École Normale Supérieure

Paris-Saclay

Plans de leçons pour l'agrégation

Auteur :
Jean Cazalis

2016 – 2017

Table des matières

Première partie

Algèbre et géométrie

Un exemple

Rapport du jury

Introduction

Plan

Développements

(i)

(ii)

Références

(i)

(ii)

(iii)

(iv)

101 : Groupe opérant sur un ensemble. Exemples et applications.

Rapport du jury

“Dans cette leçon, il faut bien dominer les deux approches de l’action de groupe : l’approche naturelle et l’approche via le morphisme du groupe agissant vers le groupe des permutations de l’ensemble sur lequel il agit. La formule des classes et ses applications immédiates sont incontournables. Des exemples de natures différentes doivent être présentés : actions sur un ensemble fini, sur un espace vectoriel (en particulier les représentations), sur un ensemble de matrices, sur des groupes ou des anneaux. Les exemples issus de la géométrie ne manquent pas (groupes d’isométries d’un solide).

S’ils le désirent, les candidats peuvent aller plus loin en décrivant les actions naturelles de $\mathrm{PGL}(2, \mathbb{F}_q)$ sur la droite projective qui donnent des injections intéressantes pour $q = 2, 3$ et peuvent plus généralement en petit cardinal donner lieu à des isomorphismes de groupes. En notant que l’injection du groupe de permutations dans le groupe linéaire par les matrices de permutations donne lieu à des représentations, ils pourront facilement en déterminer le caractère.”

Introduction

Plan

I Généralités sur les actions de groupes et étude des groupes finis

(a) Définition et premiers exemples

Définition d’une action de groupe (les deux). Action de $\mathfrak{S}_n$ sur $\{1, \dots, n\}$. Action de $\mathrm{GL}_n(k)$ sur k^n. Action transitive, action fidèle. Orbites d’une action de groupes. Retour sur les exemples. Stabilisateur, bijection.

Exemple : translation à gauche. Application : théorème de Cayley. Exemple : automorphisme intérieur, classes de conjugaison, centralisateur. Exemple : G opère par translation à gauche (par automophisme intérieur) sur G/H.

(b) La formule des classes et quelques applications

Formule des classes. Application : Si p est le plus petit facteur premier de $|G|$ alors un sous-groupe d’indice p est distingué dans G [RIS].

Application : le centre d’un p-groupes est non trivial. Application : un groupe de cardinal p^α possède des sous-groupes d’ordre p^β pour tout $\beta \leq \alpha$. Application : sous-groupes maximaux d’un p-groupe.

Application : théorèmes de Sylow.

(c) Exemple : définition des angles en dimension 2

[RIS]

II Représentations linéaires d’un groupe fini

(a) Généralités sur les représentations linéaires

Définition représentation de groupes finis. Exemples : représentation régulière d’un groupe fini, représentation du groupe $\mathfrak{S}_n$. Définition sous-représentation, représentations irréductibles. Définition somme de représentations. Représentations unitarisables, théorème de Maschke.

Opérateurs d’entrelacement, lemme de Schur. Application aux représentations irréductibles des groupes abéliens. Application : théorème de structure des groupes abéliens finis.

(b) Exemple : représentation du groupe des quaternions

III Action de $\mathrm{GL}(E)$

Action de $\mathrm{GL}(E)$ sur E. Etude.
Action par équivalence. Rang.
Action par similitude.

Action par congruence sur les formes quadratiques, groupe orthogonal général.

Développements

(i) Théorèmes de Sylow

(ii) Quaternions et $SO_3(\mathbb{R})$

Références

(i) Perrin, Cours d'algèbre

(ii) Risler & Boyer, Algèbre pour la L3

(iii) Gourdon, Algèbre

102 : Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.

Rapport du jury

"Il ne faut pas uniquement aborder cette leçon de façon élémentaire sans réellement expliquer où et comment les nombres complexes de modules 1 et les racines de l'unité apparaissent dans divers domaines des mathématiques (exponentielle complexe et ses applications, polynômes cyclotomiques, spectre de matrices remarquables, théorie des représentations). Il ne faut pas non plus oublier la partie « groupe » de la leçon : on pourra s'intéresser au relèvement du groupe unité au groupe additif des réels et aux propriétés qui en résultent. De même les sous-groupes finis de $\mathbb{S}^1$ sont intéressants à considérer dans cette leçon.

On pourra aussi s'intéresser aux groupes des nombres complexes de $\mathbb{Q}[i]$, et les racines de l'unité qui y appartiennent ; tout comme aux sous-groupes compacts de $\mathbb{C}^*$."

Introduction

Plan

I Généralités

Définition du groupe $\mathbb{U}$ comme noyau du morphisme $z \in \mathbb{C}^* \longmapsto |z| \in \mathbb{R}^*$. $\mathbb{U}$ est compact.

(a) L'exponentielle complexe

Intro du Rudin. Application à la trigonométrie : formule explicite pour le noyau de Dirichlet ou le noyau de Féjer.

(b) Sous-groupes de $\mathbb{U}$

Les sous-groupes de $\mathbb{U}$ sont les sous-groupes de $\mathbb{R}/2\pi\mathbb{Z}$ et sont soit dense soit fini. CNS. Isomorphisme entre $\mathbb{U}_n$ et $\mathbb{Z}/n\mathbb{Z}$.

Sous-groupes compacts de $\mathbb{C}^*$.

(c) Le groupe des quaternions

Blabla [PER].

II Racines de l'unité et polynômes cyclotomique

Blabla [PER]. Théorème de Kronecker et application.

III Applications

(a) Mesure d'angles

Blabla [AUD].

(b) Représentation du groupe des isométrie directes

Blabla [AUD]. Généralisation avec le groupe quaternions.

(c) Polygones réguliers constructibles

Blabla théorème de Gauss-Wantzel.

(d) Réduction d'endomorphismes

Réduction des endomorphismes normaux, des endomorphismes unitaires. Application : connexité par arcs.

Développements

(i) Théorème de Gauss-Wantzel
(ii) Théorème de Kronecker

Références

(i) Perrin, Cours d'algèbre
(ii) Audin, Géométrie
(iii) Gourdon, Algèbre

103 : Exemples de sous-groupes distingués et de groupes quotients. Applications.

Rapport du jury

"Dans cette leçon, il faut non seulement évoquer les notions de groupe quotient, de sous-groupe dérivé et de groupe simple mais surtout savoir les utiliser et en expliquer l'intérêt. On pourra utiliser des exemples issus de la géométrie, de l'arithmétique, de l'algèbre linéaire (utilisation d'espaces vectoriels quotients par exemple). La notion de produit semi-direct n'est plus au programme ; mais, lorsqu'elle est utilisée, il faut savoir la définir proprement et savoir reconnaître des situations simples où de tels produits apparaissent (le groupe diédral D n par exemple).

S'ils le désirent, les candidats peuvent poursuivre en illustrant ces notions à l'aide d'une table de caractères et décrire le treillis des sous-groupes distingués, ainsi que l'indice du sous-groupe dérivé, d'un groupe fini à l'aide de cette table."

Introduction

Plan

I Généralités sur les groupes

(a) Sous-groupes distingués, groupes quotients

Définition des classes à gauche (à droite). Propriétés. Théorème de Lagrange. Définition indice. Définition groupe distingué. Exemple : sous-groupes triviaux, sous-groupes d'un groupe abélien, $\mathfrak{A}_n$, le sous-groupe des doubles transpositions dans $\mathfrak{A}_4$.

Groupe quotient. Propriété universelle des morphismes de groupes. Exemple de $\mathbb{Z}/n\mathbb{Z}$, $PGL_n(\mathbb{K})$.

Groupe simple. Exemple : $\mathbb{Z}/p\mathbb{Z}$ (p premier).

(b) Exemples génériques de sous-groupes distingués

Centre d'un groupe. Exemple : le centre d'un groupe commutatif est le groupe tout entier, le centre de $\mathfrak{S}_n$ est trivial dès que $n \geq 3$, le centre de $\mathrm{GL}_n(\mathbb{K})$ est l'ensemble des homothéties.

Définition groupe dérivé. Quelques exemples. Propriétés du groupe dérivé. Conséquences de la simplicité de $\mathfrak{A}_n$, groupe dérivé de $\mathrm{GL}_n(\mathbb{K})$ dans les cas simples. Application : début de Frobenius-Zolotarev.

II Applications en théorie des groupes finis

(a) Étude des p-groupes

Un sous-groupe d'indice le plus nombre premier qui divise n est distingué.

Le centre d'un p-groupe est non trivial. Application : sous-groupes d'un p groupe, les sous-groupes maximaux d'un p-groupe sont d'indice p donc distingués. Application : théorème de la base de Burnside.

(b) Les théorèmes de Sylow

Définition p-Sylow. Théorèmes de Sylow et conséquences.

Application : $\mathfrak{A}_n$ pour $n \geq 5$. Conséquences.

III Le produit semi-direct

(a) Définition

Les deux définitions équivalentes du produit semi-direct. Exemples : le groupe symétrique, le groupe diédral.

(b) Groupes d'ordre pq avec p et q premiers

Automorphismes de $\mathbb{Z}/n\mathbb{Z}$. Classification des groupes d'ordre pq.

(c) Groupe d'isométrie d'une partie fini

Définition du groupe des isométries d'une partie finie. Restriction à l'étude des isométries positives à l'aide du produit semi-direct.

Isométrie d'un solide platonicien. Exemples.

Développements

(i) Simplicité de $\mathcal{A}_n$ pour $n \geq 5$

(ii) Théorème de Frobenius-Zolotarev

Références

(i) Perrin, Cours d'algèbre

(ii) Gourdon, Algèbre

(iii) Caldero & Germoni, H2D2

104 : Groupes finis. Exemples et applications.

Rapport du jury

"Dans cette leçon il faut savoir manipuler correctement les éléments de différentes structures usuelles ($\mathbb{Z}/n\mathbb{Z}, \mathfrak{S}_n$, etc.) comme par exemple, en proposer un générateur ou une famille de générateurs, savoir calculer un produit de deux permutations, savoir décomposer une permutation en produit de cycles à supports disjoints. Il est important que la notion d'ordre d'un élément soit mentionnée et comprise dans des cas simples. Le théorème de structure des groupes abéliens finis doit être connu.

Les exemples doivent figurer en bonne place dans cette leçon. Les groupes d'automorphismes fournissent des exemples très naturels dans cette leçon. On peut par exemple étudier les groupes de symétries $\mathfrak{A}_4, \mathfrak{S}_4, \mathfrak{A}_5$ et relier sur ces exemples géométrie et algèbre, les représentations ayant ici toute leur place ; il est utile de connaître les groupes diédraux.

S'ils le désirent, les candidats peuvent ensuite mettre en avant les spécificités de groupes comme le groupe quaternionique, les sous-groupes finis de $\mathrm{SU}(2)$ ou les groupes $\mathrm{GL}_n(\mathbb{F}_q)$."

Introduction

Plan

I Généralités sur les groupes finis

(a) Groupe, ordre

Définition groupe fini. Exemples : $\mathbb{Z}/n\mathbb{Z}$, $\mathfrak{S}_n$. Ordre d'un élément. Groupe cyclique, groupe monogène. Exemple : générateurs de $\mathbb{Z}/n\mathbb{Z}$, ordres des transpositions. Indice d'un sous-groupe, théorème de Lagrange, conséquences. Exemple : groupe alterné.

