un élément de \mathbb{K} . Il n'y a pas comme pour les séries deux notations différentes.

On peut employer aussi les notations :

- $\int_{[a,b]} f$
- $\int_a^b f(t) dt$

Définition :

Soit $f :]a, b] \rightarrow \mathbb{K}$ une fonction continue par morceaux avec $a \in \mathbb{R} \cup \{-\infty\}$
On dit que $\int_a^b f$ est convergente si $x \mapsto \int_x^b f$ admet une limite finie en a
Dans ce cas, on note

$$\int_a^b f = \lim_{x \rightarrow a} \int_x^b f$$

Dans le cas contraire, on dit que $\int_a^b f$ est divergente.

Proposition :

Soit $f : [a, b[\rightarrow \mathbb{K}$ une fonction continue par morceaux avec $b \in \mathbb{R}$
Si f admet une limite finie en b alors $\int_a^b f$ est convergente
On dit que c'est une intégrale faussement impropre (ou faussement généralisée).

Remarque III.2. On peut énoncer la même propriété pour la borne inférieure de l'intégrale.

Proposition :

Soit $f : [a, b[\rightarrow \mathbb{K}$ une fonction continue par morceaux, soit $c \in [a, b[$, alors $\int_a^b f$ et $\int_c^b f$ sont de même nature.

Preuve

Preuve 16 : $\forall x \in [a, b[, \int_a^x f(t) dt = \int_a^c f(t) dt + \int_c^x f(t) dt$
Et donc si \int_a^x admet une limite quand x tend vers b , alors comme $\int_c^x f(t) dt$ aussi.

Remarque III.3. La convergence de l'intégrale ne dépend que du comportement local en b de la fonction.

On peut énoncer la même propriété pour une intégrale impropre en a .

Attention : L'intégrale $\int_0^{+\infty} f$ d'une fonction f peut converger sans que la fonction tende vers 0 en $+\infty$

III.2 Intégration sur un intervalle ouvert

Définition : Intégrale sur un intervalle ouvert

Soit $f \in \mathcal{CM}([a, b[, \mathbb{R})$ avec $-\infty \leq a < b \leq +\infty$

On dit que $\int_a^b f$ est convergente si pour $c \in]a, b[$, $\int_a^c f$, $\int_c^b f$ sont toutes les deux convergentes.

On note alors $\int_a^b f = \int_a^c f + \int_c^b f = \lim_{x \rightarrow a^+, y \rightarrow b^-} \int_x^y f$

Exemple III.2. Etudier la convergence de $\int_{-\infty}^{+\infty} \sin(t) dt$

$\int_0^{+\infty} \sin(t) dt$ diverge donc $\int_{-\infty}^{+\infty} \sin(t) dt$ diverge.

III.3 Intégration des fonctions positives

Notons qu'on peut énoncer le même type de proposition pour des fonctions définies sur un intervalle semi-ouvert à gauche.

Théorème : Rappel sur les limites monotones

Avec $a < b$: soit $g \in \mathcal{F}([a, b[, \mathbb{R})$

- Si g est croissante majorée, alors g a une limite en b^-
- Si g est croissante non-majorée, alors $g \rightarrow_{b^-} +\infty$
- Si g est croissante minorée, alors g a une limite en a^+
- Si g est croissante non-minorée, alors $g \rightarrow_{a^+} -\infty$
- Si g est décroissante majorée, alors g a une limite en a^+
- Si g est décroissante non-majorée, alors $g \rightarrow_{a^+} +\infty$
- Si g est décroissante minorée, alors g a une limite en b^-
- Si g est décroissante non-minorée, alors $g \rightarrow_{b^-} -\infty$

Proposition :

Soit $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$. Soit $f \in \mathcal{CM}([a, b[, \mathbb{R})$

Si f est positive, l'intégrale $\int_a^b f$ converge si, et seulement si, $x \mapsto \int_a^x f$ est majorée.

Preuve

Proposition 17 : pour $f \geq 0$

$x \mapsto \int_a^x f$ est croissante :

Soient $x, y \in [a, b[, x < y$

Alors $\int_a^y f - \int_a^x f = \int_x^y f(t) dt \geq 0$

Donc on a la convergence si, et seulement si, $x \mapsto \int_a^x f$ est majorée par théorème des limites monotones.

Remarque III.4. L'intégrale diverge si, et seulement si, $x \mapsto \int_a^x f$ tend vers $+\infty$ quand x tend vers b

Par analogie avec la définition de la somme d'une famille de réels positifs, on peut définir la valeur de l'intégrale d'une fonction positive continue par morceaux sur $[a, b[$ dans tous les cas :

- $\int_a^b f = +\infty$ lorsque $\int_a^b f$ diverge
- $\int_a^b f \in \mathbb{R}$ ou $\int_a^b f < +\infty$ lorsque $\int_a^b f$ converge

Proposition :

Soit $a \in \mathbb{R}, b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$

Soit $f, g \in \mathcal{CM}([a, b[, \mathbb{R})$ deux fonctions à valeurs positives telles que $0 \leq f \leq g$:

- Si $\int_a^b g$ converge alors $\int_a^b f$ converge.
- Si $\int_a^b f$ diverge alors $\int_a^b g$ diverge.

Preuve

Même démonstration que pour la proposition précédente

On peut l'appeler le critère de majoration positive, comme pour les séries.

Proposition : Corollaire

Soit $a \in \mathbb{R}, b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$

Soit $f, g \in \mathcal{CM}([a, b[, \mathbb{R})$ deux fonctions à valeurs positives telles que $f =_b o(g)$ ou $f =_b \mathcal{O}(g)$:

- Si $\int_a^b g$ converge alors $\int_a^b f$ converge.
- Si $\int_a^b f$ diverge alors $\int_a^b g$ diverge.

Preuve

Démonstration équivalente à celle sur les séries.

Proposition : Corollaire

Soit $a \in \mathbb{R}, b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$

Soit $f, g \in \mathcal{CM}([a, b[, \mathbb{R})$ deux fonctions à valeurs positives telles que $f \sim_b g$

Alors les intégrales $\int_a^b f$ et $\int_a^b g$ sont de même nature

Preuve

Démonstration équivalente à celle sur les séries.

Remarque III.5. *Attention : le résultat du commentaire s'applique également pour deux fonctions négatives et équivalentes. En revanche, si les fonctions ne sont pas de signe constant, les intégrales peuvent être de nature différentes.*

III.4 Intégrales de référence de Riemann**Proposition :**

$\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si, et seulement si, $\alpha > 1$

Preuve

Par calcul de primitives.

Proposition :

$\int_0^{+\infty} e^{-\lambda t} dt$ converge si, et seulement si, $\lambda > 0$

Proposition :

$\int_0^1 \frac{dt}{t^\alpha}$ converge si, et seulement si, $\alpha < 1$

Proposition :

$\int_0^1 \ln(t)dt$ converge

Remarque III.6. On a aussi que $\int_0^1 |\ln(t)|^\alpha$ converge, mais ce n'est pas au programme, donc on doit le retrouver à partir de $|\ln(t)|^\alpha = o\left(\frac{1}{\sqrt[t]{t}}\right)$ en 0

Proposition :

Si $a < b$:

$\int_a^b \frac{dt}{|t-a|^\alpha}$ converge si, et seulement si, $\alpha < 1$

Proposition :

Si $a > b$:

$\int_b^a \frac{dt}{|b-t|^\alpha}$ converge si, et seulement si, $\alpha < 1$

Exemple III.3. Existence et calcul éventuel de $\int_1^{+\infty} \frac{dt}{(t-1)(t+2)}$ et de $\int_2^{+\infty} \frac{dt}{(t-1)(t+2)}$

$f : t \mapsto \frac{1}{(t-1)(t+2)}$ est continue donc continue par morceaux sur $]1, +\infty[$

Etude en 1 : $\frac{1}{(t-1)(t+2)} \sim \frac{1}{3(t-1)}$

f est positive, $\int_1^2 \frac{dt}{t-1}$ diverge donc $\int_1^2 f(t)dt$ diverge par critère d'équivalent positif.

donc $\int_0^{+\infty} \frac{dt}{(t-1)(t+2)}$ diverge

$f(t) \sim_{+\infty} \frac{1}{t^2}$

$\int_2^{+\infty} \frac{dt}{t^2}$ converge, f est positive et donc par critère d'équivalent positif, $\int_2^{+\infty} f(t)dt$ converge

Pour le calcul : Faisons d'abord une décomposition en éléments simples

$$\int_2^{+\infty} f(t)dt = \int_2^{+\infty} \frac{1}{3(t-1)} - \frac{1}{3(t+2)} dt$$

(L'intégrale de la somme est convergente mais l'intégrale de chaque élément simple est divergente)

Soit $x \in [2, +\infty[$: $\int_2^x \frac{1}{3(t-1)} - \frac{1}{3(t+2)} dt$

$$= \left[\frac{1}{3} \ln \left(\frac{t-1}{t+2} \right) \right]_2^x = \frac{1}{3} \ln(4) - \frac{1}{3} \ln \left(\frac{x-1}{x+2} \right)$$

Or $\ln \left(\frac{x-1}{x+2} \right) \rightarrow_{+\infty} 0$ donc $\int_2^{+\infty} f(t)dt = \frac{1}{3} \ln(4)$

Exemple III.4. Convergence de :

$$1. \int_0^{+\infty} t^{12} e^{-t} dt$$

$f \geq 0$, f est continue donc continue par morceaux sur $[0, +\infty[$

$f =_\infty o(e^{-\frac{t}{2}})$ or $f =_\infty o(\frac{1}{t^2})$

$\int_1^{+\infty} \frac{dt}{t^2}$ converge donc par critère de domination positif, $\int_0^{+\infty} f(t)dt$ converge.

$$2. \int_0^{+\infty} \frac{e^{-\sqrt{t}}}{t^{2/3}} dt$$

$f \geq 0$, f continue sur $]0, +\infty[$

En 0 : $f \sim \frac{1}{t^{2/3}}$ et donc $\int_0^1 f(t)dt$ converge par comparaison à du Riemann

En $+\infty$: $f = o\left(\frac{1}{t^2}\right)$ et on a la convergence aussi.

$$3. \int_2^{+\infty} \frac{dt}{t^\alpha \ln(t)^\beta}$$

f est continue sur $[2, +\infty[$ positive :

Si $\alpha > 1$, alors $\frac{1}{t^\alpha (\ln(t))^\beta} = o\left(\frac{1}{t^{\frac{1+\alpha}{2}}}\right)$ par croissance comparée

$\int_1^{+\infty} \frac{dt}{t^{\frac{1+\alpha}{2}}}$ converge car $\frac{1+\alpha}{2} > 1$ converge

Donc par critère de domination positif, $\int_2^{+\infty} \frac{dt}{t^\alpha (\ln(t))^\beta}$ converge

Si $\alpha < 1$, alors $\frac{1}{t^\alpha (\ln(t))^\beta} \gg \frac{1}{t^{\frac{1+\alpha}{2}}}$

Donc $\int_2^{+\infty} \frac{dt}{t^\alpha (\ln(t))^\beta}$ diverge.

Si $\alpha = 1$:

Pour $x > 2$, $\int_2^{+\infty} \frac{dt}{t (\ln(t))^\beta} = \int_{\ln(2)}^{\ln(x)} \frac{du}{u^\beta}$ (Changement de variable avec $u = \ln(t)$, $du = \frac{dt}{t}$)

Et donc $\int_2^{+\infty} \frac{dt}{t \ln(t)^\beta}$ converge si, et seulement si, $\beta > 1$

Et donc, $\int_2^{+\infty} \frac{dt}{t^\alpha \ln(t)^\beta}$ converge si, et seulement si, $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$)

4. $\int_0^{1/2} \frac{dt}{t^\alpha |\ln(t)|^\beta} dt$

On a convergence si, et seulement si, $\alpha < 1$ ou ($\alpha = 1$ et $\beta > 1$)

III.5 Intégrabilité

Définition :

Soit I un intervalle, f continue par morceaux sur I à valeurs dans \mathbb{K}
On dit que $\int_I f$ est absolument convergente si $\int_I |f|$ converge.

Remarque III.7. On peut comme dans le cadre de la théorie de la sommabilité des familles de \mathbb{K} utiliser la définition de la valeur de l'intégrale d'une fonction positive. Ainsi, la définition précédente peut s'écrire :

f est intégrable sur I si $\int_I |f| < +\infty$

Théorème :

Soit f continue par morceaux sur un intervalle I .
Si $\int_I f$ est absolument convergente alors $\int_I f$ est convergente.

Preuve

Même démonstration que pour les séries :
Pour f réelle, $0 \leq |f| - f \leq 2|f|$ et on intègre
Ensuite pour les fonctions complexes, on écrit que $|\Re(f)| \leq |f|$ et que $|\Im(f)| \leq |f|$

Proposition :

Une intégrale absolument convergente est convergente et pour f intégrable sur un intervalle I

$$|\int_I f| \leq \int_I |f|$$

III.6 Propriétés des intégrales impropres

Proposition : Propriétés des intégrales impropres

- Linéarité : si I un intervalle, f, g continues par morceaux sur I d'intégrales convergentes sur I , $\lambda \in \mathbb{K}$, alors :

$$\int_I f + \lambda g \text{ converge et } \int_I f + \lambda g = \int_I f + \lambda \int_I g$$
- Positivité : si f continue par morceaux sur I réelle positive, d'intégrale sur I convergente, alors $\int_I f \geq 0$
- Croissance : si f et g sont continues par morceaux sur I i réelles positives d'intégrales sur I convergentes avec $f \leq g$, alors $\int_I f \leq \int_I g$
- Inégalité triangulaire : si f continue par morceaux sur I intégrable sur I , alors $|\int_I f| \leq \int_I |f|$
- Positivité améliorée : si I un intervalle, f continue réelle positive sur I d'intégrale sur I convergente, alors $\int_I f = 0 \Rightarrow \forall x \in I, f(x) = 0$

Preuve

Pour la positivité améliorée : cas $I = [a, b[$

$$f \geq 0, \int_{[a,b[} f = \lim_{x \rightarrow b^-} \int_a^x f(t) dt = \sup_{x \in [a,b[} \int_a^x f(t) dt$$

(car $x \mapsto \int_a^x f(t) dt$ est croissante)

$$\text{Donc } \int_{[a,b[} f = 0 \Rightarrow \forall x \in [a, b[, 0 \leq \int_a^x f(t) dt \leq 0$$

Donc $\forall x \in [a, b[\int_a^x f(t) dt = 0$ donc par stricte positivité d'une intégrale sur un segment, f est nulle sur tout segment de $[a, b[$ donc nulle sur $[a, b[$

III.7 Relation de Chasles

Définition :

Soit $a, b \in \bar{\mathbb{R}}$, tels que $a < b$ et $f \in \mathcal{CM}(]b, a[, \mathbb{K})$ telle que $\int_b^a f$ converge.

On définit $\int_a^b f(t) dt$ par $-\int_b^a f(t) dt$ (comme pour une intégrale classique)

Proposition : Relation de Chasles

Soit I un intervalle, soit $f \in \mathcal{CM}(I, \mathbb{K})$. Soit a, b, c dans l'adhérence de I dans $\mathbb{R} \cup \{+\infty\}$

Si $\int_I f$ converge, alors $\int_a^c f, \int_c^b f$ et $\int_a^b f$ convergent et

$$\int_a^b f = \int_a^c f + \int_c^b f$$

III.8 Fonction dépendant de la borne supérieure

Proposition :

Pour $f \in \mathcal{CM}(I, \mathbb{K})$ telle que $\int_I f$ converge

Soit $a \in \bar{I}$ (l'adhérence est prise dans \mathbb{R})

Soit $F_a : \begin{cases} I & \rightarrow \mathbb{K} \\ x & \mapsto \int_a^x f \end{cases}$ la fonction intégrale dépendant de la borne supérieure

- F_a est continue sur I
- Si f est continue sur I , F_a est \mathcal{C}^1 de dérivée f

III.9 Intégration par parties

Théorème : Intégration par parties

Avec $f, g \in \mathcal{C}^1$ sur $]a, b[$, si fg a une limite en a^+ et en b^- alors $\int_a^b f(t)g'(t)dt$ et $\int_a^b f'(t)g(t)dt$ sont de même nature et en cas de convergence :

$$\int_a^b f(t)g'(t)dt = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)dt$$

Preuve

Avec x, y tels que $a < x < y < b$, alors $\int_x^y fg' = [fg]_x^y - \int_x^y f'g$

Exemple III.5. $\int_1^{+\infty} \frac{\sin(t)}{t} dt$

$t \mapsto \frac{\sin(t)}{t}$ est continue sur $[1, +\infty[$

$\frac{-\cos(t)}{t}$ tend vers 0 en l'infini : $\forall t \in [1, +\infty[, \left| \frac{-\cos(t)}{t} \right| \leq \frac{1}{t}$, d'où la limite

Donc $\int_1^{+\infty} \frac{\sin(t)}{t} dt$ est de même nature que $\int_1^{+\infty} \frac{\cos(t)}{t^2} dt$

Et $\forall t \in [1, +\infty[, \left| \frac{\cos(t)}{t^2} \right| < \frac{1}{t^2}$

Donc par théorème de comparaison positive, $\int_1^{+\infty} \frac{\cos(t)}{t^2} dt$ converge absolument

Donc $\int_1^{+\infty} \frac{\sin(t)}{t} dt$ converge et $\int_1^{+\infty} \frac{\sin(t)}{t} dt = \left[\frac{-\cos(t)}{t} \right]_1^{+\infty} - \int_1^{+\infty} \frac{\cos(t)}{t^2} dt$

III.10 Changement de variable

Théorème :

Soient a, b, α, β tels que $-\infty \leq a < b \leq +\infty, -\infty \leq \alpha < \beta \leq +\infty$

Soit $f \in]a, b[\rightarrow \mathbb{K}$ une fonction continue.

Soit $\varphi :]\alpha, \beta[\rightarrow]a, b[$ une fonction bijective, strictement croissante et de classe \mathcal{C}^1

Les intégrales $\int_a^b f(t)dt$ et $\int_\alpha^\beta (f \circ \varphi)(u)\varphi'(u)du$ sont de même nature, et en cas de convergence

$$\int_a^b f(t)dt = \int_\alpha^\beta (f \circ \varphi)(u)\varphi'(u)du$$

Remarque III.8. On remarque que $a = \lim_{x \rightarrow \alpha} \varphi(x)$ et $b = \lim_{x \rightarrow \beta} \varphi(x)$

Remarque III.9. On peut énoncer le même type de théorème avec φ strictement décroissant. On a alors dans le cas convergent :

$$\int_a^b f(t)dt = - \int_\alpha^\beta (f \circ \varphi)(u)\varphi'(u)du$$

III.11 Intégration des relations de comparaisons

Théorème : Intégration des ordres de grandeur

Soit $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$

Soit $f \in \mathcal{CM}([a, b[, \mathbb{K})$

Soit $\varphi \in \mathcal{CM}([a, b[, \mathbb{R})$ une fonction positive sur $[a, b[$

- Si φ est intégrable :
 - Si $f =_b \mathcal{O}(\varphi)$ alors f est intégrable sur $[a, b[$ et $\int_x^b f =_b \mathcal{O}(\int_x^b \varphi)$
 - Si $f =_b o(\varphi)$ alors f est intégrable sur $[a, b[$ et $\int_x^b f =_b o(\int_x^b \varphi)$
 - Si $f \sim_b \varphi$ alors f est intégrable sur $[a, b[$ et $\int_x^b f \sim_b \int_x^b \varphi$
- Si φ n'est pas intégrable :
 - Si $f =_b \mathcal{O}(\varphi)$ alors $\int_a^x f =_b \mathcal{O}(\int_a^x \varphi)$
 - Si $f =_b o(\varphi)$ alors $\int_a^x f =_b o(\int_a^x \varphi)$
 - Si $f \sim_b \varphi$ alors f n'est pas intégrable sur $[a, b[$ et $\int_a^x f \sim_b \int_a^x \varphi$

Preuve

Rappel sur la sommation des ordres de grandeurs :

$u_n = o(w_n) \Leftrightarrow u_n = w_n \varepsilon_n$ avec $(\varepsilon_n) \rightarrow 0$

Dans le cas convergent : soit $\varepsilon > 0$, prenons $n_0 \in \mathbb{N}$ tel que $\forall n \in \mathbb{N}, n \geq n_0 \Rightarrow |\varepsilon_n| \leq \varepsilon$

$$\text{Alors } \forall n \geq n_0, \left| \frac{\sum_{k=n}^{+\infty} u_k}{\sum_{k=n}^{+\infty} w_k} \right| \leq \left| \frac{\sum_{k=n}^{+\infty} \varepsilon_k w_k}{\sum_{k=n}^{+\infty} w_k} \right| \leq \varepsilon$$

f, g deux fonctions continues par morceaux sur $[a, b[$:

Si $f =_b o(g)$, prenons ψ telle que $f = \psi g$ et $\psi \rightarrow_b 0$

Si $\int_a^b g(t) dt$ converge alors on a :

Soit $\varepsilon > 0$, prenons $\alpha > 0$ tel que $\forall x \in]b - \alpha, b[, |\psi(x)| < \varepsilon$

$$\text{Pour un tel } x : \left| \frac{\int_x^b f(t) dt}{\int_x^b g(t) dt} \right| \leq \left| \frac{\int_x^b \psi(t) g(t) dt}{\int_x^b g(t) dt} \right| \leq \varepsilon \left| \frac{\int_x^b g(t) dt}{\int_x^b g(t) dt} \right| \leq \varepsilon$$

Dans le cas divergent :

Dans le cas où f et g sont continues par morceaux sur $]a, b[$, on reprend la même démonstration mais en prenant $x \in]a, a + \alpha[$

Remarque III.10. Attention : dans le cas où φ n'est pas intégrable, on ne sait rien sur l'intégrabilité de f

Remarque III.11. On peut énoncer le même type de proposition pour des fonctions définies sur un intervalle semi-ouvert à gauche.

IV Intégrale dépendant d'un paramètre

IV.1 Cas d'un paramètre entier

Théorème : de convergence dominée

Soit (f_n) une suite de fonctions continues par morceaux de I intervalle de \mathbb{R} dans \mathbb{K} . On suppose que :

- La suite (f_n) converge simplement sur I vers une fonction f continue par morceaux
- Il existe une fonction φ positive et intégrable sur I telle que

$$\forall n \in \mathbb{N}, |f_n| \leq \varphi \text{ (hypothèse de domination)}$$

Alors les fonctions f_n pour $n \in \mathbb{N}$ et la fonction f sont intégrables sur I et

$$\int_I f_n \rightarrow \int_I f$$

Preuve

On admet la preuve car le théorème repose sur la théorie d'intégration de Lebesgue.

On peut cependant le prouver dans le cas où f_n converge uniformément vers f sur tout segment de $[a, b[$

Pour tout $x \in [a, b[$, on a :

$$\left| \int_a^b f_n - \int_a^b f \right| \leq \int_a^x |f_n - f| + \int_x^b |f_n - f| \leq |x - a| \|f_n - f\|_\infty + \int_x^b (\varphi + |f|)$$

Remarque IV.1. On suppose que f est continue par morceaux pour rester dans le cadre du programme où les intégrales sont définies uniquement pour de telles fonctions. Cette hypothèse n'a pas la même importance que l'hypothèse de domination qui est cruciale dans ce théorème.

Remarque IV.2. Si l'intervalle I est borné, une fonction constante peut parfois jouer le rôle de la fonction de domination φ

Théorème : Intégration terme à terme positif

Soit I une intervalle

Soit (u_n) une suite de fonctions réelles positives continues par morceaux et intégrables sur I . On suppose que $\sum u_n$ converge simplement et que $\sum_{n=0}^{+\infty} u_n$ est continue par morceaux sur I Alors :

$$\int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n$$

Remarque IV.3. Cette égalité a lieu dans $\bar{\mathbb{R}}$, c'est à dire que dans le cas d'une série de fonction de terme général positif (u_n) , $\sum_{n=0}^{+\infty} u_n$ est intégrable sur I si, et seulement si, $\sum_{n=0}^{+\infty} \int_I u_n < +\infty$

Théorème : Intégration terme à terme

Soit I un intervalle

Soit (u_n) une suite de fonctions à valeur dans \mathbb{K} . On suppose que :

- Pour tout entier $n \in \mathbb{N}$, u_n est continue par morceaux et intégrable sur I
- La série $\sum u_n$ converge simplement et $\sum_{n=0}^{+\infty} u_n$ est continue par morceaux sur I
- La série $\sum \int_I |u_n|$ converge

Alors $\sum_{n=0}^{+\infty} u_n$ est intégrable sur I et

$$\int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n$$

Preuve

Ce théorème est encore une fois admis.

Remarque IV.4. Là encore, on suppose que f est continue par morceaux pour rester dans le cadre du programme. Cette hypothèse n'a pas la même importance que l'hypothèse de convergence de $\sum \int_I |u_n|$ qui est cruciale dans ce théorème.

Remarque IV.5. Si chaque u_n est continue et que l'on ne connaît pas d'expression simple de $\sum_{n=0}^{+\infty} u_n$, on peut parfois s'assurer que c'est une fonction continue (et donc continue par morceaux) par convergence uniforme/normale sur tout segment.

IV.2 Cas d'un paramètre réel**Théorème : échange des limites non-discrètes**

Soient I, J deux intervalles de \mathbb{R} , f une fonction définie sur $J \times I$ à valeurs dans \mathbb{K} . Soit λ_0 dans l'adhérence de J ($\in \bar{R}$). On suppose que :

- pour tout $\lambda \in J$, la fonction $t \mapsto f(\lambda, t)$ est continue par morceaux sur I
- il existe une fonction l continue par morceaux de I dans \mathbb{K} telle que pour tout $t \in I$, $\lim_{\lambda \rightarrow \lambda_0} f(\lambda, t) = l(t)$
- Il existe une fonction φ continue par morceaux positive et intégrable sur I telle que :

$$\forall (\lambda, t) \in J \times I, |f(\lambda, t)| \leq \varphi(t) \text{ (hypothèse de domination)}$$

Alors les fonctions $t \mapsto f(\lambda, t)$ (pour tout $\lambda \in J$) et la fonction l sont intégrables sur I et :

$$\lim_{\lambda \rightarrow \lambda_0} \int_I f(\lambda, t) dt = \int_I l(t) dt$$

Preuve

Pour une fonction f dans les conditions de l'énoncé, dire que $f \rightarrow_{\lambda_0} l$ équivaut à $\forall (u_n) \in J^{\mathbb{N}}, (u_n) \rightarrow \lambda_0 \Rightarrow f(u_n, t) \rightarrow l(t)$

Soit $(u_n) \in J^{\mathbb{N}}$ de limite λ_0 . On note tout $n \in \mathbb{N}$ et $t \in I$, $f_n = f(u_n, t)$

$\forall n \in \mathbb{N}, f_n \in \mathcal{CM}(I, \mathbb{K})$

$\forall t \in I, f_n(t) \rightarrow_{n \rightarrow +\infty} l(t)$ et il existe par hypothèse φ intégrable sur I telle que $\forall t \in I, \forall n \in \mathbb{N}, |l(t)| \leq \varphi(t)$

Donc par théorème de convergence dominée discret :

$$\int_I f(u_n, t) \rightarrow \int_I l(t) dt$$

Donc par caractérisation séquentielle de la convergence :

$$\int_I f(\lambda, t) \rightarrow_{\lambda \rightarrow \lambda_0} \int_I l(t) dt$$

Théorème : Autre formulation

Soit I et J deux intervalles de \mathbb{R} , $(f_\lambda)_{\lambda \in J}$ une famille de fonctions définies sur J dans \mathbb{K} . Soit λ_0 dans l'adhérence de J (dans $\bar{\mathbb{R}}$). On suppose que :

- pour tout $\lambda \in J$, la fonction f_λ est continue par morceaux sur I
- il existe une fonction l continue par morceaux de I dans \mathbb{K} telle que pour tout $t \in I$,
 $\lim_{\lambda \rightarrow \lambda_0} f_\lambda(t) = l(t)$
- il existe une fonction φ continue par morceaux, positive et intégrable sur I telle que
 $\forall (\lambda, t) \in J \times I, |f_\lambda(t)| \leq \varphi(t)$ (hypothèse de domination)

Alors les fonctions f_λ (pour $\lambda \in J$) et l sont intégrables sur I et

$$\lim_{\lambda \rightarrow \lambda_0} \int_I f_\lambda = \int_I l$$

Théorème : Continuité dominée

Soit A une partie d'un EVN de dimension finie, I un intervalle de \mathbb{R} , f une fonction définie sur $A \times I$ à valeurs dans \mathbb{R} . On suppose que :

- pour tout $x \in A$, la fonction $t \mapsto f(x, t)$ est continue par morceaux sur I
- pour tout $t \in I$, la fonction $x \mapsto f(x, t)$ est continue sur A
- il existe une fonction φ continue par morceaux, positive et intégrable sur I telle que
 $\forall (x, t) \in A \times I, |f(x, t)| \leq \varphi(t)$ (hypothèse de domination)

Alors, pour tout $x \in A$, la fonction $t \mapsto f(x, t)$ est intégrable sur I et la fonction

$$g : \begin{cases} A & \rightarrow \mathbb{K} \\ x & \mapsto \int_I f(x, t) dt \end{cases}$$

est continue.

Proposition : Extension

Si l'hypothèse de domination est satisfaite au voisinage d'un point a de A , on peut en conclure la continuité de $x \mapsto \int_I f(x, t)dt$ en a

Si A est un intervalle de \mathbb{R} , et que l'hypothèse de domination est satisfaite sur tout segment de A , alors

$$g : \begin{cases} A & \rightarrow \mathbb{K} \\ x & \mapsto \int_I f(x, t)dt \end{cases} \text{ est continue}$$

IV.3 Dérivabilité d'une intégrale à paramètres**Théorème : Dérivabilité**

Soit I et J deux intervalles de \mathbb{R} , f une fonction définie sur $J \times I$ à valeurs dans \mathbb{K} . On suppose que :

- pour tout $x \in J$, la fonction $t \mapsto f(x, t)$ est continue par morceaux et dérivable sur I
- la fonction f admet sur $J \times I$ une dérivée partielle par rapport à la première variable, $\frac{\partial f}{\partial x}$
- la fonction $\frac{\partial f}{\partial x}$ vérifie les hypothèses du théorème 36 :
 - pour tout $x \in J$, la fonction $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur I
 - pour tout $t \in I$, la fonction $x \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue sur J
 - il existe une fonction φ continue par morceaux, positive et intégrable sur I telle que

$$\forall (x, t) \in J \times I, \left| \frac{\partial f}{\partial x}(x, t) \right| \leq \varphi(t) \text{ (hypothèse de domination)}$$

Alors pour tout $x \in J$, la fonction $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est intégrable sur I ,
la fonction $g : x \mapsto \int_I f(x, t)dt$ est de classe \mathcal{C}^1 sur J et vérifie :

$$\forall x \in J, g'(x) = \int_I \frac{\partial f}{\partial x}(x, t)dt$$

Proposition : Extension

Si l'hypothèse de domination de $\frac{\partial f}{\partial x}$ est satisfaite sur tout segment de J , la conclusion du théorème subsiste.

Remarque IV.6. Comme toujours, l'hypothèse de domination est essentielle.

Théorème : Classe d'une intégrale à paramètre

Soit I et J deux intervalles de \mathbb{R} , f une fonction définie sur $J \times I$ à valeurs dans \mathbb{K} et $k \in \mathbb{N}^*$. On suppose que :

- pour tout $j \in \llbracket 0, k \rrbracket$, la fonction f admet sur $J \times I$ une dérivée partielle d'ordre j par rapport à la première variable, $\frac{\partial^j f}{\partial x^j}$
- pour tout $j \in \llbracket 0, k-1 \rrbracket$, pour tout $x \in J$, la fonction $\frac{\partial^j f}{\partial x^j}(x, t)$ est continue par morceaux et intégrable sur I
- pour tout $t \in I$, pour tout $j \in \llbracket 0, k \rrbracket$, $x \mapsto \frac{\partial^j f}{\partial x^j}(x, t)$ est continue sur J
- pour tout segment K inclus dans J , il existe une fonction φ_K continue par morceaux, positive et intégrable sur I telle que

$$\forall (x, t) \in K \times I, \left| \frac{\partial^k f}{\partial x^k}(x, t) \right| \leq \varphi_K(t) \text{ (hypothèse de domination sur tout segment)}$$

Alors, pour tout $x \in J$, la fonction $t \mapsto \frac{\partial^k f}{\partial x^k}(x, t)$ est intégrable sur I , la fonction $g : x \mapsto \int_I f(x, t) dt$ est de classe \mathcal{C}^k sur J et vérifie

$$\forall x \in J, g^{(k)}(x) = \int_I \frac{\partial^k f}{\partial x^k}(x, t) dt$$

Chapitre IX

Probabilités

I Espace probabilisé

I.1 Univers, tribu

Définition :

Une expérience aléatoire est une expérience renouvelable, et qui renouvelée dans des conditions identiques ne donne pas le même résultat à chaque renouvellement.

Définition :

On appelle univers l'ensemble des issues possibles d'une expérience aléatoire donnée. On le note en général Ω

En première année, le programme ne traite que les expériences aléatoires associées à un univers fini, on peut poser

$$\Omega = \{\omega_1, \dots, \omega_n\}$$

Un événement est une partie de Ω , c'est-à-dire un ensemble d'issues de l'expérience. L'ensemble des événements est donc $\mathcal{P}(\Omega)$

Pour tout $\omega \in \Omega$, le singleton $\{\omega\}$ est appelé événement élémentaire. On peut décrire tout événement comme réunion finie d'événements élémentaires.

On considère maintenant une expérience aléatoire ayant un nombre infini d'issues, donc avec un univers Ω infini.

Contrairement au cours de première année, tout sous-ensemble de Ω ne définit par forcément un événement.

On se limite alors aux parties d'un sous-ensemble \mathcal{A} de $\mathcal{P}(\Omega)$.

Pour pouvoir définir une probabilité, on demande à \mathcal{A} d'être une tribu, au sens de la définition suivante :

Définition :

Soit Ω l'univers d'une expérience aléatoire.

On appelle tribu sur Ω une partie de \mathcal{A} vérifiant les trois hypothèses suivantes :

1. $\Omega \in \mathcal{A}$
2. $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$ (stabilité par passage au complémentaire)
3. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{A} , $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$ (stabilité par réunion dénombrable)

Le couple (Ω, \mathcal{A}) est dit espace probabilisable.

\mathcal{A} est l'ensemble des événements (on rappelle qu'un événement est une partie de Ω)

Remarque I.1. $\mathcal{C}_\Omega(\bigcap_{n \in \mathbb{N}} A_n) = \bigcap_{n \in \mathbb{N}} \mathcal{C}_\Omega A_n$

Donc les intersections infinies sont possibles dans une tribu.

Remarque I.2. La définition d'une tribu et la proposition s'énoncent en termes probabilistes :

- Ω et \emptyset sont des événements, ce sont les événements certains et impossibles.
- L'événement contraire d'un événement est un événement.
- Une réunion finie d'ou dénombrable d'événements est un événement.
- Une intersection finie ou dénombrable d'événements est un événement.

Remarque I.3. Soit A et B deux événements et $(A_n)_{n \in \mathbb{N}}$ une suite d'événements. Comme en sup, on peut traduire les opérations élémentaires sur les ensembles en termes probabilistes et vice-versa :

Insérer le tableau.

Définition :

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.

- La suite (A_n) est dite croissante lorsque

$$\forall n \in \mathbb{N}, A_n \subset A_{n+1}$$

ie : pour tout $n \in \mathbb{N}$, la réalisation de A_n implique celle de A_{n+1}

- La suite (A_n) est dite décroissante lorsque

$$\forall n \in \mathbb{N}, A_{n+1} \subset A_n$$

ie : pour tout $n \in \mathbb{N}$, la réalisation de A_{n+1} implique celle de A_n

- La suite (A_n) est une suite d'événement deux à deux incompatibles (disjoints) lorsque :

$$\forall i, j \in \mathbb{N}, i \neq j \Rightarrow A_i \cap A_j = \emptyset$$

ie : il est impossible que deux événements d'indices différents de la suite soient réalisés simultanément

I.2 Probabilités**Définition :**

On appelle probabilité sur l'espace probabilisable (Ω, \mathcal{A}) toute application P définie sur \mathcal{A} vérifiant :

1. $\forall A \in \mathcal{A}, P(A) \in [0, 1]$
2. $P(\Omega) = 1$
3. Pour tout famille dénombrable $(A_i)_{i \in I}$ d'événements deux-à-deux incompatibles :

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i) \text{ } (\sigma\text{-additivité})$$

Dans le cas d'un univers fini, la définition précédente est compatible avec celle donnée en sup, on étend les propriétés suivantes au cas d'un univers infini :

Proposition :

Soit (Ω, \mathcal{A}) un espace probabilisé, et soit A et B deux événements. On a :

- $P(\bar{A}) = 1 - P(A)$
- $P(\emptyset) = 0$ (cas particulier du résultat précédent)
- $P(A \setminus B) = P(A) - P(A \cap B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (formule de Poincaré)
- P est croissante : Si $A \subset B$ alors $P(A) \leq P(B)$

Preuve

Toutes ces démonstrations fonctionnent exactement de la même manière sur des univers finis.

$$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) = 1$$

$$\text{D'où } P(\bar{A}) = 1 - P(A)$$

Donc, comme $\bar{\Omega} = \emptyset$, on a $P(\emptyset) = 0$

$$A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A), \text{ donc } P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\text{Si } A \subset B, \text{ comme } B = A \cup (B \setminus A) \text{ alors } P(B) = P(A) + P(B \setminus A) \geq P(A)$$

Remarque I.4. 1. Une probabilité est positive. Toutes les sommes qui interviennent s'inscrivent dans le cadre des familles sommables positives. Elles peuvent donc se calculer en identifiant une bijection entre l'ensemble sur lequel on somme et \mathbb{N} ou par théorème de sommation par paquets positifs.