(b) Sous-groupe distingué, groupe simple

Définitio sous-groupe distingué. Exemple : sous-groupes triviaux. Exemple du noyau d'un morphisme. Exemple de $\mathfrak{A}_n$.

Groupes simples. Exemples : $\mathfrak{A}_3$, $\mathfrak{A}_4$, $\mathbb{Z}/p\mathbb{Z}$.

Centre d'un groupe. Exemple : $\mathfrak{S}_n$, groupe abélien.

Groupe dérivé. Exemple [PER].

II Actions de groupes et applications

(a) Définition et premières propriétés

Définition action de groupes. Action transitive, action fidèle. Exemple : action par translation à gauche et théorème de Cayley. Orbites, stabilisateur, lien entre les orbites et les stabilisateurs. Exemple : action par automophisme intérieur.

(b) Formule des classes et applications

Formule des classes. Applications aux p-groupes. Théorème de la base de Burnside. Définition sous-groupes de Sylow. Théorèmes de Sylow.

(c) Représentation des groupes finis

Définition représentation jusqu'à théorème de Maschke.

III Exemples de groupes finis

(a) Les groupes abéliens

Théorème de structure des groupes abéliens.

(b) Le groupe symétrique

Définition. Décomposition en produit de cycles à support disjoint. Conjugaison. Signature. Théorème de Frobenius-Zolotarev. Simplicité de $\mathfrak{A}_n$. Conséquences.

(c) Le groupe linéaire sur les corps finis

(d) Le groupe des quaternions

Développements

(i) Théorèmes de Sylow
(ii) Simplicité du groupe alterné

Références

(i) Perrin, Cours d'algèbre
(ii) Kosmann-Schwarzbach, Groupes et symétries
(iii) BMP, Objectif Agrégation
(iv) Risler & Boyer, Algèbre pour la licence 3
(v) Caldero & Germoni, H2D2

105 : Groupe des permutations d'un ensemble fini. Applications.

Rapport du jury

"Parmi les attendus, il faut savoir relier la leçon avec les notions d'orbites et d'actions de groupes. Il faut aussi savoir décomposer une permutation en cycles à supports disjoints, tant sur le plan théorique (preuve du théorème de décomposition), que pratique (sur un exemple). Il est important de savoir déterminer les classes de conjugaisons du groupe symétrique par la décomposition en cycles, d'être capable de donner des systèmes de générateurs.

L'existence du morphisme signature est un résultat non trivial mais ne peut pas constituer, à elle seule, l'objet d'un développement.

Les applications sont nombreuses, il est très naturel de parler des déterminants, des polynômes symétriques ou des fonctions symétriques des racines d'un polynôme.

S'ils le désirent, les candidats peuvent aller plus loin en s'intéressant aux automorphismes du groupe symétrique, à des problèmes de dénombrement ou aux représentations des groupes des permutations."

Introduction

Plan

I Définition et premières propriétés

(a) Le groupe symétrique

Définition du groupe symétrique d'un ensemble fini. Cardinal. Cas de l'ensemble $[\![1\,,n]\!]$, isomorphisme. Centre de $\mathfrak{S}_n$.

(b) Décomposition et générateurs

Cycles, transpositions. Support. Décomposition d'une permutation en produit de cycles à support disjoints en utilisant l'action de $\mathfrak{S}_n$ sur $[\![1\,,n]\!]$ [PER]. Ordre d'une permutation. Stabilisateur d'un point.

Parties génératrices de $\mathfrak{S}_n$.

(c) Classes de conjugaison

Action de la conjugaison. Dénombrement des classes de conjugaison. Exemple sur $\mathfrak{S}_4$ [ULM].

II La signature d'une permutation

(a) Signature

Définition signature. Moyens de calculs. Signauture d'une transposition. La signature est un morphisme surjectif dans $\{-1,1\}$ Permutations paires, impaires. Théorème de Frobenius-Zolotarev.

(b) Le groupe alterné

Définition du groupe alterné. Cardinal, sous-groupe distingué. Générateurs. Simplicité et conséquences.

(c) Automorphismes de $\mathfrak{S}_n$

Étude des automorphismes de $\mathfrak{S}_n$

III Applications

(a) Géométrie

Isomorphisme entre les isométries (positives) du tétraèdre et $\mathfrak{S}_4$ ($\mathfrak{A}_4$). Isomorphisme entre les isométries (positives) du cube et $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$ ($\mathfrak{S}_4$).

Sous-groupes finis de $\mathrm{SO}_3(\mathbb{R})$.

(b) Polynômes symétriques

Définition fonctions symétriques élémentaires. Action de $\mathfrak{S}_n$ sur $\mathbb{K}[X_1, \dots, X_m]$. Définition polynôme symétrique. Exemples [Ris]. Théorème des polynômes symétriques. Application : théorème de Kronecker.

Développements

(i) Simplicité du groupe alterné
(ii) Théorème de Frobenius-Zolotarev
(iii) Théorème de Kronecker

Références

(i) Perrin, Cours d'algèbre
(ii) Ulmer, Théorie des groupes
(iii) Caldero & Germoni, H2D2
(iv) Risler & Boyer, Algèbre pour la licence 3

106 : Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $\mathrm{GL}(E)$. Applications.

Rapport du jury

"Cette leçon ne doit pas se résumer à un catalogue de résultats épars sur $\mathrm{GL}(E)$. Il est important de savoir faire correspondre les sous-groupes du groupe linéaire avec les stabilisateurs de certaines actions naturelles (sur des formes quadratiques, symplectiques, sur des drapeaux, sur une décomposition en somme directe, etc.). On doit présenter des systèmes de générateurs, étudier la topologie et préciser pourquoi le choix du corps de base est important. Les liens avec le pivot de Gauss sont à détailler.

Il faut aussi savoir réaliser $\mathfrak{S}_n$ dans $\mathrm{GL}(n, \mathbb{K})$ et faire le lien entre signature et déterminant. S'ils le désirent, les candidats peuvent aller plus loin en remarquant que la théorie des représentations permet d'illustrer l'importance de $\mathrm{GL}_n(\mathbb{C})$ et de son sous-groupe unitaire."

Introduction

Plan

I Propriétés algèbriques du groupe linéaire

(a) Définition et premières applications

Définition de $\mathrm{GL}(E)$. Représentation matricielle. Interprétation en terme de base. Dénombrement de $\mathrm{GL}_n(\mathbb{F}_p)$. Plongement de $\mathfrak{S}_n$ dans $\mathrm{GL}_n(\mathbb{F}_p)$. Application : théorème de Sylow.

Déterminant, définition de $\mathrm{SL}(E)$ (qui est un sous-groupe distingué). Premier dévissage.

(b) Générateurs du groupe linéaire

Définition des dilatations et des transvections. Propriétés. Description de la multiplication à gauche ou à droite par une dilatation ou une transvection. Générateurs de $\mathrm{GL}(E)$, de $\mathrm{SL}(E)$. Pivot de Gauss. Résolution effective de systèmes linéaires.

(c) Sous-groupes dérivé

Détermination du sous-groupe dérivé (cas faciles et cas exceptionnels). Application : théorème de Frobenius-Zolotarev.

Détermination du centre de $\mathrm{GL}(E)$. Application : $\mathrm{GL}_n(\mathbb{R})$ et $GL_n(\mathbb{C})$ ne sont pas isomorphes [FGN].

Théorème de Burnside.

II Le groupe orthogonal euclidien

Définition du groupe orthogonal euclidien, groupe des isométries positives. Exemple : réflexions, renversements, symétries orthogonales. Centre et générateurs de $\mathrm{O}(E)$, de $\mathrm{SO}(E)$.

Quaternions et $\mathrm{SO}_3(\mathbb{R})$.

III Propriétés topologiques du groupe linéaire

(a) Premières propriétés

Choix de la norme. $\mathrm{GL}_n(\mathbb{K})$ est ouvert. L'application inverse est continue et même $\mathcal{C}^\infty$.

$\mathrm{GL}_n(\mathbb{K})$ est dense. Application : AB et BA ont mêmes polynôme caractéristique.

$\mathrm{GL}_n(\mathbb{C})$ est connexe. Application : l'ensemble des projecteurs de rang p de $\mathcal{M}_n(\mathbb{C})$ est connexe. $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe. $\mathrm{SO}_n(\mathbb{R})$ est connexe par arcs.

(b) Décomposition polaire et applications

$‰_n(\mathbb{K})$ est un sous-groupe compact. Décomposition polaire dans $\mathrm{GL}(E)$. $\mathrm{GL}_n(\mathbb{K})^+$ est connexe par arcs. Application : composantes connexes dans formes quadratiques non dégénérées.

Étude des sous-groupes compacts de $\mathrm{GL}(E)$. Enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$.

IV Actions de $\mathrm{GL}(E)$

Action de $\mathrm{GL}(E)$ sur E. Etude.
Action par équivalence. Rang.
Action par similitude.
Action par congruence sur les formes quadratiques, groupe orthogonal général.

Développements

(i) Théorème de Frobenius-Zolotarev
(ii) Quaternions et $\mathrm{SO}_3(\mathbb{R})$

Références

(i) Perrin, Cours d'algèbre
(ii) Mneimné & Testard, Introduction à la théorie des groupes de Lie classiques

108 : Exemples de parties génératrices d'un groupe. Applications.

Rapport du jury

"C'est une leçon qui doit être illustrée par des exemples très variés en relation avec les groupes de permutations et les groupes linéaires ou de leurs sous-groupes. La connaissance de parties génératrices s'avère très utile dans l'analyse des morphismes de groupes ou pour montrer la connexité de certains groupes.

Tout comme dans la leçon 106, la présentation du pivot de Gauss et de ses applications est envisageable."

Introduction

Plan

I Généralités sur les parties génératrices

Définition sous-groupe engendré par une partie [GOU]. Exemple : groupe monogène, groupe de type fini, groupe libre à m éléments. Exemple de $\mathbb{Z}$, de $\mathbb{Z}/n\mathbb{Z}$.

Définition groupe dérivé. Exemples simples.

Restriction de l'étude des propriétés stables par "morphisme" aux éléments d'une partie génératrice.

II Groupes abéliens

(a) Groupes monogènes, groupes cycliques

Définition ordre d'un élément. Propriété de l'ordre. Exemple de $\mathbb{U}_n$. Morphismes de $\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{Z}/m\mathbb{Z}$. Générateurs d'un groupe cyclique. Indicatrice d'Euler. Formule d'Euler-Mobiüs. Application : le groupe multiplicatif d'un corps fini est cyclique, les corps finis sont commutatifs.

Lemme chinois. Application à la résolution d'un système de congruence. Étude des automorphismes de $\mathbb{Z}/n\mathbb{Z}$.

(b) Groupes abéliens de type fini

Théorème de structure des groupes abéliens fini. Application : il existe 6 groupes d'ordre 600. Application : il existe un élément d'ordre le PPCM des ordre. Théorème de structure des groupes abéliens de type fini.

III Groupes non abéliens

(a) Le groupe symétrique

Définition du groupe symétrique, de la signature, du groupe alterné. Décomposition en produit de cycles à support disjoint. Différentes parties génératrices.

Les 3-cycles sont conjugués dans $\mathfrak{A}_n$. Simplicité de $\mathfrak{A}_n$. Application : groupe dérivé de $\mathfrak{S}_n$, sous-groupes distingués de $\mathfrak{S}_n$, sous-groupe d'indice n.

(b) Les p-groupes

Sous-groupe maximal, sous-groupe de Frattini, éléments mous. Théorème de la base de Burnside.

IV Groupes classiques en algèbre linéaire

(a) Le groupe linéaire et le groupe spécial linéaire

Définition groupe linéaire, groupe spécial linéaire. Définition transvection, dilatation. Interprétation en termes d'opérations sur les lignes et les colonnes.

Générateurs de $\mathrm{GL}(E)$, de $\mathrm{SL}(E)$. Pivot de Gauss et application à la résolution effective d'un système linéaire. Application : $\mathrm{GL}_n(\mathbb{R})^+$ est connexe par arcs, $\mathrm{SL}_n(\mathbb{R})$ est connexe par arcs.

Sous-groupe dérivé de $\mathrm{GL}_n(\mathbb{R})$. Application : théorème de Frobenius-Zolotarev.

(b) Le groupe orthogonal

Définition du groupe orthogonal. Définition symétrie orthogonale, renversement, reflexions. Théorème de Cartan-Dieudonné. Application : classification des isométries vectorielles en dimension 2 ou 3.

Développements

(i) Quaternions et $SO_3(\mathbb{R})$

(ii) Simplicité de $\mathfrak{A}_n$

Références

(i) Perrin, Cours d'algèbre

(ii) Gourdon, Algèbre

(iii) Zavidovique, Un max de maths

(iv) BMP, Objectif Agrégation

120 : Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.

Rapport du jury

"Dans cette leçon, l'entier n n'est pas forcément un nombre premier. Il serait bon de connaître les idéaux de $\mathbb{Z}/n\mathbb{Z}$ et, plus généralement, les morphismes de groupes de $\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{Z}/m\mathbb{Z}$.

Il est nécessaire de bien maîtriser le lemme chinois et sa réciproque. S'ils le désirent, les candidats peuvent poursuivre en donnant une généralisation du lemme chinois lorsque deux éléments ne sont pas premiers entre eux, ceci en faisant apparaître le pgcd et le ppcm de ces éléments.

Il faut bien sûr savoir appliquer le lemme chinois à l'étude du groupe des inversibles, et ainsi, retrouver la multiplicativité de l'indicatrice d'Euler. Toujours dans le cadre du lemme chinois, il est bon de distinguer clairement les propriétés de groupes additifs et d'anneaux, de connaître les automorphismes, les nilpotents et les idempotents.

Enfin, il est indispensable de présenter quelques applications arithmétiques des propriétés des anneaux $\mathbb{Z}/n\mathbb{Z}$., telles que l'étude de quelques équations diophantiennes bien choisies. De même, les applications cryptographiques telles que l'algorithme RSA sont naturelles dans cette leçon.

S'ils le désirent, les candidats peuvent aller plus loin en s'intéressant au calcul effectif des racines carrées dans $\mathbb{Z}/n\mathbb{Z}$."

Introduction

Plan

I Le groupe $\mathbb{Z}/n\mathbb{Z}$

$(\mathbb{Z}, +)$ est un groupe abélien. Les sous-groupes de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$. Groupe quotient $\mathbb{Z}/n\mathbb{Z}$. Propriétés (addition, cyclicité). Classification des groupes cycliques. Ordre, générateurs de $\mathbb{Z}/n\mathbb{Z}$.

Détermination des morphismes de groupes de $\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{Z}/m\mathbb{Z}$.

II L'anneau $\mathbb{Z}/n\mathbb{Z}$

(a) Structure d'anneau et lemme chinois

$(\mathbb{Z}, +, \times)$ est un anneau commutatif. Idéaux de $\mathbb{Z}$. Anneau quotient $\mathbb{Z}/n\mathbb{Z}$. Propriétés (multiplication). Théorème chinois. Exemple d'utilisation en arithmétique.

(b) Groupe des inversibles, indicatrice d'Euler

Inversibles. CNS pour que $\mathbb{Z}/n\mathbb{Z}$ soit un corps, soit intègre. Indicatrice d'Euler. Propriétés (fonction multiplicative, formule d'Euler-Möbius). Formule d'Euler : $a^{\varphi(n)} \equiv 1 \pmod n$ (puis théorème de Fermat). Le théorème de Fermat ne caractérise pas les nombres premiers (problème 2 [GOU] sur les nombres de Carmichael).

(c) Automorphismes de $\mathbb{Z}/n\mathbb{Z}$

[Per]

III Applications

(a) La loi de réciprocité quadratique

Symbole de Legendre. Carrés de $\mathbb{F}_p^\times$. Propriétés du symbole de Legendre (multiplicativité, calcul). Loi de réciprocité quadratique. Exemple.

(b) Le théorème des deux carrés

Anneau des entiers de Gauss. Norme (et multiplicativité de cette norme). Inversibles. Structure euclidienne. CNS pour qu'un nombre premier soit somme de deux carrés. Cas général.

(c) Le cryptosystème RSA

Problème 1 [GOU]. Exemple.

Développements

(i) Loi de réciprocité quadratique (via les formes quadratiques)
(ii) Théorème des deux carrés

Références

(i) Perrin, Cours d'algèbre
(ii) Risler & Boyer, Algèbre pour la licence 3
(iii) Gourdon, Algèbre
(iv) Caldero & Germoni, Histoires hédonistes de groupes et de géométries

121 : Nombres premiers. Applications.

Rapport du jury

"Le sujet de cette leçon est très vaste. Aussi les choix devront être clairement motivés. La réduction modulo p n'est pas hors-sujet et constitue un outil puissant pour résoudre des problèmes arithmétiques simples. La répartition des nombres premiers est un résultat historique important qu'il faudrait citer. Sa démonstration n'est bien sûr pas exigible au niveau de l'agrégation.

Quelques résultats sur les corps finis et leur géométrie sont les bienvenus, ainsi que des applications en cryptographie."

Introduction

Plan

I Généralités sur les nombres premiers

(a) Définition et premières propriétés

Définition d'un nombre premier. Exemples : 2, 3, 5, 7, 97. Théorème fondamental de l'arithmétique. Exemple : $3920 = 2^4 \times 5 \times 7^2$. Exemples : nombres de Mersenne, nombres de Fermat. Test de primalité des nombres de Mersenne.

Si p ne divise pas en entier a alors a et p sont premiers entre eux. Application : $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est premier. Théorème de Fermat. Nombres pseudo-premiers et nombres de Carmichael.

Théorème de Wilson.

(b) Répartition

L'ensemble des nombres premiers est infini. Crible d'Ératosthène. Quelques propriétés de la fonction ζ de Riemann. Théorème des nombres premiers. Théorème de Dirichlet.

II Applications aux groupes finis

(a) Formules des classes, applications aux p-groupes

Formule des classes. Le centre d'un p-groupe est non trivial. Cardinaux des sous-groupes d'un p-groupe. Les sous-groupes maximaux d'un p-groupe sont d'indice p et sont distingués. Théorème de la base de Burnside.

(b) Théorèmes de Sylow

Définition d'un p-Sylow. Théorèmes de Sylow, conséquences. Application : groupes d'ordre pq pour p et q premiers.

III Les corps finis $\mathbb{F}_q$

(a) Généralités

Caractéristique d'un corps. Cardinal d'un corps fini. Morphisme de Frobenius (avec le lemme). Existence et unicité des corps finis.

(b) Polynômes irréductibles

Quelques critères d'irréductibilité (réduction sur $\mathbb{F}_p$, critère d'Eisenstein). Exemples.

Fonction de Mobius. Propriétés. Dénombrement des polynômes irréductibles de $\mathbb{F}_q[X]$.

(c) Carrés dans $\mathbb{F}_q$

Étude des carrés dans $\mathbb{F}_q$. CNS pour que -1, pour 2 soient des carrés dans $\mathbb{F}_q$. Application : théorème des deux carrés.

Symbole de Legendre. Loi de réciprocité quadratique. Application : théorème de Frobenius-Zolotarev, ou à des tests de primalité [GOU].

IV Applications

(a) Le cryptosystème RSA

(b) Construction à la règle et au compas

Définition constructibilité. Exemples. Théorème de Wantzel. Théorème de Gauss-Wantzel.

Développements

(i) Théorèmes de Sylow

(ii) Loi de réciprocité quadratique

Références

(i) Perrin, Cours d'algèbre

(ii) Gourdon, Algèbre et Analyse

122 : Anneaux principaux. Applications.

Rapport du jury

"Cette leçon n'est pas uniquement théorique, Il est possible de présenter des exemples d'anneaux principaux classiques autres que $\mathbb{Z}$ et $\mathbb{K}[X]$ (décimaux, entiers de Gauss ou d'Eisenstein), accompagnés d'une description de leurs irréductibles. Les applications en algèbre linéaire ne manquent pas et doivent être mentionnées. Par exemple, les notions de polynôme minimal sont très naturelles parmi les applications. Les anneaux euclidiens représentent une classe d'anneaux principaux importante et l'algorithme d'Euclide a toute sa place dans cette leçon pour effectuer des calculs.

S'ils le désirent, les candidats peuvent aller plus loin en s'intéressant à l'étude des réseaux, à des exemples d'anneaux non principaux, mais aussi à des exemples d'équations diophantiennes résolues à l'aide d'anneaux principaux. À ce sujet, il sera fondamental de savoir déterminer les unités d'un anneau, et leur rôle au moment de la décomposition en facteurs premiers. De même, le calcul effectif des facteurs invariants de matrices à coefficients dans certains anneaux peut être fait."

Introduction

Plan

I Généralités sur les anneaux principaux

(a) Arithmétique dans un anneau

(b) Anneaux factoriels

(c) Anneaux principaux

II Application

(a) Polynôme minimal

(b) Corps de rupture

(c) Les entiers de Gauss et le théorème des deux carrés

Développements

(i) Théorème des deux carrés

(ii) Existence d'un polynôme minimal local qui soit égal au polynôme minimal

Références

(i) Perrin, Cours d'algèbre

123 : Corps finis. Applications.

Rapport du jury

“Une construction des corps finis doit être connue et une bonne maîtrise des calculs dans les corps finis est indispensable. Les injections des divers $\mathbb{F}_q$ doivent être connues et les applications des corps finis (y compris pour $\mathbb{F}_q$ avec q non premier !) ne doivent pas être oubliées : citons par exemple l'étude de polynômes à coefficients entiers et de leur irréductibilité. Le calcul des degrés des extensions et le théorème de la base téléscopique sont incontournables. L'étude des carrés dans un corps fini et la résolution d'équations de degré 2 sont envisageables.

S'ils le désirent, les candidats peuvent aller plus loin en détaillant des codes correcteurs.”