2. Pour $n \in \mathbb{N}$, en appliquant la σ -additivité à la suite $(A_n)_{n \in \mathbb{N}}$ dont tous les termes sont vides à partir du rang $n + 1$, on retrouve la σ -additivité finie de sup :

$$\text{Pour une famille } (A_1, \dots, A_n) \text{ d'événements deux à deux incompatibles : } P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$$

3. Dans le cas d'une suite d'événements deux à deux incompatibles $(A_i)_{i \in I}$, on a :

$$\forall i \in I, P(A_i) \geq 0 \text{ et } \bigcup_{i \in I} A_i \subset \Omega \text{ donc } \sum_{i \in I} P(A_i) \leq P(\Omega) \leq 1$$

Ce qui assure la sommabilité de la famille $(P(A_i))_{i \in I}$ dans la définition d'une probabilité donnée ci-dessus.

Théorème :

Soit (Ω, \mathcal{A}, P) un espace probabilisé.

1. Pour toute suite croissante $(A_n)_{n \in \mathbb{N}}$ d'événements de \mathcal{A} :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n) \text{ (Continuité croissante)}$$

2. Pour toute suite décroissante $(A_n)_{n \in \mathbb{N}}$ d'événements de \mathcal{A} :

$$P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n) \text{ (Continuité décroissante)}$$

3. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements de \mathcal{A} :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$$

Preuve

Pour la continuité croissante : Si (A_n) croissante, alors pour tout $n \in \mathbb{N}$, $A_n \subset A_{n+1}$

Pour un $n \in \mathbb{N}$ fixé, on a $P(\bigcup_{k=0}^n A_k) = P(A_n)$

Posons $C_0 = A_0$. Pour $n \in \mathbb{N}^*$, posons $C_n = A_n \setminus A_{n-1}$

Pour tout $i, j \in \mathbb{N}$, on a : $i \neq j \Rightarrow C_i \cap C_j = \emptyset$

On va montrer que $\bigcup_{n \in \mathbb{N}} C_n = \bigcup_{n \in \mathbb{N}} A_n$

$\forall n \in \mathbb{N}$, $C_n \subset A_n$ et donc $\bigcup_{n \in \mathbb{N}} C_n \subset \bigcup_{n \in \mathbb{N}} A_n$

Réciproquement, soit $x \in \bigcup_{n \in \mathbb{N}} A_n$, on a alors $\{n \in \mathbb{N} | x \in A_n\} \neq \emptyset$

On prend $n_0 = \min\{n \in \mathbb{N} | x \in A_n\}$

Si $n_0 = 0$ alors $x \in A_0$ donc $x \in C_0$

Sinon, $x \in A_{n_0}$ et $x \notin A_{n_0-1}$ donc $x \in C_{n_0}$

Donc $\bigcup_{n \in \mathbb{N}} A_n \subset \bigcup_{n \in \mathbb{N}} C_n$

$$P(\bigcup_{n \in \mathbb{N}} C_n) = \sum_{n=0}^{+\infty} P(C_n) = P(C_0) + \sum_{n=1}^{+\infty} P(A_n) - P(A_{n-1}) = \lim_{n \rightarrow +\infty} P(A_n)$$

D'où le résultat.

Pour la continuité décroissante : On peut passer au complémentaire de la suite, et on fait sa réunion en se servant du théorème croissant.

$$P(\bigcap_{n=0}^{+\infty} A_n) = P(\bigcup_{n=0}^{+\infty} \overline{A_n}) = 1 - \lim_{n \rightarrow +\infty} P(A_n) = 1 - P(\bigcap_{n=0}^{+\infty} A_n)$$

D'où le résultat.

Pour la sous-additivité : (A_n) une suite quelconque d'éléments de la tribu \mathcal{A}

On pose $C_n = \bigcup_{k=0}^n A_k$ suite croissante

$$\text{Donc } P(\bigcup_{n \in \mathbb{N}} C_n) = \lim_{n \rightarrow +\infty} P(C_n)$$

$$\text{Pour } n \in \mathbb{N}, P(C_n) = P(A_0) + \sum_{k=0}^n P(C_k \setminus C_{k-1}) \leq \sum_{k=0}^n P(A_k)$$

$$\text{Donc par passage à la limite monotone : } \lim_{n \rightarrow +\infty} P(C_n) \leq \sum_{n \in \mathbb{N}} P(A_n)$$

Remarque I.5. Pour appliquer les théorèmes précédents à une suite $(A_n)_{n \in \mathbb{N}}$ non nécessairement monotone, on peut considérer :

1. La suite $(\bigcup_{k=0}^n A_k)_{n \in \mathbb{N}}$ qui est croissante.
2. La suite $(\bigcap_{k=0}^n A_k)_{n \in \mathbb{N}}$ qui est décroissante.

On obtient alors :

1. $P(\bigcup_{n=0}^{+\infty} A_n) = \lim_{n \rightarrow +\infty} P(\bigcup_{k=0}^n A_k)$
2. $P(\bigcap_{n=0}^{+\infty} A_k) = \lim_{n \rightarrow +\infty} P(\bigcap_{k=0}^n A_k)$

Définition :

Soit (Ω, \mathcal{A}, P) un espace probabilisé et A un événement.

- Si $A \neq \emptyset$ et $P(A) = 0$, on dit que l'événement A est négligeable ou quasi-impossible.
- Si $A \neq \Omega$ et $P(A) = 1$, on dit que l'événement A est presque sûr ou quasi-certain.

I.3 Univers finis, équiprobabilité, dénombrement

Proposition :

On suppose qu'il existe $p \in \mathbb{R}_+^*$ tel que

$$\forall \omega \in \Omega, P(\{\omega\}) = p$$

Alors :

— Ω est un ensemble fini et

$$p = \frac{1}{\text{Card}(\Omega)}$$

— Pour tout événement A ,

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

Remarque I.6. Cette formule est plus connue sous la forme :

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

On dit qu'on est en situation d'équiprobabilité.

Il s'agit d'une des rares situations où une description de Ω est importante puisqu'il faudra en dénombrer les éléments.

On se fixe un ensemble E de cardinal $n \in \mathbb{N}^*$.

On veut compter des objets fabriqués à partir d'éléments de E , par exemple des prélèvements dans une urne. Il faut modéliser ces objets par des objets mathématiques adéquats. Les deux objets mathématiques rencontrés principalement en dénombrement sont les parties (ou combinaisons) et les listes (ou uplets, suites finies, familles).

Compter des parties

Une partie de cardinal $k \in \mathbb{N}$ de E est une k -combinaison. On peut voir une k -combinaison comme un prélèvement simultané de k éléments de E ; donc sans tenir compte ni d'ordre de tirage, ni de répétition.

Proposition : Nombre de parties

Soit E un ensemble à n éléments

1. Le nombre de k -combinaisons de E est

$$\binom{n}{k} = \frac{1}{k!} n(n-1)\dots(n-k+1) = \begin{cases} \frac{n!}{k!(n-k)!} & \text{si } k \leq n \\ 0 & \text{sinon} \end{cases}$$

2. Le nombre de parties de E est

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

Compter des listes

Soit $k \in \mathbb{N}^*$. L'ensemble des k -listes d'éléments de E est le produit cartésien E^k . On distingue deux cas particuliers de listes :

- Une k -liste sans répétition (ou k -arrangement) est un élément $(x_1, \dots, x_k) \in E^k$ où $x_i \neq x_j$ si $i \neq j$. On rencontre des k -listes sans répétition quand par exemple on modélise des tirages successifs sans remise.

- Une permutation de E est une n -liste sans répétition de l'ensemble E (de cardinal n)
 On peut voir aussi une permutation de E comme une bijection de $\llbracket 1, n \rrbracket$ dans E ou une façon de réordonner les éléments de E .

Proposition : Nombre de listes

Soit E un ensemble à n éléments.

1. Soient E_1, \dots, E_k k ensembles de cardinaux respectifs $n_1, \dots, n_k \in \mathbb{N}^*$

$$\text{Card}(E_1 \times E_2 \times \dots \times E_k) = n_1 n_2 \dots n_k$$

2. Le nombre de k -listes sans répétitions d'éléments de E est $A_n^k = n(n-1)\dots(n-k+1)$
 3. Le nombre de permutations d'éléments de E est $n!$

I.4 Probabilité conditionnelle, indépendance

Les définitions et les théorèmes concernant les probabilités conditionnelles dans un univers infini sont identiques à celles d'un univers fini étudiés en sup.

Définition :

Soit (Ω, \mathcal{A}, P) un espace probabilisé et A un événement.

On peut définir la probabilité conditionnelle de A sachant B notée $P(A/B)$ ou $P_B(A)$:

- Si B un événement de probabilité non-nulle :

$$P(A/B) = P_B(A) = \frac{P(A \cap B)}{P(B)}$$

- Si $P(B) = 0$

$$P_B(A) = 0$$

Proposition :

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Soit B un événement de probabilité non-nulle.
 P_B est une probabilité sur (Ω, \mathcal{A}) appelée probabilité conditionnelle sachant B .

Preuve

P_B est définie sur \mathcal{A} et vérifie :

Soit $A \in \mathcal{A}$, on a $A \cap B \subset B$ et P croissante donc $0 \leq P(A \cap B) \leq P(B)$

Ce qui donne $P_B(A) = \frac{P(A \cap B)}{P(B)} \in [0, 1]$

On a $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

Soient A_1, A_2 deux événements incompatibles :

$$P_B(A_1 \cup A_2) = \frac{P((A_1 \cup A_2) \cap B)}{P(B)} = \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}$$

Or $(A_1 \cap B) \cap (A_2 \cap B) = (A_1 \cap A_2) \cap B = \emptyset \cap B = \emptyset$

Donc par additivité de P :

$$P_B(A_1 \cup A_2) = \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} = \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} = P_B(A_1) + P_B(A_2)$$

La σ -additivité se déduit de cette propriété par récurrence.

Et donc on a bien que P_B est une probabilité sur (Ω, \mathcal{A})

Remarque I.7. Considérer la probabilité d'un événement sachant B , c'est considérer que B est réalisé, c'est à dire limiter l'univers à B .

Proposition :

Soient A et B deux événements d'un espace probabilisé (Ω, \mathcal{A}, P) . Alors si $P(A) \neq 0$ et $P(B) \neq 0$:

$$P(A \cap B) = P(A)P_A(B) = P(B)P_B(A)$$

Théorème : Formule des probabilités composées

Soit $(A_n)_{n \in \mathbb{N}}$ une famille d'événements telle que pour tout entier n , $P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq 0$. Alors, pour tout entier n :

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P_{A_1}(A_2) \dots P_{A_1 \cap A_2 \cap \dots \cap A_{n-1}}(A_n)$$

$$P\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} P(A_1)P_{A_1}(A_2) \dots P_{A_1 \cap \dots \cap A_{n-1}}(A_n)$$

Preuve

Si $P(A_1 \cap \dots \cap A_{n-1}) = 0$, alors la formule dit que $0 = 0$, ce qui est vrai.

Supposons que $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$

Remarquons que $A_1 \cap A_2 \cap \dots \cap A_{n-1} \subset A_1 \cap A_2 \cap \dots \cap A_{n-2} \subset \dots \subset A_1$

Donc par croissance de P , les événements qui conditionnent sont tous de probabilités non-nulles et on peut appliquer la formule usuelle d'une probabilité conditionnelle.

Procédons par récurrence sur n .

Si $n = 1$: la formule s'écrit $P(A_1) = P(A_1)$, ce qui est vrai.

Si la formule est vraie pour n événements :

Soient A_1, \dots, A_{n+1} $n + 1$ événements. En appliquant l'hypothèse de récurrence aux n événements $A_1 \cap A_2, A_3, \dots, A_{n+1}$, on obtient :

$$P(A_1 \cap A_2 \cap \dots \cap A_n \cap A_{n+1}) = P(A_1 \cap A_2)P_{A_1 \cap A_2}(A_3) \dots P_{A_1 \cap A_2 \cap \dots \cap A_n}(A_{n+1})$$

$$\text{Mais } P(A_1 \cap A_2) = P(A_1)P_{A_1}(A_2)$$

D'où la récurrence.

Les définitions de l'indépendance et de l'indépendance mutuelle sont identiques à celles du cas d'un univers fini étudiées en sup, on y ajoute simplement le cas d'une famille infinie d'événements.

Définition :

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Soient A et B deux événements. On dit que A et B sont indépendants lorsque :

$$P(A \cap B) = P(A)P(B)$$

Remarque I.8. — Si $P(A) = 0$ ou $P(B) = 0$, A et B sont indépendants.

- Il s'agit de ne pas confondre les notions d'événements indépendants et incompatibles. Si A et B sont de probabilités non-nulles et disjoints, alors ils ne sont pas indépendants, puisque $P(A)P(B)$ est non-nul alors que $P(A \cap B)$ l'est (intuitivement, A et \bar{A} ne sont pas indépendants.)
- Si $P(B) \neq 0$, A et B sont indépendants si, et seulement si, $P_B(A) = P(A)$
ie : la réalisation de A est indépendante de la réalisation de B
- La notion formelle d'indépendance recouvre donc la notion intuitive, et dans de nombreux exercices, l'indépendance des événements proposés ira de soi, elle fera a priori parité des hypothèses (même si certains énoncés l'oublient, il faut alors le préciser). Par exemple, lorsque l'on lance une pièce de monnaie équilibrée 5 fois de suite, l'indépendance des événements "il y a deux fois pile exactement

parmi les trois premiers lancers" et "il y a une fois face exactement parmi les deux derniers lancers" est parfaitement intuitive, on peut la vérifier facilement par le calcul

- La notion d'indépendance formelle est plus large que la notion intuitive. Dans certains cas, l'indépendance formelle peut sembler paradoxale, seule la vérification de $P(A)P(B) = P(A \cap B)$ peut prouver l'indépendance des événements A et B .

Définition :

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Soit $(A_i)_{i \in I}$ une famille au plus dénombrable d'événements

- Les événements $(A_i)_{i \in I}$ sont indépendants deux à deux si :

$$\forall i, j \in I, i \neq j \Rightarrow P(A_i \cap A_j) = P(A_i)P(A_j)$$
- Les événements $(A_i)_{i \in I}$ sont indépendants ou mutuellement indépendants si pour toute famille finie $J \subset I$ on a :

$$P(\bigcap_{i \in J} A_i) = \prod_{i \in J} P(A_i)$$

Remarque I.9. — La notion de mutuelle indépendance est plus forte que la notion d'indépendance deux à deux. Des événements mutuellement indépendants sont deux à deux indépendants, mais la réciproque est fautive.

- Si A et B sont indépendants alors A et \overline{B} le sont aussi.
- Généralisation : Si $(A_i)_{i \in I}$ est une famille d'événements mutuellement indépendants alors les événements de toute famille $(B_i)_{i \in I}$ telle que pour tout $i \in I$, $B_i \in \{A_i, \overline{A_i}\}$ sont mutuellement indépendants.

Preuve

Si A, B indépendants :

$A = A \cap (B \cup \overline{B}) = (A \cap B) \cup (A \cap \overline{B})$, et l'union est disjointe

$P(A \cap \overline{B}) = P(A) - P(A \cap B)$

$P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B})$

I dénombrable, $(A_i)_{i \in I}$ famille d'événements mutuellement indépendants

Alors la famille $(B_i)_{i \in I}$ qui vaut A_i ou $\overline{A_i}$ est faite d'événements mutuellement indépendants. Le résultat se prouve par récurrence sur le nombre de complémentaires.

Système complet d'événements, quasi-complet d'événements

La formule des probabilités composées permet de gérer les enchaînements de phénomènes aléatoires. On s'intéresse maintenant à la modélisation d'un phénomène aléatoire qui peut être la conséquence de plusieurs causes.

La notion de système complet d'événements et la formule des probabilités totales ont été vues en sup dans le cas d'un nombre fini d'événements. Généralisons à des systèmes complets d'événements infinis.

Définition :

Un système complet d'événements est une famille $(A_i)_{i \in I}$ au plus dénombrable d'événements tels que :

- les événements sont deux à deux incompatibles ($\forall i, j \in I, i \neq j \Rightarrow A_i \cap A_j = \emptyset$)
- leur union est l'univers tout entier ($\bigcup_{i \in I} A_i = \Omega$)

Remarque I.10. Autrement dit, $(A_i)_{i \in I}$ est un système complet d'événement si, et seulement si, à chaque réalisation de l'expérience aléatoire, un et un seul des événements A_i est réalisé.

De cette définition, il sort qu'un événement et son contraire forment un système complet.

Théorème :

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Pour tout système complet d'événements $(A_i)_{i \in I}$, on a :

$$\sum_{i \in I} P(A_i) = 1$$

Définition :

Un système quasi-complet d'événements est une famille $(A_i)_{i \in I}$ au plus dénombrable d'événements tels que :

- les événements sont incompatibles deux à deux ($\forall i, j \in I, i \neq j \Rightarrow A_i \cap A_j = \emptyset$)
- leur union est presque sûre : $P(\bigcup_{i \in I} A_i) = 1$

Formule des probabilités totales

Cette formule permet de gérer les disjonctions de cas (les causes (A_i)) qui aboutissent à la réalisation d'un événement (la conséquence B)

Théorème : Formule des probabilités totales

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Soit $(A_i)_{i \in I}$ un système complet ou quasi-complet d'événements. Pour tout événement B :

$$P(B) = \sum_{i \in I} P(A_i \cap B) = \sum_{i \in I} P(A_i)P_{A_i}(B)$$

Preuve

La preuve est contenue dans l'énoncé ou presque

Pour $B \in \mathcal{A}$, on a :

$$B = B \cap \Omega = B \cap (\bigcup_{i \in I} A_i) = \bigcup_{i \in I} B \cap A_i$$

$$\text{Et donc } P(B) = \sum_{i \in I} P(B \cap A_i)$$

Pour la dernière égalité, on rappelle juste la convention que si $P(A_i) = 0$, alors $P_{A_i}(B) = 0$

Remarque I.11. La formule reste valable pour un système complet d'événements fini, le cas le plus fréquent d'utilisation étant (A, \bar{A})

Formule de Bayes

Sachant que la conséquence B est réalisée, on veut maintenant connaître la probabilité d'une des causes.

Théorème : Formule d'inversion de Bayes

Soit $(A_i)_{i \in I}$ un système complet ou quasi-complet d'événements d'un espace probabilisé (Ω, \mathcal{A}, P) . Soit B un événement de probabilité non-nulle, soit $i_0 \in I$.

Alors

$$P_B(A_{i_0}) = \frac{P(B \cap A_{i_0})}{P(B)} = \frac{P(A_{i_0})P_{A_{i_0}}(B)}{\sum_{i \in I} P(A_i)P_{A_i}(B)}$$

Preuve

Pour un événement B quelconque de probabilité non-nulle, et la formule fonctionne immédiatement en utilisant les propriétés vues plus tôt sur les probabilités conditionnelles : $P(A \cap B) = P(A)P_A(B) = P(B)P_B(A)$ et la formule des probabilités totales au dénominateur.

II Variables aléatoires discrètes

II.1 Définition d'une variable aléatoire discrète

Deux exemples de sup : on lance simultanément deux dés discernables et on choisit comme univers $\llbracket 1, 6 \rrbracket^2$, que l'on munit de la probabilité uniforme. Notons X la somme des valeurs des dés et Y le maximum des deux valeurs.

— X peut prendre toutes les valeurs entières entre 2 et 12, Y entre 1 et 6.

— On peut voir X et Y comme des fonctions de Ω dans \mathbb{R} :

$$X : (\omega_1, \omega_2) \mapsto \omega_1 + \omega_2 \text{ et } Y : (\omega_1, \omega_2) \mapsto \max(\omega_1, \omega_2)$$

On peut alors écrire $X(\Omega) = \llbracket 2, 12 \rrbracket$ et $Y(\Omega) = \llbracket 1, 6 \rrbracket$

— Notons A l'événement "la somme des 2 dés vaut 5" : $A = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$ et donc

$$P(A) = P(X = 5) = \frac{4}{36} = \frac{1}{9}$$

Dans l'égalité précédente, on a noté $(X = 5)$ l'événement A . Il convient de voir $(X = 5)$ comme une abbréviation pratique de $X^{-1}(\{5\})$.

— Notons B l'événement "la somme des 2 dés est inférieure à 4" : $B = \{(1, 1), (2, 1), (3, 1), (2, 2), (1, 3)\}$ et donc

$$P(B) = P(X \leq 4) = \frac{6}{36} = \frac{1}{6}$$

Dans l'égalité précédente, on a noté $(X \leq 4)$ l'événement B . Il convient de voir $(X \leq 4)$ comme une abbréviation pratique de $X^{-1}(]-\infty, 4])$

Définition :

Soit E un ensemble et (Ω, \mathcal{A}) un espace probabilisable.

Une application $X : \Omega \rightarrow E$ est une variable aléatoire discrète si :

- L'ensemble $X(\Omega)$ des valeurs prises par X est au plus dénombrable.
- Pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\})$, noté $(X = x)$ ou $[X = x]$, est un événement (un élément de \mathcal{A})

Lorsque $E = \mathbb{R}$, la variable X est dite réelle.

Si U est une partie de E , on note :

$$(X \in U) = \{\omega \in \Omega \mid X(\omega) \in U\} = X^{-1}(U)$$

On note aussi $[X \in U]$ ou $\{X \in U\}$

Théorème :

Si X est une variable aléatoire discrète sur l'espace probabilisable (Ω, \mathcal{A}) , alors la suite $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements de Ω

Preuve

Pour $U \in X(\Omega)$, $X \in U = \bigcup_{x \in U} (X = x)$.

D'où le théorème.

II.2 Loi d'une variable aléatoire discrète

Définition :

Soit X une variable aléatoire discrète sur un espace probabilisé, à valeurs dans un ensemble E . La loi P_X de X est la donnée de :

- l'ensemble des valeurs prises par X appelé univers image : $X(\Omega)$
- les probabilités élémentaires : $p_x = P(X = x)$ pour tout $x \in X(\Omega)$

Remarque II.1. On note $X \sim Y$ lorsque les variables X et Y suivent la même loi ($P_X = P_Y$)

Théorème :

Si X une variable aléatoire discrète à valeurs dans un ensemble E ,

- la famille $(P(X = x))_{x \in E}$ est sommable de somme 1 :

$$\sum_{x \in E} P(X = x) = 1$$

- par σ -additivité, pour toute partie U de $X(\Omega)$,

$$P(X \in U) = P(\bigcup_{x \in U} (X = x)) = \sum_{x \in U} P(X = x)$$

- $(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$ est un espace probabilisé.

Réciproquement, si $(p_x)_{x \in E}$ est une famille sommable de réels positifs de somme 1 alors il existe une variable aléatoire discrète X telle que pour tout $x \in E$, $P(X = x) = p_x$

Preuve

$\Omega = \bigcup_{x \in X(\Omega)} (X = x)$ dénombrable, et alors $P(\Omega) = 1 = \sum_{x \in X(\Omega)} P(X = x)$

Si $(U_n)_{n \in \mathbb{N}} \in \mathcal{P}(X(\Omega))^{\mathbb{N}}$ disjoints deux à deux :

$$\begin{aligned} P_X\left(\bigcup_{n \in \mathbb{N}} U_n\right) &= P(X^{-1}\left(\bigcup_{n \in \mathbb{N}} U_n\right)) \\ &= P\left(\bigcup_{n \in \mathbb{N}} (X^{-1}(U_n))\right) \\ &= \sum_{n \in \mathbb{N}} P(X^{-1}(U_n)) \\ &= \sum_{n \in \mathbb{N}} P_X(U_n) \end{aligned}$$

Pour vérifier que $(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$ est un espace probabilisé, on se sert du fait que $X(\Omega)$ est dénombrable.

Proposition :

Soient X une variable aléatoire discrète à valeurs dans un ensemble E et f une fonction de E dans un ensemble F définie sur $X(\Omega)$

La fonction $Y = f \circ X : \begin{cases} \Omega & \rightarrow F \\ \omega & \mapsto f(X(\omega)) \end{cases}$

Proposition :

Soit X une variable aléatoire discrète et $Y = f(X)$ où $f : X(\Omega) \rightarrow F$
On a :

$$\forall y \in Y(\Omega), P(Y = y) = \sum_{x \in X(\Omega), f(x)=y} P(X = x)$$

II.3 Indépendance et variables aléatoires discrètes**Définition :**

X et Y sont deux variables aléatoires discrètes indépendantes lorsque pour toutes parties A et B de $X(\Omega)$ et $Y(\Omega)$ respectivement :

$$P((X \in A) \cap (Y \in B)) = P(X \in A)P(Y \in B)$$

On note $X \perp Y$

Théorème :

Deux variables aléatoires discrètes X et Y sont dites indépendantes si, et seulement si,

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), P((X = x) \cap (Y = y)) = P(X = x)P(Y = y)$$

ie : $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants.

Preuve

Ce théorème est admis.

Démontrons-le quand même.

L'indépendance implique la propriété.

Supposons maintenant : $\forall x \in X(\Omega), \forall y \in Y(\Omega), P(X = x \cap Y = y) = P(X = x)P(Y = y)$

Soit $A \subset \mathcal{P}(X(\Omega))$ et $B \subset \mathcal{P}(Y(\Omega))$

Alors $P(X \in A \cap Y \in B) = P((\bigcup_{a \in A} (X = a)) \cap (\bigcup_{b \in B} (Y = b)))$

$$= P\left(\bigcup_{a \in A, b \in B} (X = a) \cap (Y = b)\right) = \sum_{a \in A, b \in B} P((X = a) \cap (Y = b)) = \sum_{a \in A, b \in B} P(X = a)P(Y = b)$$

$$= \left(\sum_{a \in A} P(X = a)\right) \left(\sum_{b \in B} P(Y = b)\right)$$

$$= P(X \in A)P(Y \in B)$$

Définition :

Les variables aléatoires discrètes X_1, X_2, \dots, X_n sont mutuellement indépendantes lorsque pour toutes parties A_1, A_2, \dots, A_n de $X_1(\Omega), X_2(\Omega), \dots, X_n(\Omega)$:

$$P\left(\bigcap_{i=1}^n (X_i \in A_i)\right) = \prod_{i=1}^n P(X_i \in A_i)$$

Théorème :

Les variables aléatoires discrètes X_1, X_2, \dots, X_n sont (mutuellement) indépendantes si, et seulement si, pour tout n -uplet (x_1, x_2, \dots, x_n) de $X_1(\Omega) \times X_2(\Omega) \times \dots \times X_n(\Omega)$:

$$P\left(\bigcap_{i=1}^n (X_i = x_i)\right) = \prod_{i=1}^n P(X_i = x_i)$$

Preuve

Ce théorème est aussi admis.

Si on ne l'admettait pas, on devrait faire une récurrence avec la propriété vérifiée précédemment.

Définition :

Une suite de variables aléatoires discrètes $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires discrètes (mutuellement) indépendantes lorsque pour toute partie finie $I = \{i_1, \dots, i_n\}$ de \mathbb{N} , les variables $X_{i_1}, X_{i_2}, \dots, X_{i_n}$ sont (mutuellement) indépendantes.

Proposition :

Soient X et Y deux variables aléatoires discrètes définies sur Ω à valeurs dans E .

Soient f et g deux applications de E dans F définies respectivement sur $X(\Omega)$ et $Y(\Omega)$

Si X et Y sont indépendantes alors $f(X)$ et $g(Y)$ le sont aussi.

Proposition :

Soient X_1, \dots, X_n n variables aléatoires discrètes définies sur Ω à valeurs dans E .

Soient f_1, \dots, f_n n applications de E dans F .

Si X_1, \dots, X_n sont (mutuellement) indépendantes alors $f_1(X_1), \dots, f_n(X_n)$ le sont aussi.

Preuve

On admet le premier résultat. Le second en découle par récurrence.

Proposition : Lemme des coalitions

Soient X_1, X_2, \dots, X_n n variables aléatoires discrètes définies sur Ω à valeurs dans E .

Soient f et g deux applications respectivement de E^m dans F et de E^{n-m} dans F .

Si X_1, X_2, \dots, X_n sont (mutuellement) indépendantes alors $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ le sont aussi.

Ce théorème s'étend à plus de deux coalitions.

Preuve

Admis

II.4 Espérance d'une variable aléatoire discrète réelle ou complexe

Rappel de sup : dans le cas où Ω est fini et $\mathcal{A} = \mathcal{P}(\Omega)$, la moyenne des valeurs prises par X sur Ω est :

$$m = \frac{\sum_{\omega \in \Omega} X(\omega)}{\text{Card}(\Omega)}$$

- Si P est la probabilité uniforme sur Ω , $\forall \omega \in \Omega$, $P(\{\omega\}) = \frac{1}{\text{Card}(\Omega)}$ et $m = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\})$
- Si P est une probabilité quelconque, $\sum_{\omega \in \Omega} X(\omega)P(\{\omega\})$ est la moyenne des valeurs de $X(\omega)$ pondérée par la probabilité de ω . Cette moyenne est appelée espérance de X , notée $E(X)$

En regroupant les valeurs de ω pour lesquelles X prend une valeur fixée, on obtient une autre expression de $E(X)$:

$$E(X) = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) = \sum_{x \in X(\Omega)} x \sum_{\omega \in X^{-1}(\{x\})} P(\{\omega\}) = \sum_{x \in X(\Omega)} xP(X^{-1}(x)) = \sum_{x \in X(\Omega)} xP(X = x)$$

C'est cette dernière expression qui est utilisée pour définir l'espérance dans le cas général.

Définition :

Soit X une variable aléatoire discrète à valeurs dans $[0, +\infty]$.

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x)$$

Avec la convention $xP(X = x) = 0$ lorsque $X = +\infty$ et $P(X = +\infty) = 0$

Définition :

Soit X une variable aléatoire discrète à valeurs réelles ou complexes.

On dit que X admet une espérance (ou est d'espérance finie) lorsque la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas :

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x)$$

X est dite centrée lorsque $E(X) = 0$

Remarque II.2. La définition de l'espérance s'étend sans difficulté aux variables aléatoires à valeurs dans un espace vectoriel de dimension finie, en passant par les variables aléatoires coordonnées. L'expression

$$E(X) = \sum_{x \in X(\Omega)} P(X = x)x$$

reste valable (où x est un vecteur, $P(X = x)$ un scalaire)

Proposition :

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$. On a :

$$E(X) = \sum_{n \in \mathbb{N}} P(X \geq n)$$

Preuve

$$\begin{aligned}
E(X) &= \sum_{n \in \mathbb{N}} nP(X = n) \\
&= \sum_{n \in \mathbb{N}} \sum_{k=0}^{n-1} P(X = n) \\
&= \sum_{0 \leq k < n, (k, n) \in \mathbb{N}^2} P(X = n) \\
&= \sum_{k \in \mathbb{N}} \sum_{n < k} P(X = n) \\
&= \sum_{k \in \mathbb{N}} P(X > k) \\
&= \sum_{k \in \mathbb{N}^*} P(X \geq k)
\end{aligned}$$

Calculer l'espérance d'une variable aléatoire réelle à partir de sa définition nécessite donc de connaître sa loi, ce qui peut être délicat lorsque la variable est de la forme $f(X)$. Le théorème suivant permet de calculer l'espérance de $f(X)$ à partir de la loi de X :

Théorème : Transfert

Soit X une variable aléatoire discrète à valeurs réelles ou complexes et $f : X(\Omega) \rightarrow \mathbb{C}$

La variable aléatoire réelle $f(X)$ est d'espérance finie si, et seulement si, la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.

Dans ce cas :

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x)$$

Remarque II.3. On admet aussi que la formule de transfert s'applique aux couples, au n -uplets de variables aléatoires.

Par exemple pour deux variables aléatoires discrètes X et Y à valeurs réelles ou complexes et la fonction $f : (x, y) \mapsto xy$:

— XY est d'espérance finie si, et seulement si, la famille $(f(x, y)P((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est sommable.

— Dans le cas sommable :

$$E(XY) = \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} f(x, y)P((X, Y) = (x, y)) = \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} xyP(X = x, Y = y)$$

Théorème : Propriétés de l'espérance

Soient X, Y deux variables aléatoires discrètes à valeurs réelles ou complexes et d'espérance finie.

— $\forall \lambda, \mu \in \mathbb{C}, \lambda X + \mu Y$ est d'espérance finie et

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y) \text{ (linéarité)}$$

On peut généraliser cette propriété avec n variables aléatoires réelles discrètes d'espérance finie, qu'elles soient indépendantes ou non.

Dans la suite, X et Y sont à valeurs réelles.

— Si $X \geq 0$ alors $E(X) \geq 0$ (Positivité)

— Si $X \geq 0$ et $E(X) = 0$ alors $(X = 0)$ est presque sûr (stricte positivité)

— Si $X \leq Y$ alors $E(X) \leq E(Y)$ (croissance)

Preuve

Linéarité de l'espérance : Si X et Y admettent une espérance :

$X + Y$ a une espérance si, et seulement si, $\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |x + y| P(X = x, Y = y) < +\infty$

$$\begin{aligned} \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |x + y| P(X = x, Y = y) &\leq \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (|x| + |y|) P(X = x, Y = y) \\ &\leq \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |x| P(X = x) + \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |y| P(Y = y) \\ &\leq \sum_{x \in X(\Omega)} |x| \sum_{x \in X(\Omega)} P(X = x, Y = y) + \sum_{y \in Y(\Omega)} P(X = x, Y = y) \end{aligned}$$

Or $(X = x) = \bigcup_{y \in Y(\Omega)} (X = x) \cap (Y = y)$ et réciproquement.

D'où la majoration du terme précédent :

$$\leq \sum_{x \in X(\Omega)} |x| P(X = x) + \sum_{y \in Y(\Omega)} |y| P(Y = y) < +\infty$$

Avec la sommabilité, en se débarrassant de la valeur absolue du début, on remplace les inégalités larges par des égalités, et on a que $E(X + Y) = E(X) + E(Y)$.

Se fait de même pour la multiplication par un scalaire.

Pour la stricte positivité :

$E(X) = 0 \Rightarrow (X = 0)$ presque sûrement

$$E(X) = \sum_{x \in X(\Omega)} x P(X = x) = 0$$

Comme $x P(X = x) \geq 0$, alors $\forall x \in X(\Omega), x = 0$ ou $P(X = x) = 0$

$$\text{Donc } P(X \neq 0) = \sum_{x \in X(\Omega), x \neq 0} x P(X = x) = 0$$

Remarque II.4. Dans le cadre du théorème précédent :

- On a en particulier que $\forall \lambda, \mu \in \mathbb{C}, \lambda X + \mu$ admet une espérance et $E(\lambda X + \mu) = \lambda E(X) + \mu$
- Pour X réelle et $a, b \in \mathbb{R}$, si $a \leq X \leq b$ alors $a \leq E(X) \leq b$

Théorème : Critère de majoration positif

Soit X une variable aléatoire discrète à valeurs réelles ou complexes. Soit Y une variable aléatoire discrète à valeurs réelles.

Si on a :

— Y est d'espérance finie

— $|X| \leq Y$

Alors X est d'espérance finie.

Théorème : Espérance d'un produit de variables indépendantes

Soit X, Y deux variables aléatoires discrètes à valeurs réelles ou complexes.

Si on a :

- X et Y d'espérance finie
(X et Y admettent une espérance serait plus adapté)
- X et Y indépendantes

Alors XY est d'espérance finie et

$$E(XY) = E(X)E(Y)$$

Preuve

$$\begin{aligned} E(|XY|) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |xy| P(X=x, Y=y) \\ &= \sum_{x,y} |xy| P(X=x)P(Y=y) = \sum_{x \in X(\Omega)} |x| P(X=x) \sum_{y \in Y(\Omega)} |y| P(Y=y) \\ &= E(|X|)E(|Y|) < +\infty \text{ par hypothèse} \end{aligned}$$

On a donc que $(xyP(X=x)P(Y=y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable donc par sommation par paquets :

$$E(XY) = \sum_{x \in X(\Omega)} xP(X=x) \sum_{y \in Y(\Omega)} yP(Y=y) = E(X)E(Y)$$

Théorème : Espérance d'un produit de n variables indépendantes

Soit X_1, \dots, X_n des variables aléatoires discrètes à valeurs réelles ou complexes.