Introduction

Plan

I Généralités sur les corps finis

(a) Définition des corps finis

Le corps $\mathbb{F}_p$ pour p premier. Sous-corps premier, définition et propriétés. Caractéristique. Morphisme de Frobenius.

Existence et unicité de $\mathbb{F}_q$.

(b) Structure des corps finis

Le groupe multiplicatif $\mathbb{F}_q^\times$ est cyclique et isomorphe à $\mathbb{Z}/(q-1)\mathbb{Z}$. Théorème de l'élément primitif sur $\mathbb{F}_q$. Tout sous-groupe du groupe multiplicatif d'un corps fini est cyclique. Tout corps fini est commutatif.

Étude des sous-corps de $\mathbb{F}_q$ [Ris].

II Polynômes dans $\mathbb{F}_q$

(a) Polynômes irréductibles sur un corps fini

Critère de réduction. Exemple d'application. CNS d'irréductibilité [PER]. Exemple. X^4+1 est irréductible sur $\mathbb{Z}$ mais pas dans $\mathbb{F}_p$ pour tout p. Critère d'irréductibilité par passage à l'extension. Exemple.

Polynômes cyclotomiques. Irréductibilité. Comportement de Φ_n sur $\mathbb{F}_q$.

Dénombrement des polynômes irréductibles de $\mathbb{F}_q$.

(b) Étude des automorphismes

Étude du groupe de Galois de $\mathbb{F}_q$ sur $\mathbb{F}_p$ [Ris].

(c) Polynômes à plusieurs variables

Classification des formes quadratiques sur les corps finis.

Théorème de Chevalley-Warning. Application : théorème de Ginszbourg-Erdös-Sziv.

III Les carrés dans $\mathbb{F}_q$

(a) Étude des carrés dans $\mathbb{F}_q$

Motivation : la classification des formes quadratiques sur $\mathbb{F}_q$.

Étude des carrées de $\mathbb{F}_q$. Il existe une infinité de nombres premiers de la forme $4m+1$. Théorème des deux carrés.

(b) Le symbole de Legendre et ses propriétés

Définition du symbole de Legendre. Propriétés (multiplicativité, calcul pratique, etc). Loi de réciprocité quadratique. Applications.

Le symbole de Legendre est le seul morphisme non trivial de $\mathbb{F}_p^\times$ dans $\{-1, 1\}$. Application : théorème de Frobenius-Zolotarev.

Développements

(i) Dénombrement des polynômes irréductibles de $\mathbb{F}_q[X]$

(ii) Loi de réciprocité quadratique

Références

(i) Perrin, Cours d'algèbre

(ii) Risler & Boyer, Algèbre pour la licence 3

(iii) BMP, Objectif Agrégation

125 : Extensions de corps. Exemples et applications.

Rapport du jury

"Le théorème de la base téléscopique et ses applications à l'irréductibilité de certains polynômes, ainsi que les corps finis sont incontournables. De même il faut savoir calculer le polynôme minimal d'un élément algébrique dans des cas simples, notamment pour quelques racines de l'unité. La leçon peut être illustrée par des exemples d'extensions quadratiques et leurs applications en arithmétique, ainsi que par des extensions cyclotomiques.

S'ils le désirent, les candidats peuvent s'aventurer en théorie de Galois."

Introduction

Plan

I Généralités sur les extensions de corps

(a) Extension de corps, degré

Définition extension de corps. Exemples : $\mathbb{R} \hookrightarrow \mathbb{C}$, $\mathbb{R} \hookrightarrow \mathbb{R}(X)$. Degré d'une extension, extension finie, cas des corps finis. Théorème de la base télescopique, multiplicativité du degré. Extension engendrée par une partie, extension de type fini, extension monogène.

(b) Éléments algèbrique, extension algèbrique

Éléments algèbriques, transcendants. Polynôme minimal. Exemples. Propriétés. Extension algébrique. L'ensemble des éléments algébriques est un sous-corps algébrique. Corps algébriquement clos. Exemples : $\mathbb{C}$, clôture de $\mathbb{Q}$.

II Adjonction de racines

(a) Corps de rupture, corps de décomposition

Existence et unicité du corps de rupture. Exemple explicite d'un corps de rupture. Conservation de l'irréductibilité par extension de corps (lorsque le degré de l'extension est premier avec le degré du polynôme).

Existence et unicité du corps de décomposition. Exemples. Application : polynômes cyclotomiques.

(b) Clôture algèbrique

Clôture algébrique. Exemple de $\mathbb{C}$, de $\overline{\mathbb{Q}}$.

III Les corps finis

(a) Caractéristique et cardinalité

Caractéristique d'un corps, sous-corps premier, morphisme de Frobenius. Cardinal d'un corps fini. Existence et unicité des corps finis.

Le groupe multiplicatif d'un corps fini et cyclique. Théorème de Wedderburn.

Inclusion des corps finis.

(b) Polynômes irréductibles sur un corps fini

Critère d'irréductibilité par réduction et application pour les corps finis. Contre-exemple de la réciproque.

Étude du groupe de Galois de $\mathbb{F}_q$ sur $\mathbb{F}_p$.

Fonction de Mobiüs. Propriétés. Dénombrement des polynômes irréductibles sur un corps fini. Existence.

IV Application : construction à la règle et au compas

Constructibilité à la règle et au compas. Exemples de points constructibles. Théorème de Wantzel. Impossibilité de la trissection de l'angle, de la quadrature du cercle, de la duplication du cube.

Constructibilité des racines de l'unité, des polygones réguliers. Théorème de Gauss-Wantzel.

Développements

(i) Théorème de Gauss-Wantzel

(ii) Dénombrement des polynômes irréductibles sur un corps fini

Références

(i) Perrin, Cours d'algèbre

(ii) Risler & Boyer, Algèbre pour la licence 3

141 : Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.

Rapport du jury

“La présentation du bagage théorique permettant de définir corps de rupture, corps de décomposition, ainsi que des illustrations dans différents types de corps (réel, rationnel, corps finis) sont inévitables. Les corps finis peuvent être illustrés par des exemples de polynômes irréductibles de degré 2, 3, 4 sur $\mathbb{F}_2$ ou $\mathbb{F}_3$. Il est nécessaire de présenter des critères d’irréductibilité de polynômes et des polynômes minimaux de quelques nombres algébriques.

Il faut savoir qu’il existe des corps algébriquement clos de caractéristique nulle autres que $\mathbb{C}$; il est bon de savoir montrer que l’ensemble des nombres algébriques sur le corps $\mathbb{Q}$ des rationnels est un corps algébriquement clos. Le théorème de la base téléscopique, ainsi que les utilisations arithmétiques (utilisation de la divisibilité) que l’on peut en faire dans l’étude de l’irréductibilité des polynômes, est incontournable.”

Introduction

Plan

I Polynômes irréductibles

(a) Irréductibles de $A[X]$, A anneau intègre

Définition polynôme irréductible. Définition du contenu, éléments primitifs. Lemme de Gauss.

(b) Irréductibles de $K[X]$, K corps

Les remarques préliminaires de Perrin. Algorithme de Berlekamp.

(c) Cas où $K = \mathrm{Frac}(A)$, A anneau factoriel

Irréductibles de $A[X]$. Critère d’Eisenstein. Critère de réduction. Critères d’irréductibilité par passage à l’extension. Exemples.

II Adjonction de racines

(a) Extension algèbrique

Théorème de la base télescopique.

(b) Corps de rupture, corps de décomposition

(c) Corps algébriquement clos

III Applications

(a) Polynômes cyclotomiques

(b) Polynômes irréductibles de $\mathbb{F}_q[X]$

(c) Construction à la règle et au compas

Développements

(i) Dénombrement des polynômes irréductibles d’un corps fini

(ii) Théorème de Gauss-Wantzel

Références

(i) Perrin, Cours d'algèbre
(ii) Demazure, Cours d'algèbre

150 : Exemples d'actions de groupes sur les espaces de matrices.

Rapport du jury

"Dans cette leçon il faut présenter différentes actions (congruence, similitude, équivalence, ...) et dans chaque cas on pourra dégager d'une part des invariants (rang, matrices échelonnées réduites, ...), d'autre part des algorithmes comme le pivot de Gauss. On peut aussi, si l'on veut aborder un aspect plus théorique, faire apparaître à travers ces actions quelques décompositions célèbres ; on peut décrire les orbites lorsque la topologie s'y prête.

S'ils le désirent, les candidats peuvent travailler sur des corps finis et utiliser le dénombrement dans ce contexte."

Introduction

Plan

I Action par équivalence

(a) Définition et premières propriétés

Définition de la relation d'équivalence. Lien avec les endomorphismes (deux matrices équivalentes représentent le même endomorphisme dans des bases différentes). Conservation du rang.

(b) Opérations élémentaires

Définition des matrices de transvection, de dilatation. Explication de leur action. Application : trouver le rang d'une matrice à l'aide d'opérations élémentaire sur les lignes ou les colonnes. Pivot de Gauss.

Description des orbites de l'action par le rang. Application : A et ${}^tA$ ont même rang. Application : les matrices de dilatation et de transvection engendrent $\mathrm{GL}_n(\mathbb{R})$, les matrices de dilatation engendrent $\mathrm{SL}_n(\mathbb{R})$. Application : $\mathrm{GL}_n(\mathbb{R})^+$ est connexe par arcs.

II Action par similitude

(a) Définition et premières propriétés

Définition de l'action. Lien avec les endomorphismes. Quelques invariants (rang, déterminant, trace, polynôme caractéristique). Application à la définition du polynôme caractéristique d'un endomorphisme. Cas des matrices de taille 2 dans $\mathbb{C}$. Ne caractérisent pas les orbites dans le cas général (contre exemples).

(b) Invariants de similitude et réduction de Frobenius

Espace cycliques, endomorphismes cycliques, matrices compagnon. Existence et unicité des invariants de similitude. Réduction de Frobenius. Application : réduction de Jordan. Application : $A \sim {}^tA$. Application : extension de corps, cas de la dimension 2 et de la dimension 3 [BMP].

(c) Cas de l'action de $\mathrm{O}_n(\mathbb{R})$

Conservation des matrices symétriques. Théorème spectral, de diagonalisation pseudo simultanée. Application : log convexité du déterminant, l'exponentielle induit un homéomorphisme de $\mathcal{S}_n$ sur $\mathcal{S}_n^{++}$. Application : matrice carrée, décomposition polaire.

III Action par congruence

(a) Définition et premières propriétés

Définition de l'action [Seg]. Lien avec les formes quadratiques (changement de base). Premières propriétés (conservation du rang, conservation du déterminant modulo un carré, conservation des matrices symétriques et antisymétriques).

Exemple des matrices diagonales (multiplication des coefficients par un carré, permutation des coefficients) [Seg].

Comparaison de l'action par congruence et de l'action similitude.

(b) Réduction de Gauss

Opérations symétriques élémentaires (en particulier, on reste dans la même orbite par opération symétrique élémentaire). Réduction de Gauss dans le cas général. Exemples. Réduction de Gauss des matrices symétriques sur $\mathbb{K}$ avec $\mathrm{car}(\mathbb{K}) \neq 2$ et $\mathbb{C}$.

Développements

(i) Existence et unicité des invariants de Frobenius

(ii) Loi de réciprocité quadratique

Références

(i) Gourdon, Algèbre

(ii) Perrin, Cours d'algèbre

(iii) Seguins Pazzis, Invitation aux formes quadratiques

(iv) Caldero & Germoni, H2D2

(v) BMP, Objectif Agrégation

151 : Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.

Rapport du jury

"Dans cette leçon, il est important de présenter les résultats fondateurs de la théorie des espaces vectoriels de dimension finie en ayant une idée de leurs preuves. Ces théorèmes semblent simples car ils ont été très souvent pratiqués, mais leur preuve demande un soin particulier. Il est important de savoir justifier pourquoi un sous-espace vectoriel d'un espace vectoriel de dimension finie est aussi de dimension finie. Les diverses notions et caractérisations du rang trouvent leur place dans cette leçon. Les applications sont nombreuses, on peut par exemple évoquer l'existence de polynômes annulateurs ou alors décomposer les isométries en produits de réflexions.

S'ils le désirent, les candidats peuvent déterminer des degrés d'extensions dans la théorie des corps ou s'intéresser aux nombres algébriques."

Introduction

Plan

I Théorie de la dimension

(a) Familles libres et familles génératrices

Définition de l'application linéaire associée à une famille de vecteurs. Définition famille génératrice, famille libre. Exemple d'une droite vectorielle. Exemple de $\mathbb{K}[X]$.

Propriétés des familles libres et des familles génératrices [RDO]. Image par une application linéaire.

(b) Dimension d'un espace vectoriel

Définition base. Exemple sur $\mathbb{K}[X]$. Théorème de la base incomplète. Espaces vectoriels de dimension finie. Unicité de la dimension [RDO].

Propriétés de la dimension (deux espaces vectoriels sont isomorphes si et seulement s'ils ont la même dimension, dimension d'une produit d'espaces vectoriels, dimension d'une somme directe, dimension de l'espace des applications linéaires).

Dimension d'un sous-espace vectoriel. Codimension d'un sous espace vectoriel.

(c) Application en théorie des groupes

Sous-groupe de Frattini, éléments mous. Théorème de la base de Burnside.

II Rang d'une application linéaire

(a) Définition et premières propriétés

Rang d'une famille de vecteurs, rang d'une application linéaire. Théorème du rang. Formule de Grassman. Application : caractérisation d'une somme directe.

Caractérisation des isomorphismes. Propriétés du rang par composition à gauche ou à droite. Application : polynômes interpolateurs de Lagrange.

(b) Action par équivalence

Définition de l'action par équivalence. Rang d'une matrice. Conservation du rang. Le rang caractérise les orbites. Application : le rang d'une matrice est égal au rang de sa transposée. Application : densité des matrices inversibles de $\mathrm{GL}_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Application : si $\operatorname{car} \mathbb{K} \neq 2$ alors toute matrice est somme de deux matrices inversibles.

(c) Calcul du rang

Le rang est indépendant vis-à-vis du corps de base.
Semi-continuité inférieure du rang. L'ensemble des matrices de rang inférieur à r est un fermé.
Définition matrices de transvection, de dilatation. Pivot de Gauss, détermination effective du rang.

III Applications en théorie des corps

(a) Degré d'une extension de corps

Définition extension de corps. Théorème de la base téléscopique. Multiplicativité du degré. Exemple : $\mathbb{C}$ sur $\mathbb{R}$.

Définition extension monogène. Éléments algébriques, transcendants. Exemple : clôture algébrique de $\mathbb{Q}$ n'est pas de dimension finie.

(b) Construction à la règle et au compas

Définition points constructibles. Théorème de Wantzel. Définition polygone régulier constructible. Théorème de Gauss-Wantzel.

Développements

(i) Théorème de la base de Burnside
(ii) Théorème de Gauss-Wantzel

Références

(i) Gourdon, Algèbre
(ii) Perrin, Cours d'algèbre
(iii) BMP, Objectif Agrégation
(iv) Ramis, Deschamps & Odoux, Algèbre 1

152 : Déterminant. Exemples et applications.

Rapport du jury

"Dans cette leçon, il faut commencer par définir correctement le déterminant. Il est possible d'entamer la leçon en disant que le sous-espace des formes n-linéaires alternées sur un espace de dimension n est de dimension 1, toutefois, il est essentiel de savoir le montrer. Le plan doit être cohérent ; si le déterminant n'est défini que sur $\mathbb{R}$ ou $\mathbb{C}$, il est délicat de définir $\det(A - XI_n)$ avec A une matrice carrée. L'interprétation du déterminant comme volume est essentielle.

Le calcul explicite est important, mais le jury ne peut se contenter que d'un déterminant de Vandermonde ou d'un déterminant circulant. Les opérations élémentaires permettant de calculer des déterminants, avec des illustrations sur des exemples, doivent être présentées. Il serait bien que la continuité du déterminant trouve une application, ainsi que son caractère polynomial. Pour les utilisations des propriétés topologiques, on n'ommetra pas de préciser le corps de base sur lequel on se place.

S'ils le désirent, les candidats peuvent s'intéresser aux calculs de déterminant sur $\mathbb{Z}$ avec des méthodes multimodulaires ; de plus, le résultant et les applications simples à l'intersection ensembliste de deux courbes algébriques planes peuvent trouver leur place dans cette leçon."

Introduction

Plan

I Généralités sur le déterminant

(a) Formes multilinéaires, formes alternées

Définition formes multilinéaires, formes alternées, formes antisymétriques (équivalence en caractéristique différente de 2). Exemple. Antisymétrisation d'une forme linéaire.

(b) Déterminant d'un endomorphisme

La dimension des formes n-linéaires alternées sur un espace vectoriel de dimension finie est 1. Définition du déterminant d'une famille de vecteurs dans une base. Formule explicite. Déterminant d'un endomorphisme. Propriétés du déterminant (caractérisation d'une famille liée, morphisme, etc).

(c) Déterminant d'une matrice

Déterminant d'une matrice à coefficients dans un anneau. Formule explicite. Le déterminant est un polynôme. Application : deux matrices réelles semblables sur $\mathbb{C}$ le sont aussi sur $\mathbb{R}$. Propriétés. Mineurs et cofacteurs. Application : caractérisation de Sylvester.

II Exemples de calcul de déterminants

(a) Opérations élémentaires

Développement selon une ligne ou une colonne. Comatrice. Application : formule d'inversion d'une matrice et critère pour qu'une matrice soit inversible dans un anneau (exemple de $\mathbb{Z}$, ce critère dépend de l'anneau, exemple de matrice dans $\mathbb{Z}$ inversible dans $\mathbb{Q}$ mais pas dans $\mathbb{Z}$). Application : continuité de l'inverse. Inversion des matrices de taille 2×2, voire 3×3.

Formule du déterminant par blocs.

Matrices de transvection, de dilatation, de permutation. Décomposition LU.

(b) Quelques déterminants classiques et applications

Déterminant de Vandermonde. Application : interpolation de Lagrange.

Déterminant de Cauchy. Application : théorème de Muntz.

Déterminants circulant.

Déterminant d'une matrice compagnon.

III Interprétation géométrique du déterminant

Volume d'une partie mesurable, d'un parallélépipède. Formule du changement de variables.
Inégalité d'Hadamard. Déterminant de Gram et applications. Volume d'un ellipsoïde.
Log-concavité du déterminant. Ellipsoïde de John-Loewner. Application aux sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.

IV Applications

(a) Topologie

$\mathrm{GL}_n(\mathbb{K})$ est ouvert ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs, ce n'est pas le cas de $\mathrm{GL}_n(\mathbb{R})$.

(b) Algèbre

Polynôme caractéristique et quelques propriétés.
Le déterminant est un morphisme de groupe. Application : théorème de Frobenius-Zolotarev.
Résultant et quelques propriétés.

Développements

(i) Théorème de Frobenius-Zolotarev
(ii) Ellipsoïde de John-Loewner

Références

(i) BMP, Objectif Agrégation
(ii) Gourdon, Algèbre
(iii) FGN, Algèbre 2 et 3
(iv)

153 : Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.

Rapport du jury

"Cette leçon ne doit pas être un catalogue de résultats autour de la réduction qui est ici un moyen pour démontrer des théorèmes ; les polynômes d'endomorphismes doivent y occuper une place importante. Il faut consacrer une courte partie de la leçon à l'algèbre $\mathbb{K}[u]$, connaître sa dimension sans hésiter ; s'ils le désirent, les candidats peuvent ensuite s'intéresser à ses propriétés globales.

Les liens entre réduction d'un endomorphisme u et la structure de l'algèbre $\mathbb{K}[u]$ sont importants, tout comme ceux entre les idempotents et la décomposition en somme de sous-espaces caractéristiques. Il faut bien préciser que, dans la réduction de Dunford, les composantes sont des polynômes en l'endomorphisme, et en connaître les conséquences théoriques et pratiques.

L'aspect *applications* est trop souvent négligé. On attend d'un candidat qu'il soit en mesure, pour une matrice simple de justifier la diagonalisabilité et de déterminer un polynôme annulateur (voire minimal). Il est souhaitable que les candidats ne fassent pas la confusion entre diverses notions de multiplicité pour une valeur propre λ donnée (algébrique ou géométrique). Enfin, rappelons que pour calculer A^k , il n'est pas nécessaire en général de réduire A (la donnée d'un polynôme annulateur de A suffit souvent)."

Introduction

Définir ce qu'est la réduction du point de vu des actions de groupes. Donner l'objectif de la réduction, à savoir la décomposition de l'espace en sous-espaces stables où l'endomorphisme considéré est "simple". Faire le lien avec le lemme des noyaux.

Plan

Soit $\mathbb{K}$ un corps (commutatif) et E un $\mathbb{K}$-espace vectoriel de dimension finie n. Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

I Polynômes d'endomorphisme

(a) La sous-algèbre $\mathbb{K}[f]$

Définition de l'action de $\mathbb{K}[X]$ sur f. Définition de l'application $\mathrm{ev}_f : P \in \mathbb{K}[X] \longmapsto P(f) \in \mathcal{L}(E)$. Définition locale en x. Définition de la sous-algèbre $\mathbb{K}[f]$, de $E_{f,x}$. Identification de $\mathcal{L}(E)$ et de $\mathcal{M}_n(\mathbb{K})$, transposition des définitions au cas des matrices. Exemple sur les matrices triangulaires supérieures. Commutativité de $\mathbb{K}[X]$. Stabilité par f de $\operatorname{Ker} P(f)$ et de $\operatorname{Im} P(f)$.

(b) Idéal annulateur et polynôme minimal

Étude du noyau de ev_f, définition du polynôme minimal de f, et du polynôme local de f en x. Exemple : projecteurs et symétries. Définitions endomorphismes nilpotents, indice de nilpotence. Sous-espaces stables et polynôme minimal (ppcm). Caractérisation des inversibles par le polynôme minimal. Application au calcul de l'inverse, des puissances de f. Isomorphisme induit par le passage au quotient de ev_f, dimension de $\mathbb{K}[f]$. Base de $\mathbb{K}[f]$. Base de $E_{f,x}$. Application : construction matricielle de $\mathbb{C}$.