Si on a :

- X_1, \dots, X_n d'espérance finie
- X_1, \dots, X_n (mutuellement) indépendantes

Alors la variable $X_1 \dots X_n$ est d'espérance finie et

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i)$$

Pre H_n : $\forall X_1, \dots, X_n$ variables aléatoires complexes mutuellement indépendantes et possédant une espérance, alors $X_1 \dots X_n$ possède une espérance et $E(X_1 \dots X_n) = E(X_1) \dots E(X_n)$

On a vu H_2 dans la preuve précédente

On suppose H_n , prenons X_1, \dots, X_{n+1} complexes mutuellement indépendantes

Par H_n , $(X_2 \dots X_{n+1})$ admet une espérance et $E(X_2 \dots X_{n+1}) = E(X_2) \dots E(X_{n+1})$

Par lemme des coalitions, X_1 et $X_2 \dots X_{n+1}$ sont indépendantes

Par H_2 , $X_1(X_2 \dots X_{n+1})$ admet une espérance et :

$$E(X_1(X_2 \dots X_{n+1})) = E(X_1)E(X_2 \dots X_{n+1}) = \prod_{i=1}^n E(X_i)$$

Remarque II.5. La réciproque de cette proposition est fausse, l'égalité $E(XY) = E(X)E(Y)$ ne suffit pas à garantir l'indépendance des variables X et Y .

Par exemple, X la variable aléatoire de loi uniforme sur $\{-1, 0, 1\}$ et $Y = X^2$

On peut vérifier que $E(XY) = E(X)E(Y)$ mais $P((X, Y) = (0, 0)) = \frac{1}{3}$ tandis que $P(X=0)P(Y=0) = \frac{1}{9}$

II.5 Variance d'une variable aléatoire discrète réelle, écart type

Définition :

X une variable aléatoire complexe, on dit que X admet un moment d'ordre k pour $k \in \mathbb{N}^*$ si $E(X^k)$ existe

Théorème :

Si X admet un moment d'ordre $k + 1$ alors X admet un moment d'ordre k .

Preuve

On suppose que X admet un moment d'ordre $k + 1$ (ie $E(X^{k+1}) < +\infty$)

Alors $E(|X|^k) = \sum_{x \in X(\Omega)} |x|^k P(X = x)$

$$\begin{aligned} E(|x|^k) &\leq \sum_{x \in X(\Omega), |x| \geq 1} |x|^k P(X = x) + \sum_{x \in X(\Omega), |x| < 1} |x|^k P(X = x) \\ &\leq E(|X|^{k+1}) + 1 \end{aligned}$$

Proposition :

Soit X une variable aléatoire discrète à valeurs réelles.

Si X^2 est d'espérance finie alors, pour tout polynôme f de degré au plus 2, la variable aléatoire $f(X)$ est d'espérance finie.

Par exemple, X , $X(X - 1)$ sont d'espérance finie.

Preuve

On suppose que X^2 est d'espérance finie.

Par linéarité, la variable $Y = \frac{1}{2}(X^2 - 1)$ est aussi d'espérance finie.

De plus, on a $|X| \leq Y$, donc par critère de majoration positif, X est d'espérance finie.

On pose $f : x \mapsto a_0 + a_1x + a_2x^2$ où $a_0, a_1, a_2 \in \mathbb{R}$

D'après ce qui précède, X^2 , X et la variable constante égale à 1 sont d'espérance finie, donc par linéarité $f(X) = (a_0 + a_1X + a_2X^2)$ est d'espérance finie.

Définition : Variance et ecart-type

Soit X une variable aléatoire discrète à valeurs réelles.

Si X^2 est d'espérance finie, on définit la variance de X par

$$V(X) = E((X - E(X))^2)$$

et son écart-type par

$$\sigma(X) = \sqrt{V(X)}$$

Remarque II.6. La définition de la variance a un sens, car si X^2 est d'espérance finie, $m = E(X)$ existe et $(X - m)^2 = X^2 - 2mX + m^2$ qui est un polynôme du second degré en X d'espérance finie.

La définition de l'écart-type a aussi un sens, car par positivité de l'espérance, la variance est un réel positif.

Proposition : de la variance

Soit X une variable aléatoire discrète à valeurs réelles.

On suppose que X^2 est d'espérance finie.

1. Köning-Huygens : formule pratique pour la variance :

$$V(X) = E(X^2) - E(X)^2$$

2. Soient $a, b \in \mathbb{R}$

$$V(aX + B) = a^2 V(X)$$

En particulier, la variance est invariante par translation

3. $V(X)$ est nulle si, et seulement si, $P(X = E(X)) = 1$

ie : X est presque sûrement constante.

Preuve

Pour le premier point :

$$\text{On a } V(X) = E((X - E(X))^2) = E(X^2 - 2E(X)X + E(X)^2) = E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - E(X)^2$$

Pour le deuxième point :

$$\text{On a } V(aX + b) = E((aX + b - E(aX + b))^2) = E((aX + b - aE(X) - b)^2) = E(a^2(X - E(X))^2) = a^2 V(X)$$

Preuve

Pour le troisième point : si $V(X) = 0$

$$\text{On pose } m = E(X), \text{ d'après le théorème de transfert : } V(X) = E((X - E(X))^2) = E((X - m)^2) = \sum_{n=0}^{+\infty} (x_n - m)^2 P(X = x_n)$$

$$\text{Donc } \forall n \in \mathbb{N}, 0 \leq (x_n - m)^2 P(X = x_n) \leq \sum_{n=0}^{+\infty} (x_n - m)^2 P(X = x_n) = 0$$

$$\text{Et donc } \forall n \in \mathbb{N}, (x_n - m)^2 P(X = x_n) = 0$$

Or il existe n_0 tel que $P(X = x_{n_0}) \neq 0$ (sinon la somme des probabilités ne vaudrait pas 1)

$$\text{Donc pour un tel } n_0, x_{n_0} = m, \text{ car sinon } (x_{n_0} - m)^2 P(X = x_{n_0}) \neq 0$$

Comme les x_n sont deux à deux distincts : $\forall n \in \mathbb{N} \setminus \{n_0\}, x_n \neq m$

$$\text{D'où } \forall n \in \mathbb{N} \setminus \{n_0\}, P(X = x_n) = 0$$

$$\text{Comme } ((X = x_n))_{n \in \mathbb{N}} \text{ est un système complet d'événements : } P(X = x_{n_0}) = 1 - \sum_{n \in \mathbb{N}, n \neq n_0} P(X = x_n) = 1$$

$$\text{D'où } P(X = x_{n_0}) = P(X = m) = P(X = E(X)) = 1$$

Réciproquement, on suppose que $P(X = E(X)) = 1$

Comme $P(X = E(X)) \neq 0$, $E(X) \in X(\Omega)$, et il existe alors $n_0 \in \mathbb{N}$ tel que $x_{n_0} = E(X)$

On a que pour $n \in \mathbb{N} \setminus \{n_0\}, (X = x_n) \subset \overline{(X = x_{n_0})}$, et par croissance de P :

$$0 \leq P(X = x_n) \leq P(\overline{(X = x_{n_0})}) = 1 - P(X = x_{n_0}) = 1 - 1 = 0$$

$$\text{Donc } \forall n \in \mathbb{N} \setminus \{n_0\}, P(X = x_n) = 0$$

On en déduit par théorème de transfert :

$$V(X) = E((X - E(X))^2) = \sum_{n=0}^{+\infty} (x_n - E(X))^2 P(X = x_n) = 0$$

Définition :

Soit X une variable aléatoire discrète à valeurs réelles telle que X^2 est d'espérance finie et telle que $\sigma(X) > 0$

La variable $\frac{X}{\sigma(X)}$ a un écart-type égal à 1. Elle est appelée réduite de X .

La variable $\frac{X-E(X)}{\sigma(X)}$ a une espérance nulle et un écart-type égal à 1. Elle est appelée variable centrée réduite associée à X .

II.6 Covariance, coefficient de corrélation linéaire

Dans ce paragraphe, X et Y sont deux variables aléatoires discrètes à valeurs réelles.

Proposition : Inégalité de Cauchy-Schwarz

Si X^2 et Y^2 sont d'espérance finie, alors XY l'est aussi et

$$E(XY)^2 \leq E(X^2)E(Y^2)$$

Avec égalité si, et seulement si, $P(X = 0) = 1$ ou il existe $\lambda \in \mathbb{R}$ tel que $P(Y = \lambda X) = 1$

Proposition :

Si les variables aléatoires X^2 et Y^2 sont d'espérance finie alors, pour toute fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ polynomiale de degré au plus 2, la variable aléatoire $f(X, Y)$ est d'espérance finie.

Définition :

Si les variables aléatoires X^2 et Y^2 sont d'espérance finie alors on peut définir la covariance de X et Y par :

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

Preuve

Il suffit de démontrer que $\sum |xy|P(X = x, Y = y)$ est d'espérance finie.

$\forall x \in X(\Omega), \forall y \in Y(\Omega), |xy| \leq \frac{1}{2}(|x|^2 + |y|^2)$

Donc $E(|XY|) < +\infty$ donc $E(XY)$ admet une espérance

De plus X admet une variance donc une espérance. De même pour Y .

Donc $(X - E(X))(Y - E(Y))$ admet une espérance et :

$$E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y)$$

Remarque II.7. La définition de la covariance a un sens, car si X^2 et Y^2 sont d'espérance finie, $m = E(X)$ et $m' = E(Y)$ existent et $(X - m)(Y - m') = XY - m'X - mY + mm'$ qui est un polynôme du second degré en X et Y , et donc d'espérance finie.

La covariance est l'espérance du produit des variables centrées.

Remarque II.8. La covariance est une forme bilinéaire symétrique positive pour l'espace vectoriel des variables aléatoires réelles.

On dispose donc de l'inégalité de Cauchy-Schwarz.

Si E est un \mathbb{R} -ev et φ une forme bilinéaire symétrique sur E , alors :

$$\forall x, y \in E, |\varphi(x, y)| \leq \sqrt{\varphi(x, x)}\sqrt{\varphi(y, y)}$$

On fixe $(x, y) \in E$ et on considère : $P : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ t & \mapsto \varphi(x + ty, x + ty) \end{cases}$

Par bilinéarité : $P(t) = \varphi(x, x) + 2t\varphi(x, y) + t^2\varphi(y, y)$

P est une fonction polynomiale positive.

Si $\varphi(y, y) > 0$: P est polynomiale de degré 2, positive si son discriminant est négatif.

Or P est positive par positivité de φ

Donc $\varphi(x, y)^2 - \varphi(x, x)\varphi(y, y) \leq 0$

Si $\varphi(y, y) = 0$: P est affine. Comme φ est positive, alors P aussi. Donc P est constante (affine et positive, elle vaut $\varphi(x, x)$ qui est alors positive et $\varphi(x, y) = 0$).

Donc $\varphi(x, y) \leq \sqrt{\varphi(x, x)}\sqrt{\varphi(y, y)}$

Définition :

Si X et Y possèdent des variances non-nulles, on peut définir la corrélation de X et Y :

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

D'après Cauchy-Schwarz, on a : $-1 \leq \text{corr}(X, Y) \leq 1$

Proposition : de la covariance

Si les variables aléatoires X^2 et Y^2 sont d'espérance finie alors :

$$\text{--- cov}(X, X) = V(X)$$

--- Köning-Huygens : formule pratique pour la covariance :

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y)$$

--- Si X et Y sont indépendantes alors

$$\text{cov}(X, Y) = 0$$

--- La covariance est une forme linéaire positive, symétrique et bilinéaire (c'est "presque" un produit scalaire)

Proposition : Variance d'une somme

Si les variables aléatoires X^2 et Y^2 sont d'espérance finie alors $(X + Y)^2$ aussi avec

$$V(X + Y) = V(X) + 2\text{cov}(X, Y) + V(Y)$$

Si de plus les variables X et Y sont indépendantes :

$$V(X + Y) = V(X) + V(Y)$$

Proposition : Variance d'une somme

Si X_1, \dots, X_n sont n variables telles que X_1^2, \dots, X_n^2 sont d'espérance finie alors $(X_1 + \dots + X_n)^2$ l'est aussi avec :

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j)$$

Si de plus les variables X_1, \dots, X_n sont indépendantes **deux à deux** alors :

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k)$$

Preuve

$$\begin{aligned} V\left(\sum_{i=1}^n X_i\right) &= \text{cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) \\ &= \sum_{i=1}^n \text{cov}\left(X_i, \sum_{j=1}^n X_j\right) = \sum_{i \leq n, j \leq n} \text{cov}(X_i, X_j) \\ &= \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j) \end{aligned}$$

La loi binomiale $\mathcal{B}(n, p)$ est lme nombre de succès pour une expérience aléatoire de probabilité de succès p répétée n fois, où les expériences sont indépendantes.

Si on appelle X_i la variable aléatoire indicatrice du succès à la i -ème tentative, si $X \rightsquigarrow \mathcal{B}(n, p)$ alors $X = \sum_{i=1}^n X_i$

Les X_i sont indépendantes, $X_i \rightsquigarrow \mathcal{B}(1, p)$, $E(X_i) = p$, $E(X_i^2) = 1$ et $V(X_i) = p - p^2 = p(1 - p)$

$E(X) = np$ et $V(X) = np(1 - p)$

La loi hypergéométrique $\mathcal{H}(N, n, p)$: elle modélise le tirage sans remise de n boules dans une urne contenant N boules, donc Np noires.

X est le nombre de boules noires tirées. L'univers de X est $\llbracket 0, n \rrbracket$.

Pour $k \in X(\Omega)$: $P(X = k) = \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}$

Notons X_i la variable indicatrice du succès au rang i , $X_i \rightsquigarrow \mathcal{B}(1, p)$

On a $X = \sum_{i=1}^n X_i$, et donc $E(X) = np$

Pour i, j , on a que : $P(X_i = 1, X_j = 1) = \frac{Np(Np-1) \binom{n-2}{n-2} (n-2)!}{\binom{N}{n} n!} = \frac{p(Np-1)}{N-1}$

D'où $\text{cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \frac{p(Np-1)}{N-1} - p^2 = \frac{p(p-1)}{N-1}$

$V(X) = np(1 - p) + n(n-1) \frac{p(p-1)}{N-1} = np(1 - p) \left(1 - \frac{n-1}{N-1}\right)$

III Lois usuelles

III.1 Lois usuelles finies

Dans toute la suite on considère $n \in \mathbb{N}^*$, $p \in [0, 1]$ et $q = 1 - p$

Définition :

Une variable aléatoire discrète X suit une loi de Bernoulli de paramètre p si

$$X(\Omega) = \{0, 1\} \text{ et } P(X = 1) = p$$

On note $X \sim \mathcal{B}(p)$

Proposition :

Si $X \sim \mathcal{B}(p)$ alors

$$E(X) = P \text{ et } V(X) = pq$$

En particulier, le paramètre d'une loi de Bernoulli est son espérance.

Preuve

On a $E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = p$

Mais si $X(\Omega) \subset \{0, 1\}$ alors $X^2 = X$. Donc :

$$V(X) = E(X^2) - E(X)^2 = p - p^2 = p(1 - p) = pq$$

Définition :

Une variable aléatoire discrète X suite une loi uniforme sur l'ensemble fini non-vide $K \subset \mathbb{R}$ si

$$X(\Omega) = K \text{ et } \forall k \in K, P(X = k) = \frac{1}{\text{Card}(K)}$$

On note $X \sim \mathcal{U}(K)$

Proposition :

Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$ alors

$$E(X) = \frac{n+1}{2} \text{ et } V(X) = \frac{n^2-1}{12}$$

Preuve

Pour l'espérance : $E(X) = \sum_{i=1}^n n \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$

$$E(X^2) = \sum_{i=1}^n i^2 \frac{1}{n} = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\text{D'où } V(X) = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{4n^2+2n+4n+2-3n^2-6n-3}{12} = \frac{n^2-1}{12}$$

Définition :

Une variable aléatoire discrète X suit une loi binomiale des paramètres n et p si $X(\Omega) = \llbracket 0, n \rrbracket$ et

$$\forall k \in \llbracket 0, n \rrbracket, P(X = k) = \binom{n}{k} p^k q^{n-k}$$

$X \sim \mathcal{B}(n, p)$

Proposition : Modèle loi binomiale

Si X est le nombre de succès lors de la répétition de n épreuves de Bernoulli indépendantes de même paramètre p alors X suit la loi $\mathcal{B}(n, p)$

Formellement :

On pose pour $k \in \llbracket 1, n \rrbracket$ X_k la variable aléatoire discrète qui vaut 1 si la k -ième épreuve est un succès et 0 sinon. Si :

- pour tout $k \in \llbracket 1, n \rrbracket$, $X_k \sim \mathcal{B}(p)$
- X_1, X_2, \dots, X_n sont mutuellement indépendantes
- $X = X_1 + X_2 + \dots + X_n$

Alors

$$X \sim \mathcal{B}(n, p)$$

En outre

$$E(X) = np \text{ et } V(x) = npq$$

Preuve

X est une somme de n termes qui valent 0 ou 1, donc $X(\Omega) \subset \llbracket 0, n \rrbracket$

Soit $k \in \llbracket 0, n \rrbracket$

On a l'union d'événements disjoints deux à deux : $(X = k) = \bigcup_{I \subset \llbracket 1, n \rrbracket, \text{Card}(I)=k} \bigcap_{i \in I} (X_i = 1) \cap \bigcap_{i \in \bar{I}} (X_i = 0)$

Donc, par additivité : $P(X = k) = \sum_{I \subset \llbracket 1, n \rrbracket, \text{Card}(I)=k} P\left(\bigcap_{i \in I} (X_i = 1) \cap \bigcap_{i \in \bar{I}} (X_i = 0)\right)$

Donc $P(X = k) = p^k q^{n-k} \sum_{I \subset \llbracket 1, n \rrbracket, \text{Card}(I)=k} 1 = \binom{n}{k} p^k q^{n-k}$

Par linéarité de l'espérance on a donc $E(X) = E(X_1) + E(X_2) + \dots + E(X_n) = np$

Par indépendance mutuelle donc deux à deux des X_i : $V(X) = V(X_1) + V(X_2) + \dots + V(X_n) = npq$

III.2 Lois usuelles infinies**Définition :**

On suppose que $p \in]0, 1[$

Une variable aléatoire discrète X suit une loi géométrique de paramètre p si $X(\Omega) = \mathbb{N}^*$ et

$$\forall n \in \mathbb{N}^*, P(X = n) = pq^{n-1}$$

On note $X \sim \mathcal{G}(p)$

Proposition :

Si X suite une loi géométrique de paramètre p alors pour tout $k \in \mathbb{N}^*$

$$P(X > k) = (1 - p)^k$$

Proposition : Espérance et variance d'une loi géométrique

Si la variable aléatoire discrète X suit une loi $\mathcal{G}(p)$ alors X possède une espérance et une variance qui valent

$$E(X) = \frac{1}{p} \text{ et } V(X) = \frac{q}{p^2}$$

Preuve

$$\text{Espérance : } p \sum_{i=0}^{+\infty} npq^{n-1} = \left(\frac{p}{1-q} \right)' = \frac{1}{p}$$

$$\text{Variance : } V(X) = E(X^2 - X) + E(X)(E(X))^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2} = \frac{q}{p^2}$$

Proposition : Modèle loi géométrique

Si X est le rang du premier succès dans une suite illimitée d'épreuves de Bernoulli indépendantes de même paramètre p alors X suit la loi $\mathcal{G}(p)$

Formellement :

On pose pour tout $k \in \mathbb{N}^*$ X_k la variable aléatoire discrète qui vaut 1 si la k -ième épreuve est un succès et 0 sinon.

Si :

- pour tout $k \in \mathbb{N}^*$, $X_k \sim \mathcal{B}(p)$
- la suite (X_n) est une suite de variables mutuellement indépendantes
- $X = \min\{n \in \mathbb{N}^*, X_n = 1\}$

Alors $X \sim \mathcal{G}(p)$

Exemple III.1. L'exemple le plus classique est le temps d'attente d'un premier pile lors des lancers successifs d'une pièce qui donne pile avec probabilité p .

Définition : Loi de Poisson

Une variable aléatoire discrète X suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ si $X(\Omega) = \mathbb{N}$ et

$$\forall n \in \mathbb{N}, P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$$

On note $X \sim \mathcal{P}(\lambda)$

Proposition : Espérance et variance d'une loi de Poisson

Si la variable aléatoire discrète X suit une loi $\mathcal{P}(\lambda)$ alors X possède une espérance et une variance et

$$E(X) = V(X) = \lambda$$

Preuve

$$E(X) = \sum_{n \in \mathbb{N}} ne^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n \in \mathbb{N}^*} \frac{\lambda^n}{(n-1)!} = \lambda e^{-\lambda} \sum_{k \in \mathbb{N}} \frac{\lambda^k}{k!} = \lambda$$

$$E(X^2 - X) = \sum_{n \in \mathbb{N}} n(n-1)e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{k \geq 2} \frac{\lambda^k}{(k-2)!} = \lambda^2$$

$$\text{Donc } V(X) = E(X^2 - X) + E(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

IV Fonctions génératrices

Soit X une variable aléatoire discrète à valeurs dans \mathbb{N}

Définition :

On note R_X le rayon de convergence de la série entière $\sum P(X = n)t^n$

La fonction génératrice d'une variable aléatoire X à valeurs dans \mathbb{N} est définie pour tout $t \in]-R_X, R_X[$ par

$$G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} P(X = n)t^n$$

Exemple IV.1. *Fonction génératrice de $X \rightsquigarrow \mathcal{U}([1, n])$:*

$$G_X : t \mapsto \frac{1}{n} \sum_{k=1}^n t^k = \frac{1}{n} t \frac{1-t^n}{1-t}$$

Fonction génératrice de $X \rightsquigarrow \mathcal{B}(n, p)$:

$$G_X : t \mapsto \sum_{k=1}^n \binom{n}{k} p^k (1-p)^{n-k} t^k = (pt + 1 - p)^n$$

Fonction génératrice de $X \rightsquigarrow \mathcal{P}(\lambda)$:

$$G_X : t \mapsto \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k t^k}{k!} = e^{-\lambda + \lambda t}$$

Fonction génératrice de $X \rightsquigarrow \mathcal{G}(p)$:

$$G_X : t \mapsto \sum_{n=1}^{+\infty} (1-p)^{n-1} p t^n = \frac{pt}{(1 - (1-p)t)}$$

Proposition :

La série $\sum P(X = n)t^n$ converge pour $t = 1$ donc le rayon de convergence R_X de la série entière $\sum P(X = n)t^n$ est au moins égal à 1

Preuve

On a par définition que $\sum_{k \in \mathbb{N}} P(X = k) = 1$, donc le rayon est au moins 1.

Proposition :

La série entière $\sum P(X = n)t^n$ converge normalement sur $[-1, 1]$ donc G_X est continue sur $[-1, 1]$

Preuve

La convergence normale est assurée par la majoration de $|P(X = n)t^n|$ par $P(X = n)$ sur $[-1, 1]$
La continuité découle de celle de $t \mapsto t^n$

Théorème :

Soit X une variable aléatoire à valeurs dans \mathbb{N} .

1. La loi de X est entièrement caractérisée par la connaissance de sa fonction génératrice G_X ie : $G_X : t \mapsto \sum_{n=0}^{+\infty} a_n t^n$ est la fonction génératrice de X si, et seulement si, $\forall n \in \mathbb{N}, P(X = n) = a_n$
2. $G_X(1) = 1$
3. X admet une espérance si, et seulement si, G_X est dérivable en 1, avec dans ce cas

$$E(x) = G'_X(1)$$

4. X admet une variance si, et seulement si, G_X est deux fois dérivable en 1, avec dans ce cas :

$$E(X(X-1)) = G''_X(1)$$

et donc

$$V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$$

Preuve

La première propriété découle de l'unicité d'un DSE.

La deuxième propriété découle de la définition d'une probabilité.

Pour la troisième propriété :

Dans les hypothèses de l'énoncé, on suppose $E(X) < +\infty$ (ce qu'on peut écrire puisque X est positive)

Alors la série $\sum nP(X = n)$ converge, et : $\forall t \in [-1, 1], |nP(X = n)t^n| \leq nP(X = n)$

Donc :

- $\forall n \in \mathbb{N}, t \mapsto P(X = n)t^n$ est \mathcal{C}^1
- $\sum t \mapsto P(X = n)t^n$ converge simplement sur $[-1, 1]$
- $\sum t \mapsto nP(X = n)t^{n-1}$ converge normalement donc uniformément sur $[-1, 1]$

Donc par théorème de dérivation des séries de fonctions, G_X est dérivable sur $[-1, 1]$ de dérivée

$$t \mapsto \sum_{n=1}^{+\infty} nP(X = n)t^{n-1}$$

Donc $G'_X(1) = E(X)$

Réciproquement : Procédons par contraposée, en supposant $E(X) = +\infty$ (possible comme X positive)

$$\forall t \in [0, 1[, G'_X(t) = \sum_{n=1}^{+\infty} nP(X = n)t^{n-1}$$

G'_X est croissante sur $[0, 1[$, donc G'_X a une limite finie ou infinie en 1^-

Et comme $nP(X = n)t^{n-1}$ positive sur $[0, 1[$:

$$\forall N \in \mathbb{N}, G'_X(t) \geq \sum_{n=1}^N n(P(X = n))t^{n-1}$$

Passage à la limite quand $t \rightarrow 1^-$:

$$\lim_{t \rightarrow 1^-} G'_X(t) \geq \sum_{n=1}^{+\infty} nP(X = n)$$

Or $\left(\sum_{k=0}^N kP(X = k) \right)$ est la suite des sommes partielles d'une série divergente de terme général positif.

Donc $\sum_{n=1}^N nP(X = n) \rightarrow_{N \rightarrow +\infty} +\infty$

Donc $G'_X \rightarrow_{t \rightarrow 1^-} +\infty$

Donc G_X est positive en 1 et $G'_X \rightarrow_{t \rightarrow 1^-} +\infty$, et donc par théorème de prolongement \mathcal{C}^1 , G_X n'est pas dérivable en 1.

La dernière propriété se prouve de même que la troisième, mais avec la dérivée seconde, en distant que $V(X) = E(X^2 - X) + E(X) - (E(X))^2 = G''_X(1) + G'_X(1) - G'_X(1)^2$.

Remarque IV.1. $G_X = G_Y$ si, et seulement si, X et Y ont même loi.

Théorème :

Si X et Y sont deux variables aléatoires indépendantes à valeurs dans \mathbb{N} alors :

$$\forall t \in [-1, 1], G_{X+Y} = G_X G_Y$$

Soit $n \in \mathbb{N}, n \geq 2$. Si X_1, \dots, X_n sont des variables aléatoire réelles mutuellement indépendantes alors :

$$G_{X_1+\dots+X_n} = G_{X_1} \dots G_{X_n}$$

Preuve

Soit $t \in [-1, 1]$:

$$G_{X+Y}(t) = \sum_{n \in \mathbb{N}} P(X+Y=n)t^n = \sum_{n \in \mathbb{N}} \sum_{i+j=n} P(X=i, Y=j)t^n$$

$$= \sum_{n \in \mathbb{N}} \sum_{i+j=n} P(X=i)P(Y=j)t^n \text{ par indépendance}$$

$$= \left(\sum_{i \in \mathbb{N}} P(X=i)t^i \right) \left(\sum_{j \in \mathbb{N}} P(Y=j)t^j \right) \text{ par produit de Cauchy de deux séries absolument convergentes}$$

$$= G_X(t)G_Y(t)$$

Pour l'extension du théorème, on se contente de faire une récurrence en se servant de la propriété avec deux variables.

V Couples de variables aléatoires réelles

V.1 Loi conjointe, lois marginales

Définition :

Soient X et Y deux variables aléatoires discrètes réelles sur un même espace probabilisé (Ω, \mathcal{A}, P) .

1. On appelle couple des variables X et Y , et on note $Z = (X, Y)$ l'application

$$Z : \begin{cases} \Omega & \rightarrow X(\Omega) \times Y(\Omega) \\ \omega & \mapsto (X(\omega), Y(\omega)) \end{cases}$$

2. La loi du couple $Z = (X, Y)$ est appelée loi conjointe, elle est définie par :

- $Z(\Omega) = X(\Omega) \times Y(\Omega)$
- Les probabilités élémentaires $P(X = x, Y = y) = P((X = x) \cap (Y = y))$ pour tous $(x, y) \in X(\Omega) \times Y(\Omega)$

On a bien sûr

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), P(X = x, Y = y) \geq 0 \text{ et } \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} P(X = x, Y = y) = 1$$

3. Les lois de X et Y sont appelées lois marginales du couple $Z = (X, Y)$

Si la loi conjointe du couple $Z = (X, Y)$ est connue, alors les lois marginales de X et Y le sont aussi :

- On détermine la loi de X en appliquant la formule des probabilités totales avec le système complet d'événements $([Y = y])_{y \in Y(\Omega)}$:

$$\forall x \in X(\Omega), P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y)$$

- On détermine la loi de Y en appliquant la formule des probabilités totales avec le système complet d'événements $([X = x])_{x \in X(\Omega)}$:

$$\forall y \in Y(\Omega), P(Y = y) = \sum_{x \in X(\Omega)} P(X = x, Y = y)$$

La réciproque est évidemment fautive, la connaissance des lois marginales de X et Y ne permet pas de déterminer la loi conjointe du couple $Z = (X, Y)$

Remarque V.1. Dans le cas de variables aléatoires finies on peut noter :

$$X(\Omega) = \{x_1, x_2, \dots, x_r\} \text{ et } Y(\Omega) = \{y_1, y_2, \dots, y_s\}$$

Et notons pour alléger, si $i \in \llbracket 1, r \rrbracket$ et $j \in \llbracket 1, s \rrbracket$:

$$p_i = P(X = x_i), q_j = P(Y = y_j) \text{ et } p_{i,j} = P(X = x_i, Y = y_j)$$

On peut alors se représenter la loi du couple (X, Y) et leurs lois marginales dans un tableau à double-entrée.

Remarque V.2. On rappelle que X et Y sont indépendantes si, et seulement si,

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), P(X = x, Y = y) = P(X = x)P(Y = y)$$

Donc X et Y sont indépendantes si, et seulement si, tous les coefficients du tableau de la loi du couple sont le produit des probabilités de X et Y situées en marge du tableau.

V.2 Lois conditionnelles d'un couple

Définition :

Soit X une variable aléatoire réelle discrète sur (Ω, \mathcal{A}, P) et A un événement de probabilité non-nulle. La loi conditionnelle de X sachant A est la donnée de :

- $X(\Omega)$
- $\forall x \in X(\Omega), P_A(X = x)$

Elle est notée $X_{/A}$

Définition :

Soit $Z = (X, Y)$ un couple de variables aléatoires réelles discrètes sur (Ω, \mathcal{A}, P)

La loi de Z est donnée par :

1. Les lois de X conditionnées par Y sont les lois de X conditionnées par les événements $[Y = y]$ pour tout $y \in Y(\Omega)$
Plus précisément, pour $y \in Y(\Omega)$ fixé tel que $P(Y = y) \neq 0$, la loi de X sachant $[Y = y]$ est définie par :
 - La donnée de $X(\Omega)$
 - Les nombres $P_{Y=y}(X = x) = \frac{P(X=x, Y=y)}{P(Y=y)}$ pour tout $x \in X(\Omega)$
2. Les lois de Y conditionnées par X sont les lois de Y conditionnées par les événements $[X = x]$ pour tout $x \in X(\Omega)$
Plus précisément, pour $x \in X(\Omega)$ fixé tel que $P(X = x) \neq 0$, la loi de Y sachant $[X = x]$ est définie par :
 - La donnée de $Y(\Omega)$
 - Les nombres $P_{X=x}(Y = y) = \frac{P(X=x, Y=y)}{P(X=x)}$ pour tout $y \in Y(\Omega)$

Remarque V.3. Souvent, lorsque l'on étudie un couple de variables aléatoires réelles :

1. On détermine la loi de X
2. Puis pour tout $x \in X(\Omega)$, on détermine la loi de Y conditionnée par l'événement $X = x$
3. On déduit la loi conjointe du couple $Z = (X, Y)$ en utilisant que pour tout $x \in X(\Omega)$ et $y \in Y(\Omega)$

$$P(X = x, Y = y) = P(X = x)P_{X=x}(Y = y)$$

4. Et enfin on détermine la loi de Y

VI Résultats probabilistes asymptotiques

VI.1 Inégalités de Markov et Bienaymé-Tchebychev

Théorème : Inégalité de Markov

Soit X une variable aléatoire réelle discrète positive d'espérance finie, on a :

$$\forall a > 0, P(X \geq a) \leq \frac{E(X)}{a}$$

Preuve

Avec $a \in \mathbb{R}_+$ fixé :

$$\begin{aligned} E(X) &= \sum_{x \in X(\Omega)} xP(X=x) \geq \sum_{x \in X(\Omega), x \geq a} xP(X=x) \text{ (comme } X \text{ est positive)} \\ &\geq a \sum_{x \in X(\Omega), x \geq a} P(X=x) \geq aP(X \geq a) \end{aligned}$$

D'où l'inégalité $P(X \geq a) \leq \frac{E(X)}{a}$

Théorème : Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle discrète telle que X^2 est d'espérance finie, on a :

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

En passant à l'événement contraire :

$$\forall \varepsilon > 0, P(|X - E(X)| < \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}$$

Preuve

Pour $\varepsilon > 0$:

$$\begin{aligned} E((X - E(X))^2) &= \sum_{x \in X(\Omega)} (x - E(X))^2 P(X=x) \\ &\geq \varepsilon^2 \sum_{x \in X(\Omega), |x - E(X)| \geq \varepsilon} P(X=x) \\ &\geq \varepsilon^2 P(|X - E(X)| \geq \varepsilon) \end{aligned}$$

D'où l'inégalité.

VI.2 Loi faible des grands nombres

Théorème : Loi faible des grands nombres

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables réelles indépendantes de même loi, de variance finie.

En notant $S_n = \sum_{k=1}^n X_k$, $m = E(X_1)$ et $\sigma = \sigma(X_1)$, on a :

$$1. \forall \varepsilon > 0, P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Cette inégalité doit être démontrée à chaque utilisation d'après le programme

$$2. \forall \varepsilon > 0, P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \rightarrow_{n \rightarrow +\infty} 0$$

Preuve

On a ici que $\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n E(X_i) = m$

Par indépendance, $V(S_n) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{\sigma^2}{n}$

Remarque VI.1. La loi faible des grands nombres est utilisée lorsque l'on cherche à comparer la probabilité p d'un événement A et la fréquence d'apparition de cet événement quand l'on répète un grand nombre de fois l'expérience aléatoire associée à A

On suppose les expériences indépendantes.

Pour $k \in \mathbb{N}^*$, à la k -ième réalisation de l'expérience, on note X_k la variable aléatoire réelle indicatrice de l'événement A . ($(X_k = 1)$ si A est réalisé lors de cette expérience, $(X_k = 0)$ sinon)

Pour tout $k \in \mathbb{N}^*$, X_k suit la loi de Bernouilli $\mathcal{B}(p)$, d'espérance $m = p$, d'écart-type $\sigma = \sqrt{p(1-p)}$

Pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^n X_k$ est le nombre de fois où l'événement A a été observé pendant les n premières expériences et la fréquence d'apparition de A lors de ces n expériences est :

$$F_n = \frac{S_n}{n}$$

On a :

$$\forall \varepsilon > 0, P(|F_n - p| \leq \varepsilon) = P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}$$

Et :

$$\forall \varepsilon > 0, P(|F_n - p| \geq \varepsilon) \rightarrow_{n \rightarrow +\infty} 0$$

D'où :

$$\forall \varepsilon > 0, P(|F_n - p| < \varepsilon) \rightarrow_{n \rightarrow +\infty} 1$$

L'événement " F_n tend vers p est quasi-certain.

Lorsque la valeur de p , n'est pas connue, on majore souvent $\frac{p(1-p)}{n\varepsilon^2}$ par $\frac{1}{4n\varepsilon^2}$ en utilisant :

$$\boxed{\forall p \in [0, 1], 0 \leq p(1-p) \leq \frac{1}{4}}$$

Chapitre X

Espaces préhilbertiens

Dans tout le cours, E est un \mathbb{R} -ev.

I Préliminaires

Définition :

Soit E un \mathbb{R} -ev, on dit qu'une application $\varphi : E \times E \rightarrow \mathbb{R}$ est :

- une forme bilinéaire si : $\forall x \in E, y \mapsto \varphi(x, y) = \varphi(x, \cdot)$ est linéaire et $\varphi(\cdot, x)$ est linéaire.
- symétrique si :

$$\forall x, y \in E, \varphi(x, y) = \varphi(y, x)$$

- positive si :

$$\forall x \in E, \varphi(x, x) \geq 0$$

- définie si :

$$\forall x \in E, \varphi(x, x) = 0 \Rightarrow x = 0$$

Définition : Produit scalaire

Une forme bilinéaire symétrique définie positive est un produit scalaire.

Remarque I.1. L'expression matricielle du produit scalaire, avec $X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$ est $X^A T Y$ avec A la matrice d'une forme bilinéaire symétrique. (ses coefficients sont les $\varphi(e_i, e_j)$)

Comme φ est symétrique, $\varphi(e_i, e_j) = \varphi(e_j, e_i)$ donc A est une matrice symétrique.

Proposition : Identités polaires

φ une forme bilinéaire symétrique.

$$\forall x, y \in E, \varphi(x, y) = \frac{1}{2}(\varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y))$$

$$= \frac{1}{4}(\varphi(x + y, x + y) - \varphi(x - y, x - y))$$

φ est donc entièrement caractérisée par l'application $u \mapsto \varphi(u, u)$

Définition :

Un espace préhilbertien réel est un espace vectoriel avec un produit scalaire.

Un espace euclidien est un espace préhilbertien réel de dimension finie.

Remarque I.2. On note $\langle ., . \rangle$ ou $(. | .)$ le produit scalaire dans un espace préhilbertien.

Un produit scalaire induit une norme d'expression $\|x\| = \sqrt{\langle x, x \rangle}$

Théorème : Cauchy-Schwarz

E un \mathbb{R} -v et φ une forme bilinéaire symétrique positive

Alors $\forall x, y \in E, |\varphi(x, y)| \leq \varphi(x, x)\varphi(y, y)$

Dans un espace préhilbertien réel E ,

$$\forall x, y \in E, |\langle x, y \rangle| \leq \|x\| \|y\|$$

avec le cas d'égalité si, et seulement si, x et y sont colinéaires.

Preuve

L'inégalité a déjà été prouvée dans le chapitre de proba.

Si x et y sont colinéaires : si $x = 0$ alors $|\langle x, y \rangle| = 0 = \|x\| \|y\|$

Sinon, on peut écrire $y = \alpha x$ pour $\alpha \in \mathbb{R}$, et donc :

$$|\langle x, y \rangle| = |\alpha \langle x, x \rangle| = |\alpha| \|x\|^2 = \|x\| |\alpha x| = \|x\| \|y\|$$

Réciproquement, on suppose que $|\langle x, y \rangle| = \|x\| \|y\|$

Considérons P le polynôme qui à $t \in \mathbb{R}$ associe $\|x + ty\|^2 = \|x\|^2 + 2t\langle x, y \rangle + t^2\|y\|^2$

Ce polynôme est de degré 2 et de discriminant nul par hypothèse.

Donc il existe un unique $t_0 \in \mathbb{R}$ tel que $\|x + t_0 y\|^2 = 0$

Par définition de la norme, $x = -t_0 y$ et les deux vecteurs sont donc colinéaires.

Théorème : Inégalité triangulaire

Soient $x, y \in E$, alors :

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$$

$$(\|x\| + \|y\|)^2 = \|x\|^2 + \|y\|^2 + 2\|x\| \|y\|$$

Donc par Cauchy-Schwarz, $|\langle x, y \rangle| \leq \|x\| \|y\|$ et on en déduit l'inégalité triangulaire :

$$\|x + y\| \leq \|x\| + \|y\|$$

Avec cas d'égalité si x et y sont colinéaires et de même sens.

II Orthogonaux

Définition : Orthogonalité

- pour $u, v \in E$, u et v sont orthogonaux, noté $u \perp v$, si $\langle u, v \rangle = 0$
- deux sev F et G de E sont orthogonaux si $\forall x \in F, \forall y \in G, \langle x, y \rangle = 0$
- pour $A \subset E$, l'orthogonal de A est l'ensemble $A^\perp = \{x \in E | \forall a \in A, \langle a, x \rangle = 0\}$

Proposition : propriétés des orthogonaux

— Pour A, B inclus dans E , $A \subset B \Rightarrow B^\perp \subset A^\perp$

—

$$\forall A \subset E, A^\perp = \text{vect}(A)^\perp$$

— A^\perp est un sev de E

— F, G sev de E :

$$F^\perp \cap G^\perp = (F + G)^\perp$$

$$F^\perp + G^\perp \subset (F \cap G)^\perp$$

(réciproque fausse)

—

$$F \subset (F^\perp)^\perp$$

Preuve

$$F^\perp \cap G^\perp = (F + G)^\perp$$

$$F^\perp \cap G^\perp = (F \cup G)^\perp = \text{vect}(F \cup G)^\perp = (F + G)^\perp$$

$$F^\perp + G^\perp \subset (F \cap G)^\perp$$

$$F \cap G \subset F$$

$$F^\perp \subset (F \cap G)^\perp$$

$$F \cap G \subset G$$

$$G^\perp \subset (F \cap G)^\perp$$

$$F \subset (F^\perp)^\perp$$

Si $x \in F$, alors $\forall y \in F^\perp, \langle x, y \rangle = 0$. Donc $x \in (F^\perp)^\perp$

Remarque II.1. En dimension finie, si F est un sev de E de base (u_1, \dots, u_p) , alors :

$$\forall x \in E, x \in F^\perp \Leftrightarrow \begin{cases} \langle x, u_1 \rangle = 0 \\ \dots \\ \langle x, u_p \rangle = 0 \end{cases}$$

Exemple II.1. $E = \mathcal{C}([0, 1], \mathbb{R})$, muni du produit scalaire $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$

$F = \mathbb{R}[X]$ un sev de E

Si $f \in F^\perp$, on considère $(P_n) \in \mathbb{R}[X]^\mathbb{N}$ qui converge uniformément vers f (qui existe par Stone-Weierstrass)

Comme f est bornée continue sur le segment $[0, 1]$, alors $(f \times P_n) \rightarrow f^2$

Donc $\int_0^1 f(t)P_n(t)dt \rightarrow \int_0^1 f^2(t)dt$ par convergence uniforme.

Or $f \in F^\perp$ donc $\forall n \in \mathbb{N}, \langle f, P_n \rangle = 0$

Donc $\int_0^1 f^2 = 0$, f^2 est positive continue sur le segment $[0, 1]$ donc par stricte positivité, $\forall x \in [0, 1], f(x) = 0$

Donc $f = 0$, donc $F^\perp \subset \{0\}$

Donc $F^\perp = \{0\}$

Remarque II.2. Si F et G sont deux sev de E

F et G sont orthogonaux si, et seulement si, $F \subset G^\perp$

Définition : Familles orthogonales

Soit $(x_i)_{i \in I}$ une famille de vecteurs d'un espace préhilbertien E
 (x_i) est une famille orthogonale si :

$$\forall i, j \in I, i \neq j \Rightarrow \langle x_i, x_j \rangle = 0$$

(x_i) est une famille orthonormale si :
 $(x_i)_{i \in I}$ est orthogonale et $\forall i \in I, \|x_i\| = 1$

Proposition :

Toute famille orthogonale de vecteurs non nuls est une famille libre.

Preuve

$(u_i)_{i \in I}$ une famille orthogonale.

Soit J fini, $J \subset I$

Soit $(\lambda_j)_{j \in J} \in \mathbb{R}^J$

On suppose que $\sum_{j \in J} \lambda_j u_j = 0$

Prenons $j_0 \in J$

$$\langle u_{j_0}, \sum_{j \in J} \lambda_j u_j \rangle = 0$$

$$\text{Donc } \lambda_{j_0} \langle u_{j_0}, u_{j_0} \rangle = 0$$

$$\text{Comme } \langle u_{j_0}, u_{j_0} \rangle \neq 0, \text{ donc } \lambda_{j_0} = 0$$

Remarque II.3. En particulier dans un espace de dimension n , une famille orthonormale de n vecteurs est une base orthonormale (BON).

De l'intérêt dans un espace euclidien de travailler dans une BON :

Soit E euclidien et $B = (e_1, \dots, e_n)$ une BON

alors avec $\forall x \in E, x = \sum_{i=1}^n x_i e_i$

et $\forall y \in E, y = \sum_{i=1}^n y_i e_i$

$$\langle x, y \rangle = \sum_{i,j \in \llbracket 1, n \rrbracket} x_i y_j \langle e_i, e_j \rangle$$

$$= \sum_{i=1}^n x_i y_i = X^T Y$$

$\|x\|^2 = \sum_{i=1}^n x_i^2$ est vrai si, et seulement si, on a une BON

$\forall i \in \llbracket 1, n \rrbracket, x_i = \langle x, e_i \rangle$ est vrai si, et seulement si, on a une BON

Théorème :

Soit E un espace euclidien et F un sev de E

Alors

$$F \oplus F^\perp = E$$

(ie : F^\perp est un supplémentaire de F)

Preuve

$$\forall x \in E, x \in F \cap F^\perp \Rightarrow \langle x, x \rangle = 0$$

Donc $x = 0$

Donc $F \cap F^\perp = \{0\}$ (ce qui est valable même en dimension infinie)

On note n la dimension de E et p la dimension de F . On se donne une base de F (u_1, \dots, u_p)

$$\text{On considère } \varphi : \begin{cases} E & \rightarrow \mathbb{R}^p \\ x & \mapsto (\langle x, u_1 \rangle, \dots, \langle x, u_p \rangle) \end{cases}$$

Donc $\varphi \in \mathcal{L}(E, \mathbb{R}^p)$

On a que $F^\perp = \ker \varphi$

$$\text{Considérons } \varphi|_F \text{ ie } \varphi|_F : \begin{cases} F & \rightarrow \mathbb{R}^p \\ x & \mapsto \varphi(x) \end{cases}$$

$$\text{Alors } \ker \varphi|_F = \ker \varphi \cap F = \{0\}$$

Donc $\varphi|_F$ est une application injective de l'espace F de dimension p dans \mathbb{R}^p de dimension p
donc $\varphi|_F$ est bijective donc surjective

Donc φ est surjective

$$\text{Donc } \text{rg}(\varphi) = p$$

Donc par théorème du rang, $\dim \ker \varphi = n - p$

$$\text{Donc } \dim F^\perp + \dim F = \dim E \text{ et } F + F^\perp \text{ est directe}$$

$$\text{Donc } F \oplus F^\perp = E$$

Proposition : Corollaire

Tout espace euclidien possède une BON.

Preuve

On le fait par récurrence forte sur la dimension :

H_n : tout espace de dimension n ou moins possède une BON.

Dans un espace de dimension $n + 1$, en sachant H_n , on pourra prendre un espace de dimension 1 et son orthogonal de dimension n , on réutilise l'hypothèse de récurrence pour avoir une BON de l'orthogonal et on norme le vecteur de base de la droite.

Théorème : Extension

Soit E préhilbertien réel et F un sev de E de dimension finie. Alors :

$$F \oplus F^\perp = E$$

Preuve

Dans le cadre de l'énoncé, on sait déjà que $F \cap F^\perp = \{0\}$

Soit $x \in E$, on considère $G_x = \text{Vect}(x) + F$

G est euclidien. F est un sev de G_x et $x \in G_x$

Donc $G_x = F \oplus F^{\perp_{G_x}}$ (l'orthogonal dans G_x)

$$\text{Donc } \exists u, v \in F \times F^{\perp_{G_x}}, x = u + v$$

$$\forall f \in F, \langle v, f \rangle = 0$$

$$\text{Donc } v \in F^\perp$$

$$\text{Donc } E \subset F + F^\perp$$

Théorème : Méthode de Schmidt

Cette méthode permet de "redresser une boîte à chaussures écrasée", c'est à dire construire une base orthogonale.

Soit E un espace préhilbertien, $(u_i)_{i \in I}$ famille libre de E avec $I \subset \mathbb{N}$

Alors on peut construire par récurrence une famille orthogonale $(v_n)_{n \in I}$ qui vérifie :

$$\forall n \in I, \text{Vect}(u_0, \dots, u_n) = \text{Vect}(v_0, \dots, v_n)$$

(v_n) est définie par la relation :

$$\begin{cases} v_0 = u_0 \\ v_{n+1} = u_{n+1} - \sum_{i=1}^n \frac{\langle u_{n+1}, v_i \rangle}{\|v_i\|^2} v_i \end{cases}$$

Pour rendre cette famille orthonormale, il suffit de prendre la famille et de diviser chaque vecteur par sa norme

Preuve

Pour montrer que la famille est orthogonale, on procède par récurrence pour montrer que l'espace engendré est bien le même à chaque nouvel élément.

Proposition :

Si E est euclidien, F et G deux sev de E .

$$(F \cap G)^\perp = F^\perp + G^\perp$$

$$F = (F^\perp)^\perp$$

Preuve

On a déjà les inclusions, on se sert juste de l'égalité des dimensions en dimension finie.

III Projections orthogonales

III.1 Généralités

Définition :

Dans tous les cas où $F \oplus F^\perp = E$, on peut définir :

- la projection orthogonale sur F (la projection sur F parallèlement à F^\perp)
- la symétrie orthogonale par rapport à F

Théorème : de la meilleure approximation

Soit E un espace préhilbertien et F un sev de dimension finie.

Alors pour tout x de E ,

$d(x, F)$ est atteinte en un unique vecteur de F , le projeté orthogonal de x sur F

Preuve

$$d(x, F) = \inf_{y \in F} d(x, y)$$

Dans les conditions de l'énoncé, soit $y \in F$, on désigne p la projection orthogonale sur F

$$\|x - y\|^2 = \|x - p(x) + p(x) - y\|^2$$

Or $x - p(x) \in F^\perp$ et $p(x) - y \in F$, donc les deux sont orthogonaux

Donc $\|x - y\|^2 = \|x - p(x)\|^2 + \|p(x) - y\|^2$ (par pythagore vectoriel)

$$\text{Donc } \|x - y\|^2 \geq \|x - p(x)\|^2$$

$$\text{et } \|x - y\|^2 = \|x - p(x)\|^2 \Leftrightarrow \|p(x) - y\|^2 = 0 \Leftrightarrow y = p(x)$$

III.2 Calcul pratique du projeté orthogonal

Proposition :

F un sev de E de dimension finie p , (u_1, \dots, u_p) une base de F

$x \in E$, $p(x)$ son projeté orthogonal sur F

Donc (px) est entièrement défini par les équations :

$$(1) \begin{cases} p(x) \in F \\ x - p(x) \in F^\perp \end{cases}$$

Preuve

$$(1) \Leftrightarrow \begin{cases} \exists (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p, p(x) = \sum_{i=1}^p \lambda_i u_i \\ \langle x - p(x), u_1 \rangle = 0 \\ \dots \\ \langle x - p(x), u_p \rangle = 0 \end{cases}$$

$$\begin{cases} \Leftrightarrow \exists (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p, p(x) = \sum_{i=1}^p \lambda_i u_i \\ \sum_{j=1}^p \langle u_1, u_j \rangle \lambda_j = \langle x, u_1 \rangle \\ \dots \\ \sum_{j=1}^p \langle u_p, u_j \rangle \lambda_j = \langle x, u_p \rangle \end{cases}$$

On note (S) ce dernier système, et $(S) \Leftrightarrow G(u_1, \dots, u_p) \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} \langle x, u_1 \rangle \\ \dots \\ \langle x, u_p \rangle \end{pmatrix}$

Où $G(u_1, \dots, u_p) \in \mathcal{M}_p(\mathbb{R})$ tel que $G(u_1, \dots, u_p) = (\langle u_i, u_j \rangle)_{(i,j) \in \llbracket 1, p \rrbracket^2}$

Dans le cas où (u_1, \dots, u_p) est une BON de F :

$$p(x) = \langle x, u_1 \rangle u_1 + \dots + \langle x, u_p \rangle u_p$$

Si (u_1, \dots, u_p) est seulement une base orthogonale :

$$p(x) = \frac{\langle x, u_1 \rangle}{\|u_1\|^2} u_1 + \dots + \frac{\langle x, u_p \rangle}{\|u_p\|^2} u_p$$

III.3 Exemples d'utilisation du projeté orthogonal

Exemple III.1. *Les séries de Fourier* : soit $E = \mathcal{C}_T(\mathbb{R})$ l'ensemble des fonctions continues T périodiques.

On munit E du produit scalaire $\forall f, g \in E, \langle f, g \rangle = \frac{1}{T} \int_0^T f(t)g(t)dt$

Pour $n \in \mathbb{N}$, on considère $c_n : \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ t \mapsto \cos\left(\frac{2\pi n}{T}t\right) \end{cases}$

Pour $n \in \mathbb{N}^*$, on considère $s_n : \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ t \mapsto \sin\left(\frac{2\pi n}{T}t\right) \end{cases}$

Pour $n \in \mathbb{N}$, on note $F_n = \text{Vect}((c_k)_{0 \leq k \leq n}, (s_k)_{1 \leq k \leq n})$

$\forall n \in \mathbb{N}, c_n \in E, \forall n \in \mathbb{N}^*, s_n \in E$

$\forall n, p \in \mathbb{N}, \langle c_n, c_p \rangle = \frac{1}{T} \int_0^T \cos\left(\frac{2\pi n}{T}t\right) \cos\left(\frac{2\pi p}{T}t\right) dt$

$= \frac{1}{2T} \int_0^T \cos\left(\frac{2\pi(n+p)}{T}t\right) + \cos\left(\frac{2\pi(n-p)}{T}t\right) dt = 0$

De même, $\forall n, p, \langle c_n, s_p \rangle = 0$ et $\forall n, p, n \neq p \Rightarrow \langle s_n, s_p \rangle = 0$

Donc pour tout $n \in \mathbb{N}$ $(c_0, \dots, c_n, s_1, \dots, s_n)$ est une base orthogonale (famille orthogonale de vecteurs non-nuls) de F_n .

$\langle c_0, c_0 \rangle = 1$

$\forall n \in \mathbb{N}^*, \langle c_n, c_n \rangle = \frac{1}{2}$

$\forall n \in \mathbb{N}^*, \langle s_n, s_n \rangle = \frac{1}{2}$

Pour $f \in E$, et $n \in \mathbb{N}$, on pose :

$$\begin{aligned} S_n(f) &= \frac{\langle f, c_0 \rangle^2}{\|c_0\|^2} c_0 + \sum_{k=1}^n \frac{\langle f, c_k \rangle}{\|c_k\|^2} c_k + \frac{\langle f, s_k \rangle}{\|s_k\|^2} s_k \\ &= \langle f, c_0 \rangle c_0 + \sum_{k=1}^n 2\langle f, c_k \rangle c_k + 2\langle f, s_k \rangle s_k \end{aligned}$$

$S_n(f)$ est la meilleure approximation de f dans F_n

$\|f\|^2 = \|f - S_n(f)\|^2 + \|S_n(f)\|^2$

où $\|S_n(f)\|^2 = \langle f, c_0 \rangle^2 + 2 \sum_{k=1}^n \langle f, c_k \rangle^2 + \langle f, s_k \rangle^2$

$(\|S_n(f)\|^2)$ est la suite des sommes partielles d'une série à termes positifs majorée par $\|f\|^2$, donc $\|S_n(f)\|^2$ converge en $+\infty$ et $\|S_\infty(f)\|^2 \leq \|f\|^2$ (inégalité de Parseval)

Exemple III.2. Régression linéaire : On dispose de p points de \mathbb{R}^n

On note $\begin{cases} x_{1,\cdot} = x_{1,1}, x_{1,2}, \dots, x_{1,n} & z_1 \\ x_{2,\cdot} = x_{2,1}, x_{2,2}, \dots, x_{2,n} & z_2 \\ \dots & \dots \\ x_{p,\cdot} = x_{p,1}, \dots, x_{p,n} & z_p \end{cases}$

On cherche le "meilleur modèle" affine expliquant z , ie les "meilleurs" coefficients a_0, \dots, a_n tels que $a_0 + a_1 x_{i,1} + \dots + a_n x_{i,n}$ approxime le mieux possible z_i

On va chercher :

$$\inf_{(a_0, \dots, a_n) \in \mathbb{R}^{n+1}} \left(\sum_{i=1}^p z_i - (a_0 + a_1 x_{i,1} + \dots + a_n x_{i,n}) \right)^2$$

On munit \mathbb{R}^p du produit scalaire canonique :

$\sum_{i=1}^p (z_i - (a_0 + a_1 x_{i,1} + \dots + a_n x_{i,n}))^2 = \|z - (a_0 u + a_1 x_{\cdot,1} + \dots + a_n x_{\cdot,n})\|^2$

Où $u = \begin{pmatrix} 1 \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$ et $z = \begin{pmatrix} z_1 \\ \cdot \\ \cdot \\ z_p \end{pmatrix}$

Si $(x_{\cdot,1}, \dots, x_{\cdot,n}, u)$ est libre, le problème a une unique solution $\tilde{a}_0, \tilde{a}_1, \dots, \tilde{a}_n$ telle que $\tilde{a}_0 u + \tilde{a}_1 x_{\cdot,1} + \dots + \tilde{a}_n x_{\cdot,n}$ est le projeté orthogonal de z sur $F = \text{Vect}(u, x_{\cdot,1}, \dots, x_{\cdot,n})$

IV Endormorphismes particuliers des espaces euclidiens

Définition :

Soit E un \mathbb{K} -ev. On appelle dual de E l'ensemble $E^* = \mathcal{L}(E, \mathbb{K})$ des formes linéaires sur E .

Proposition :

E^* est un \mathbb{K} ev. Si E est de dimension finie, la dimension de E^* est de dim finie, et $\dim E^* = \dim E$

Théorème : de représentation de Reese (cas euclidien)

Soit E un espace euclidien, l'application :

$$\psi : \begin{cases} E & \rightarrow E^* \\ a & \mapsto \varphi_a : \begin{cases} E & \rightarrow \mathbb{K} \\ x & \mapsto \langle a, x \rangle \end{cases} \end{cases} \quad \text{est un isomorphisme}$$

Preuve

ψ est linéaire : immédiat ($\psi(a+b) = \psi(a) + \psi(b)$ ie $\forall x \in E, (\psi(a+b))(x) = \psi(a)(x) + \psi(b)(x)$)

Soit $a \in \ker \psi$, alors $\forall x \in E, \langle a, x \rangle = 0$

Donc $a \in E^\perp$

Donc $a = 0$

Donc ψ est une application linéaire injective de E dans l'espace de même dimension E^* , donc ψ est bijective.

Définition : Adjoint

Soit E un espace euclidien et $u \in \mathcal{L}(E)$.

On appelle adjoint de u l'application de E dans E u^* telle que :

$$\forall x, y \in E \times E, \langle u(x), y \rangle = \langle x, u^*(y) \rangle$$

Remarque IV.1. Le fait que cette définition soit correcte découle du théorème de représentation des formes linéaires (théorème précédent)

En effet pour $y \in E$ fixé : $x \mapsto \langle u(x), y \rangle \in E^*$

Donc $\exists! u^*(y), \forall x \in E, \langle y, u(x) \rangle = \langle u^*(y), x \rangle$

Proposition : propriétés de l'ajoint

Soit E euclidien, $u, v \in \mathcal{L}(E)$.

—

$$u^* \in \mathcal{L}(E)$$

—

$$(u^*)^* = u$$

—

$$(u \circ v)^* = v^* \circ u^*$$

— $u \rightarrow u^*$ est linéaire, ie $(u+v)^* = u^* + v^*$ et $\forall \lambda \in \mathbb{R}, (\lambda u)^* = \lambda u^*$

Preuve

$y, z \in E$
 Soit $x \in E$, $\langle x, u^*(y+z) \rangle = \langle u(x), y+z \rangle$
 $= \langle u(x), y \rangle + \langle u(x), z \rangle = \langle x, u^*(y) \rangle + \langle x, u^*(z) \rangle$
 Donc $\langle x, u^*(y+z) \rangle = \langle x, u^*(y) + u^*(z) \rangle$
 Donc $\forall x \in E$, $\langle x, u^*(y+z) - (u^*(y) + u^*(z)) \rangle = 0$
 Donc $u^*(y+z) - u^*(y) - u^*(z) \in E^\perp$
 De même, $\forall \lambda \in \mathbb{R}, \forall y \in E, u^*(\lambda y) = \lambda u^*(y)$
 $\forall (x, y) \in E \times E, \langle u^*(x), u \rangle = \langle x, u(y) \rangle$ donc $(u^*)^* = u$
 Soit $x \in E$ et $y \in E$
 $\langle u \circ v(x), y \rangle = \langle v(x), u^*(y) \rangle = \langle x, v^* \circ u^*(y) \rangle$
 Donc $(u \circ v)^* = v^* \circ u^*$
 La dernière propriété est immédiate.

Théorème :

Soit E un espace euclidien. Soit B une BON de E . Soit $u \in \mathcal{L}(E)$
 On note $A = \text{Mat}_B(u)$
 Alors $\text{Mat}_B(u^*) = A^T$

Preuve

Pour $x \in E$, on note $X = \text{Mat}_B(x) \in \mathcal{M}_{n,1}(\mathbb{R})$
 Pour $y \in E$, on note $Y = \text{Mat}_B(y) \in \mathcal{M}_{n,1}(\mathbb{R})$
 $\forall x, y \in E, \langle u(x), Y \rangle = (AX)^T Y = X^T (A^T Y)$ (par propriétés de la transposée)
 Si on note $B = \text{Mat}_B(u^*)$, alors $\langle x, u^*(y) \rangle = X^T B Y$
 Donc $\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R}), X^T (A^T Y - B Y) = 0$
 Donc $A^T = B$

Proposition :

Soit E un espace euclidien, $u \in \mathcal{L}(E)$, F un sev de E .
 Si F est stable par u alors F^\perp est stable par u^*

Preuve

Soit $y \in F^\perp$. Alors $\forall x \in F, \langle x, y \rangle = 0$
 Or $\forall x \in F, u(x) \in F$
 Donc $\forall x \in F, \langle u(x), y \rangle = \langle x, u^*(y) \rangle = 0$
 Donc $u^*(y) \in F^\perp$
 Donc F^\perp est stable par u^* .

V Isométries vectorielles

Définition :

Soit E un espace euclidien et $f \in \mathcal{L}(E)$. On dit que f est une isométrie si

$$\forall x \in E, \|f(x)\| = \|x\|$$

Proposition : Caractérisation

Soit $f \in \mathcal{L}(E)$ avec E euclidien muni d'une base \mathcal{B} orthonormée, f est une isométrie si, et seulement si, elle vérifie l'une des propriétés suivantes :

1.

$$\forall x \in E, \|f(x)\| = \|x\|$$

2.

$$\forall x, y \in E, \langle f(x), f(y) \rangle = \langle x, y \rangle$$

3. $f(\mathcal{B})$ est une base orthonormée.4. $f \in \mathcal{GL}(E)$ et $f^* = f^{-1}$ **Preuve**

$1 \Rightarrow 2$: on suppose que f conserve la norme.

$$\begin{aligned} \text{Soit } x, y \in E, \text{ alors } \langle f(x), f(y) \rangle &= \frac{1}{2} (\|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2) = \\ &= \frac{1}{2} (\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2) \\ &= \frac{1}{2} (\|x+y\|^2 - \|x\|^2 - \|y\|^2) = \langle x, y \rangle \end{aligned}$$

$2 \Rightarrow 3$: immédiat, une isométrie conserve le produit scalaire.

$3 \Rightarrow 1$: On suppose que $f(\mathcal{B})$ est une base orthonormée. On note $\mathcal{B} = (e_1, \dots, e_n)$

$$\forall x \in E, x = \sum_{i=1}^n x_i e_i \text{ et } \|x\|^2 = \sum_{i=1}^n x_i^2 \text{ car } \mathcal{B} \text{ est une base orthonormale.}$$

$$f(x) = \sum_{i=1}^n x_i f(e_i), \text{ donc } \|f(x)\|^2 = \sum_{i=1}^n x_i^2 \text{ car } f(\mathcal{B}) \text{ est une base orthonormale.}$$

Pour la 4 :

Si f est une isométrie : l'image d'une base par un isomorphisme est une base

Donc f est bijective

$$\forall x, y \in E, \langle x, f(y) \rangle = \langle f(f^{-1}(x)), f(y) \rangle = \langle f^{-1}(x), y \rangle$$

$$\text{Donc } f^* = f^{-1}$$

Réciproquement, on suppose que $f \in \mathcal{GL}(E)$ et $f^* = f^{-1}$

Soit $x \in E$:

$$\|f(x)\|^2 = \langle f(x), f(x) \rangle = \langle x, f^* \circ f(x) \rangle = \langle x, x \rangle = \|x\|^2$$

Remarque V.1. Attention : on emploie parfois le terme d'endomorphisme orthogonal pour désigner les isométries.

Ce vocabulaire est trompeur :

— Une projection orthogonale n'est pas un endomorphisme orthogonal (sauf l'identité)

— Une symétrie orthogonale est un endomorphisme orthogonal. En effet, notons $x \in E, x = x_F + x_{F^\perp}$

$$\text{Donc } s(x) = x_F - x_{F^\perp}$$

$$\|x\|^2 = \|x_F\|^2 + \|x_{F^\perp}\|^2 = \|x_F\|^2 + \|-x_{F^\perp}\|^2 = \|s(x)\|^2$$

$s \in \mathcal{L}(E)$ conserve la norme, c'est une isométrie.

Exemple V.1. Soit E un espace euclidien et s une symétrie. Montrez que s est une isométrie si, et seulement si, s est une symétrie orthogonale.

On a déjà la première implication par la remarque précédente, raisonnons par contraposée :

Si s est une symétrie non-orthogonale :

f est une symétrie par rapport à F parallèlement à G . s n'est pas orthogonale, donc F et G ne sont pas orthogonaux.

Donc on peut noter $f \in F, g \in G$ tels que $\langle f, g \rangle \neq 0$

$$\|f + g\|^2 = \|f\|^2 + 2\langle f, g \rangle + \|g\|^2$$

$$\|f - g\|^2 = \|f\|^2 - 2\langle f, g \rangle + \|g\|^2$$

Et donc $\|s(f+g)\| \neq \|f+g\|$
 Donc s n'est pas une isométrie.

V.1 Isométries directes, indirectes

Proposition :

Soit E un espace euclidien, u une isométrie. Alors :

$$(\det u) \in \{-1, 1\}$$

On dit que u est une isométrie directe si $\det(u) = 1$, indirecte sinon.

Preuve

Si u une isométrie.

$$\det(u) = \det(u^*) = \det(u^{-1}) = \frac{1}{\det(u)}$$

Proposition :

- L'ensemble des isométries est un groupe pour la loi \circ
 On l'appelle le groupe orthogonal noté $\mathcal{O}(E)$.
- L'ensemble des isométries directes est un groupe pour la loi \circ .
 On l'appelle le groupe spécial orthogonal, noté $\mathcal{SO}(E)$ ou $\mathcal{O}_+(E)$.
- L'ensemble des isométries indirectes **n'est pas** un groupe pour la loi \circ .
 On le note $\mathcal{O}_-(E)$

Preuve

On démontre que $\mathcal{O}(E)$ est un sous-groupe de $(\mathcal{GL}(E), \circ)$

Si $u, v \in \mathcal{O}(E)$, alors :

$$\forall x \in E, \|u \circ v(x)\| = \|u(v(x))\| = \|v(x)\| = \|x\|$$

Donc $u \circ v \in \mathcal{O}(E)$

$$\forall x \in E, \|u^{-1}(x)\|^2 = \langle u^*(x), u^{-1}(x) \rangle = \langle x, (u^*)^{-1}u^{-1}(x) \rangle$$

$$= \langle x, u \circ u^{-1}(x) \rangle = \langle x, x \rangle$$

Donc $\mathcal{O}(E)$ est un sous-groupe de $\mathcal{GL}(E)$

"Vous revenez pas à Matusalem" - Chakroun, pour dire de ne pas prouver que quelque chose est un groupe et de prouver que c'est un sous-groupes parce que c'est quand même plus simple.

V.2 Orientation de l'espace

Définition :

E un espace euclidien, on définit la relation sur les bases de cet espace :

$$B \mathcal{R} B' \Leftrightarrow \det_B(B') > 0$$

\mathcal{R} est une relation d'équivalence.

Il y a deux classes d'équivalence.

Preuve

Soit $B = (e_1, e_2, \dots, e_n)$ une base, alors $B' = (-e_1, e_2, \dots, e_n)$ n'est pas dans la classe de B

Soit B'' une base, si B'' n'est pas dans la classe de B

Alors $\det_{B'} B'' = \det_B B'' \det_{B'} B' B > 0$ ($\det_B(B'') < 0$ et $\det_{B'}(B') < 0$) donc B'' est dans la classe de B' .

Orienter l'espace, c'est choisir une classe d'équivalence ie choisir la classe des bases directes.

Proposition :

Soit E un espace euclidien orienté, \mathcal{B} une base orthonormale directe (BOND)

$u \in \mathcal{L}(E)$ est une isométrie directe si, et seulement si, $u(\mathcal{B})$ est une BOND.

$$\det_B(u(e_1), \dots, u(e_n)) = \det u \det_B(e_1, \dots, e_n)$$

Définition : Produit mixte

Soit E un espace euclidien orienté et \mathcal{B} une base orthonormale directe.

L'application $(x_1, \dots, x_n) \mapsto \det_{\mathcal{B}}(x_1, \dots, x_n)$ ne dépend pas de \mathcal{B}

On l'appelle produit mixte noté $[x_1, \dots, x_n]$

Preuve

\mathcal{B} une base orthonormale directe.

$$\det_{\mathcal{B}'}(x_1, \dots, x_n) = \det_{\mathcal{B}}(x_1, \dots, x_n) \det_{\mathcal{B}'}(\mathcal{B})$$

L'application linéaire qui envoie \mathcal{B}' sur \mathcal{B} est une isométrie directe.

Donc $\det_{\mathcal{B}'} \mathcal{B} = +1$

VI Matrices orthogonales

VI.1 Généralités

Définition :

On dit qu'une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est orthogonale si elle vérifie l'une des propriétés équivalentes suivantes :

1.

$$M^T M = I$$

2.

$$M M^T = I$$

3. $M \in \mathcal{GL}_n(\mathbb{R})$ et $M^T = M^{-1}$

4. Les vecteurs colonnes de M forment une base orthonormale pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$

5. Les vecteurs lignes de M forment une base orthonormale pour le produit scalaire canonique de $\mathcal{M}_{1,n}(\mathbb{R})$

Preuve

Les seules implications à montrer sont $1 \Leftrightarrow 4$, puisque $1 \Leftrightarrow 2 \Leftrightarrow 3$ et $4 \Leftrightarrow 5$

Soit $M \in \mathcal{M}_n(\mathbb{R})$

On considère C_1, \dots, C_n les vecteurs colonnes de M

$$\forall j \in \llbracket 1, n \rrbracket, C_j = \begin{pmatrix} M_{1,j} \\ \vdots \\ M_{n,j} \end{pmatrix}$$

$$\text{Pour } j, k \in \llbracket 1, n \rrbracket, \langle C_j, C_k \rangle = \sum_{i=1}^n M_{i,j} M_{i,k} = \sum_{i=1}^n M_{j,i}^T M_{i,k} = (M^T M)_{j,k}$$

Donc $M^T M = I \Leftrightarrow (C_1, \dots, C_n)$ est une BON.

Exemple VI.1. $\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$

$$\begin{pmatrix} \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{\sqrt{5}}{\sqrt{14 \times 5}} & \frac{-6}{\sqrt{14 \times 5}} & \frac{5}{\sqrt{14 \times 5}} \end{pmatrix}$$

Proposition :

La matrice de passage entre deux BON est une matrice orthogonale.

Proposition :

Le déterminant d'une matrice orthogonale est égal à 1 ou à -1

Définition :

Pour $u \in \mathcal{L}(E)$, on a équivalence entre :

- u est une isométrie vectorielle
- la matrice associée à u dans toute BON est orthogonale
- il existe une BON dans laquelle la matrice associée à u est orthogonale

VI.2 Classification des matrices orthogonales du plan

$\mathcal{O}_2(\mathbb{R})$ l'ensemble des matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$

$$\mathcal{O}_2(\mathbb{R}) = \{R_\theta | \theta \in \mathbb{R}\} \cup \{S_\theta | \theta \in \mathbb{R}\}$$

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

$$S_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$$

$\mathcal{SO}_2(\mathbb{R}) = \mathcal{O}_2^+(\mathbb{R}) = \{R_\theta | \theta \in \mathbb{R}\}$ (ensemble des rotations d'angle θ , ensemble des matrices de $\mathcal{O}_2(\mathbb{R})$ de déterminant 1)

$$\mathcal{O}_2^-(\mathbb{R}) = \{S_\theta | \theta \in \mathbb{R}\}$$
 (ensemble des matrices de $\mathcal{O}_2(\mathbb{R})$ de déterminant -1)

Proposition :

$\forall(\theta, \varphi) \in \mathbb{R}^2 :$

—

$$R_\theta R_\varphi = R_{\theta+\varphi}$$

—

$$R_\theta^{-1} = R_{-\theta} = R_\theta^T$$

— $\mathcal{SO}_2(\mathbb{R})$ est un groupe commutatif, en particulier :

$$\forall \theta, \varphi \in \mathbb{R}, R_\theta R_\varphi R_\theta^{-1} = R_\varphi$$

Dans ce cas, R_θ est la matrice de passage d'une BOND dans une BOND.

Preuve

Se prouvent par du calcul matriciel, et un peu de trigonométrie.

S_θ est une symétrie orthogonale par rapport à une droite

$$S_\theta \times \begin{pmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{pmatrix} = \begin{pmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{pmatrix}$$

$$S_\theta \times \begin{pmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{pmatrix} = - \begin{pmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{pmatrix}$$

Donc $S_p(S_\theta) = \{-1, 1\}$ Et $E_1(S_\theta)$ est orthogonal à $E_{-1}(S_\theta)$

VI.3 Classification des matrices orthogonales d'un espace de dimension 2

(où E_2 un espace euclidien de dimension 2)

D'après l'étude de $\mathcal{O}_2(\mathbb{R})$, on peut définir dans E_2 la rotation d'angle θ :

Définition :

Dans E_2 euclidien de dimension 2, on définit la rotation d'angle θ comme r_θ l'endomorphisme dont la matrice dans une BOND est R_θ

Remarque VI.1. Donc $\mathcal{SO}(E_2) = \{r_\theta | \theta \in \mathbb{R}\}$

$\mathcal{O}^-(E_2)$ est alors l'ensemble des symétries orthogonales de E_2 par rapport à des droites.