(c) Le lemme des noyaux

Lemme des noyaux. Les projecteurs sont des polynômes en f. Noyaux itérés. Exemples : projecteurs et symétries. Application : existence de $x \in E$ tel que $\mu_{f,x} = \mu_f$.

II Espaces propres, diagonalisibilité et trigonalisibilité

(a) Spectre et espaces propres

Définition valeurs propres, vecteurs propres, espaces propres, spectre. Les espaces propres sont en somme directe. Polynômes et valeurs propres. Caractérisation des valeurs propres par les racines du polynôme minimal. Exemple sur $\mathbb{C}$. Commutation et espaces propres.

(b) Le polynôme caractéristique de f

Définition du polynôme caractéristique pour une matrice puis pour un endomorphisme. Degré. Sous-espaces stables et polynôme caractéristique. Exemple : matrices triangulaires par blocs, endomorphisme de rang 2. Caractérisation des valeurs propres par les racines du polynôme caractéristique. Théorème de Cayley-Hamilton. Conséquences.

(c) Critères polynomiaux de diagonalisibilité et de trigonalisibilité, applications

Définition endomorphisme diagonalisable. Critère via les polynômes annulateurs (en particulier χ_f). Application : endomorphisme induit par un sev stable. Exemple : endomorphisme de rang 1 et trace. Application aux sous-groupes finis de $\mathrm{GL}_n(\mathbb{C})$. Critère via le polynôme caractéristique. Définition trigonalisibilité. Exemple endomorphismes nilpotents. Critère de trigonalisibilité. Exemple sur $\mathbb{C}$. Décomposition de Dunford sur $\mathbb{C}$. Théorème de co-diagonalisation, de co-trigonalisibilité. Application : diagonalisibilité par blocs [FGN, p.103].

III Réduction d'endomorphismes

(a) Espaces caractéristiques et décomposition de Dunford

Définition sous-espaces caractéristiques. Théorème de Décomposition de Dunford. Application exponentielle d'endomorphismes, limite de e^{tM}.

(b) Espaces cycliques et réduction de Frobenius

Définitions espaces f-cycliques, endomorphismes cycliques. Définition matrices compagnons, propriétés. Théorème des invariants de similitudes. Réduction de Frobenius. Application à la décomposition de Jordan, application : ${}^tA \sim A$.

Développements

(i) Décomposition de Dunford
(ii) Existence et unicité des invariants de similitudes

Références

(i) Gourdon, Algèbre
(ii) Mneimné et Mansuy, Réduction des endomorphismes
(iii) Caldero et Germoni, H2G2, tome I
(iv) FGN, Algèbre 2

154 : Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.

Rapport du jury

"Dans cette leçon, il faut présenter des propriétés de l'ensemble des sous-espaces stables par un endomorphisme. Des études détaillées sont les bienvenues, par exemple le cas d'une matrice diagonalisable ou le cas d'une matrice nilpotente d'indice maximum. La décomposition de Frobenius trouve tout à fait sa place dans cette leçon. Il ne faut pas oublier d'examiner le cas des sous-espaces stables par des familles d'endomorphismes. Ceci peut déboucher par exemple sur des endomorphismes commutant entre eux ou sur la théorie des représentations."

Introduction

Insister sur l'endomorphisme induit par un sev stable. Principe de la réduction : exemple du lemme des noyaux. Faire le lien entre commutation et sev stables.

Plan

Soit $\mathbb{K}$ un corps (commutatif) et E un $\mathbb{K}$-espace vectoriel de dimension finie n. Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

I Sous-espaces stables par un endomorphisme

(a) Définitions et premières propriétés

Définition d'un sev f-stable. Endomorphisme induit. Exemple : $\operatorname{Ker} P(f)$ est stable pour tout $P \in \mathbb{K}[X]$. Si F est f-stable et g-stable alors F est stable par $f+g$ et $f \circ g$. Conjugaison et stabilité. Application à une rotation. Représentation matricielle de la stabilité. Sous-espaces stables et polynôme minimal, polynôme caractéristique.

(b) Exemples de sous-espaces stables

Exemples projecteurs et symétries. Les homothéties sont les seuls endomorphismes stabilisant tous les sev. Sous-espaces stables d'un endomorphismes nilpotents. Polynôme minimal local, sous-espaces cyclique de f en x, base de $E_{f,x}$.

(c) Le lemme des noyaux

Lemme des noyaux. Les projecteurs sont des polynômes en f. Noyaux itérés. Exemples : projecteurs et symétries. Application : existence de $x \in E$ tel que $\mu_{f,x} = \mu_f$. Sous-espaces caractéristiques. Projecteurs caractéristiques. Application : Décomposition de Dunford (existence).

II Stabilité et commutativité

(a) Commutation

Définition commutation, exemples $\operatorname{Ker} P(f)$ et $\operatorname{Im} P(f)$. Le centre de $\mathcal{L}(E)$ est l'ensemble des homothéties.

(b) Réduction simultanée

Théorèmes de co-diagonalisation et de co-trigonalisation. Application : diagonalisibilité par blocs [FGN, p.103]. Application : Décomposition de Dunford (unicité).

(c) Endomorphismes normaux

On suppose E hermitien ou euclidien. Définition endomorphismes normaux. Exemples : endomorphismes auto-adjoints, antisymétriques, isométries ou unitaires. Si F est u-stable alors $F^\perp$ est u^*-stable. Si u est normal et F u-stable, alors $F^\perp$ est u-stable.

[Peut être enlevé] On suppose E hermitien. Caractérisation des endomorphismes normaux (version endomorphisme et version matricielle).

On suppose E euclidien. Réduction des endomorphismes normaux dans $\mathbb{R}$. Application aux isométries. Application : $\mathrm{SO}_n(\mathbb{R})$ puis $\mathrm{GL}_n^+(\mathbb{R})$ est cpa. Application : Composantes connexes de $Q(E)$. Application : $\exp : \mathcal{A}_n(\mathbb{R}) \longmapsto \mathrm{SO}_n(\mathbb{R})$ est surjective.

III Espaces cycliques et invariants de similitudes

Définitions endomorphismes cycliques. Définition matrices compagnons, propriétés. Théorème des invariants de similitudes. Réduction de Frobenius. Application à la décomposition de Jordan, application : ${}^tA \sim A$.

IV Représentations linéaires d'un groupe fini

Définition représentation de groupes finis. Exemples : représentation régulière d'un groupe fini, représentation du groupe $\mathfrak{S}_n$, représentation du groupe des quaternions. Définition sous-représentation, représentations irréductibles. Définition somme de représentations. Représentations unitarisables, théorème de Maschke.

[TROUVER REF] Bonus : opérateurs d'entrelacement, lemme de Schur. Application aux représentations irréductibles des groupes abéliens. Application : théorème de structure des groupes abéliens finis.

Développements

(i) Existence et unicité des invariants de similitudes

(ii) Décomposition de Dunford

Références

(i) Gourdon, Algèbre

(ii) Mneimné et Mansuy, Réduction des endomorphismes

(iii) Caldero et Germoni, H2G2, tome I

(iv) FGN, Algèbre 2

155 : Endomorphismes diagonalisables en dimension finie.

Rapport du jury

“Dans cette leçon, on attend des exemples naturels d’endomorphismes diagonalisables et des critères de diagonalisabilité. On peut étudier certaines propriétés topologiques en prenant le soin de donner des précisions sur le corps $\mathbb{K}$ et la topologie choisie pour $\mathcal{M}_n(\mathbb{K})$. Le calcul de l’exponentielle d’un endomorphisme diagonalisable est immédiat une fois que l’on connaît les valeurs propres et ceci sans diagonaliser la matrice, par exemple à l’aide des projecteurs spectraux. Sur les corps finis, on a des critères spécifiques de diagonalisabilité. On peut dénombrer les endomorphismes diagonalisables, ou possédant des propriétés données, liées à la diagonalisation.

S’ils le désirent, les candidats peuvent s’intéresser aux liens qui peuvent aussi être fait avec la théorie des représentations et la transformée de Fourier rapide.”

Introduction

Motivation : résolution de système (différentiels) linéaires. S’incrit dans le cadre plus général de la réduction. Parler de la réduction. Action de $\mathrm{GL}_n(\mathbb{R})$ par conjugaison puis par congruence.

Plan

Soit $\mathbb{K}$ un corps (commutatif) et E un espace vectoriel de dimension finie n. Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

I Les outils pour la diagonalisation et premières applications

(a) Spectre et espaces propres

Définition valeurs propres, vecteurs propres, espaces propres, spectre. Les espaces propres sont en somme directe. Polynômes et valeurs propres. Commutation et espaces propres. Définition endomorphisme diagonalisable. Condition suffisante de diagonalisabilité : avoir n valeurs propres différentes.

(b) Polynômes d’endomorphismes

Définition du polynôme minimal de f. Caractérisation des valeurs propres par les racines du polynôme minimal.

Définition du polynôme caractéristique pour une matrice puis pour un endomorphisme. Degré. Caractérisation des valeurs propres par les racines du polynôme caractéristique. Exemple : polynôme caractéristique d’un endomorphisme nilpotent. Le spectre de f est fini. Application : $\mathrm{GL}(E)$ est dense dans $\mathcal{L}(E)$ si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Sous-espaces stables et polynôme caractéristique. Exemple : matrices triangulaires par blocs. Théorème de Cayley-Hamilton. Conséquences.

Lemme des noyaux. Application : Décomposition de Dunford (existence). Calcul de l’exponentielle d’un endomorphisme, à l’aide des projecteurs caractéristiques.

(c) Critères polynomiaux de diagonalisibilité

Condition suffisante de diagonalisabilité par le polynôme caractéristique. Critère via le polynôme caractéristique. Exemple : matrice avec que des 1.

Critère via les polynômes annulateurs (en particulier χ_f). Application : endomorphisme induit par un sev stable est diagonalisable. Exemple : endomorphisme de rang 1 et trace. Application aux sous-groupes finis de $\mathrm{GL}_n(\mathbb{C})$ et Théorème de Burnside [FGN p.185]. Théorème de co-diagonalisation. Application : diagonalisibilité par blocs [FGN, p.103]. Application : Décomposition de Dunford (unicité). Densité des matrices diagonalisables dans $\mathcal{M}_n(\mathbb{C})$.

II Le cadre euclidien/hermitien

E est maintenant un espace euclidien ou hermitien.

(a) Endomorphismes auto-adjoints

Définition endomorphisme auto-adjoint. Théorème spectral. Orthogonalisation simultanée. Théorème : racine carrée, décomposition polaire (homéomorphisme). Application : $\mathrm{GL}_n^+(\mathbb{R})$ est cpa. Application : enveloppe convexe de $\mathrm{GL}_n(\mathbb{R})$. Application : exp induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^+ + (\mathbb{R})$.

(b) Endomorphismes normaux

On suppose E hermitien. Définition endomorphismes normaux. Exemples : endomorphismes auto-adjoints, antisymétriques, isométries ou unitaires. Si F est u-stable alors $F^\perp$ est u^*-stable. Si u est normal et F u-stable, alors $F^\perp$ est u-stable. Caractérisation des endomorphismes normaux (version endomorphisme et version matricielle).