Proposition :

Si u et v sont deux vecteurs normés de E_2 il existe une unique rotation r_θ telle que $v = r_\theta(u)$

On appellera angle entre u et v cette rotation (θ est la mesure de cet angle défini à 2π près)

On retrouve la propriété de Chasles sur les angles.

Preuve

u et v sont normés

$u^\perp = Vect(w)$ où $\|w\| = 1$ Donc la droite $Vect(w)$ il y a seulement deux vecteurs normés : w et $-w$

Donc il y a exactement 2 BON de premier vecteur u , (u, w) et $(u, -w)$

Une seule d'entre eux est directe.

Quitte à échanger w et $-w$, on peut considérer que (u, w) est une BOND.

Donc $v \in Vect(u, w)$, $v = au + bw$ où $a^2 + b^2 = 1$

L'unique rotation r telle que $r(u) = v$ a pour matrice dans $(u, w) : \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

Remarque VI.2. Par définition pour 2 vecteurs non nuls u, v , on appellera angle de u et v l'angle entre $\frac{u}{\|u\|}$ et $\frac{v}{\|v\|}$

Proposition :

L'angle de deux vecteurs est caractérisé par sa mesure notée $\widehat{u, v}$ définie à 2π près, donc caractérisée par $\cos(\widehat{u, v})$ et $\sin(\widehat{u, v})$

$$\cos(\widehat{u, v}) = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

$$\sin(\widehat{u, v}) = \frac{[u, v]}{\|u\| \|v\|}$$

Preuve

u, v non nuls.

On considère $(\frac{u}{\|u\|}, w)$ l'unique BOND de premier vecteur $\frac{u}{\|u\|}$

On note $\theta = \widehat{u, v}$

$$\frac{v}{\|v\|} = \cos \theta \frac{u}{\|u\|} + \sin \theta w$$

$$\text{Donc } \langle \frac{v}{\|v\|}, \frac{u}{\|u\|} \rangle = \cos \theta$$

$$\left[\frac{u}{\|u\|}, \frac{v}{\|v\|} \right] = \begin{vmatrix} 1 & \cos \theta \\ 0 & \sin \theta \end{vmatrix} = \sin \theta$$

Exemple VI.2. R^2 euclidien orienté (la base canonique est une BOND)

Donner une mesure de l'angle entre $u = (5, 3)$ et $v = (1, -2)$

$$\cos(\widehat{u, v}) = \frac{5-6}{\sqrt{34}\sqrt{5}} = \frac{-1}{\sqrt{34}\sqrt{5}}$$

$$\sin \theta = \frac{-10-3}{\sqrt{34}\sqrt{5}} = \frac{-13}{\sqrt{34}\sqrt{5}}$$

Donc $\theta = -\arccos\left(\frac{-1}{\sqrt{34}\sqrt{5}}\right)$ (comme le produit mixte de u et v est négatif, ce qui permet de choisir le signe)

VI.4 Classification des matrices orthogonales d'un espace euclidien orienté de dimension 3

Pour $u \in \mathcal{O}(E_3)$, on note $F = \ker(u - id)$

Si $\dim F = 3$: alors $u = id$, donc u est nécessairement dans $\mathcal{O}_+(E_3)$

Si $\dim F = 2$, alors $\dim F^\perp = 1$, l'endomorphisme induit par u sur F^\perp est égal à id_{F^\perp} ou $-id_{F^\perp}$

Si u_{F^\perp} était égal à id_{F^\perp} , alors u serait l'identité. Donc nécessairement, $u_{F^\perp} = -id_{F^\perp}$

Donc u est la symétrie orthogonale par rapport à F

Si $\dim F = 1$, alors $\dim F^\perp = 2$. Notons v l'endomorphisme induit par u sur F^\perp .

v est donc soit une symétrie orthogonale par rapport à une droite, soit une rotation du plan. Si v était une symétrie orthogonale, on aurait $\dim F = 2$, donc v est une rotation du plan F^\perp .

Donc si on choisit w un vecteur qui oriente F et qu'on choisit u, v dans F^\perp tels que (u, v) soit une base orientée de F^\perp et que (w, u, v) soit une base directe de E_3 , on a alors que u est la rotation d'axe w orienté par w d'angle θ .

Si $\dim F = 0$, alors ça veut dire que 1 n'est pas valeur propre, donc -1 est valeur propre. u est donc la composée d'une symétrie orthogonale et d'une rotation.

Exemple VI.3. Prenons $u \in \mathcal{L}(E_3)$ avec \mathcal{B} une BOND de E_3 telle que $\text{Mat}_{\mathcal{B}}(u) = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ 2 & -2 & -1 \\ 1 & 2 & -2 \end{pmatrix}$

On peut faire les produits scalaires des colonnes les unes avec les autres, et on obtiendra que les produits scalaires sont nuls. Donc $A \in \mathcal{O}_3(\mathbb{R})$, soit $u \in \mathcal{O}(E_3)$

Le calcul du déterminant de A donne 1, donc $u \in \mathcal{SO}(E_3)$, et $u \neq \text{id}$. Donc u est une rotation d'axe d'angle θ orientée par w .

w est solution de $(\text{id} - u)(w) = 0$. On échelonne le système, et on obtient que w est solution de

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

Choisissons $w = 3e_1 + e_2 + e_3$, qu'on norme en $e'_1 = \frac{1}{\|w\|}w = \frac{1}{\sqrt{11}}w$

Pour le choix du deuxième élément de la base, on a juste besoin d'un élément de F^\perp , donc on peut choisir le vecteur normé qu'on veut. On prend $e'_2 = \frac{1}{\sqrt{10}}(e_1 - 3e_2)$

On n'a plus de choix pour le troisième élément de la base cependant. Pour déterminer un vecteur qui soit orthogonal aux deux précédents, on peut utiliser le produit vectoriel, et on a $e'_3 = \frac{1}{\sqrt{10\sqrt{11}}}(3e_1 + e_2 - 10e_3)$

On note $\mathcal{B}' = (e'_1, e'_2, e'_3)$

On a alors $\text{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$

(La matrice de passage de \mathcal{B} dans \mathcal{B}' est la matrice d'une BOND dans une BOND, donc son inverse est égal à sa transposée.)

On a que $\text{tr}(A) = \text{tr}(A')$, donc $\frac{-2}{3} = 1 + 2 \cos \theta$.

D'où $\cos \theta = \frac{-5}{6}$

Et donc $\theta = \pm \arccos\left(\frac{-5}{6}\right)$

Pour déterminer le signe, on fait le produit mixte de $w, e'_2, r_{w,\theta}(e'_2)$, qui vaut $\|w\| \sin \theta$ dans la base \mathcal{B}' .

Dans la base \mathcal{B} d'origine, on a que leur produit mixte vaut $\begin{vmatrix} 3 & \frac{1}{\sqrt{10}} & \frac{-5}{3\sqrt{10}} \\ 1 & \frac{-3}{\sqrt{10}} & \frac{3\sqrt{10}}{8} \\ 1 & 0 & \frac{-5}{3\sqrt{10}} \end{vmatrix} > 0$

Donc $\sin \theta > 0$

Donc $\theta = + \arccos\left(\frac{-5}{6}\right)$

VII Réduction des isométries dans un espace euclidien

Théorème :

Soit E un espace euclidien et $u \in \mathcal{O}(E)$ alors il existe une base orthonormale \mathcal{B} telle que $\text{Mat}_{\mathcal{B}}(u)$ est diagonale par bloc et chaque bloc est de la forme :

$$\begin{bmatrix} 1 \\ -1 \\ R_{\theta} \end{bmatrix}$$

pour $\theta \in \mathbb{R}$

ie chaque bloc est soit I , soit $-I$, soit une rotation.

Remarque VII.1. *Que l'isométrie soit directe ou non dépend du nombre de -1 dans cette matrice*

Les seules valeurs propres réelles possibles pour une isométrie sont 1 et -1 comme elle conserve la norme.

Preuve

Lemme préparatoire : tout endomorphisme d'un \mathbb{R} -ev E de dimension finie possède une droite stable ou un plan stable.

Si E un \mathbb{R} -ev de dimension finie, et $u \in \mathcal{L}(E)$:

- Si u possède une valeur propre réelle λ
Alors on dispose de $x \in E, x \neq 0$ tel que $u(x) = \lambda x$
Donc $\text{Vect}(x)$ est stable par u
- Si u n'a pas de valeur propres réelles : la décomposition en facteurs premiers de π_u dans $\mathbb{R}[X]$ ne comporte que des facteurs de degré 2 irréductibles.
 $\pi_u = (X^2 + aX + b)Q$ où $Q \in \mathbb{R}[X]$
On a donc $(u^2 + au + bid) \circ Q(u) = 0$
Par l'absurde, si $\ker(u^2 + au + bid) = \{0\}$, alors $u^2 + au + bid$ serait injectif. Comme on est en dimension finie, il serait bijectif.
Donc $Q(u) = 0$, et donc Q est annulateur de u avec $\deg Q < \deg \pi_u$
D'où la contradiction
On peut donc prendre $x \in E$ non-nul tel que $x \in \ker(u^2 + au + bid)$
 $(x, u(x))$ est libre car u n'a pas de valeurs propres réelles, donc $F = \text{Vect}(x, u(x))$ est un plan, $u(x) \in F$ et $u(u(x)) = -au(x) - bx$ donc dans F
Donc F stable par u

Récurrence sur $n \in \mathbb{N}^*$, H_n : Si E est un espace de dimension inférieure ou égale à n alors il existe une BON B de E dans laquelle la matrice est de la bonne forme.

H_1 et H_2 sont vérifiées

Soit $n \in \mathbb{N}^*$, on suppose H_n

Soit E un espace euclidien de dimension $n + 1$, soit $u \in \mathcal{O}(E)$:

- Si u possède une valeur propre réelle λ : cette valeur propre est 1 ou -1 . Alors $E_\lambda(u)$ est stable par u donc $E_\lambda(u)^\perp$ est stable par u^* donc $E_\lambda(u)^\perp$ est stable par u
On prend B_1 une BON de $E_\lambda(u)$, et on note u_1 l'endomorphisme induit par u sur $E_\lambda(u)$.
Alors $\text{Mat}_{B_1}(u_1) = \lambda \text{id}$
Par H_n , on prend B_2 une BON de $E_\lambda(u)^\perp$ telle que l'endomorphisme induit par u sur $E_\lambda(u)^\perp$ a une matrice de la forme recherchée.
Donc dans la BON $B = (B_1, B_2)$, $\text{Mat}_B(u)$ est de la forme recherchée.
- Si u n'a pas de valeurs propres réelles : il existe un plan stable de u , donc soit u induit sur ce plan est une symétrie orthogonale, soit c'est une rotation. Si c'était une symétrie orthogonale, sa matrice n'aurait que des 1 et des -1 sur la diagonale dans une BON, mais u n'a pas de valeurs propres réelles donc c'est la matrice d'une rotation qu'on doit utiliser et on concatène avec une BON de l'espace orthogonal (l'endomorphisme induit dessus est de la bonne forme par H_n)

VIII Les endomorphismes auto-adjoints

Définition :

Soit E un espace euclidien et $u \in \mathcal{L}(E)$.

On dit que u est auto-adjoint si $u^* = u$

ie $\forall x, y \in E, \langle u(x), y \rangle = \langle x, u(y) \rangle$

Proposition :

Soit E un espace euclidien de dimension n , soit \mathcal{B} une base orthonormale de E , soit $u \in \mathcal{L}(E)$
 u est auto-adjoint si $\text{Mat}_{\mathcal{B}}(u) \in \mathcal{S}_n(\mathbb{R})$

Remarque VIII.1. On dit parfois que u est *symétrique* pour désigner un endomorphisme auto-adjoint.

Exemple VIII.1. Un projecteur orthogonal est un endomorphisme *symétrique*/auto-adjoint.

Théorème : Théorème spectral

Tout endomorphisme auto-adjoint d'un espace euclidien est diagonalisable dans une BON.

Preuve

Lemme : si E est euclidien et $u \in \mathcal{L}(E)$ auto-adjoint, alors u possède une valeur propre réelle.

On considère \mathcal{B} une base orthonormale de E et $S = \text{Mat}_{\mathcal{B}}(u) \in \mathcal{S}_n(\mathbb{R})$.

On considère v l'endomorphisme de \mathbb{C}^n canoniquement associé à S

Alors $\chi_v = \chi_S = \chi_u$

Soit λ une valeur propre de v

On a donc $X \in \mathcal{M}_{n,1}(\mathbb{C})$, $X \neq 0$ tel que $SX = \lambda X$

$\bar{X}^T SX = \lambda \bar{X}^T X$

mais on a aussi $\bar{X}^T SX = (S\bar{X})^T X = (\bar{S}\bar{X})^T X = \bar{\lambda} \bar{X}^T X$

Et $\bar{X}^T X \in \mathbb{R}_+^*$

Donc $\lambda \in \mathbb{R}$

Or $v \in \mathcal{L}(\mathbb{C}^n)$ donc χ_v est scindé, donc v possède au moins une valeur propre qui est réelle. Donc S possède une valeur propre réelle.

Donc u en possède une.

Par récurrence forte, montrons que $\forall n \geq 1$, H_n : tout endomorphisme auto-adjoint d'un espace euclidien de dimension inférieure ou égale à n est diagonalisable en BON.

H_1 est vérifiée.

Soit $n \in \mathbb{N}^*$, on suppose H_n

Soit E un espace euclidien de dimension $n + 1$ et $u \in \mathcal{S}(E)$ (l'ensemble des endomorphismes auto-adjoints/symétriques)

Soit λ une valeur propre réelle de u (qui existe d'après le lemme)

On note $F = E_{\lambda}(u)$

F est stable par u , donc F^{\perp} est stable par u^* donc par u et l'endomorphisme u' induit par u sur F^{\perp} est auto-adjoint

$\dim F \geq 1$ donc $\dim F^{\perp} \leq n$

Donc par H_n , il existe \mathcal{B}' une BON de F^{\perp} telle que $\text{Mat}_{\mathcal{B}'} u'$ est diagonale.

On prend \mathcal{B}_1 une BON de F

$\mathcal{B} = (\mathcal{B}_1, \mathcal{B}')$ est une BON de E et $\text{Mat}_{\mathcal{B}} u$ est donc diagonale.

Ce qui conclut la récurrence.

Proposition : théorème spectral matriciel

Toute matrice symétrique réelle est orthogonalement semblable à une matrice diagonale.

ie $\forall S \in \mathcal{S}_n(\mathbb{R}), \exists D \in \mathcal{D}_n(\mathbb{R}), \exists P \in \mathcal{O}_n(\mathbb{R}), D = P^{-1}SP = P^TSP$

Preuve

Soit $S \in \mathcal{S}_n(\mathbb{R})$

Soit $u \in \mathcal{L}(\mathbb{R}^n)$ canoniquement associé à S , on munit \mathbb{R}^n du produit scalaire canonique (ie sa base canonique est une BON).

On note \mathcal{B} la base canonique de \mathbb{R}^n

u est auto-adjoint.

Donc d'après le théorème spectral, u est diagonalisable dans une BON \mathcal{B}'

La matrice P de passage de \mathcal{B} dans \mathcal{B}' est orthogonale et on a :

$$\text{Mat}_{\mathcal{B}'}(u) = P^{-1}SP = P^TSP \text{ où } \text{Mat}_{\mathcal{B}'} \in \mathcal{D}_n(\mathbb{R})$$

IX Endomorphismes auto-adjoints et formes bilinéaires symétriques

Théorème :

Soit E un espace euclidien et φ une fbs sur E . Il existe un unique endomorphisme auto-adjoint u tel que :

$$\forall x, y \in E, \varphi(x, y) = \langle x, u(y) \rangle$$

Preuve

Soit $y \in E$, on note ψ_y la forme linéaire $x \mapsto \varphi(x, y)$

Par théorème de représentation des formes linéaires, il existe un unique vecteur noté $u(y)$ tel que $x \mapsto \langle u(y), x \rangle = \psi_y$

On a montré l'existence d'une unique application de E dans E u vérifiant $\forall x, y \in E, \langle u(x), y \rangle = \varphi(x, y)$

Montrons que u est linéaire et $u = u^*$

Soient $x, x' \in E$, soit $\lambda \in \mathbb{R}$, soit $y \in E$

$$\langle u(\lambda x + x'), y \rangle = \varphi(\lambda x + x', y) = \lambda \varphi(x, y) + \varphi(x', y)$$

$$= \lambda \langle u(x), y \rangle + \langle u(x'), y \rangle = \langle \lambda u(x) + u(x'), y \rangle$$

$$\text{Donc } \langle u(\lambda x + x') - \lambda u(x) - u(x'), y \rangle = 0$$

$$\text{Donc } u(\lambda x + x') - \lambda u(x) - u(x') \in E^\perp$$

$$\text{Donc } u \in \mathcal{L}(E)$$

$$\text{De plus, } \forall x, y \in E, \varphi(x, y) = \varphi(y, x)$$

$$\text{Donc } \langle u(x), y \rangle = \langle x, u(y) \rangle$$

$$\text{Donc } u = u^*$$

Définition :

Soit E un espace euclidien, $u \in \mathcal{S}(E)$ un endomorphisme auto-adjoint.

On dit que :

- u est positif si $(x, y) \mapsto \langle u(x), y \rangle$ est une forme bilinéaire symétrique positive. On note l'ensemble de ces endomorphismes $\mathcal{S}_+(E)$.
- u est défini positif si $(x, y) \mapsto \langle u(x), y \rangle$ est un produit scalaire. On note l'ensemble de ces endomorphismes $\mathcal{S}_{++}(E)$.

Définition :

Soit $S \in \mathcal{S}_n(\mathbb{R})$.

On dit que :

- S est positive si $(X, Y) \mapsto X^T S Y$ est une forme bilinéaire symétrique positive sur $\mathcal{M}_{n,1}(\mathbb{R})$. On note l'ensemble de ces matrices $\mathcal{S}_n^+(\mathbb{R})$.
- u est défini positif si $(X, Y) \mapsto X^T S Y$ est un produit scalaire sur $\mathcal{M}_{n,1}(\mathbb{R})$. On note l'ensemble de ces endomorphismes $\mathcal{S}_n^{++}(\mathbb{R})$.

Théorème :

Soit E un espace euclidien et $u \in \mathcal{S}(E)$ alors :

- u est positif si, et seulement si, $S_p(u) \subset \mathbb{R}_+$
- u est défini positif si, et seulement si, $S_p(u) \subset \mathbb{R}_+^*$

Preuve

u est auto adjoint

Par théorème spectral, on peut trouver une BON (e_1, \dots, e_n) de vecteurs propres.

Les valeurs propres de u $\lambda_1, \dots, \lambda_n$

$$\forall x \in E, x = \sum_{i=1}^n x_i e_i$$

$$\langle x, u(x) \rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{i=1}^n \lambda_i x_i e_i \right\rangle$$

$$= \sum_{i=1}^n \lambda_i \langle x_i, x_i \rangle = \sum_{i=1}^n \lambda_i x_i^2$$

Ensuite, on raisonne par double implication :

Si $S_p(u) \subset \mathbb{R}_+$: Alors $\langle x, u(x) \rangle \geq 0$ comme on a une somme de termes positifs, donc u est positif

Si u est défini positif : Alors $\forall x \in E, \langle x, u(x) \rangle \geq 0$. En particulier, $\langle e_i, u(e_i) \rangle = \lambda_i \times 1^2$ (la base est normée), donc tous les λ_i sont positifs ou nuls.

De même pour u défini positif.

Exemple IX.1. $E = \mathbb{R}^3$ avec (e_1, e_2, e_3) la base canonique.

φ la fbs définie par :

$$\forall x = x_1 e_1 + x_2 e_2 + x_3 e_3, \varphi(x, x) = x_1^2 + 2x_2^2 + x_3^2 + x_1 x_2 - x_1 x_3$$

Pour se mettre dans le cadre, on munit E du produit scalaire canonique. Déterminer si φ est positive ou positive définie.

$$\text{On pose } \varphi(x, x) = X^T S X, \text{ et on obtient alors } S = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 2 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}$$

$$\text{On a } \chi_S = X^3 - 4X^2 + \left(\frac{7}{4} + 2 + \frac{3}{4}\right)X - \left(2 - \frac{1}{2} - \frac{1}{4}\right) = X^3 - 4X^2 + \frac{9}{2}X - \frac{5}{4}$$

On peut répondre à la question en dérivant le polynôme, en trouvant les racines de ce polynôme du second degré pour avoir un tableau de variations.

Cela permet de savoir que les racines sont strictement positives, et donc que φ est un produit scalaire.

X Hors-programme

X.1 Racine carrées des matrices symétriques définies positives

Théorème :

Soit $S \in \mathcal{S}_n^{++}(\mathbb{R})$, alors il existe une unique racine carrée définie positive de S .

Preuve

Analyse : soit s l'endomorphisme canoniquement associé à S dans \mathbb{R}^n muni du produit scalaire et de la base canonique.

On pose $\delta \in \mathcal{S}^{++}(\mathbb{R})$ tel que $\delta \circ \delta = s$

On a alors $\delta \circ s = \delta^3 = s \circ \delta$

Donc s et δ commutent, et les espaces propres de s sont stables par δ .

Soit $\lambda \in S_p(s)$, on note s_λ l'endomorphisme induit par s sur $E_\lambda(s)$ et δ_λ l'endomorphisme induit par δ sur $E_\lambda(s)$

Prenons $\mu \in S_p(\delta_\lambda)$ et x un vecteur propre de δ_λ associé à la valeur propre μ

$s_\lambda(x) = \lambda x$ et $\delta_\lambda(x) = \mu x$

Donc $s_\lambda(x) = \delta_\lambda \circ \delta_\lambda(x) = \mu^2 x$

Donc $\mu^2 = \lambda$

δ est définie positive donc $\mu > 0$, d'où $\mu = +\sqrt{\lambda}$

δ_λ est auto-adjoint et diagonalisable, il a dnc 1 valeur propre sur chaque espace propre.

Donc $\delta_\lambda(x) = \sqrt{\lambda}x$

Et comme $E = \bigoplus_{\lambda \in S_p(s)} E_\lambda(s)$, δ est unique.

Synthèse : Soit $s \in \mathcal{S}^{++}(\mathbb{R}^n)$, notons δ l'endomorphisme tel que $\forall \lambda \in S_p(s) \forall x \in E_\lambda(s), \delta(x) = \sqrt{\lambda}x$

On a bien que δ est un endomorphisme.

Pour $\lambda \in S_p(s)$, notons \mathcal{B}_λ une base orthonormale de $E_\lambda(s)$.

$\mathcal{B} = (\mathcal{B}_{\lambda_1}, \dots, \mathcal{B}_{\lambda_p})$ est une base orthonormale de E telle que $Mat_{\mathcal{B}}(\delta) \in \mathcal{S}_n(\mathbb{R})$

Donc δ est bien auto-adjoint.

On a de plus que $S_p(\delta) = \{\sqrt{\lambda} | \lambda \in S_p(s)\}$, donc $S_p(\delta) \subset \mathbb{R}_+^*$

Donc $\delta \in \mathcal{S}^{++}(\mathbb{R}^n)$

X.2 Matrices de Gramm d'une famille de vecteurs

Définition :

E un \mathbb{R} -ev euclidien de dimension n ou un espace préhilbertien.

(u_1, \dots, u_p) une famille de E .

$G(u_1, \dots, u_p) = (\langle u_i, u_j \rangle)_{(i,j) \in \llbracket 1, p \rrbracket^2}$

(immédiatement, $G \in S_p(\mathbb{R})$)

Proposition :

$F = Vect(u_1, \dots, u_p)$ avec $\dim F = q \leq p$

\mathcal{B} une BON de F , $\mathcal{B} = (f_1, \dots, f_q)$

$A = Mat_{\mathcal{B}}(u_1, \dots, u_p) \in \mathcal{M}_{q,p}(\mathbb{R})$

Alors $A = (\langle u_j, f_i \rangle)_{(i,j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket}$

De plus, $A^T A = G(u_1, \dots, u_p)$

Preuve

$$\begin{aligned}
A^T A &= \sum_{k=1}^q A_{i,k}^T A_{k,j} = \sum_{k=1}^q A_{k,i} A_{k,j} \\
&= \sum_{k=1}^q \langle u_i, f_k \rangle \langle u_j, f_k \rangle = \langle u_i, u_j \rangle \\
&\text{(puisque } f_1, \dots, f_q \text{ est une BON de } F)
\end{aligned}$$

Proposition : Corollaire

(u_1, \dots, u_p) et $G(u_1, \dots, u_p)$ sont de même rang.

Preuve

Soit A la matrice de (u_1, \dots, u_p) dans une BON de F
 $rg(A) = rg(u_1, \dots, u_p)$ et $G(u_1, \dots, u_p) = A^T A$
 Soit $X \in \mathcal{M}_{p,1}(\mathbb{R})$
 $X \in \ker A \Rightarrow X \in \ker(A^T A)$
 Réciproquement :
 Si $X \in \ker(A^T A)$, alors $A^T A X = 0$
 Donc $X^T A^T A X = 0$
 D'où $(AX)^T (AX) = 0$ qui est le produit scalaire canonique de AX avec lui-même.
 Par définition du produit scalaire, $AX = 0$ donc $X \in \ker A$
 Donc $\ker A = \ker A^T A$
 $A \in \mathcal{M}_{q,p}(\mathbb{R})$ et $A^T A \in \mathcal{M}_p(\mathbb{R})$
 $a \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ canoniquement associée à A , $b \in \mathcal{L}(\mathbb{R}^p)$ canoniquement associée à $A^T A$
 $\ker a = \ker b$ donc par théorème du rang, $rg(a) = rg(b)$
 D'où le résultat.

Proposition :

$G(u_1, \dots, u_p) \in \mathcal{S}_p^+(\mathbb{R})$
 et $G(u_1, \dots, u_p) \in \mathcal{S}_p^{++}(\mathbb{R})$ si, et seulement si, (u_1, \dots, u_p) est libre

Preuve

Rappel : $S \in \mathcal{S}_p(\mathbb{R}) : S \in \mathcal{S}_p^+(\mathbb{R}) \Leftrightarrow S_p(S) \subset \mathbb{R}_+ \Leftrightarrow \forall X \in \sqrt{\cdot}, \infty(\mathbb{R}), X^T S X \geq 0$
 On prend $A \in \mathcal{M}_{q,p}(\mathbb{R})$ telle que $G(u_1, \dots, u_p) = A^T A$
 Soit $X \in \mathcal{M}_{p,1}(\mathbb{R})$, alors $X^T G(u_1, \dots, u_p) X = (AX)^T (AX) \geq 0$
 Donc $G(u_1, \dots, u_p) \in \mathcal{S}_p^+(\mathbb{R})$
 Le rang donne l'équivalence.

Remarque X.1. On a le problème réciproque : Si $S \in \mathcal{S}_p^{++}(\mathbb{R})$, E un espace euclidien de dimension $n \geq p$ (ou un espace préhilbertien). Existe-t-il p vecteurs de (u_1, \dots, u_p) tels que $G(u_1, \dots, u_p) = S$?

On fixe \mathcal{B} une BON de E et on considère $s \in \mathcal{L}(E)$ tel que $\text{Mat}_{\mathcal{B}}(s) = S$

Comme \mathcal{B} est une BON, alors $s \in S^{++}(E)$

Par théorème spectral, on considère $\mathcal{B}' = (e'_1, \dots, e'_p)$ une BON de E des vecteurs propres de s .

Alors $\text{Mat}_{\mathcal{B}'} = \text{Diag}(\lambda_1, \dots, \lambda_p)$, et $\forall i \in \llbracket 1, p \rrbracket, \lambda_i > 0$

Soit $\Delta = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_p})$

X.3 Familles isogonales

Dans un espace euclidien E

Définition :

Une famille $(u_i)_{1 \leq i \leq p}$ d'un espace euclidien est dite isogonale si :

$$\begin{cases} \forall i \in \llbracket 1, p \rrbracket, \|u_i\| = 1 \\ \forall i, j \in \llbracket 1, p \rrbracket, i \neq j \Rightarrow \|u_i - u_j\| = d \end{cases}$$

Exemple X.1. On a $\|u_i - u_j\|^2 = \|u_i\|^2 + \|u_j\|^2 - 2\langle u_i, u_j \rangle = 2(1 - \langle u_i, u_j \rangle) = d^2$

On note $K = \langle u_i, u_j \rangle$ pour $i \neq j$, qui vaut donc $\frac{2-d^2}{2}$

On a $G(u_1, \dots, u_p) = S_K$ une matrice avec deux blocs K au-dessus de la diagonale des 1 (la matrice est symétrique).

Les valeurs propres de S_K sont $1 - K$, qui est d'ordre $p - 1$. Comme la trace correspond à la somme des valeurs propres, on déduit que la dernière valeur propre est $x = p - (p - 1)(1 - K)$

S_K est symétrique positive si, et seulement si, $1 - K \geq 0$ et $p - (p - 1)(1 - K) \geq 0$

$$\Leftrightarrow \frac{-1}{p-1} \leq K \leq 1$$

Cas $K = \frac{-1}{p-1}$:

S_K est de rang $p - 1$ (une valeur propre est nulle, mais toutes les autres sont positives), si (u_1, \dots, u_p) est une rotation, $rg(u_1, \dots, u_p) = p - 1$

Dans E de dimension n , trouver $n + 1$ vecteurs (u_1, \dots, u_{n+1}) des vecteurs tels que $G(u_1, \dots, u_{n+1}) = S_n = \begin{pmatrix} 1 & \frac{-1}{n} \\ \frac{-1}{n} & 1 \end{pmatrix}$

La matrice S_n de $\mathcal{S}_n(\mathbb{R})$ est symétrique définie positive, on peut bien trouver (u_1, \dots, u_n) telle que $G(u_1, \dots, u_n) = S_n$

On cherche $u_{n+1} : u_{n+1} = \sum_{i=1}^n \alpha_i u_i$, où $\forall j \in \llbracket 1, n \rrbracket, \langle u_{n+1}, u_j \rangle = \frac{-1}{n}$

$$= \sum_{i=1}^n \alpha_i \langle u_i, u_j \rangle = \frac{-1}{n} \sum_{1 \leq i \leq n, i \neq j} \alpha_i + \alpha_j$$

Si on prend $u_{n+1} = -\sum_{i=1}^n u_i$, on a bien le résultat.

$$\text{Vérifions que } \langle u_{n+1}, u_{n+1} \rangle = \langle -\sum_{i=1}^n u_i, -\sum_{i=1}^n u_i \rangle = \sum_{1 \leq i, j \leq n} \langle u_i, u_j \rangle = -\frac{1}{n}n(n-1) + n = 1$$

On a donc cette formule pour les familles isogonales maximales.

Chapitre XI

Equations différentielles linéaires

Dans le cadre de ce chapitre, E est un \mathbb{K} -ev de dimension finie (\mathbb{K} est \mathbb{R} ou \mathbb{C}) et I est un intervalle de \mathbb{R} .

On s'intéresse aux équations différentielles dont les solutions sont des fonctions de I dans E .

Exemple .1. $y' + xy = e^x$ est une équation différentielle linéaire du premier ordre, scalaire, normalisée (le coefficient devant y' vaut 1)

Ses solutions sont les fonctions $y \in \mathcal{D}(I, \mathbb{K})$ telles que $\forall x \in \mathbb{R}, y'(x) + xy(x) = e^x$

$y' + xy^2 = e^x$ n'est pas linéaire.

$y'y + 3y = e^x$ n'est pas linéaire non plus.

$\begin{cases} x'(t) = 3x(t) + t^2y(t) + e^t \\ y'(t) = 2x(t) + 3y(t) - e^t \end{cases}$ est une équation différentielle linéaire, mais non scalaire.

$y'' + t^2y' + (e^t + 1)y = \sin(t)$ est une équation différentielle linéaire du premier ordre scalaire normalisée.

I Définitions

Définition : Premier ordre

Soit I un intervalle, soit E un \mathbb{K} -ev de dimension finie.

Soit a une application continue de I dans $\mathcal{L}(E)$

Soit b une application continue de I dans E

On appelle $y' + a \cdot y = b$ équation différentielle linéaire du premier ordre normalisée.

Ses solutions sont les fonctions $y \in \mathcal{D}(I, E)$ vérifiant $\forall t \in I, y'(t) + a(t) \cdot y(t) = b(t)$

Définition : Traduction matricielle

Avec les notations ci-dessus, on fixe une base \mathcal{B} de E .

$\forall t \in I, Y(t) = \text{Mat}_{\mathcal{B}}y(t) \Rightarrow Y'(t) = \text{Mat}_{\mathcal{B}}y'(t) \in \mathcal{M}_{n,1}(\mathbb{K})$

$A(t) = \text{Mat}_{\mathcal{B}}a(t) \in \mathcal{M}_n(\mathbb{K})$

$B(t) = \text{Mat}_{\mathcal{B}}b(t) \in \mathcal{M}_{n,1}(\mathbb{K})$

L'équation devient :

$$Y'(t) = A(t)Y(t) + B(t)$$

Exemple I.1. $\begin{cases} x'(t) = 2tx(t) + 3t^2y(t) + \sin(t) \\ y'(t) = (3t + 1)x(t) + \frac{1}{1+t^2}y(t) + t \end{cases}$
Alors $Y'(t) = \begin{pmatrix} 2t & 3t^2 \\ (3t + 1) & \frac{1}{1+t^2} \end{pmatrix} Y(t) + \begin{pmatrix} \sin(t) \\ t \end{pmatrix}$

Exemple I.2. Exemple de vectorialisation d'une équation du second ordre :

$$(E) : y''(t) + 3ty'(t) + \sin(t)y(t) = t^2$$

$$\text{On pose } Y(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}$$

$$Y \text{ solution de } (E) \Leftrightarrow Y'(t) = \begin{pmatrix} 0 & 1 \\ -\sin(t) & -3t \end{pmatrix} Y(t) + \begin{pmatrix} 0 \\ t^2 \end{pmatrix}$$

On transforme donc cette équation différentielle linéaire scalaire du second ordre en une équation différentielle linéaire vectorielle du premier ordre.

Définition : Equations différentielles linéaires normalisées scalaires d'ordre p

L'équation $y^{(p)} + a_{p-1}y^{(p-1)} + \dots + a_0y = b$ où :

$\forall i \in \llbracket 0, p-1 \rrbracket$, a_i est une application continue de I dans \mathbb{K}

et b est une application continue de I dans \mathbb{K}

est une équation différentielle linéaire normalisée scalaire d'ordre p

Ses solutions sont les fonctions $y \in \mathcal{D}_p(I, \mathbb{K})$ vérifiant :

$$\forall t \in I, y^{(p)}(t) + a_{p-1}(t)y^{(p-1)}(t) + \dots + a_0(t)y(t) = b(t)$$

Proposition : Vectorialisation

Toute équation différentielle linéaire normalisée scalaire (éqdlns) d'ordre p est équivalente à une équation différentielle linéaire d'ordre 1 dont les solutions sont à valeurs dans \mathbb{K}^p

Preuve

$(E) : y^{(p)} + a_{p-1}y^{(p-1)} + \dots + a_0y = b$ une telle équation.

Pour $y \in \mathcal{D}_p(I, \mathbb{K})$, on pose $z = (y, y', \dots, y^{(p-1)})$

On a $z \in \mathcal{D}_1(I, \mathbb{K})$.

y est solution de (E) si, et seulement si, $\forall t \in I, z'(t) = (y'(t), \dots, y^{(p)}(t)) = a(t) \cdot z(t) + b_2(t)$

$$\text{où } a(t) : \begin{cases} \mathbb{K}^p \rightarrow \mathbb{K}^p \\ (x_0, \dots, x_p) \mapsto (x_1, \dots, x_{p-1}, -\sum_{i=1}^{p-1} a_i(t)x_i) \end{cases}$$

$$\text{et } b_2(t) = (0, \dots, 0, b(t)) \in \mathbb{K}^p$$

Proposition : Traduction matricielle

On va noter $Y(t) = \begin{pmatrix} y(t) \\ \vdots \\ y^{(p-1)}(t) \end{pmatrix}$, et dans ce cas la vectorialisation sera :

$$Y'(t) = \begin{pmatrix} y'(t) \\ \vdots \\ y^{(p)}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & -a_1 & -a_2 & \dots & -a_{p-1} \end{pmatrix} \begin{pmatrix} y(t) \\ \vdots \\ y^{(p-1)}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$$

II Généralités sur l'équation du premier ordre

Définition : Equation homogène associée

Avec $(E_q) : y' + a \cdot y = b$
 (où $a \in \mathcal{C}(I, \mathcal{L}(E))$ et $b \in \mathcal{C}(I, E)$)
 L'équation homogène associée est :

$$(H) : y' + a \cdot y = 0$$

Proposition :

S_H , l'ensemble des solutions de (H) , est un \mathbb{K} -ev.

Preuve

On peut montrer que c'est un sous-espace vectoriel de $\mathcal{D}(I, E)$

En effet, 0 est nécessairement solution de (H)

Et si on prend y_1, y_2 deux solutions et $\lambda \in \mathbb{K}$, par linéarité de $a(t)$ on obtient que $y_1 + \lambda y_2$ est solution de (H) aussi.

Proposition :

S_E , l'ensemble des solutions de (E_q) , est un espace affine.

ie si y_0 est une solution particulière de (E_q) , alors on a que $S_{E_q} = y_0 + S_H = \{y_0 + y_H | y_H \in S_H\}$

Preuve

Soit y_0 une solution de (E_q)

Soit $y \in \mathcal{D}(I, E)$

y est solution $\Leftrightarrow \forall t \in I, y'(t) + a(t) \cdot y(t) = y'_0(t) + a(t) \cdot y_0(t)$

$$\Leftrightarrow \forall t \in I, (y - y_0)'(t) + a(t) \cdot ((y - y_0)(t)) = 0$$

$$\Leftrightarrow y - y_0 \in S_H$$

Théorème : Superposition

(E_{q_1}) et (E_{q_2}) deux équations de même équation homogène associé et $\lambda \in \mathbb{K}$:

$$(E_{q_1}) : y' + a(t) \cdot y = b_1$$

$$(E_{q_2}) : y' + a(t) \cdot y = b_2$$

$$(E_{q_+}) : y' + a(t) \cdot y = b_1 + b_2$$

$$(E_{q_\lambda}) : y' + a(t) \cdot y = \lambda b_1$$

Si $y_1 \in S_{E_{q_1}}$ et $y_2 \in S_{E_{q_2}}$, alors :

$$y_1 + y_2 \in S_{E_{q_+}}$$

et

$$\lambda y_1 \in S_{E_{q_\lambda}}$$

Preuve

Immédiate, en se servant du fait que les ensembles de solutions sont des espaces affines.

Théorème : Cauchy-Lipschitz linéaire

Soit I un intervalle de \mathbb{R}

Soit E un \mathbb{K} -ev de dimension finie

Soit (E_q) une équation différentielle linéaire normalisée $y' + a \cdot y = b$

Soit $t_0 \in I$, $y_0 \in E$, alors :

Il existe une unique solution f de (E_q) vérifiant :

$$f(t_0) = y_0$$

Remarque II.1. Toutes les conditions sont importantes :

Si on n'est pas sur un intervalle : $(E_q)y' = 0$ sur \mathbb{R}^* a comme solutions les fonctions qui valent une constante sur \mathbb{R}_+^* et une autre constante sur \mathbb{R}_-^* . Il y a donc un \mathbb{R} -ev de dimension 2 de solutions.

Si l'équation n'est pas normalisée (ou normalisable) : $(E_q) = t^3 y' - y = 0$ (on cherche des solutions réelles)

Sur $I = \mathbb{R}_+^*$ ou \mathbb{R}_-^* , l'équation est normalisable et $S_I = \{t \mapsto \alpha e^{-\frac{1}{2t^2}} \mid \alpha \in \mathbb{R}\}$

Quelles sont les solutions sur \mathbb{R} ? On procède par analyse-synthèse.

Soit f une solution sur \mathbb{R} , alors $f|_{\mathbb{R}_+^*}$ et $f|_{\mathbb{R}_-^*}$ sont solutions.

Donc il existe $\alpha, \beta \in \mathbb{R}$ tels que : $\forall t \in \mathbb{R}_+^*, f|_{\mathbb{R}_+^*}(t) = \alpha e^{-\frac{1}{2t^2}}$ et $\forall t \in \mathbb{R}_-^*, f(t) = \beta e^{-\frac{1}{2t^2}}$

Par continuité de f en 0, on a $f(0) = 0$, et donc aucune condition en plus.

Synthèse : Pour tout $\alpha, \beta \in \mathbb{R}$, l'application $f : t \mapsto \begin{cases} \alpha e^{-\frac{1}{2t^2}} & t > 0 \\ \beta e^{-\frac{1}{2t^2}} & t < 0 \\ 0 & t = 0 \end{cases}$ est solution de (E_q)

Si on ajoute $t_0, y_0 \in \mathbb{R}$ tel que $f(t_0) = y_0$ alors : si $t_0 < 0$ alors on peut toujours faire varier α , si $t_0 > 0$ on pourra faire varier β ; dans les deux cas on aura une infinité de solutions. Si $t_0 = 0$ et $y_0 \neq 0$, on n'a aucune solution. Donc il n'y a ni existence ni unicité de la solution.

Exemple II.1. Exemple d'équation non-linéaire : $(E_q) : y' = 1 + y^2$ sur \mathbb{R}

Soit f une solution de (E_q) sur un intervalle I

Alors $\frac{f'}{1+f^2} = 1$

Donc il existe $K \in \mathbb{R}$ tel que $\forall t \in I, \arctan f(x) = x - k$

Donc $\forall x \in I, f(x) = \tan(x - k)$

Les solutions ne sont pas un espace affine.

Proposition : Corollaire

Si I est un intervalle et $(H) : y' = a \cdot y$ une équation différentielle linéaire **normalisée**
L'ensemble des solutions de (H) sur I à valeurs dans E est un \mathbb{K} -ev de dimension $\dim E$

Preuve

On fixe t_0 et on prend $\varphi : \begin{cases} S_H(I) & \rightarrow E \\ f & \mapsto f(t_0) \end{cases}$
 φ est un isomorphisme, donc $\dim E = \dim S_H(I)$

Définition : Wronskien

Soit I un intervalle, E un \mathbb{K} -ev de dimension p , \mathcal{B} une base de E

Soit (H) une équation différentielle linéaire homogène normalisée $(H) : y' + a \cdot y = 0$

Pour (u_1, \dots, u_p) une famille de p fonctions solutions de (H) , on définit leur wronskien relativement à la base \mathcal{B} :

$$w(u_1, \dots, u_p) : \begin{cases} I & \rightarrow E \\ t & \mapsto \det_{\mathcal{B}}(u_1(t), \dots, u_p(t)) \end{cases}$$

(Ou $w(u_1, \dots, u_p) = \det_{\mathcal{B}}(u_1, \dots, u_p)$)

Proposition :

$w(u_1, \dots, u_p) = 0$ si, et seulement si, $\exists t \in I, w(u_1(t), \dots, u_p(t)) = 0$

Preuve

Si $\forall t \in I, w(u_1(t), \dots, u_p(t)) = 0$, alors il existe $t \in I$ tel que $w(u_1(t), \dots, u_p(t)) = 0$

Réciproquement, supposons qu'il existe $t_0 \in I$ tel que $\det_{\mathcal{B}}(u_1(t_0), \dots, u_p(t_0)) = 0$

On a alors $(u_1(t_0), \dots, u_p(t_0))$ famille liée de E

On a donc $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ non tous nuls tels que $\lambda_1 u_1(t_0) + \dots + \lambda_p u_p(t_0) = 0$

Donc $\lambda_1 u_1 + \dots + \lambda_p u_p$ est une solution de (H) qui s'annule en t_0

Donc par Cauchy-Lipschitz, $\lambda_1 u_1 + \dots + \lambda_p u_p = 0$ (comme 0 est déjà solution qui s'annule en t_0 et qu'on a l'unicité de la solution du problème de Cauchy)

Donc $\forall t \in I, (u_1(t), \dots, u_p(t))$ est liée.

Proposition : Autre formulation

(u_1, \dots, u_p) est une base de S_H si, et seulement si, $\exists t \in I, (u_1(t), \dots, u_p(t))$ est une base de E .
 $\Leftrightarrow \forall t \in I, (u_1(t), \dots, u_p(t))$ est une base de E

Exemple II.2. $(H) : t^2 y'' + ty' + y = 0$ sur \mathbb{R}_+^* :

\mathbb{R}_+^* est un intervalle, (H) est une équation différentielle linéaire normalisable donc S_H est un \mathbb{K} -ev de dimension 2.

Cherchons les solutions de la forme $t \mapsto t^\alpha$.

On constate que $f_1 : t \mapsto t^i = e^{i \ln(t)}$ et $f_2 : t \mapsto t^{-i} = e^{-i \ln(t)}$ sont solutions.

$$\text{Pour } t \in \mathbb{R}_+^*, w(f_1(t), f_2(t)) = \begin{vmatrix} f_1(t) & f_2(t) \\ f_1'(t) & f_2'(t) \end{vmatrix} = \begin{vmatrix} e^{i \ln(t)} & e^{-i \ln(t)} \\ \frac{i}{t} e^{i \ln(t)} & -\frac{i}{t} e^{-i \ln(t)} \end{vmatrix} = \frac{-2i}{t}$$

Donc $w(f_1, f_2) \neq 0$

f_1 et f_2 sont donc indépendantes

Donc dans \mathbb{C} , $S_H(\mathbb{R}_+^*) = \text{Vect}(f_1, f_2) = \{\lambda_1 f_1 + \lambda_2 f_2 \mid \lambda_1, \lambda_2 \in \mathbb{C}\}$

Pour les solutions à valeurs réelles, on peut prendre $g_1 = \frac{f_1 + f_2}{2} = \cos(\ln(t))$ et $g_2 = \frac{f_1 - f_2}{2i} = \sin(\ln(t))$ qui sont deux solutions indépendantes et à valeurs réelles, donc une base des solutions réelles.

Théorème : Variation des constantes

Soit I un intervalle, (E_q) une équation différentielle linéaire normalisée $y' = a \cdot y + b$

(H) l'équation homogène associée.

Si (u_1, \dots, u_p) est une base de S_H

Alors (E_q) possède une solution particulière de la forme $t \mapsto \lambda_1(t) + \dots + \lambda_p(t)u_p(t)$

où $\lambda_1, \dots, \lambda_p$ sont des fonctions dérivables à valeurs dans \mathbb{K}

Preuve

(u_1, \dots, u_p) base de S_H , pour $\lambda_1, \dots, \lambda_p \in \mathcal{D}(I, \mathbb{K})$, on considère $v = \lambda_1 u_1 + \dots + \lambda_p u_p$

v est dérivable et $v' = \lambda_1' u_1 + \lambda_1 u_1' + \dots + \lambda_p' u_p + \lambda_p u_p' = \sum_{i=1}^p (\lambda_i' u_i + \lambda_i u_i')$

(1) : v est solution de (E_q) si, et seulement si, $\forall t \in I, \sum_{i=1}^p (\lambda_i'(t) u_i(t) + \lambda_i(t) u_i'(t)) = a(t) \cdot$

$$\left(\sum_{i=1}^p \lambda_i(t) u_i(t) \right) + b(t)$$

$$= \sum_{i=1}^p \lambda_i(t) a(t) u_i(t) + b(t)$$

$$(1) \Leftrightarrow \sum_{i=1}^p \lambda_i'(t) u_i(t) = b(t)$$

ce qui vient du fait que $\forall i \in \llbracket 1, p \rrbracket, u_i'(t) = a(t) \cdot u_i(t)$

On a vu que $\forall t \in I, (u_1(t), \dots, u_p(t))$ est une base de E

Donc pour $t \in I$, on peut écrire $b(t) = \sum_{i=1}^p b_i(t) u_i(t)$ où $(b_1(t), \dots, b_p(t)) \in \mathbb{K}^p$ sont les coordonnées de $b(t)$ dans la base $(u_1(t), \dots, u_p(t))$

$$\text{Donc (1)} \Leftrightarrow \begin{cases} \lambda_1' = b_1 \\ \vdots \\ \lambda_p' = b_p \end{cases}$$

On admet que les b_i sont continus, et alors le système a une solution.

Exemple II.3. Avec $(E_q) : y'' + y = \frac{1}{\cos(t)}$ sur $] -\frac{\pi}{2}, \frac{\pi}{2}[$, cherchons les solutions réelles.

Les solutions de (H) ($y'' + y = 0$) sont $S_H = \{t \mapsto \alpha \cos(t) + \beta \sin(t) \mid \alpha, \beta \in \mathbb{R}\}$

On recherche une solution particulière y de (E) telle que :

$$\forall t \in] -\frac{\pi}{2}, \frac{\pi}{2}[, y(t) = \alpha(t) \cos(t) + \beta(t) \sin(t) \text{ et } y'(t) = \alpha(t)(-\sin(t)) + \beta(t) \cos(t)$$

On a alors que $\alpha'(t) \cos(t) + \beta'(t) \sin(t) = 0$

$$\text{Et } y''(t) = \alpha(t)(-\cos(t)) + \beta(t)(-\sin(t)) + \alpha'(t)(-\sin(t)) + \beta'(t) \cos(t)$$

$$\text{Et on a alors que } \alpha'(t)(-\sin(t)) + \beta'(t) \cos(t) = \frac{1}{\cos(t)}$$

$$\text{Donc } \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\cos t} \end{pmatrix}$$

$$\text{Donc } \alpha'(t) = -\tan(t) \text{ et } \beta'(t) = 1$$

$$\text{Et donc, à constante près : } \alpha(t) = \ln(\cos t) \text{ et } \beta(t) = t$$

Théorème : Résolution de l'équation différentielle linéaire normalisée à coefficients constants

Notons $\forall t \in I, y'(t) = a \cdot y(t)$, où $a \in \mathcal{L}(E)$

L'ensemble des solutions de cette équation est : $\{t \mapsto \exp(ta) \cdot x | x \in E\}$

plus précisément, la solution du problème de Cauchy $\begin{cases} y' = a \cdot y \\ y(t_0) = x_0 \end{cases}$ est $t \mapsto \exp((t - t_0)a) \cdot x_0$

Preuve

Rappel sur $\exp(ta - t_0a) \cdot x_0 = (\exp(ta) \circ (\exp(-t_0a))) \cdot x_0$: pour le démontrer, on passe aux matrices, et on le prouve coordonnées par coordonnées en utilisant un binôme de Newton et un produit de Cauchy.

Si $\varphi : \begin{cases} \mathbb{R} \rightarrow \mathbb{R}^p \\ t \mapsto \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_p(t) \end{pmatrix} \end{cases}$, on dit que φ est dérivable si toutes ses coordonnées le sont.

Notons $f : \begin{cases} \mathbb{R} \rightarrow \mathcal{L}(E) \\ t \mapsto \exp(ta) \end{cases}$

Il faut montrer que f est dérivable et que $f' : t \mapsto (\exp(ta) \circ a)x = a \cdot f(t)$, puis dans une deuxième partie montrer que toutes les équations différentielles linéaires à coefficients constants ont cette solution.

Utilisons d'abord le théorème de dérivation terme à terme :

On note $f(t) = \sum_{i=1}^{+\infty} u_n(t)$

$u_n = \frac{t^n}{n!} a^n \in \mathcal{C}^1(I, E)$

$\sum u_n$ converge simplement

$\sum u'_n$ converge uniformément

Donc f est \mathcal{C}^1 et $f' = \sum u'_n$

$\|\cdot\|$ une norme d'algèbre, $\forall t \in \mathbb{R}, \|\frac{t^n}{n!} a^n\| \leq \frac{(\|t\| \|a\|)^n}{n!}$

On a $\forall t \in \mathbb{R}, \forall n \in \mathbb{N}^*, u'_n(t) = \frac{t^{n-1}}{(n-1)!} a^n$

Et donc $\forall [\alpha, \beta] \subset \mathbb{R}, \forall t \in [\alpha, \beta], \forall n \in \mathbb{N}^*, \|u_n(t)\| \leq \frac{\max(|\alpha|, |\beta|) \|a\|^n}{n!}$

Donc $f'(t) = \sum_{n=1}^{+\infty} \frac{t^{n-1}}{(n-1)!} a^n = a \circ \exp(ta) = \exp(ta) \circ a$

Prouvons le reste : posons $\psi : \begin{cases} \mathbb{R} \rightarrow E \\ t \mapsto \varphi(t) \cdot x \end{cases}$ pour $\varphi : \mathbb{R} \rightarrow \mathcal{L}(E)$

Alors ψ est dérivable et $\forall t \in \mathbb{R}, \psi'(t) = \varphi'(t) \cdot x$, donc φ admet un DL au premier ordre.

$\psi(x+h) = \varphi(t+h) \cdot x = (\varphi(t) + h\varphi'(t) + h\varepsilon(h)) \cdot x$

$= \varphi(t) \cdot x + h\varphi'(t) \cdot x + h\varepsilon(h) \cdot x$

et $\|\varepsilon(h) \cdot x\| \leq \|\varepsilon(h)\| \|x\|$ qui tend vers 0 quand h tend vers 0

Exemple II.4. $\begin{cases} x' = 2x + y + z \\ y' = x + 2y + z \\ z' = x + y + 2z \end{cases}$

On a $Y' = AY$ avec $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ et $Y = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

On a immédiatement que 1 est vp d'ordre 2, donc 4 est vp d'ordre 1 et A est diagonalisable.

D est semblable à A avec $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, avec la matrice de passage $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$

On aura $A = PDP^{-1}$, et donc $\exp(tA) = P \begin{pmatrix} \exp(t) & 0 & 0 \\ 0 & \exp(t) & 0 \\ 0 & 0 & \exp(4t) \end{pmatrix} P^{-1}$

On peut poser $Y = PZ$, avec $Z = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$

Alors $PZ' = APZ \Leftrightarrow Z' = P^{-1}APZ = DZ$

Donc $Z = \exp(tD) \cdot X$

Donc $Y = P \begin{pmatrix} \alpha e^t \\ \beta e^t \\ \gamma e^{4t} \end{pmatrix}$

III Topo sur les exponentielles

III.1 Rappel sur les familles sommables

Soit $(a_i)_{i \in I} \in \mathbb{C}^I$ avec I dénombrable est. (a_i) est sommable si $\sum |a_i| < +\infty$

Alors pour toute bijection $\sigma : \mathbb{N} \rightarrow I$, on a $\sum_{i \in I} a_i = \sum_{n \in \mathbb{N}} a_{\sigma(n)}$ ($\sum_{n \in \mathbb{N}} a_{\sigma(n)}$ ne dépend pas de la bijection σ)

Si $(I_j)_{j \in J}$ est une partition de I dénombrable, et $(a_i)_{i \in I} \in \mathbb{C}^I$, si $\sum_{i \in I} |a_i| < +\infty$, alors $\sum_{i \in I} a_i = \sum_{j \in J} \sum_{k \in I_j} a_k$

Extension pour $(a_i) \in E^I$ où E un \mathbb{K} -ev de dimension finie, où $\|\cdot\|$ est une norme (équivalente aux autres normes comme on est en dimension finie) :

Si $\sum_{i \in I} \|a_i\| < +\infty$, alors la famille est sommable et pour toute bijection $\sigma : \mathbb{N} \rightarrow I$, on a que $\sum_{n \in \mathbb{N}} a_{\sigma(n)}$ ne dépend pas de la bijection choisie et $\sum_{i \in I} a_i = \sum_{n \in \mathbb{N}} a_{\sigma(n)}$

On le prouve en prenant une base (e_1, \dots, e_p) de E et en se munissant de la $\|\cdot\|_\infty$

$\forall i \in I, a_i = [a_i]^1 e_1 + \dots + [a_i]^p e_p$ et $\forall i \in I, \forall j \in \llbracket 1, p \rrbracket, |[a_i]^j| \leq \|a_i\|_\infty$

III.2 Exponentielle dans une algèbre de dimension finie

Définition : Exponentielle

Si E est une algèbre de dimension finie, on note :

$$\exp(u) = \sum_{n \in \mathbb{N}} \frac{1}{n!} u^n$$

On munit E d'une norme d'algèbre, alors $\|\frac{1}{n!} u^n\| \leq \frac{\|u\|^n}{n!}$ et donc $\sum_{n \in \mathbb{N}} \frac{1}{n!} u^n$ est sommable.

Proposition : Morphisme d'algèbre

Si $u, v \in E$ tels que $uv = vu$
 $\left(\frac{1}{p!} u^p \frac{1}{q!} v^q\right)_{p, q \in \mathbb{N}}$ est sommable et :

$$\begin{aligned} \exp(u) \exp(v) &= \sum_{p, q \in \mathbb{N}^2} \frac{1}{p!} u^p \frac{1}{q!} v^q = \sum_{s \in \mathbb{N}} \frac{1}{s!} \sum_{p+q=s} \frac{s!}{p!q!} u^p v^q \\ &= \sum_{s \in \mathbb{N}} \frac{1}{s!} \sum_{p+q=s} \binom{s}{p} u^p v^q = \sum_{s \in \mathbb{N}} \frac{1}{s!} (u+v)^s = \exp(u+v) \end{aligned}$$

Proposition : Changement de base

Si $M \in \mathcal{M}_n(\mathbb{K})$, $P \in \mathcal{GL}_n(\mathbb{K})$:

Alors $\exp(P^{-1}MP) = P^{-1}\exp(M)P$

Preuve

On voit l'égalité pour tous les rangs (P et P^{-1} se compensent) et ensuite on se sert de la continuité de $M \mapsto P^{-1}M$

Définition : Série entière à coefficient matriciel

On définit l'application :

$$\varphi \begin{cases} \mathbb{R} & \rightarrow \mathcal{M}_n(\mathbb{K}) \\ t & \mapsto \exp(tA) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} A^n \end{cases}$$

$$\text{On a } \varphi'(t) = \sum_{n=1}^{+\infty} \frac{nt^{n-1}}{n!} A^n = \sum_{n=0}^{+\infty} A \frac{t^n}{n!} A^n = A \exp(tA) = \exp(tA)A$$

Chapitre XII

Calcul différentiel

.1 Conditions du chapitre

Les fonctions considérées dans ce chapitre sont définies sur un ouvert d'un \mathbb{R} -espace vectoriel normé de dimension finie et à valeurs dans un \mathbb{R} -espace vectoriel normé de dimension finie.

Dans tout ce chapitre, sauf mention contraire, E et F désignent deux \mathbb{R} -espaces vectoriels normés de dimension finies dont les normes sont notées $\|\cdot\|$. U désigne un ouvert de E .

Si nécessaire, E sera euclidien.

.2 Rappels sur la continuité

Définition : Limite

Pour $l \in F$ et $a \in \bar{A}$ avec $A \subset E$, $f \rightarrow_a l$ si $\forall \varepsilon \in \mathbb{R}_+^*, \exists \alpha \in \mathbb{R}_+^*, \forall x \in A, x \in \mathcal{B}(a, \alpha) \Rightarrow f(x) \in \mathcal{B}(l, \varepsilon)$

Proposition :

En dimension finie, une fonction respecte une propriété locale si, et seulement si, toutes ses fonctions coordonnées les respectent.

Remarque .1. Pour déterminer si une fonction f est continue en a , on se sert de $f(x, \lambda x)$ en le faisant tendre vers a pour obtenir une limite.

Si f est continue, c'est la seule limite et on peut trouver une majoration par quelque chose qui tend vers 0.

Sinon, on cherche à trouver une autre limite (par exemple avec $f(a - (x, x + x^{2023})))$)

Exemple .1. Soit $f : (x, y) \mapsto \frac{2y}{(x^2+y^2)}$ et vaut 0 en $(0, 0)$

Déjà, f n'est pas continue en 0 : pour $\lambda_1, \lambda_2 \in \mathbb{R}_+, \lambda_1 \neq \lambda_2$:

$$\lim_{x \rightarrow 0} f(x, \lambda_1 x) = \frac{\lambda_1}{1+\lambda_1} \neq \frac{\lambda_2}{1+\lambda_2} = \lim_{x \rightarrow 0} f(x, \lambda_2 x)$$

Définition : o

Pour E, F, G \mathbb{R} -ev de dimension finie, $A \subset E$.

Soit $f : A \rightarrow F$ et $g : A \rightarrow G$

On dit que :

$$f =_a o(g) \text{ si } \exists \varepsilon \in \mathcal{F}(A, \mathbb{R}_+), \varepsilon \rightarrow_a 0 \text{ et } \|f\| = \varepsilon \|g\|$$

Si g ne s'annule pas au voisinage de a , on peut écrire :

$$f =_a o(g) \text{ si } \frac{\|f\|}{\|g\|} \rightarrow_a 0$$

Proposition :

Si on a $x = (x_1, \dots, x_n)$, $\alpha_1, \dots, \alpha_n \in \mathbb{N}$ et $\gamma < \alpha_1 + \alpha_2 + \dots + \alpha_n$, alors :

$$(x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}) =_0 o(\|x\|^\gamma)$$

I Applications différentiables

I.1 Dérivée selon un vecteur

Définition :

Soit f une fonction définie d'un ouvert U dans F , a un point de U et v un vecteur de E .

On dit que f admet une dérivée selon le vecteur v au point a si la fonction $t \mapsto f(a + tv)$ est dérivable en 0.

Si elle existe, cette dérivée est appelée dérivée de f selon le vecteur v au point a et est notée $D_v f(a)$:

$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$$

Exemple I.1. Soit f une fonction définie d'un ouvert U dans F , a un point de U . La fonction f admet une dérivée selon le vecteur 0 au point a et $D_0 f(a) = 0$

Exemple I.2. $f : (x, y) \mapsto x^2 y + e^{xy}$

Trouver la dérivée en $(0, 0)$ suivant $(1, 1)$:

Pour $t \in \mathbb{R}$, posons $g = f(t, t) = t^3 + e^{t^2}$

Cette fonction est dérivable de dérivée $t \mapsto 3t^2 + 2te^{t^2}$

Donc $D_{(1,1)} f(0, 0) = g'(0) = 0$

Remarque I.1. La fonction $t \mapsto f(a + tv)$ est définie pour tout $t \in \mathbb{R}$ tel que $a + tv \in U$. Comme U est un ouvert de E contenant a , elle est bien définie sur un voisinage de 0 dans \mathbb{R} .

Remarque I.2. Une fonction peut avoir des dérivées selon tout vecteur en un point et ne pas être continue en ce point.

Remarque I.3. Si f possède une dérivée suivant v en a , alors f possède une dérivée selon αv pour tout $\alpha \in \mathbb{R}^*$

On suppose que f possède une dérivée en a suivant v :

La fonction $g : t \mapsto f(a + tv)$ est dérivable en 0 et $D_v f(a) = g'(0)$

Considérons $h : t \mapsto f(a + t\alpha v) = g(\alpha t)$

h est dérivable en 0 et $h'(0) = \alpha g'(0)$
 Donc $D_{\alpha v}f(a) = \alpha D_v f(a)$

Exemple I.3. Soit $f : (x, y) \mapsto \frac{xy}{(x^2+y^2)^2}$ et qui vaut 0 en $(0, 0)$.

Cette fonction n'est pas continue en 0 : $\lim_{x \rightarrow 0} f(x, x) = +\infty$

Cette fonction admet des dérivées en $(0, 0)$ suivant $(1, 0)$ et $(0, 1)$:

$$D_{(1,0)}f(1, 0) = 0 \text{ et } D_{(0,1)}f(0, 0) = 0$$

Exemple I.4. Soit $f : (x, y) \mapsto \frac{x^2 y}{(x^4 + y^2)}$ et qui vaut 0 en $(0, 0)$

Pour tout $x \in \mathbb{R}$, on a $f(x, x^2) = \frac{x^4}{2x^4} = \frac{1}{2} \rightarrow_0 \frac{1}{2}$

Donc f n'est pas continue en $(0, 0)$

Si $(\alpha, \beta) \in \mathbb{R}^2 \neq (0, 0)$:

$$f(t\alpha, t\beta) = \frac{t^3 \alpha^3 \beta}{t^4 \alpha^4 + t^2 \beta^2}$$

Si $\beta = 0$: $\forall t \in \mathbb{R}, f(t\alpha, t\beta) = 0$ donc $D_{(\alpha,0)}f(0, 0)$ existe et vaut 0

Sinon : $f(t\alpha, t\beta) = t \frac{\alpha^2}{\beta} + o(t)$

Donc $D_{(\alpha,\beta)}f(0, 0)$ existe et vaut $\frac{\alpha^2}{\beta^2}$

Remarque I.4. On constate que même si une fonction admet une dérivée suivant n'importe quel vecteur en un point, ça ne veut pas dire qu'elle est continue en ce point.

I.2 Dérivées partielles

Définition :

On suppose ici que $E = \mathbb{R}^n$. Soit f une fonction définie d'un ouvert U dans F et $a = (a_1, \dots, a_n)$ un point de U .

Si elle existe, on appelle i -ème dérivée partielle de f en a la dérivée de f selon le vecteur $(0, \dots, 0, 1, 0, \dots, 0)$ (avec le 1 en i -ème place) au point a

et on note $\frac{\partial f}{\partial x_i}(a)$ ou $\partial_i f(a)$

Définition :

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Soit f une fonction définie d'un ouvert U dans F et a un point de U .

Si elle existe, on appelle i -ème dérivée partielle de f dans a dans la base \mathcal{B} la dérivée de f selon le vecteur e_i au point a

et on note $\frac{\partial f}{\partial x_i}(a)$ ou $\partial_i f(a)$

Remarque I.5. Lorsqu'une base $\mathcal{B} = (e_1, \dots, e_n)$ de E est fixée, on sait que l'application :

$$\Phi : \begin{cases} \mathbb{R}^n & \rightarrow E \\ (x_1, \dots, x_n) & \mapsto \sum_{i=1}^n x_i e_i \end{cases}$$

est un isomorphisme. À une application f de U dans E on peut associer l'application :

$$f_{\mathcal{B}} : \begin{cases} \Phi^{-1}(U) & \rightarrow F \\ (x_1, \dots, x_n) & \mapsto f\left(\sum_{i=1}^n x_i e_i\right) \end{cases}$$

Comme Φ est une application linéaire et que \mathbb{R}^n est de dimension finie, Φ est continue et donc $\Phi^{-1}(U)$ est un ouvert de \mathbb{R} .

On autorise l'identification entre f et $f_{\mathcal{B}}$, et donc on s'autorisera à identifier $f(x)$ et $f(x_1, \dots, x_n)$ où $x = \sum_{i=1}^n x_i e_i$

Remarque I.6. Si $E = \mathbb{R}^n$ et que \mathcal{B} est la base canonique de \mathbb{R}^n alors $f = f_{\mathcal{B}}$ et la i -ème dérivée partielle de f en a est la i -ème dérivée partielle de f en a dans la base canonique \mathcal{B} (i.e. la dérivée selon le i -ème vecteur de la base canonique)

I.3 Différentielle

Définition :

Soit f une fonction définie d'un ouvert U dans F et a un point de U .

On dit que f est différentiable au point a s'il existe une application $u \in \mathcal{L}(E, F)$ telle qu'au voisinage de 0 :

$$f(a + h) = f(a) + u(h) + o(h)$$

Dans ce cas, une telle application linéaire u est unique, on la note $df(a)$ et on l'appelle différentielle de f en a .

On l'appelle aussi application linéaire tangente à f en a .

Remarque I.7. La notation $o(h)$ désigne une fonction négligeable devant $\|h\|$

Proposition :

Il y a unicité de la différentielle d'une application différentiable.

Preuve

Supposons l'existence de deux applications linéaires u et v telles que pour h dans un certain voisinage de 0, on ait :

$$f(a + h) - f(a) = u(h) + o(h) = v(h) + o(h)$$

Considérons un vecteur non nul $x \in E$ et t un réel positif.

On obtient pour t dans un voisinage à droite de 0 :

$$(u - v)(tx) = \|tx\|\varepsilon(t)$$

Avec $\varepsilon \rightarrow_0 0$

On en déduit immédiatement que $\|(u - v)(x)\| = \|x\|\|\varepsilon(t)\|$

Par passage à la limite en 0, on obtient $(u - v)(x) = 0$.

La fonction $u - v$ s'annule donc sur E , donc $u = v$

Note : pour alléger l'écriture, il est d'usage de noter $df(a) \cdot h$ au lieu de $(df(a))(h)$.

On utilise parfois $df_a(h)$, mais le problème de cette notation, c'est qu'elle ne coïncide pas bien avec la notation dans \mathbb{R}

Remarque I.8. On dit que f admet un développement limité d'ordre 1 en a s'il existe $u \in \mathcal{L}(E, F)$ telle qu'au voisinage de 0 :

$$f(a + h) = f(a) + u(h) + o(h)$$

Une fonction f est différentiable au point a si, et seulement si, elle admet un développement limité d'ordre 1 en a .

Dans ce cas, $h \mapsto f(a) + df(a) \cdot h$ est une approximation affine de f au voisinage de a .

Remarque I.9. Comme la notion de continuité, la notion de différentiabilité en un point est une notion locale :

Soit f une fonction définie sur un ouvert U , $a \in U$ et V un ouvert de U contenant a .

La fonction f est différentiable en a si, et seulement si, $f|_V$ est différentiable en a .

Proposition :

Si $E = \mathbb{R}$, soit f une fonction définie de l'ouvert U dans F et $a \in U$. La fonction f est différentiable en a si, et seulement si, f est dérivable en a et dans ce cas $df(a)$ est l'application $h \mapsto hf'(a)$.
On a en particulier $f'(a) = df(a) \cdot 1$.

Preuve

C'est une des définitions de la dérivabilité : f est dérivable en a si, et seulement si, f possède un développement limité à l'ordre 1 en a :

$$f(a + h) = f(a) + \lambda h + o(h)$$

$h \mapsto \lambda h$ est linéaire. La définition de la dérivabilité est donc celle de la différentiabilité dans le cas particulier d'un \mathbb{R} -espace vectoriel de dimension 1

Définition :

Soit f une fonction définie d'un ouvert U dans F . Si f est différentiable en tout point de U , on dit que f est différentiable sur U et on appelle différentielle de f sur U l'application :

$$df : \begin{cases} U & \rightarrow \mathcal{L}(E, F) \\ a & \mapsto df(a) \end{cases}$$

Remarque I.10. Si f est différentiable :

$$df \in \mathcal{F}(U, \mathcal{L}(E, F))$$

$$\forall a \in U, df(a) \in \mathcal{L}(E, F)$$

Exemple I.5.

$$f : \begin{cases} \mathcal{M}_n(\mathbb{R})^2 & \rightarrow \mathcal{M}_n(\mathbb{R}) \\ (A, B) & \mapsto AB \end{cases}$$

Soit $(A, B) \in \mathcal{M}_n(\mathbb{R})^2$

$$f((A, B) + (H, K)) = (A + H)(B + K) = AB + HB + AK + HK$$

On a que $(H, K) \mapsto HB + AK \in \mathcal{L}(\mathcal{M}_n(\mathbb{R})^2, \mathcal{M}_n(\mathbb{R}))$.

Il suffit alors de montrer que $HK = o(\|(H, K)\|)$ pour obtenir que f est différentiable en (A, B)

On prend $\|\cdot\|$ une norme d'algèbre sur $\mathcal{M}_n(\mathbb{K})$ et on prend la norme infinie associée sur $\mathcal{M}_n(\mathbb{K}) \times \mathcal{M}_n(\mathbb{K})$ (le max des normes d'algèbres des deux)

$$\frac{\|HK\|}{\|(H, K)\|} \leq \frac{\|H\|\|K\|}{\max(\|H\|, \|K\|)} \leq \|(H, K)\|$$

Donc $HK = o(\|(H, K)\|)$

Donc f est différentiable en AB et $df(A, B) \cdot (H, K) = HB + AK$

Proposition :

Soit f une fonction définie d'un ouvert U dans F et $a \in U$. Si f est différentiable en a , alors f est continue en a .

Preuve

Si f est différentiable en a , il existe une application linéaire u (continue car F est de dimension finie) telle que, pour h dans un voisinage de 0 :

$$f(a+h) - f(a) = u(h) + o(h)$$

Par continuité de u , $f(a+h) - f(a) \rightarrow_{h \rightarrow 0} 0$

Donc par définition, f est continue en a

Proposition :

Soit f une fonction définie d'un ouvert U dans F et $a \in U$.

Si f est différentiable en a , alors f est dérivable en a selon tout vecteur $v \in E$ et :

$$D_v f(a) = df(a) \cdot v$$

Preuve

Supposons f différentiable en a

Soit v un vecteur de E , alors pour tout t réel dans un voisinage de 0 , on a :

$$f(a+tv)(a) = f(a) + df(a) \cdot (tv) + o(t) = f(a) + tdf(a) \cdot v + o(t)$$

Donc $t \mapsto f(a+tv)$ est dérivable en 0 de dérivée $df(a) \cdot v$

Ce qui conclut la preuve.

Remarque I.11. La réciproque est fausse.

Proposition : Corollaire

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Soit f une fonction définie d'un ouvert U dans F et $a \in U$.

Si f est différentiable en a , alors f admet des dérivées partielles (dans la base \mathcal{B}) et pour $v = \sum_{i=1}^n v_i e_i \in E$:

$$D_v f(a) = df(a) \cdot v = \sum_{i=1}^n v_i \partial_i f(a)$$

Preuve

Pour $i \in \llbracket 1, n \rrbracket$, la i -ème dérivée partielle est la dérivée suivant le vecteur e_i .

On déduit immédiatement de la proposition précédente l'existence des n dérivées partielles et par linéarité de $df(a)$, pour tout vecteur $v = \sum_{i=1}^n v_i e_i$:

$$D_v f(a) = df(a) \cdot \left(\sum_{i=1}^n v_i e_i \right) = \sum_{i=1}^n v_i df(a) \cdot e_i = \sum_{i=1}^n v_i \partial_i f(a)$$

I.4 Matrice jacobienne

Proposition :

Soit $\mathcal{B} = (e_1, \dots, e_m)$ une base de E et $\mathcal{B}' = (e'_1, \dots, e'_n)$ une base de F .

Soit f une fonction définie d'un ouvert U dans F différentiable sur U . Notons f_1, \dots, f_n les fonctions composantes de f dans la base \mathcal{B}' .