Développements

(i) Décomposition de Dunford

(ii) L'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$

Références

(i) Gourdon, Algèbre

(ii) Mneimné et Mansuy, Réduction des endomorphismes

(iii) FGN, Algèbre 2

156 : Exponentielle de matrices. Applications.

Rapport du jury

“Bien que ce ne soit pas une leçon d’analyse, il faut toutefois pouvoir justifier clairement la convergence de la série exponentielle. La distinction entre le cas réel et complexe doit être clairement évoqué. Les questions de surjectivité ou d’injectivité doivent être abordées. Par exemple la matrice $A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ est-elle l’exponentielle d’une matrice à coefficients réels ? La matrice définie par blocs $B = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}$ est-elle l’exponentielle d’une matrice à coefficients réels ? La décomposition de Dunford multiplicative (décomposition de Jordan) de $\exp(A)$ trouve toute son utilité dans cette leçon. Notons que l’exponentielle fait bon ménage avec la décomposition polaire dans bon nombre de problèmes sur les sous-groupes du groupe linéaire. L’étude du logarithme (quand il est défini) trouve toute sa place dans cette leçon. Si l’on traite du cas des matrices nilpotentes, on pourra invoquer le calcul sur les développements limités. Les applications aux équations différentielles doivent être évoquées sans constituer l’essentiel de la leçon. On pourra par exemple faire le lien entre réduction et comportement asymptotique, mais le jury déconseille aux candidats de proposer ce thème dans un développement.

S’ils le désirent, les candidats peuvent s’aventurer vers les sous-groupes à un paramètre du groupe linéaire (on peut alors voir si ces sous-groupes constituent des sous-variétés fermées de $\mathrm{GL}_n(\mathbb{R})$, $\mathbb{R}^q$) ou vers les algèbres de Lie.”

Introduction

Plan

I Exponentielle de matrices

(a) Définition et premières propriétés

Définition de l’exponentielle d’une matrice. Exemple pour une matrice diagonale, triangulaire supérieure. L’exponentielle d’une matrice A est un polynôme en A, donc commute avec A.

Propriétés de l’application exponentielle (majoration de la norme, continuité) [GOU]. Conjugaison. Application : $\det(\exp A) = e^{\operatorname{tr} A}$.

Exponentielle de la somme de deux matrices qui commutent [GOU]. Contre-exemple dans le cas où il n’y a pas commutation. Conséquence : exp est à valeurs dans $\mathrm{GL}_n(\mathbb{K})$ (expression de l’inverse de $\exp(A)$). Formule généralisé de l’exponentielle de la somme de deux matrices quelconque [GOU].

(b) Inversion de l’exponentielle

L’exponentielle n’est pas injective. Exemple : calcul des solutions de $\exp A = I_2$ dans $\mathcal{M}_n(\mathbb{C})$ et $\mathcal{M}_n(\mathbb{R})$ en dimension 2 [MT] puis en dimension quelconque [FGN]. L’exponentielle est injective sur l’ensemble des matrices diagonalisables.

Calcul de la différentielle de l’exponentielle en 0 et inversion locale.

Image de l’exponentielle de $\mathcal{M}_n(\mathbb{C})$ [Zav]. Application : $GL_n(\mathbb{C})$ est connexe (par arcs). L’exponentielle n’est pas surjective dans $GL_n(\mathbb{R})$ [GOU].

Définition logarithme de matrice nilpotente. L’exponentielle induit un homéomorphisme de $\mathcal{N}$ (matrices nilpotentes) sur $\mathcal{U}$ (matrices unipotentes) [BMP].

L’exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$. Décomposition polaire. Application : $\mathrm{GL}_n(\mathbb{R}) \simeq \mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{\frac{n(n+1)}{2}}$ [MT]. On a la même chose pour $\mathbb{C}$.

(c) Calcul pratique de l’exponentielle

Sous-espaces caractéristiques, projecteurs caractéristiques. Décomposition de Dunford. Décomposition de Dunford multiplicative. Calcul pratique de l’exponentielle [GOU]. Exemple : exo 1 [GOU]. Application : CNS pour que $\lim e^{tM} = 0$, exo 2 [GOU]

II Applications

(a) Équations différentielles linéaires à coefficients constants

Dérivation de $t \longmapsto \exp(tA)$. Solutions d'une équation différentielle linéaire à coefficients constants. Application : équation matricielle [GOU].

Théorème de Liapunov [ROU].

(b) Sous-groupes à un paramètre

Morphismes de groupes dérivables (puis continus) de $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{C}), \times)$. Applications : groupes à un paramètre dans $\mathcal{M}_n(\mathbb{R})$ [FGN].

Développements

(i) L'exponentielle induit un homéomorphisme sur $\mathcal{S}_n(\mathbb{R})$.

(ii) Équation matricielle

Références

(i) Gourdon, Algèbre et Analyse

(ii) FGN, Algèbre 2

(iii) Mneimé & Testard, Introduction à la théorie des groupes de Lie classiques

(iv) BMP, Objectif Agrégation

(v) Rouvière, PGCD

157 : Endomorphismes trigonalisables. Endomorphismes nilpotents.

Rapport du jury

"L'utilisation des noyaux itérés est fondamentale dans cette leçon, ceci pour déterminer si deux matrices nilpotentes sont semblables par exemple. Il est intéressant de présenter des conditions suffisantes de trigonalisation simultanée ; l'étude des endomorphismes cycliques a toute sa place dans cette leçon. Notons que l'étude des nilpotents en dimension 2 débouche naturellement sur des problèmes de quadriques et que l'étude sur un corps fini donne lieu à de jolis problèmes de dénombrement.

S'ils le désirent, les candidats peuvent aussi présenter la décomposition de Frobenius."

Introduction

Plan

I Généralités sur la réduction

(a) Polynômes d'endomorphismes

Définition polynôme minimal d'un endomorphisme (local et global). Propriétés. $\mathbb{K}$-algèbre d'un endomorphisme (locale et globale). Définition endomorphismes nilpotents.

Lemme des noyaux. Les projecteurs sont des polynômes en f. Application : existence de $x \in E$ tel que $\mu_{f,x} = \mu_f$.

(b) Valeurs propres, espaces propres

Définition valeur propre, espace propre. Les espaces propres sont en somme directe. Les valeurs propres sont les racines du polynôme minimal. Définition polynôme caractéristique. Les valeurs propres sont les racines du polynôme caractéristique. Théorème de Cayley-Hamilton.

II Endomorphismes trigonalisables

(a) Critères de trigonalisation

Définition endomorphisme trigonalisable. Lecture du spectre sur les coefficients diagonaux. Critère de trigonalisiblité, de co-trigonalisibilité. Exemple de matrices 2×2 ou 3×3. Application : densité des endomorphismes diagonalisables dans $\mathcal{M}_n(\mathbb{C})$. Application : démonstration du théorème de Cayley-Hamilton sur $\mathcal{M}_n(\mathbb{C})$.

(b) Espaces caractéristiques et décomposition de Dunford

Définition sous-espaces caractéristiques. Théorème de Décomposition de Dunford. Application exponentielle d'endomorphismes, limite de e^{tM}.

III Endomorphismes nilpotents

Exemples de matrices nilpotentes. Exemples d'endomorphismes nilpotents [BMP]. L'espace vectoriel engendré par les endomorphismes nilpotents est le noyau de la trace. Indice de nilpotence (qui est plus petit que n d'après Cayley-Hamilton). Somme de deux nilpotents qui commutent. Caractérisation des endomorphismes nilpotents [BMP] (en particulier, u est nilpotent ssi $\operatorname{tr}(u^k) = 0$ pour tout k). Application : théorème de Burnside.

Exemple des blocs de Jordan [BMP]. Noyaux itérés et décomposition de Jordan d'un endomorphisme nilpotent. CNS pour que deux endomorphismes nilpotents soient semblables.

IV Espaces cycliques et invariants de similitudes

Polynômes locaux. Définition espaces cycliques. Bases, dimension. Définitions endomorphismes cycliques. Définition matrices compagnons, propriétés. Théorème des invariants de similitudes. Réduction de Frobenius. Application à la décomposition de Jordan, application : ${}^tA \sim A$. Mettre en évidence le cas nilpotent.

Développements

(i) Décomposition de Dunford

(ii) Existence et unicité des invariants de similitudes

Références

(i) BMP, Objectif Agrégation

(ii) Gourdon, Algèbre

(iii) FGN, Algèbre 2

158 : Matrices symétriques réelles, matrices hermitiennes.

Rapport du jury

"Le théorème spectral est indispensable dans cette leçon sans toutefois être un développement consistant. La notion de signature doit être présentée ainsi que son unicité dans la classe de congruence d'une matrice symétrique réelle. L'action du groupe linéaire et du groupe orthogonal sur l'espace des matrices symétriques peut donner un cadre naturel à cette leçon. Le lien avec les formes quadratiques et les formes hermitiennes est incontournable. La partie réelle et la partie imaginaire d'un produit hermitien définissent des structures sur l'espace vectoriel réel sous-jacent. L'orthogonalisation simultanée est un résultat important de cette leçon. Il faut en connaître les applications géométriques aux quadriques."

Introduction

Plan

I L'espace des matrices symétriques réelles, des matrices hermitiennes

Définition matrices symétriques, matrices antisymétriques, matrices hermitiennes, matrices antihermitiennes. Exemples : exemples simples, A^*A, exemple de la hessienne. Propriétés ($\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$, $\mathcal{H}_n(\mathbb{C}) = \mathcal{S}_n(\mathbb{C}) \oplus i\mathcal{A}_n(\mathbb{R})$, explicitation d'une base, dimension) [GOU]. $\mathcal{H}_n(\mathbb{C})$ n'est pas un $\mathbb{C}$-espace vectoriel.

II Liens avec l'algèbre bilinéaire et sesquilinéaire

(a) Formes quadratiques, formes hermitiennes

Définition forme bilinéaire, forme sesquilinéaire. Exemples simples en dimension finie ou infinie.

Formule bilinéaire symétrique, antisymétrique, forme sesquilinéaire hermitienne. Formule de polarisation. Exemples.

(b) Écriture matricielle

Écriture matricielle en dimension finie. Exemple. Formule de changement de bases. Action par congruence.

(c) Matrices positives, définies positives

Définition forme quadratiques positives ou définie positive (de même pour les formes hermitiennes). Exemple de A^*A.

III Lien avec l'adjoint d'un endomorphisme

Définition adjoint d'un endomorphisme. Représentation matricielle. Endomorphisme auto-adjoint. Les valeurs propres d'un endomorphisme auto-adjoint, donc d'une matrice symétrique ou hermitienne, sont réelles.

IV Diagonalisation des matrices symétriques et des matrices hermitiennes

(a) Réduction et signature

Définition signature, théorème de réduction. Quelques exemples. Application : lemme de Morse. Étude locale de courbe en un point critique. Critère de Sylvester.

(b) Théorème spectral

Théorème spectral. Pseudo réduction simultanée. Log-concavité du déterminant. Ellipsoïde de John-Loewner. L'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.

V Applications

(a) Racine carrée

Existence et unicité de la racine carrée d'une matrice symétrique (hermitienne) positive.