Soit $a \in U$. La matrice dans les bases \mathcal{B} et \mathcal{B}' de l'application linéaire $df(a)$ est :

$$J_f(a) = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \cdots & \partial_m f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \cdots & \partial_m f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_n(a) & \partial_2 f_n(a) & \cdots & \partial_m f_n(a) \end{pmatrix}$$

Preuve

Pour tout $j \in \llbracket 1, n \rrbracket$:

$$df(a) \cdot (e_j) = \sum_{i=1}^n df_i(a) \cdot (e_j) e'_i = \sum_{i=1}^n \partial_j f_i(a) e'_i$$

Définition : Matrice Jacobienne

Si $E = \mathbb{R}^m$ et $F = \mathbb{R}^n$. Soit f une fonction définie d'un ouvert U dans F différentiable sur U .

Soit $a \in U$. La matrice, dans les bases canoniques, de l'application linéaire $df(a)$ est :

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \cdots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}$$

et est appelée matrice jacobienne de f en a .

Exemple I.6.

$$\varphi : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ (r, \theta) & \mapsto (r \cos \theta, r \sin \theta) = (x, y) \end{cases}$$

C'est une application surjective. On admet qu'elle est différentiable, et on calcule sa jacobienne :

$$\frac{\partial x}{\partial r} = \cos \theta$$

$$\frac{\partial y}{\partial r} = \sin \theta$$

$$\frac{\partial x}{\partial \theta} = -r \sin \theta$$

$$\frac{\partial y}{\partial \theta} = r \cos \theta$$

$$\text{Donc } J_\varphi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

Pour ne pas se tromper en l'écrivant, on peut écrire :

avec $(r, \theta) \mapsto x(r, \theta)$, on prend $dx = \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta$, qui est la somme des expressions sur la ligne x , et les éléments différentiels $dr, d\theta$ indiquent dans quelle colonne ils vont.

On note aussi $dy = \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta$

Et alors $\begin{pmatrix} dx \\ dy \end{pmatrix} = J_\varphi(r, \theta) \begin{pmatrix} dr \\ d\theta \end{pmatrix}$

II Opérations sur les applications différentiables

II.1 Différentielle d'une combinaison linéaire d'applications différentiables

Proposition :

Soit f et g deux fonctions définies d'un ouvert U dans F , différentiables sur U . Soit λ, μ deux réels. La fonction $\lambda f + \mu g$ est différentiable sur U et :

$$d(\lambda f + \mu g) = \lambda df + \mu dg$$

Preuve

Soit $a \in U$

Soit $h \in E$ dans un voisinage de 0

$$\begin{aligned} (\lambda f + \mu g)(a + h) &= (\lambda(f(a) + df(a) \cdot h + o(h)) + \mu(g(a) + dg(a) \cdot h + o(h))) \\ &= (\lambda f + \mu g)(a) + (\lambda df + \mu dg)(a) \cdot h + o(h) \end{aligned}$$

Donc $\lambda f + \mu g$ est différentiable en a de différentielle en a $(\lambda df + \mu dg)(a)$.

II.2 Différentielle de $B(f, g)$ où B est bilinéaire et f et g sont deux applications différentiables

Proposition :

Soit F, G, H des espaces vectoriels normés de dimension finie. Soit B une application bilinéaire définie de $F \times G$ dans H .

Il existe $C \in \mathcal{R}_+$ tel que :

$$\forall (x, y) \in F \times G, \|B(x, y)\| \leq C\|x\|\|y\|$$

Preuve

En dimension finie, une application bilinéaire est continue (chaque composante est polynômiale en les coordonnées).

Le produit cartésien des disques unités de F et G est compact comme produit de compacts.

Donc B est bornée par une constante C sur ce compact.

On en déduit :

Pour $(x, y) \in F \times G$ non-nuls, $\left\| B\left(\frac{1}{\|x\|}x\right)\left(\frac{1}{\|y\|}y\right) \right\| \leq C$

Soit, en utilisant la bilinéarité de B et la positivité de la norme :

$$\|B(x, y)\| \leq C\|x\|\|y\|$$

L'égalité reste valide pour x ou y nuls.

Ce qui conclut la preuve.

Proposition :

Soit E, F, G, H quatre espaces vectoriels normés de dimension finie et U un ouvert de E . Soit f (respectivement g) une fonction définie de U dans F (respectivement G) et différentiable sur U .

Soit B une application bilinéaire définie de $F \times G$ dans H .

La fonction $B(f, g) : x \mapsto B(f(x), g(x))$ est différentiable sur U et :

$$d(B(f, g)) = B(df, g) + B(f, dg)$$

Preuve

Soit $a \in U$

Soit $h \in E$ dans un voisinage de 0.

$$B(f(a+h), g(a+h)) = B(f(a) + df(a) \cdot h + o(h), g(a) + dg(a) \cdot h + o(h))$$

$$B(f(a), g(a)) + B(df(a) \cdot h, dg(a) \cdot h) + o(h)$$

Ce qui conclut la preuve.

On a utilisé la proposition précédente pour montrer que les termes restants étaient bien négligeables devant h . Par exemple il existe une constante C et une constante M telles que :

$$\|B(df(a) \cdot h, dg(a) \cdot h)\| \leq C\|df(a) \cdot h\|\|dg(a) \cdot h\| \leq M\|h\|^2$$

Remarque II.1. La proposition s'étend à la différentiation de $M(f_1, \dots, f_n)$ lorsque M est n -linéaire et que les f_i sont différentiables.

II.3 Différentielle d'une composée d'applications différentiables

Proposition :

Soit E, F, G trois espaces vectoriels normés de dimension finie et U un ouvert de E .
 Soit f une fonction définie de U dans F , différentiable sur U . Soit g une fonction définie de V dans G , avec V un ouvert de F contenant $f(U)$, différentiable sur V .
 La fonction $g \circ f$ est différentiable sur U et, pour tout $a \in U$,

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

Preuve

Soit $a \in U$

Soit $h \in E$ dans un voisinage de 0

$$\begin{aligned} (g \circ f)(a + h) &= g(f(a + h)) \\ &= g(f(a) + df(a) \cdot h + o(h)) \\ &= g(f(a)) + dg(f(a)) \cdot (df(a) \cdot h + o(h)) + o(h) \\ &= g(f(a)) + (dg(f(a)) \circ df(a)) \cdot h + dg(f(a)) \cdot (o(h)) + o(h) \\ &= g(f(a)) + (dg(f(a)) \circ df(a)) \cdot h + o(h) \end{aligned}$$

On a bien le résultat annoncé.

Remarque II.2. Avec les notations ci-dessus, si f est à valeurs réelles et g une fonction définie sur un intervalle réel, la formule se ré-écrit :

$$d(g \circ f)(a) = g'(f(a))f'(a)$$

Proposition : Cas d'une algèbre

Si F est une algèbre

Soient f, g deux applications différentiables en a définies de U dans F

Alors fg est différentiable en a et :

$$d(fg) : h \mapsto (df(a) \cdot h)g(a) + f(a)(dg(a) \cdot h)$$

Preuve

Soit $h \in E$:

$\varepsilon_1, \varepsilon_2$ des fonctions de E dans \mathbb{R} qui tendent vers 0 quand h tend vers 0

$$\begin{aligned} fg(a + h) &= f(a + h)g(a + h) \\ &= (f(a) + df(a) \cdot h + \varepsilon_1(h)\|h\|)(g(a) + dg(a) \cdot h + \varepsilon_2(h)\|h\|) \\ &= f(a)g(a) + (df(a) \cdot h)g(a) + f(a)(dg(a) \cdot h) \\ &\quad + \varepsilon_1(h)\|h\|[g(a) + (dg(a) \cdot h)] + \varepsilon_2(h)\|h\|[f(a) + df(a) \cdot h] \\ &= fg(a) + (df(a) \cdot h)g(a) + f(a)(dg(a) \cdot h) + o(h) \end{aligned}$$

II.4 Règle de la chaîne

Proposition : Règle de la chaîne

Soit m, p, n des entiers naturels non nuls.

Soit f une application différentiable sur U un ouvert de \mathbb{R}^m et à valeurs dans \mathbb{R}^p .

Soit g une application différentiable sur V un ouvert de \mathbb{R}^p contenant $f(U)$ et à valeurs dans \mathbb{R}^n .

On note f_1, \dots, f_p les fonctions composantes de f et on pose $h = g \circ f$.

On a donc le schéma suivant :

$$\begin{array}{ccccc} \mathbb{R}^m & \xrightarrow{f} & \mathbb{R}^p & \xrightarrow{g} & \mathbb{R}^n \\ (x_1, \dots, x_m) & \mapsto & (f_1, \dots, f_p) & \mapsto & (g_1, \dots, g_n) \end{array}$$

Soit $x = (x_1, \dots, x_m) \in U$. Pour $j \in \llbracket 1, m \rrbracket$ et $i \in \llbracket 1, n \rrbracket$:

Avec des jacobiniennes, on a $J_{g \circ f}(x) = J_g(f(x))J_f(x)$ et donc :

$$J_g(f(x)) = \begin{pmatrix} \cdot & \cdot & \cdot \\ \frac{\partial g_i}{\partial f_1}(f(x)) & \cdots & \frac{\partial g_i}{\partial f_p}(f(x)) \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdots & \frac{\partial f_1}{\partial x_j}(x) & \cdots \\ \cdot & \cdot & \cdot \\ \cdots & \frac{\partial f_p}{\partial x_j}(x) & \cdots \end{pmatrix}$$

Ce qui donne bien :

$$\begin{aligned} \frac{\partial (g \circ f)_i}{\partial x_j}(x) &= \sum_{k=1}^p \frac{\partial g_i}{\partial f_k}(f(x)) \cdot \frac{\partial f_k}{\partial x_j}(x) \\ \partial_j h(x) &= \sum_{k=1}^p \partial_k g(f(x)) \partial_j f_k(x) \end{aligned}$$

Ou encore sous forme plus mnémotechnique :

$$\frac{\partial h}{\partial x_j}(x) = \sum_{i=1}^p \frac{\partial f_i}{\partial x_j}(x) \frac{\partial g}{\partial f_i}(f(x))$$

Et enfin de manière encore plus concise mais abusive (les g_i désignent à la fois les fonctions coordonnées de g et de $g \circ f$ et les points d'application ne sont pas indiqués) :

pour tout $i \in \llbracket 1, n \rrbracket$ et $j \in \llbracket 1, m \rrbracket$:

$$\frac{\partial g_i}{\partial x_j} = \sum_{k=1}^p \frac{\partial f_k}{\partial x_j} \frac{\partial g_i}{\partial f_k}$$

Preuve

On note (e_1, \dots, e_m) la base canonique de \mathbb{R}^m et (u_1, \dots, u_p) la base canonique de \mathbb{R}^p
 Soit $j \in \llbracket 1, m \rrbracket$ et $x \in \mathbb{R}^m$

$$dh(x) \cdot e_j = d(g \circ f)(x) \cdot (e_j) = dg(f(x)) \cdot (df(x) \cdot e_j)$$

On écrit $df(x) \cdot e_j$ dans la base canonique de \mathbb{R}^p :

$$df(x) \cdot e_j = \sum_{k=1}^p (df_k(x) \cdot e_j) u_k$$

Et par linéarité de $dg(f(x))$:

$$dh(x) \cdot e_j = \sum_{k=1}^p (df_k(x) \cdot e_j) dg(f(x)) \cdot (u_k)$$

Ce qui est exactement la formule proposée.

On pouvait aussi, si on veut éviter le raisonnement intrinsèque, écrire le produit des jacobiniennes des applications.

Exemple II.1. *Exprimons la composée(/divergence) en polaire :*

$$\begin{array}{ccccc} \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (r, \theta) & \mapsto & (x, y) = (r \cos \theta, r \sin \theta) & \mapsto & f(x, y) \end{array}$$

$$\text{On note } \varphi : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ (r, \theta) & \mapsto (r \cos \theta, r \sin \theta) = (x, y) \end{cases} \text{ et } f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto f(x, y) \end{cases}$$

On suppose les deux différentiables (on montrera plus tard que φ l'est)

Pour $r, \theta \in \mathbb{R}^2$:

$$\frac{\partial(f \circ \varphi)}{\partial r}(r, \theta) = \frac{\partial f}{\partial x} \frac{\partial \varphi}{\partial r}(r, \theta) + \frac{\partial f}{\partial y} \frac{\partial \varphi}{\partial \theta}(r, \theta) = \frac{\partial f}{\partial x}(x, y) \cos \theta + \frac{\partial f}{\partial y}(x, y) \sin \theta$$

$$\frac{\partial(f \circ \varphi)}{\partial \theta}(r, \theta) = \frac{\partial f}{\partial x}(x, y)(-r \sin \theta) + \frac{\partial f}{\partial y}(x, y)r \cos \theta$$

III Applications de classe \mathcal{C}^1

III.1 Définition et caractérisation

Définition :

Une application f d'un ouvert U dans F est dite de classe \mathcal{C}^1 si elle est différentiable sur U et si df est continue sur U

$$\text{(i.e. } \begin{array}{ccc} U & \rightarrow & \mathcal{L}(E, F) \\ x & \mapsto & df(x) \end{array} \text{ continue)}$$

Théorème :

Soit \mathcal{B} une base de E . Soit f une application définie d'un ouvert U dans F .

L'application f est de classe \mathcal{C}^1 sur U si, et seulement si, les dérivées partielles relativement à la base \mathcal{B} existent en tout point de U et sont continues sur U .

Preuve

Hors-programme :

Si f est \mathcal{C}^1 :

Alors f admet une dérivée suivant un vecteur quelconque u en tout point x , $df(x) \cdot u$.

L'application $x \mapsto df(x) \cdot u$ est la composée de $df : x \mapsto df(x)$ et de $\varphi : l \mapsto l(u)$ ($l \in \mathcal{L}(E, F)$)

df est continue (comme f est \mathcal{C}^1) et φ est continue (linéaire en dimension finie).

Donc $\varphi \circ df$ est continue, et donc pour tout point x , les dérivées partielles selon tout élément d'une base sont continues.

Réciproquement, on suppose que f admet p dérivées suivant les vecteurs e_1, \dots, e_p continues sur U .

Soit $x = x_1 e_1 + \dots + x_p e_p$ et $h = h_1 e_1 + \dots + h_p e_p$ avec $x + h \in U$

$$\begin{aligned} f(x+h) - f(x) &= f(x_1 + h_1, \dots, x_p + h_p) - f(x_1, x_2 + h_2, \dots, x_p + h_p) \\ &\quad + f(x_1, x_2 + h_2, \dots, x_p + h_p) - f(x_1, x_2, x_3 + h_3, \dots, x_p + h_p) \\ &\quad + \dots \\ &\quad + f(x_1, \dots, x_{p-1}, x_p + h_p) - f(x) \\ &= h_1 \partial_1 f(x_1, x_2 + h_2, \dots, x_p + h_p) + h_1 \varepsilon_1(h_1) \\ &\quad + h_2 \partial_2 f(x_1, x_2, x_3 + h_3, \dots, x_p + h_p) + h_2 \varepsilon_2(h_2) \\ &\quad + \dots \\ &\quad + h_p \partial_p f(x_1, \dots, x_p) + h_p \varepsilon_p(h_p) \end{aligned}$$

où $\varepsilon_1, \dots, \varepsilon_p$ sont des fonctions qui tendent vers 0 en 0.

Par hypothèse, $\forall i \in \llbracket 1, p \rrbracket$, $\partial_i f$ est continue.

Donc $\partial_1 f(x_1, x_2 + h_2, \dots, x_p + h_p) = \partial_1 f(x_1, \dots, x_p) + \varphi_1(h_1, h_2, \dots, h_p)$ avec $\varphi_1 \rightarrow_0 0$

On raisonne de même pour les autres dérivées partielles.

Donc $f(x+h) - f(x) = \sum_{i=1}^p h_i \partial_i f(x) + \sum_{i=1}^p h_i (\varepsilon_i(h_i) + \varphi_i(h))$

Donc c'est la somme d'une application linéaire en h et d'un $o(h)$.

Donc f est différentiable, et $\forall x \in U, \forall h \in E, df(x) \cdot h = \sum_{i=1}^p h_i \partial_i f(x)$

df est une application de U dans $\mathcal{L}(E, F)$

C'est la somme des applications $g_i : x \mapsto (h \mapsto h_i \partial_i f(x))$, composées de projections en dimension finie et de $\partial_i f$ qui est continue.

Donc df est \mathcal{C}^1 , donc df est continue.

Remarque III.1. Dire que df est continue, c'est dire que J_f est continue, donc que chacune des applications coordonnées de J_f est continue.

Exemple III.1.

$$f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ (x, y) & \mapsto (2x^3 y e^{x+y}, \arctan(xy)) \end{cases}$$

Montrer que f est différentiable :

f admet deux dérivées partielles sur \mathbb{R}^2 :

Pour $(x, y) \in \mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(x, y) = \left(2x^2 y e^{x+y} (3 + x), \frac{y}{1+x^2 y^2} \right)$$

$$\frac{\partial f}{\partial y}(x, y) = \left(2x^3 e^{x+y} (1 + y), \frac{x}{1+x^2 y^2} \right)$$

$(x, y) \mapsto \frac{\partial f}{\partial x}(x, y)$ est continue par opérations usuelles.

$(x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$ est continue par opérations usuelles.

Donc f est \mathcal{C}^1 .

$$J_f(x) = \begin{pmatrix} 2x^2ye^{x+y}(3+x) & 2x^3e^{x+y}(1+y) \\ \frac{y}{1+x^2y^2} & \frac{x}{1+x^2y^2} \end{pmatrix}$$

III.2 Opérations algébriques sur les applications de classe \mathcal{C}^1

Proposition :

Soit f et g deux fonctions d'un ouvert U dans F de classe \mathcal{C}^1 . Soit λ, μ deux réels.
La fonction $\lambda f + \mu g$ est de classe \mathcal{C}^1 sur U .

Proposition :

Soit E, F, G, H quatre espaces vectoriels normés de dimension finie et U un ouvert de E .
Soit f une fonction définie de U dans F de classe \mathcal{C}^1 .
Soit g une fonction définie de U dans G de classe \mathcal{C}^1 .
Soit B une application bilinéaire définie de $F \times G$ dans H .
La fonction $B(f, g) : x \mapsto B(f(x), g(x))$ est de classe \mathcal{C}^1 sur U .

Preuve

Immédiat :

Pour $i \in \llbracket 1, n \rrbracket$, la i -ème dérivée partielle de $\lambda f + \mu g$ existe et vaut $\lambda \partial_i f + \mu \partial_i g$

Elle est donc continue par opération usuelle.

Proposition :

Soit f et g deux applications de classe \mathcal{C}^1 d'un ouvert U dans \mathbb{R} . La fonction fg est \mathcal{C}^1 .

Preuve

Immédiat :

Pour $i \in \llbracket 1, n \rrbracket$, la i -ème dérivée partielle de fg existe et vaut $(\partial_i f)g + f\partial_i g$

Elle est donc continue par opération usuelle.

Proposition :

Supposons que F est un espace vectoriel euclidien. Soit f et g deux applications de classe \mathcal{C}^1 de U dans F .

La fonction $\begin{cases} U & \rightarrow \mathbb{R} \\ t & \mapsto \langle f(t), g(t) \rangle \end{cases}$ est \mathcal{C}^1

Preuve

Immédiat :

Pour $i \in \llbracket 1, n \rrbracket$, la i -ème dérivée partielle de $\langle f, g \rangle$ existe et vaut $\langle \partial_i f, g \rangle + \langle f, \partial_i g \rangle$

Elle est donc continue par opération usuelle.

Proposition :

Soit E, F, G trois espaces vectoriels normés de dimension finie et U un ouvert de E .
 Soit f une fonction définie de U dans F de classe \mathcal{C}^1 sur U .
 Soit g une fonction définie de V un ouvert de F contenant $f(U)$ dans G , g de classe \mathcal{C}^1 sur V .
 La fonction $g \circ f$ est de classe \mathcal{C}^1 sur U .

Preuve

Immédiat :

Pour $i \in \llbracket 1, n \rrbracket$, la i -ème dérivée partielle de $g \circ f$ existe et vaut

$$\partial_i(g \circ f) = \sum_{k=1}^{\dim F} (\partial_k g \circ f) \partial_i f_k$$

Elle est donc continue par opération usuelle.

IV Dérivées partielles d'ordre supérieur

IV.1 Dérivées partielles d'ordre k

Définition :

Soit \mathcal{B} une base de E et f une application d'un ouvert U dans F .
 Si f admet une i -ème dérivée partielle $\partial_i f$ sur U , on dit que f admet une dérivée partielle seconde d'indices i, j si $\partial_i f$ admet une j -ème dérivée partielle et on la note alors $\partial_j \partial_i f$.

On utilise aussi la notation $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et on l'appelle dérivée partielle seconde par rapport aux variables x_i, x_j .

Par récurrence on définit la notion de dérivée partielle d'ordre k d'indices i_1, \dots, i_k ou dérivée partielle d'ordre k par rapport aux variables x_{i_1}, \dots, x_{i_k} , que l'on note $\partial_{i_k} \dots \partial_{i_1} f$ ou $\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}$.

Définition :

Une application f définie d'un ouvert U dans F est dite de classe \mathcal{C}^k sur U si toutes ses dérivées partielles d'ordre k existent et sont continues sur U .

Elle est dite de classe \mathcal{C}^∞ si elle est de classe \mathcal{C}^k pour tout $k \in \mathbb{N}^*$.

IV.2 Théorème de Schwarz

Théorème : Schwarz

Soit \mathcal{B} une base de E et soit f définie d'un ouvert U dans F .
 Si f est de classe \mathcal{C}^2 alors :

$$\forall i, j \in \llbracket 1, n \rrbracket^2, \partial_i \partial_j f = \partial_j \partial_i f$$

Preuve

Admis

IV.3 Opérations sur les applications de classe \mathcal{C}^k

Dans ce paragraphe, $k \in \mathbb{N} \cup \{+\infty\}$

Proposition :

Soit f et g deux fonctions de U dans F de classe \mathcal{C}^k . Soit λ, μ deux réels.
La fonction $\lambda f + \mu g$ est de classe \mathcal{C}^k sur U .

Preuve

Admis

Proposition :

Soit E, F, G, H quatre espaces vectoriels normés de dimension finie et U un ouvert de E .
Soit f une fonction définie de U dans F de classe \mathcal{C}^k .
Soit g une fonction définie de U dans G de classe \mathcal{C}^k .
Soit B une application bilinéaire définie de $F \times G$ dans H .
La fonction $B(f, g)$ est de classe \mathcal{C}^k sur U .

Preuve

Admis

Proposition :

Soit E, F, G trois espaces vectoriels normés de dimension finie et U un ouvert de E .
Soit f une fonction définie de U dans F de classe \mathcal{C}^k sur U .
Soit g une fonction définie de V un ouvert de F contenant $f(U)$ dans G et de classe \mathcal{C}^k sur V .
La fonction $g \circ f$ est de classe \mathcal{C}^k sur U .

Preuve

Admis

V Cas des applications numériques

Se limiter au cadre de $E \rightarrow \mathbb{R}$ n'est pas un problème, une application de E dans F avec F de dimension n est simplement n applications de E dans \mathbb{R} .

V.1 Gradient

Dans tout ce paragraphe, E est un espace vectoriel euclidien (et $F = \mathbb{R}$)

Définition :

On appelle dual de E et on note \mathcal{E}^* l'ensemble $\mathcal{L}(E, \mathbb{R})$

Définition :

Soit f une fonction définie et différentiable sur un ouvert U et à valeurs dans \mathbb{R} . Soit $a \in U$. On appelle gradient de f en a et on note $\nabla f(a)$ l'unique vecteur de E tel que :

$$\forall h \in E, df(a) \cdot h = \langle \nabla f(a), h \rangle$$

On rencontre parfois aussi la notation $\text{Grad}f(a)$

Proposition : Interprétation géométrique du gradient

Pour h un vecteur de E , $D_h f(a) = df(a) \cdot h = \langle \nabla f(a), h \rangle$.

Si $\nabla f(a) \neq 0$, il est colinéaire et de même sens que le vecteur unitaire selon lequel la dérivée de f en a est maximale.

Preuve

Notons $v = \frac{1}{\|\nabla f(a)\|} \nabla f(a)$ le vecteur normé colinéaire à $\nabla f(a)$ de même sens.

$$D_v f(a) = df(a) \cdot v = \|\nabla f(a)\|$$

Considérons un vecteur unitaire u .

$$D_u f(a) = df(a) \cdot u = \langle \nabla f(a), u \rangle$$

Et par inégalité de Cauchy-Schwarz :

$$|D_u f(a)| \leq \|\nabla f(a)\| \leq D_v f(a)$$

Proposition :

Soit f une fonction définie et différentiable sur un ouvert U et à valeurs dans \mathbb{R} . Soit $a \in U$.

Dans (e_1, \dots, e_n) une base orthonormale de E , $\nabla f(a)$ s'écrit $\nabla f(a) = \sum_{i=1}^n \partial_i f(a) \cdot e_i$

Preuve

C'est une simple écriture de la formule de différentiation :

Pour $h \in E$, $h = h_1 e_1 + \dots + h_n e_n$:

$$df(a) \cdot h = \sum_{i=1}^n \partial_i f(a) h_i = \left\langle \sum_{i=1}^n \partial_i f(a) \cdot e_i, h \right\rangle$$

V.2 Dérivée le long d'un arc**Définition :**

On appelle arc paramétré une application définie sur un intervalle I de \mathbb{R} à valeur dans \mathbb{R}^p .

Cette notion modélise la notion de courbe dans un espace de dimension p décrite par une fonction horaire de parcours.

Si γ est un arc paramétré, l'image de γ est appelé son support.

Exemple V.1. La fonction γ définie sur \mathbb{R} à valeur dans \mathbb{R}^2 euclidien par :

$$\gamma : t \mapsto \begin{cases} x = \cos t \\ y = \sin t \end{cases}$$

γ est un arc paramétré dont le support est le cercle centré en 0 de rayon 1.

Proposition :

Soit γ un arc paramétré dérivable sur un intervalle ouvert I et $t_0 \in I$.

Si $\gamma'(t_0)$ est un vecteur non nul, alors le support de γ admet pour tangente au point $\gamma(t_0)$ la droite dirigée par $\gamma'(t_0)$ passant par le point $\gamma(t_0)$.

Preuve

On peut considérer que γ est à valeur dans un espace vectoriel normé.

On a tout simplement :

$$\frac{1}{h}(\gamma(t_0 + h) - \gamma(t_0)) \xrightarrow{h \rightarrow 0} \gamma'(t_0)$$

Donc le vecteur normé $\frac{1}{\|\gamma(t_0) - \gamma(t_0 + h)\|}(\gamma(t_0) - \gamma(t_0 + h))$ converge vers $\frac{1}{\|\gamma'(t_0)\|}\gamma'(t_0)$ en 0.

Exemple V.2. $\gamma : t \mapsto \begin{cases} \cos^3(t) \\ \sin^3(t) \end{cases}$ un astroïde

On note $M(t)$ le point en $t \in \mathbb{R}$ du support de l'arc. On étudie les symétries du support.

— $\forall t \in \mathbb{R}, M(t + 2\pi) = M(t)$, donc on peut se restreindre à un intervalle de largeur 2π

— $\forall t \in \mathbb{R}, M(t + \pi) = -M(t) = S_0(M(t))$, donc on peut se restreindre à un intervalle de largeur π et faire la symétrie

— $\forall t \in \mathbb{R}, M\left(\frac{\pi}{2} - t\right) = S_{y=x}(M(t))$

— $\forall t \in \mathbb{R}, M(-t) = S_{y=0}(M(t))$

Donc cette courbe est symétrique par rapport à l'origine, par rapport à l'axe des abscisses (et donc aussi par rapport à l'axe des ordonnées par composition) et par rapport à la droite $y = x$ et $y = -x$

On n'a donc besoin de l'étudier que sur $\left[0, \frac{\pi}{4}\right]$, et on l'étudie sur $\left[0, \frac{\pi}{2}\right]$

$$\forall t \in \left[0, \frac{\pi}{2}\right], \gamma'(t) = \begin{cases} -3 \cos^2 t \sin t \\ 3 \sin^2 t \cos t \end{cases}$$

On constate que $x'(t)$ est négative sur la première moitié de l'intervalle et $y'(t)$ positive sur le même ensemble

étude locale en 0 :

$$x(t) =_0 \left(1 - \frac{t^2}{2} + o(t^3)\right)^3 = 1 - \frac{3t^2}{2} + o(t^3)$$

$$y(t) =_0 \left(t - \frac{t^3}{6} + o(t^3)\right)^3 = t^3 + o(t^3)$$

$$\text{Donc } \overrightarrow{OM}(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t^2 \begin{pmatrix} -\frac{3}{2} \\ 0 \end{pmatrix} + t^3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + o(t^3)$$

Remarque V.1. La proposition précédente peut être étendue au cas où γ est dérivable sur un intervalle d'intérieur non vide. On parlera alors de demi-tangente en $\gamma(t_0)$ si t_0 est une extrémité de l'intervalle.

Proposition : Dérivée le long d'un arc

Soit f une fonction définie d'un ouvert U dans F . Soit γ une application définie sur un intervalle d'intérieur non-vide I de \mathbb{R} et à valeurs dans U .

Si γ est dérivable en t et si f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et :

$$(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$$

Preuve

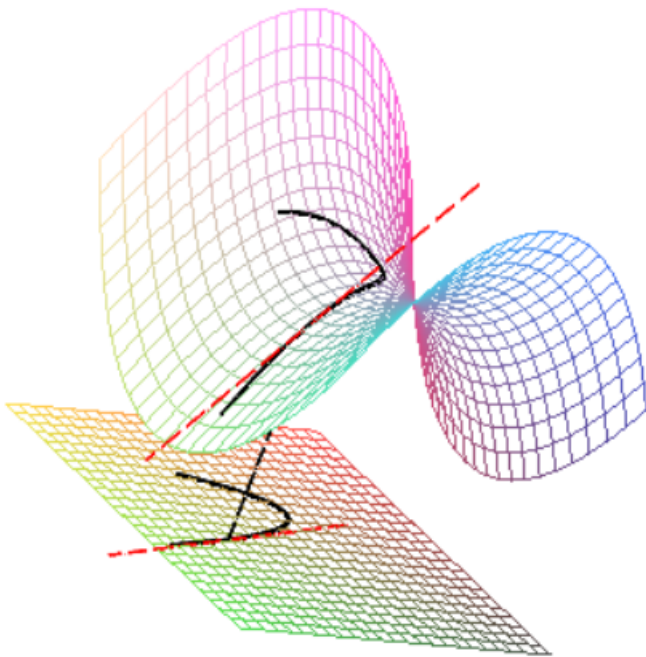
Il s'agit simplement de l'application de la formule donnant la différentielle d'une composée dans un cas particulier.

Pour un réel h au voisinage de 0 :

$$\begin{aligned} h(f \circ \gamma)'(t) &= d(f \circ \gamma)(t) \cdot h \\ &= df(\gamma(t)) \cdot (d\gamma(t) \cdot (h)) \\ &= df(\gamma(t)) \cdot (h\gamma'(t)) \\ &= h df(\gamma(t)) \cdot (\gamma'(t)) \end{aligned}$$

Remarque V.2. *Interprétation géométrique :*

Si $\gamma'(t) \notin \ker(df(\gamma(t)))$, la tangente au point de paramètre t de l'arc paramétré $f \circ \gamma$ est dirigé par l'image d'un vecteur directeur de la tangente à γ au point de paramètre t par l'application $df(\gamma(t))$



Remarque V.3. *Cas particulier fondamental :*

Soit $a \in U$, soit $h \in E$ et γ l'application définie par $\gamma(t) = a + th$.

Si f est différentiable en a , alors l'application $f \circ \gamma$ est dérivable en 0 et :

$$(f \circ \gamma)'(0) = df(a) \circ h = D_h f(a)$$

V.3 Intégration le long d'un arc d'une application de classe C^1

Proposition :

Si f est une application de classe C^1 d'un ouvert U dans F , si γ est une application de classe C^1 de $[0, 1]$ dans Ω , si $\gamma(0) = a$, $\gamma(1) = b$, alors :

$$f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt$$

Preuve

Dans les conditions de l'énoncé, $t \mapsto df(\gamma(t)) \cdot \gamma'(t)$ est la dérivée de la fonction $g : t \mapsto f(\gamma(t))$
 Donc :

$$\int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt = g(1) - g(0) = f(b) - f(a)$$

Remarque V.4. La valeur de $\int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt$ ne dépend pas de l'application γ mais uniquement de ses valeurs en 0 et 1.

Remarque V.5. Pour $E = \mathbb{R}^2$ ou \mathbb{R}^3 :

En physique, on dit qu'un champ de vecteurs V dérive d'un potentiel s'il existe une fonction f différentiable telle que pour tout $h \in E$ et $a \in E$, $df(a) \cdot h = \langle V(a), h \rangle$.

Dans ce cas, $\int_0^1 \langle \nabla f(\gamma(t)), \gamma'(t) \rangle dt$ dépend uniquement des valeurs de $f(\gamma(0))$ et $f(\gamma(1))$

Proposition :

Si f est une application d'un ouvert U dans F .

Si U est connexe par arcs, la fonction f est constante sur U si, et seulement si, elle est différentiable sur U et si $df = 0$

Preuve

Hors-programme sauf pour les fonctions \mathcal{C}^1 :

Si f constante, immédiat.

Pour U convexe et $f \in \mathcal{C}^1$ avec $df = 0$:

Pour tout $a, b \in U$

$$f(b) - f(a) = \int_0^1 df(a + t(b-a)) \cdot (b-a) = 0$$

V.4 Vecteurs tangents à une partie d'un espace normée de dimension finie

Définition :

Si X est une partie de E et x un point de X , un vecteur v de E est tangent à X en x s'il existe $\varepsilon > 0$ et un arc γ défini sur $] -\varepsilon, \varepsilon[$, dérivable en 0, à valeurs dans X , tels que $\gamma(0) = x$ et $\gamma'(0) = v$
 On note $T_x X$ l'ensemble des vecteurs tangents à X en x .

Exemple V.3. L'ensemble des vecteurs tangents en un point à un ouvert de E est E lui-même puisque pour tout $x \in U$ et pour tout vecteur u de E , on peut définir sur $] -\varepsilon, \varepsilon[$ l'arc $\gamma : t \mapsto x + tu$ (dont le support est une partie de la droite dirigée par u)
 $\gamma(0) = x$ et γ' est constante égale à u .

Définition :

Si E est un espace vectoriel euclidien.

Soit f une fonction définie d'un ouvert U dans \mathbb{R} .

Soit $w \in \mathbb{R}$.

On dit que $X = \{x \in U | w = f(x)\}$ est une ligne de niveau de f .

Proposition :

$T_x X$ est un espace vectoriel

Preuve

Le vecteur nul est tangent : on prend γ constant valant x .

Si $v \in T_x X$ avec γ arc associé et $\lambda \in \mathbb{R}$, alors $\gamma(\lambda t)$ est un arc paramétré qui vaut x en 0 et $\gamma'(0) = \lambda v$. Donc $\lambda v \in T_x X$.

Si $v_1, v_2 \in T_x X$ associés aux arcs γ_1, γ_2 , alors $\gamma = \frac{1}{2}(\gamma_1(2t) + \gamma_2(2t))$ défini sur l'intervalle de définition le plus petit est un arc paramétré, $\gamma(0) = x$ et $\gamma'(0) = v_1 + v_2$

Donc on a bien un sev de E .

Proposition :

Si E est un espace vectoriel euclidien.

Soit f une fonction définie d'un ouvert U dans \mathbb{R} , différentiable.

Soit X une ligne de niveau de f .

Les vecteurs tangents à X au point x_0 de X annulent la différentielle de f en x_0 . Ce sont donc des vecteurs orthogonaux au gradient de f en x_0 .

Preuve

Tout vecteur u tangent à X en x_0 est la dérivée en 0 d'un arc paramétré γ dérivable sur un voisinage de 0, à valeurs dans X et tel que $\gamma(0) = x_0$

Pour t au voisinage de 0, $f(\gamma(t)) = f(x_0)$

Donc la fonction $t \mapsto f(\gamma(t))$ est constante au voisinage de 0 donc de dérivée nulle.

On en déduit, toujours pour t au voisinage de 0 :

$$df(\gamma(t)) \cdot \gamma'(t) = 0$$

En particulier pour $t = 0$: $df(x_0) \cdot \gamma'(0) = 0$, c'est à dire :

$$\langle \nabla f(x_0), u \rangle = 0$$

On admettra le théorème plus fort suivant :

Théorème :

Soit E un espace vectoriel euclidien. Soit f une fonction définie d'un ouvert U dans \mathbb{R} , de classe \mathcal{C}^1 .

Soit X une ligne de niveau de f . Soit $x_0 \in X$ de X .

Si $df(x_0) \neq 0$, alors :

$$T_{x_0} X = \ker(df(x_0)) = (\nabla f(x_0))^\perp$$

V.5 Optimisation

C'est quoi l'optimisation d'un problème qui a plusieurs paramètres ? C'est trouver les meilleurs paramètres. Pour ça, il faut un critère, qu'on modélise par une fonction f à plusieurs variables (les paramètres) et à valeurs réelles. Si f représente un coût, on tente de trouver un minimum, et si f représente une fonction de performance on tente de trouver un maximum.

Conditions de premier ordre

Définition :

Soit f une fonction définie sur $A \subset E$ et à valeurs dans \mathbb{R} .

On dit que f présente un maximum local en un point a de A s'il existe un voisinage V de a dans A tel que :

$$\forall x \in V, f(x) \leq f(a)$$

On dit que f présente un minimum local en un point a de A s'il existe un voisinage V de a dans A tel que :

$$\forall x \in V, f(x) \geq f(a)$$

On dit que f admet un extremum local en un point a de A si elle admet un maximum ou un minimum local en a .

Remarque V.6. On dit que f présente un maximum (global) en un point a de A si :

$$\forall x \in U, f(x) \leq f(a)$$

On définit de même la notion de minimum (global) et d'extremum (global).

Remarque V.7. Attention, un extremum global n'est local que si $x_0 \in \overset{\circ}{A}$.

Par exemple, $f : \begin{cases} [-1, 1] & \rightarrow \mathbb{R} \\ x & \mapsto x^2 \end{cases}$ a un extremum global en 1 mais il n'est pas local.

Définition :

Soit f une fonction définie sur un ouvert U et à valeurs dans \mathbb{R} .

Le point $a \in U$ est appelé point critique de f si f est différentiable en a et si $df(a) = 0$.

Remarque V.8. Si E est un espace euclidien, un point a est un point critique si $\nabla f(a) = 0$.

Théorème :

Soit f une fonction définie sur un ouvert U et à valeurs dans \mathbb{R} et $a \in U$.

Si f admet un extremum local en a et si f est différentiable en a , alors $df(a) = 0$.

(i.e. a est un point critique de f)

Preuve

Supposons $df(a)$ non nul.

On peut alors trouver un vecteur u tel que $df(a) \cdot u \neq 0$.

On peut alors écrire pour $t \in \mathbb{R}$ suffisamment petit :

$$f(a + tu) - f(a) = df(a) \cdot tu + o(t) \sim_0 t df(a) \cdot u$$

Donc $f(a + tu) - f(a)$ change de signe en 0 et f n'a donc pas d'extrémum en a .

On a montré le résultat par contraposée.

Notez l'importance du fait que U est ouvert dans la démonstration.

Exemple V.4. $x \mapsto x^3$ admet un point critique en 0 mais ce n'est pas un extremum local. Ainsi, la réciproque est fausse.

Exemple V.5. Avec $(x, y) \mapsto x^2 + y^2$
 $(0, 0)$ est un point critique

$$\begin{aligned} f(h, k) - f(0, 0) &= h^2 + k^3 \\ &= h^2 + o(\|(h, k)\|^2) = (h, k) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} + o(\|(h, k)\|^2) = \|(h, k)\|^2 \left(\frac{h^2}{h^2 + k^2} + o(1) \right) \end{aligned}$$

Donc si on suit la méthode avec les DL telle quelle, on aurait que c'est un minimum local, mais ça ne prend pas k en compte. Or, sur $(0, k)$, f change de signe avec k , donc $f(0, 0) = 0$ n'est pas un minimum local.

Remarque V.9. Si on doit étudier les extrema d'une fonction f définie sur A partie de E et à valeurs dans F :

1. On se place sur $\overset{\circ}{A}$. On étudie si f est différentiable sur $\overset{\circ}{A}$ et on détermine les points critiques.
2. Les points où f admet un extremum sont à chercher dans $A \setminus \overset{\circ}{A}$, les points de $\overset{\circ}{A}$ où f n'est pas différentiable et les points critiques.

Proposition :

Si f est une fonction numérique définie sur l'ouvert U , si X est une partie de U , si la restriction de f à X admet un extremum local en x et si f est différentiable en x , alors $df(x)$ s'annule en tout vecteur tangent à X en x .

Preuve

Supposons par contraposée que $df(x)$ ne s'annule pas sur $T_x X$.

Alors il existe $v \in T_x X$ tel que $df(x) \cdot v \neq 0$

On a donc un arc paramétré dérivable $\gamma :]-\alpha, \alpha[\rightarrow X$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$

Pour $t \in]-\alpha, \alpha[$:

$$f(\gamma(t)) - f(\gamma(0)) = df(\gamma(0)) \cdot \gamma'(0)t + o(t)$$

$$f(\gamma(t)) - f(x) = tdf(x) \cdot v + o(t)$$

Donc $f(\gamma(t)) - f(x) \sim tdf(x) \cdot v$

Donc $f(\gamma(t)) - f(x)$ restreint change de signe, et donc f_X n'a pas d'extrémum en x .

Théorème : optimisation sous une contrainte

Si f et g sont des fonctions numériques définies et de classe ∞^1 sur l'ouvert Ω de E , si X est l'ensemble des zéros de g , si $x \in X$ et $dg(x) \neq 0$ et si la restriction de f à X admet un extremum local en x , alors $df(x)$ est colinéaire à $dg(x)$.

Preuve

Ici, on a que $T_x X = \ker dg(x)$ (théorème admis)

$df(x)$ et $dg(x)$ sont des formes linéaires telles que $\ker dg(x) \subset \ker df(x)$

Et donc $df(x) = \lambda dg(x)$

$X = g^{-1}(\{0\})$, $dg(x) \neq 0$ et $f|_X$ admet une différentielle en x

Ici, $T_x X = \ker dg(x)$ et de plus, $\ker dg(x) \subset \ker df(x)$ par le théorème précédent, comme $f|_X$ admet un extrémum en x

Si $df(x) = 0$: $df(x) = 0dg(x)$

Sinon : $\ker df(x)$ est un hyperplan (parce que $df(x)$ est une forme linéaire), donc $\ker df(x) = \ker dg(x)$ puisque $\ker dg(x)$ est de dimension 1 en tant que gradient d'un vecteur

Exemple V.6. Cherchons l'aire maximale d'un rectangle à périmètre fixé.

$g(x, y) \mapsto 2x + 2y - p$ représente la différence avec le périmètre (ses zéros sont l'ensemble X).

La fonction f qu'on doit maximiser est $f : (x, y) \mapsto xy$

On se sert du théorème pour trouver les extrémums locaux.

$$\nabla g(x, y) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$\nabla f(x, y) = \begin{pmatrix} y \\ x \end{pmatrix}$$

Si $f|_X$ admet un extrémum local en (x, y) alors : $\frac{x}{2} = \frac{y}{2}$

De plus, $x + y = \frac{p}{2}$

Donc $x = y = \frac{p}{4}$

Donc si on a un extrémum local, celui-là est de cette forme. Et dans ce cas, le rectangle est un carré.

Conditions de second ordre

Définition :

Soit f une fonction de classe \mathcal{C}^2 sur un ouvert U de \mathbb{R}^n euclidien, à valeurs réelles. Soit $x \in U$. La matrice hessienne de f en x est la matrice symétrique :

$$H_f(x) = (\partial^2 f_{i,j}(x))_{(i,j) \in \llbracket 1, n \rrbracket^2}$$

Proposition : Formule de Taylor-Young à l'ordre 2

Soit f une fonction de classe \mathcal{C}^2 sur un ouvert U de \mathbb{R}^n euclidien, à valeurs réelles. Soit $x \in U$.

$$f(x+h) =_{h \rightarrow 0} f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x) \cdot h, h \rangle + o(\|h\|^2)$$

$$f(x+h) =_{h \rightarrow 0} f(x) + \nabla f(x)^\top h + \frac{1}{2} h^\top H_f(x) \cdot h + o(\|h\|^2)$$

Proposition :

- Si f est une fonction de classe \mathcal{C}^2 sur un ouvert de \mathbb{R}^n et si f admet un minimum local en x , alors x est un point critique de f et $H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$
- Si f est une fonction de classe \mathcal{C}^2 sur un ouvert de \mathbb{R}^n , si x est point critique de f et si $H_f(x) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet un minimum local strict en x .

Remarque V.10. — La proposition précédente s'applique aux maxima en remplaçant $H_f(x)$ par $-H_f(x)$

- En particulier, en un point critique x de f tel que $H_f(x) \notin \mathcal{S}_n^+(\mathbb{R})$ et $-H_f(x) \notin \mathcal{S}_n^+(\mathbb{R})$, on peut conclure que f n'admet pas d'extrémum en x .
- La condition x est point critique de f et $H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$ n'est pas suffisante pour déterminer si f admet un extrémum local en x (voir par exemple $(x, y) \mapsto x^2$ et $(x, y) \mapsto x^2 + y^2$)

Proposition : Autre formulation

Si $f : U \subset E \rightarrow \mathbb{R}$ est \mathcal{C}^2 avec x_0 un point critique de f .

- Si $H_f(x_0) \in \mathcal{S}_p^{++}(\mathbb{R})$, f a un minimum local.
- Si $-H_f(x_0) \in \mathcal{S}_p^{++}(\mathbb{R})$, f a un maximum local.
- Si $H_f(x_0)$ possède une valeur propre strictement négative ($H_f \notin \mathcal{S}_p^+(\mathbb{R})$), alors f n'a pas de minimum en x_0 .
- Si $H_f(x_0)$ possède une valeur propre strictement positive ($-H_f \notin \mathcal{S}_p^+(\mathbb{R})$), alors f n'a pas de maximum en x_0 .

Preuve

$H_f(x_0) \in \mathcal{S}_p^{++}(\mathbb{R})$

Les valeurs propres de $H_f(x_0)$ sont $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$

Il existe $P \in \mathcal{O}_n(\mathbb{R})$ tel que $P^T H_f(x_0) P = \text{diag}(\lambda_1, \dots, \lambda_p)$

Pour $h \in E$, on pose $h = Ph'$

$$f(x_0 + h) - f(x_0) = \|h\|_2^2 \left(\frac{1}{2} \frac{\sum_{i=1}^p \lambda_i h_i'^2}{\sum_{i=1}^p h_i'^2} + o(\|h\|_2^2) \right)$$

$$\text{Or } \frac{1}{2} \frac{\sum_{i=1}^p \lambda_i h_i'^2}{\sum_{i=1}^p h_i'^2} \geq \frac{1}{2} \lambda_1$$

Donc on a bien que c'est un minimum local, puisque positif sur un certain voisinage.

Pour le troisième cas : On suppose que $H_f(x_0)$ possède une valeur propre $\lambda < 0$

On note v un vecteur propre associé.

Pour $t \in]-\alpha, \alpha[$ avec $\alpha > 0$, on a :

$$\begin{aligned} f(x_0 + tv) - f(x_0) &= \frac{1}{2} (tv)^T H_f(x_0) (tv) + o(t^2) \\ &= \frac{1}{2} t^2 v^T v \lambda + o(t^2) \\ &\sim \left(\frac{1}{2} v^T v \lambda \right) t^2 \end{aligned}$$

Et $\lambda < 0$, donc il n'y a pas de minimum en ce point.

Exemple V.7. Cherchons les extrémums de $f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto (2x + y)e^{-x^2+2y^2} \end{cases}$

Pour tous $(x, y) \in \mathbb{R}^2$, f admet deux dérivées partielles :

$$\frac{\partial f}{\partial x} = (2 - (2x + y)2x)e^{-(x^2+2y^2)}$$

$$\frac{\partial f}{\partial y} = (1 - (2y + y)4y)e^{-(x^2+2y^2)}$$

$\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues donc f est \mathcal{C}^1 donc différentiable.

Soit $(x, y) \in \mathbb{R}^2$.

$$\begin{aligned}
 (x, y) \text{ est critique si, et seulement si, } & \begin{cases} 2x(2x + y) = 2 \\ 4y(2x + y) = 1 \end{cases} \\
 \Leftrightarrow & \begin{cases} 2x + y \neq 0, x \neq 0, y \neq 0 \\ \frac{1}{x} = \frac{1}{4y} \\ x(2x + y) = 1 \end{cases} \\
 \Leftrightarrow & \begin{cases} x = 4y \\ 4y(9y) = 1 \end{cases} \\
 \Leftrightarrow & \begin{cases} x = \frac{4}{6} = \frac{2}{3} \\ y = \frac{1}{6} \end{cases} \quad \text{ou} \quad \begin{cases} x = \frac{-2}{3} \\ y = \frac{-1}{6} \end{cases}
 \end{aligned}$$

Si f avec un extrémum en $u_0 = (x_0, y_0)$, alors u_0 est critique.

étude locale en $(x_0, y_0) = (\varepsilon \frac{2}{3}, \varepsilon \frac{1}{6})$ avec $\varepsilon = \pm 1$

Avec (h, k) au voisinage de 0, et on veut obtenir un $\mathcal{O}(\|(h, k)\|^3)$, pour avoir le signe du $o(h)$

$$\begin{aligned}
 f(x_0 + h, y_0 + k) &= (2x_0 + y_0 + 2h + k)e^{-(x_0+h)^2 - 2(y_0+h)^2} \\
 &= (2x_0 + y_0 + 2h + k)e^{-x_0^2 - 2y_0^2}e^{-2kx_0 - h^2 - 4ky_0 - 2k^2} \\
 &= (2x_0 + y_0 + 2h + k)e^{-x_0^2 - 2y_0^2}(1 - 2kx_0 - h^2 - 4ky_0 - 2k^2 \\
 &\quad + \frac{1}{2}(4h^2x_0^2 + 16k^2y_0^2 + 16hky_0x_0) + o(\|(h, k)\|^2))
 \end{aligned}$$

$$\begin{aligned}
 f(x_0 + h, y_0 + k) - f(x_0, y_0) &= e^{-x_0^2 - 2y_0^2}h(2 - 2x_0(x_0 + y_0)) + e^{-x_0^2 - 2y_0^2}k(1 - 4y_0(2x_0 + y_0)) \\
 &\quad + e^{-x_0^2 - 2y_0^2}h^2(-4x_0 + (2x_0 + y_0)[-1 + 2x_0^2]) \\
 &\quad + e^{-x_0^2 - 2y_0^2}k^2(-4y_0 + (2x_0 + y_0)(-2 + 8y_0)) \\
 &\quad + e^{-x_0^2 - 2y_0^2}hk(8y_0 - 2yx_0 + (2x_0 + y_0)(8x_0y_0)) \\
 &\quad + o(\|(h, k)\|^2)
 \end{aligned}$$

Les termes en h et en k sont censés s'annuler, en retrouvant les équations qui ont donné les points critiques. C'est bien le cas ici, donc on n'a pas fait d'erreurs de calculs là-dessus. On prend ensuite $\varepsilon = 1$

$$\begin{aligned}
 f(x_0 + h, y_0 + k) &= e^{-x_0^2 - 2y_0^2}\left(\frac{-17}{6}h^2 - \frac{10}{3}k^2 - \frac{4}{3}hk\right) + o(\|(h, k)\|^2) \\
 &= \frac{e^{-x_0^2 - 2y_0^2}}{6}(-17h^2 - 20k^2 - 8hk) + o(\|(h, k)\|^2)
 \end{aligned}$$

On retrouve le même développement que pour une forme bilinéaire symétrique. Par identification, on trouve S cette matrice :

$$-17h^2 - 20k^2 - 8hk = (h, k) \begin{pmatrix} -17 & -4 \\ -4 & -20 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

S possède deux valeurs propres négatives telles que $\lambda_1 < \lambda_2 < 0$ (comme $\lambda_1\lambda_2 = \det(S)$, $\det(S)$ positif donc elles sont de même signe, et de plus $\lambda_1 + \lambda_2 = \text{tr}(S)$, et comme $\text{tr}(S)$ est négative les deux sont négatives.)

Il existe une matrice $P \in \mathcal{O}_2(\mathbb{R})$, $P^{-1}SP = P^TSP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

Si on pose pour $(h, k) \in \mathbb{R}^2$:

$$\begin{pmatrix} h \\ k \end{pmatrix} = P \begin{pmatrix} h' \\ k' \end{pmatrix}$$

$$\begin{aligned} (h, k)S \begin{pmatrix} h \\ k \end{pmatrix} &= (h', k')P^TSP \begin{pmatrix} h' \\ k' \end{pmatrix} \\ &= (h', k') \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} h' \\ k' \end{pmatrix} \\ &= \lambda_1 h'^2 + \lambda_2 k'^2 \end{aligned}$$

Et donc, avec $K = \frac{e^{-x_0^2 - 2y_0^2}}{6}$:

$$f(x_0 + h, y_0 + k) = -K \|(h, k)\|_2^2 \left(\frac{(-\lambda_1)h'^2 + (-\lambda_2)k'^2}{\|(h', k')\|_2} + o(1) \right)$$

Or $\|(h, k)\|_2 = \|(h', k')\|$ parce que P est une isométrie.

Et donc $\frac{(-\lambda_1)h'^2 + (-\lambda_2)k'^2}{\|(h', k')\|} \geq (-\lambda_2) > 0$

Donc le DL est de signe négatif en (x_0, y_0) .

Donc f possède un maximum local en $(\frac{2}{3}, \frac{1}{6})$.

On pourra prouver de la même manière, en notant que le DL est positif en $(-x_0, -y_0)$, que c'est un minimum local.

En travaillant sur la fonction de manière globale, on pourra prouver qu'il s'agit en fait d'extrémums globaux.

Exemple V.8.

$$f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto (2x + y)e^{-x^2 - 2y^2} \end{cases}$$

Alors $\frac{\partial f}{\partial x}(x, y) = (2 - (2x + y)2x)e^{-x^2 - 2y^2}$ et $\frac{\partial f}{\partial y}(x, y) = (1 - (2x + y)4y)e^{-x^2 - 2y^2}$

$$\frac{\partial^2 f}{\partial x^2}(x, y) = (-8x - 2y - (2 - (2x + y)2x)2x)e^{-x^2 - 2y^2}$$

$$\frac{\partial^2 f}{\partial y^2}(x, y) = (8x - 8y - (1 - 2(x + y)4y)4y)e^{-x^2 - 4y^2}$$

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = (-8y - (1 - (2x + y)4y)2x)e^{-x^2 - 2y^2}$$

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = (-2y - (2 - (2x + y)2x)4y)e^{-x^2 - 2y^2}$$

Exemple V.9. Donnons le développement limité en ordre 2 du f précédent en $(\frac{2}{3}, \frac{1}{6})$

On pose $K = e^{-x_0^2 - 2y_0^2}$

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) = K \left(-8\frac{2}{3} - \frac{2}{6} \right) = \frac{-17}{3}K$$

$$\frac{\partial^2 f}{\partial y^2}(x_0, y_0) = K \left(-8\frac{2}{3} - \frac{8}{4} \right) = \frac{-20}{3}K$$

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{-4}{3}K$$

Donc :

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = -\frac{1}{2} \frac{K}{3} \begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} 17 & 4 \\ 4 & 20 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} + o(\|(h, k)\|^2)$$

VI Hors programme

VI.1 Fonctions k fois différentiables

Définition :

Soit $f : U \subset E \rightarrow F$

f est k fois différentiable pour $k \geq 1$ si f est $(k-1)$ fois différentiable et $d^{k-1}f$ est différentiable.

f est \mathcal{C}^k si f est k fois différentiable et $d^k f$ est continue.

Exemple VI.1. Si f deux fois différentiable, pour $x \in U, h, k \in E$

$$d^2 f(x) \in \mathcal{L}(E, \mathcal{L}(E, F))$$

$$d^2 f(x) \cdot H \in \mathcal{L}(E, F)$$

$$d^2 f(x) \cdot h \cdot k = (d^2 f(x) \cdot h) \cdot k \in F$$

Par théorème de Schwarz, si f est 2 fois différentiable alors $\forall h, k \in E, d^2 f(x) \cdot h \cdot k = d^2 f(x) \cdot k \cdot h$

Si (e_1, \dots, e_n) est une base de E

On sait que $\partial_j f(x) = df(x) \cdot e_j$

Et donc $\partial_i \partial_j f(x) = (t \mapsto df(x + te_i) \cdot e_j)'(0) = d^2 f(x) \cdot e_i \cdot e_j$

En effet, $df(x + te_i) \cdot e_j = (df(x) + d^2 f(x) \cdot e_i + o(t)) \cdot e_j = df(x) \cdot e_j + td^2 f(x) \cdot e_i \cdot e_j + o(t)$

Si $h = h_1 e_1 + \dots + h_n e_n$ et $k = k_1 e_1 + \dots + k_n e_n$

$$\begin{aligned} d^2 f(x) \cdot h \cdot k &= d^2 f(x) \cdot \left(\sum_{i=1}^n h_i e_i \right) \cdot \left(\sum_{j=1}^n h_j e_j \right) \\ &= \sum_{1 \leq i, j \leq n} h_i k_j d^2 f(x) \cdot e_i \cdot e_j &= \sum_{1 \leq i, j \leq n} h_i k_j \partial_i \partial_j f(x) \end{aligned}$$

Si $F = \mathbb{R}$

$d^2 f(x) \cdot h \cdot k = H^T H_f(x) K$ avec $H = \text{Mat}_{\mathcal{B}}(h)$ et $K = \text{Mat}_{\mathcal{B}}(k)$

La formule de Taylor à l'ordre k devient :

Avec U un ouvert connexe, $a, b \in U$.

Si f est une fonction \mathcal{C}^{k+1}

$$\varphi : \begin{cases} [0, 1] & \rightarrow F \\ t & \mapsto f(a + (b-a)t) \end{cases}$$

Alors φ est \mathcal{C}^{k+1}

$$\forall t \in [0, 1], \varphi'(t) = df(a + t(b-a)) \cdot (b-a)$$

$$\varphi''(t) = d^2 f(a + t(b-a)) \cdot (b-a) \cdot (b-a) \quad \dots \varphi^p(t) = d^p f(a + t(b-a)) \cdot (b-a)^p$$

Et donc $\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2} \varphi''(0) + \dots + \varphi^{(k)}(0) + \int_0^1 \frac{(1-t)^k}{k!} d^{k+1} f(a + t(b-a)) \cdot (b-a)^{k+1} dt$

Ce qui nous donne finalement :

$$f(b) = f(a) + df(a) \cdot (b-a) + \frac{1}{2} d^2 f(a) \cdot (b-a) \cdot (b-a) + \dots + \frac{1}{k!} d^k f(a) \cdot (b-a)^k + o(h)$$

Donc pour revenir au premier ordre $f(a+h) = f(a) + df(a) \cdot h + \int_0^1 (1-t) d^2 f(a+ht) \cdot h \cdot h dt$

VI.2 Démo hors-programme : fonction constante différentiable

Définition :

Soit U un ouvert de E connexe par arc.

Si $f : U \rightarrow F$ est différentiable, alors :

$$f \text{ constante} \Leftrightarrow \forall x \in U, df(x) = 0$$

Preuve

On a déjà prouvé le sens direct, et le sens indirect quand U était un ouvert convexe.

On suppose U connexe par arcs non-vide et que $\forall x \in U, df(x) = 0$.

On choisit $x_0 \in U$

On considère $A = \{x \in U | f(x) = f(x_0)\}$

On va montrer que $A = U$

$A = f^{-1}(\{f(x_0)\})$, or f est continue, donc A est l'image réciproque du fermé $\{f(x_0)\}$ par une application continue.

Donc A est un fermé relatif de U .

Soit $x \in A$, comme U est ouvert on prend $\alpha > 0$ telle que $\mathcal{B}(x, \alpha) \subset U$

Or $\mathcal{B}(x, \alpha)$ est convexe

Donc $\forall y \in \mathcal{B}(x, \alpha), f(y) - f(x) = \int_0^1 df(x + t(y - x)) \cdot (y - x) dt = 0$

Donc $\mathcal{B}(x, \alpha) \subset A$

Donc A est un ouvert.

Dans un ouvert connexe par arcs, un ensemble qui est un ouvert et un fermé en même temps est soit vide, soit l'ensemble tout entier.

Pour le montrer, supposons par l'absurde que $\mathcal{C}_U A \neq \emptyset$

On a donc $y_0 \in \mathcal{C}_U A$

Comme U est connexe par arcs, on a une arc continu de x_0 à y_0 :

$$\gamma \begin{cases} [0, 1] & \rightarrow U \\ t & \mapsto \gamma(t) \end{cases}$$

où $\gamma(0) = x_0 \in A$ et $\gamma(1) = y_0 \in \mathcal{C}_U A$

On peut donc considérer $t_0 \in [0, 1]$ tel que $t_0 = \sup\{t \in [0, 1] | \gamma(t) \in A\}$

$\mathcal{C}_U A$ est ouvert donc $t_0 < 1$ ($\exists \beta \in \mathbb{R}_+^*, \mathcal{B}(y_0, \beta) \subset \mathcal{C}_U A$)

Et $\gamma(t_0) \in A$: il existe $t_n \in [0, 1]^{\mathbb{N}}$ tel que $t_n \rightarrow t_0$ et $\forall n \in \mathbb{N}, \gamma(t_n) \in A$ par caractérisation séquentielle du sup ; donc par continuité de $f \circ \gamma$, $f \circ \gamma(t_0) = f(x_0)$

Or A est ouvert, donc il existe $\delta \in \mathbb{R}_+^*$ tel que $\mathcal{B}(\gamma(t_0), \delta) \subset A$

Ce qui nie la définition de t_0

D'où la contradiction.

VI.3 Différentielle d'un inverse

Proposition : Dans \mathbb{R}

Si f dérivable bijective d'un intervalle I sur un intervalle J , si f' ne s'annule pas alors f^{-1} est dérivable et :

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

Proposition : Dans le cas général

Si f est différentiable bijective d'un ouvert U sur un ouvert V , et si f^{-1} est différentiable
Alors $\dim E = \dim F$ et :

$$\begin{cases} \forall x \in U, df^{-1}(f(x)) = df(x)^{-1} \\ \forall x \in V, df^{-1}(x) = df(f^{-1}(x)) \end{cases}$$

Théorème :

Si f est différentiable bijective d'un ouvert U sur un ouvert V

Soit $x \in U$ tel que $df(x)$ est bijective.

Alors f^{-1} est différentiable en $f(x)$ et :

$$df^{-1}(f(x)) = (df(x))^{-1}$$

Exemple VI.2.

$$\varphi : \begin{cases} \mathbb{R}_+^* \times]0, 2\pi[& \rightarrow \mathbb{R}^2 \setminus D_+ \\ (r, \theta) & \mapsto (x, y) = (r \cos \theta, r \sin \theta) \end{cases}$$

Avec D_+ la demi-droite $\{(x, y) \in \mathbb{R}^2 \mid y = 0, x \geq 0\}$. C'est le seul moyen d'avoir une application en polaire qui soit bijective.

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

$$\text{Et } J_{\varphi^{-1}}(r \cos \theta, r \sin \theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \frac{-\sin \theta}{r} & \frac{\cos \theta}{r} \end{pmatrix}$$

$$\psi : \begin{cases} \mathbb{R}_+^* \times]0, 2\pi[\times]0, 2\pi[& \rightarrow \mathbb{R}^3 \setminus P^+ \\ (r, \theta, \varphi) & \mapsto (x, y, z) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) \end{cases}$$

Qui est bijective à l'exception d'un demi-plan.

Chapitre XIII

Groupes et anneaux

I Rappels de sup

Définition : LCI

Sur un ensemble E , une lci \star est une application :

$$\star : \begin{cases} E \times E & \rightarrow E \\ (x, y) & \mapsto x \star y \end{cases}$$

On dit qu'elle est associative si :

$$\forall a, b, c \in E, a \star (b \star c) = (a \star b) \star c$$

On dit qu'elle est commutative si :

$$\forall a, b \in E, a \star b = b \star a$$

Définition : Neutre

Un élément neutre est un élément e tel que :

$$\forall x \in E, x \star e = e \star x = x$$

Proposition : Unicité du neutre

Il n'y a qu'un seul élément neutre dans un groupe

Preuve

Si e et e' étaient des neutres, alors $e \star e' = e = e'$. D'où l'unicité.

Définition : Symétrique

Le symétrique x' de $x \in E$ est un élément tel que, si la lci admet un neutre e :

$$x \star x' = x' \star x = e$$

On note x^{-1} le symétrique de x .

Proposition : Unicité du symétrique

Si la loi est associative, alors x admet un unique symétrique

Preuve

Supposons que $x \in E$ admette deux symétriques, x' et x'' .

$$x' \star x \star x'' = (x' \star x) \star x'' = x'' = x' \star (x \star x'') = x'$$

Définition : Groupe

Un groupe est un ensemble qui est muni d'une loi associative, qui possède un neutre et où tout élément admet un symétrique.

Un groupe commutatif ou abélien est un groupe où la loi est commutative.

Un ensemble muni d'une loi associative avec un neutre mais où tout élément n'est pas symétrisable est appelé un monoïde.

Proposition :

Si G un monoïde si $a, b \in G$ sont symétrisables, alors ab est symétrisable et $(ab)^{-1} = b^{-1}a^{-1}$

Preuve

$$\begin{aligned} abb^{-1}a^{-1} &= aea^{-1} \\ &= aa^{-1} = e \end{aligned}$$

Donc on a exhibé un inverse de ab , qui est unique.

Définition : Sous-groupe

Si G est un groupe pour la loi \star de neutre e :

H est un sous-groupe de G si :

$$\begin{cases} \forall (x, y) \in H, x \star y \in H \\ H \neq \emptyset (\Leftrightarrow e \in H) \\ \forall x \in H, x^{-1} \in H \end{cases}$$

Remarque I.1. Notations : en général, on les prend de manières multiplicatives ou additives. On réserve les notations additives pour les groupes commutatifs la plupart du temps.

On note $\star \cdot$ de manière multiplicative. On note 1 le neutre le plus souvent.

Pour $a \in G$, on note : $\forall n \in \mathbb{N}^*, a^n = a \cdot \dots \cdot a$ n fois.

Par convention, on a $a^0 = 1$

Pour $n \in \mathbb{N}^*$, $a^{-n} = a^{-1} \star \dots \star a^{-1}$ n fois.

On note $\langle a \rangle = \{a^n | n \in \mathbb{Z}\}$ le groupe engendré par a . C'est le plus petit groupe contenant a .

On note $\star \top$ de manière additive. On note 0 le neutre le plus souvent, et $-a$ le symétrique.

Pour $a \in G$, on note : $\forall n \in \mathbb{N}^*, na = a \top \dots \top a$ n fois.

Par convention, on a $0a = 0$

Pour $n \in \mathbb{N}^*$, $-na = -a \star \dots \star -a$ n fois.

On note $\langle a \rangle = \{na | n \in \mathbb{Z}\}$ le groupe engendré par a . C'est le plus petit groupe contenant a .

Exemple I.1. Exemples fondamentaux : $(\mathbb{Z}, +)$ est un groupe, mais pas \mathbb{N}

Si G un groupe et A un ensemble quelconque, alors $\mathcal{F}(A, G)$ est un groupe pour la loi $(f, g) \mapsto fg = (a \mapsto f(a) \star g(a))$

Avec $(G_1, \star_1), (G_2, \star_2)$ deux groupes, alors $G_1 \times G_2$ est un groupe, muni de la loi $(g_1, g_2) \star (g'_1, g'_2) = (g_1 \star_1 g'_1, g_2 \star_2 g'_2)$

Par exemple, $(\mathbb{Z}, +) \times (\mathbb{R}_+^*, \cdot)$ est un groupe et $(k, x) \cdot (l, y) = (k + l, xy)$ est sa loi.

Avec A un ensemble, $\mathcal{B}(A, A)$ l'ensemble des bijections de A dans A , aussi noté $S_A (= \mathfrak{S}(A))$ quand A est fini, est un groupe pour la composition.

Définition :

Avec (G, \star) un groupe et \mathcal{R} une relation d'équivalence sur G

On note $\mathcal{G}/\mathcal{R} = \{\overline{g_{\mathcal{R}}} | g \in G\}$ où $\overline{g_{\mathcal{R}}} = \{g' \in G | g' \mathcal{R} g\}$

Si \mathcal{R} est compatible avec \star alors on peut définir une loi naturelle sur $G_{\mathcal{R}}$

$$\cdot : \begin{cases} G_{\mathcal{R}} \times G_{\mathcal{R}} & \rightarrow G_{\mathcal{R}} \\ \overline{x}, \overline{y} & \mapsto \overline{x \star y} \end{cases}$$

Et dans ce cas $\varphi : \begin{cases} G & \rightarrow G_{\mathcal{R}} \\ x & \mapsto \overline{x} \end{cases}$ est un morphisme surjectif, qu'on appelle la projection.

Exemple I.2. Dans $(\mathbb{Z}, +)$, avec $n \in \mathbb{N}, n \geq 2$, on note $\mathcal{R} \equiv n$ la congruence modulo n définie par :

$$x \mathcal{R} y \Leftrightarrow x - y \equiv 0[n] \Leftrightarrow (x - y) \in n\mathbb{Z}$$

On appelle ce groupe le groupe quotient.

On peut généraliser cette notion : Si G un groupe, et H un sous-groupe de G , on définit \mathcal{R}_H sur G par :

$$\forall x, y \in G, x \mathcal{R}_H y \Leftrightarrow xy^{-1} \in H$$

C'est bien une relation d'équivalence :

$\forall x \in G, xx^{-1} = 1$ et H est un sous-groupe de G donc $1 \in H$ donc \mathcal{R}_H est réflexive.

$\forall x, y \in G, xy^{-1} \in H$ donne que $(xy^{-1})^{-1} \in H$, or $(xy^{-1})^{-1} = yx^{-1}$, donc $y \mathcal{R}_H x$, donc \mathcal{R}_H est symétrique.

$\forall x, y, z \in G, x \mathcal{R}_H y$ et $y \mathcal{R}_H z$ donne que $xy^{-1} \in H$ et $yz^{-1} \in H$, donc $xy^{-1} = xy^{-1}yz^{-1} \in H$, donc \mathcal{R}_H est transitive.

Malheureusement, cette relation n'est pas toujours compatible. C'est vrai si la loi est commutative.

On peut quand même avoir ce résultat :

Soit $y \in G$, soit $x \in G$.

$$\begin{aligned} x \in \overline{y_{\mathcal{R}_H}} &\Leftrightarrow xy^{-1} \in H \\ &\Leftrightarrow \exists h \in H, xy^{-1} = h \\ &\Leftrightarrow \exists h \in H, x = hy \\ &\Leftrightarrow x \in Hy \end{aligned}$$

Et donc $\overline{y_{\mathcal{R}_H}} = Hy = \{hy | h \in H\}$

Donc le cardinal d'une classe selon \mathcal{R}_H est le cardinal de H .

Donc le cardinal de G est le cardinal de H multiplié par le nombre de classes.

On peut en déduire le théorème (hors-programme) de Lagrange :

Si G est un groupe fini et H un sous-groupe de G , alors $\text{card} H | \text{card} G$

On le prouve en considérant $\mathcal{R}_H : x\mathcal{R}_Hy \Leftrightarrow xy^{-1} \in H$

Et donc $\forall y \in G, \text{card} \overline{y\mathcal{R}_H} = \text{card} H$

Donc $\text{card} G = \text{card}(G/\mathcal{R}_H) \text{card} H$

Revenons sur la condition de compatibilité de \mathcal{R}_H :

$$\forall (x, y) \in G^2, HxHy = \{h_1xh_2y \mid (h_1, h_2) \in H^2\} \subset Hxy$$

L'autre inclusion est toujours vraie : $\forall x, y \in G, \forall h \in H, hxy = hxy \in Hxy$ et $1 \in H$ car H est un sous-groupe de G

Si le groupe est commutatif, alors : $\forall x, y \in G, \forall h_1, h_2 \in H, h_1xh_2y = (h_1h_2)xy$ et $h_1h_2 \in H$ car H est un sous-groupe de G .

Les sous-groupes qui fonctionnent sont appelés les groupes distingués.

II Groupes engendrés par une partie non-vide

Proposition : Intersection

Si G un groupe

Si $(H_i)_{i \in I}$ une famille de sous-groupes de G

Alors $(\cap H_i)_{i \in I}$ est un sous-groupe de E

Preuve

$\forall i \in I, e \in H_i$ donc $e \in (\cap H_i)_{i \in I}$

Si $x, y \in G$ tels que $\forall i \in I, x, y \in H_i$ alors $\forall i \in I, xy \in H_i$ comme c'est un sous-groupe donc $xy \in (\cap H_i)_{i \in I}$

Donc c'est bien un sous-groupe de G .

Remarque II.1. On peut donc définir un groupe engendré par un élément a comme l'intersection de tous les sous-groupes de G contenant a .

Définition :

Si $A \neq \emptyset$, alors le plus petit sous-groupe de G contenant A est l'intersection de tous les sous-groupes de G contenant H .

C'est le groupe engendré par A , noté $\langle A \rangle$

Définition : Groupe monogène

On note $\langle \{a\} \rangle = \langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$

Preuve

On prouve que c'est un groupe.

Ensuite, on prouve que tout groupe contenant a contient aussi les itérés de a .

Définition :

Si $a, b \in G$:

Si G est commutatif, alors $\langle \{a, b\} \rangle = \{a^p b^q \mid (p, q) \in \mathbb{Z}^2\}$

Sinon, on a $\langle \{a, b\} \rangle = \{w \in G \mid \exists n \in \mathbb{N}, w = x_1 \dots x_n \text{ et } \forall i \in \llbracket 1, n \rrbracket, x_i \in \{a, b, a^{-1}, b^{-1}\}\}$

Remarque II.2. *Ainsi, on ne traitera pas les groupes engendrés par plus d'un élément, puisqu'il est très difficile de les manipuler.*