(b) Décomposition polaire

Théorème de décomposition polaire (existence et unicité). Application : composantes connexes de $\mathrm{GL}_n(\mathbb{R})$, composantes connexes de l'ensemble des formes quadratiques non dégénérées. Application : enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$.

Développements

(i) Lemme de Morse

(ii) L'exponentielle induit un homéomorphisme sur $\mathcal{S}_n(\mathbb{R})$

Références

(i) Gourdon, Algèbre

(ii) Rouvière, PGCD

(iii) FGN, Algèbre 3

(iv) Szpirglas, Algèbre L3

A FAIRE 159 : Formes linéaires et dualité en dimension finie. Exemples et applications.

Rapport du jury

"Il est important de bien placer la thématique de la dualité dans cette leçon ; celle-ci permet de mettre en évidence des correspondances entre un morphisme et son morphisme transposé, entre un sous-espace et son orthogonal (canonique), entre les noyaux et les images ou entre les sommes et les intersections. Bon nombre de résultats d'algèbre linéaire se voient dédoublés par cette correspondance. Les liens entre base duale et fonctions de coordonnées doivent être parfaitement connus. Savoir calculer la dimension d'une intersection d'hyperplans via la dualité est important dans cette leçon.

L'utilisation des opérations élémentaires sur les lignes et les colonnes permet facilement d'obtenir les équations d'un sous-espace vectoriel ou d'exhiber une base d'une intersection d'hyperplans. Cette leçon peut être traitée sous différents aspects : géométrique, algébrique, topologique ou analytique. Il faut que les développements proposés soient en lien direct avec la leçon. Enfin rappeler que la différentielle d'une fonction à valeurs réelles est une forme linéaire semble incontournable."

Introduction

Plan

Développements

(i)

(ii)

Références

(i)

(ii)

(iii)

(iv)

160 : Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).

Rapport du jury

"Dans cette leçon, le caractère euclidien de l'espace est essentiel pour que l'endomorphisme soit remarquable. Le théorème spectral pour les auto-adjoints et la réduction des endomorphismes orthogonaux sont des résultats incontournables. Le lemme des noyaux ou la décomposition de Dunford ne sont pas des développements adaptés à cette leçon. En revanche, l'utilisation du fait que l'orthogonal d'un sous-espace stable par un endomorphisme est stable par l'adjoint doit être mis en valeur. De même la réduction de endomorphismes normaux peut être évoquée."

Introduction

Plan

I Endomorphismes orthogonaux dans un espace euclidien

(a) Propriétés du groupe orthogonal

Définitions équivalentes. Groupe spécial orthogonal. $\mathrm{O}(E)$ est un groupe compact. Représentation matricielle. Exemple : réflexions, renversements, symétries orthogonales.
Centre et générateurs de $\mathrm{O}(E)$, de $\mathrm{SO}(E)$.

(b) Le groupe des quaternions

Définition de l'espace des quaternions, de la norme, du groupe des quaternions. Paramétrisation de $\mathrm{SO}_3(\mathbb{R})$ à l'aide du groupe des quaternions.

(c) Réduction des endomorphismes orthogonaux

Si un sous-espace est stable par f alors son orthogonal est aussi stable par f. Théorème de réduction des endomorphismes orthogonaux. Application : $\mathrm{SO}(E)$ est connexe par arcs. Composantes connexes de $\mathrm{O}(E)$. Forme matricielle des isométries du plan et de l'espace.

II Endomorphismes auto-adjoints, endomorphismes antisymétriques

(a) Adjoint d'un endomorphisme

Définition de l'adjoint d'un endomorphisme. Interprétation matricielle. Calcul d'adjoint. Quelques propriétés de l'application adjoint (involution, somme, multiplication par un scalaire, composition, conservation du rang, du déterminant). Si un sous-espace est stable par f alors son orthogonal est stable par f^*. Propriété sur le noyau, sur l'image. Application : si $A^*A = 0$ alors $A = 0$ [GRI].

(b) Endomorphisme auto-adjoint, endomorphisme antisymétrique

Définition endomorphisme auto-adjoint, antisymétrique. Exemple de $f \circ f^*$. Représentation matricielle. Si F est stable par un endomorphisme auto-adjoint (ou symétrique) alors $F^\perp$ l'est aussi. Les sous-espaces propres d'un endomorphisme auto-adjoint sont deux à deux orthogonaux. Les valeurs propres d'un endomorphisme auto-adjoint sont réelles. $\mathcal{L}(E) = \mathcal{A}(E) \oplus \mathcal{S}(E)$. Application : exo 3 p.236 [GOU].

(c) Théorème spectral

Théorème spectral, théorème de diagonalisation pseudo simultanée. Contre-exemple dans le cas des matrices symétriques complexes [GRI].
Application : log concavité du déterminant. Application : ellipsoïde de John-Loewner. Application aux sous-groupes compacts.
Application : racine carrée puis décomposition polaire. Application : $GL_n(\mathbb{R})^+$ est connexe par arcs. Application : enveloppe convexe de $\mathrm{O}(E)$.
Application : l'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.

III Endomorphismes normaux [Facultatif]

Définition endomorphisme normal. Propriétés. Étude en dimension 2 et dans le cas réel. Réduction des endomorphismes normaux réels. Application à la réduction des endomorphismes antisymétriques.

Développements

(i) Quaternions et $\mathrm{SO}_3(\mathbb{R})$

(ii) L'exponentielle induit un homéomorphisme sur $\mathcal{S}_n(\mathbb{R})$

Références

(i) Gourdon, Algèbre

(ii) Perrin, Cours d'algèbre

(iii) Grifone, Algèbre linéaire

(iv) FGN, algèbre 2

161 : Isométries d'un espace affine euclidien de dimension finie. Applications en dimensions 2 et 3.

Rapport du jury

"La classification des isométries en dimension 2 et 3 est exigible. Il faut savoir prouver qu'une isométrie est affine, pouvoir donner des générateurs du groupe des isométries affines, et savoir composer des isométries affines. En dimension 3, il faut savoir classifier les rotations et connaître les liens avec la réduction. On peut aussi penser aux applications aux isométries laissant stables certains objets en dimension 2 et 3."

Introduction

Plan

I Le groupe orthogonal

(a) Définition et premières propriétés

Définition du groupe orthogonal euclidien comme le stabilisateur d'une forme quadratique définie positive. Définitions équivalentes. Interprétation matricielle. Compacité. Groupe spécial orthogonal. Exemples : symétries orthogonales, réflexions, renversements, rotations.

Générateurs de $\mathrm{O}(E)$, de $\mathrm{SO}(E)$.

(b) Réduction des endomorphismes orthogonaux

Théorème de réduction des endomorphismes orthogonaux. Application : $\mathrm{O}(E)$ a deux composantes connexes. Décomposition polaire. Application : enveloppe convexe de $\mathrm{O}(E)$.

II Groupe des isométries dans un espace euclidien de dimension fini

(a) Le groupe des isométries

Définition des isométries affines, structure de groupe. Exemple : translations, refléxions affines, renversements affines, etc. Déplacements, anti-déplacements.

Des trucs sur les points fixes.

(b) Structure des isométries

Théorème de Cartan-Dieudonné. Théorème de structure des isométries affines. Forme réduite des isométries [AUD].

III Classification des isométries du plan et de l'espace

(a) Dans le plan

Voir [AUD].

(b) Dans l'espace

Voir [AUD]. Le groupe des quaternions et $\mathrm{SO}_3(\mathbb{R})$.

IV Groupe d'isométries d'une partie ...

(a) ... finie

Groupe des isométries conservant une partie finie. Centre de symétrie. On peut se ramener à l'étude des déplacements.

(b) ... en dimension 2

Définition du groupe diédral [Per].

(c) ... en dimension 3

Tétraèdre, cube.

Développements

(i) Quaternions et $\mathrm{SO}_3(\mathbb{R})$
(ii) Enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$

Références

(i) Perrin, Cours d'algèbre
(ii) Audin, Géométrie
(iii) Caldero & Germoni, H2D2

170 : Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.

Rapport du jury

"Il faut tout d'abord noter que l'intitulé implique implicitement que le candidat ne doit pas se contenter de travailler sur $\mathbb{R}$. Le candidat pourra parler de la classification des formes quadratiques sur le corps des complexes et sur les corps finis. L'algorithme de Gauss doit être énoncé et pouvoir être pratiqué sur une forme quadratique simple.

Les notions d'isotropie et de cône isotrope sont un aspect important de cette leçon. On pourra rattacher cette notion à la géométrie différentielle."

Introduction

Plan

I Définition et premières propriétés

(a) Formes bilinéaires symétriques et formes quadratiques

Définition d'une forme bilinéaire sur un espace vectoriel quelconque. Exemple [GOU]. Écriture matricielle en dimension finie. Discriminant. Forme bilinéaire non dégénérée, noyau, rang. Exemples simples. Forme bilinéaire symétrique, forme quadratique. Exemples simples ou pas. Forme polaire, identité de polarisation.

(b) Orthogonalité et isotropie

Vecteurs orthogonaux, orthogonal d'une partie, propriétés [PER]. Interprétation matricielle. Exemple, contre exemples.

Vecteur isotrope, cône isotrope. Interprétation matricielle. Espace totalement isotrope. SETIM. Indice. Propriétés. Formes anisotropes.

II Classification des formes quadratiques

(a) Bases q-orthogonales

Définition base q-orthogonale, existence d'une base orthogonale. Changement de base, relation de congruence (d'équivalence), écriture matricielle, relation de congruence. Stabilité de quelques propriétés (rang, indice). Algorithme de Gauss. Exemple. Interprétation matricielle.

(b) Classification

Théorème de classification des formes quadratiques dans le cas d'un corps algébriquement clos, dans le cas de $\mathbb{R}$, dans le cas d'un corps fini. Quelques critères de classification (signature, discriminant). Relation entre signature, rang et indice.

Application : loi de réciprocité quadratique.

III Formes quadratiques réelles

(a) Signature d'une forme quadratique réelle

Quelques calculs de signature [GOU]. Lemme de Morse. Étude locale de courbes en un point critique. Classification des coniques.

(b) Formes quadratiques positives

Inégalité de Cauchy-Schwarz, inégalité de Minkowski. Norme induite. Composantes connexes de l'ensemble des formes quadratiques non dégénérées. Ellipsoïde de John-Loewner. Applications sur les sous-groupes compacts de $GL_n(\mathbb{R})$.

Développements

(i) Composantes connexes de l'ensemble des formes quadratiques non dégénérées
(ii) Loi de réciprocité quadratique (via les formes quadratiques)

Références

(i) Perrin, Cours d'algèbre
(ii) Gourdon, Algèbre
(iii) FGN, Algèbre 2
(iv) Rouvière, PGCD
(v) description

171 : Formes quadratiques réelles. Coniques. Exemples et applications.

Rapport du jury

“Dans cette leçon, la loi d’inertie de Silvester doit être présentée ainsi que l’orthogonalisation simultanée. L’algorithme de Gauss doit être énoncé et pouvoir être expliqué sur une forme quadratique de $\mathbb{R}^3$; le lien avec la signature doit être clairement énoncé et la signification géométrique des deux entiers r et s composant la signature d’une forme quadratique réelle doit être expliqué. La différentielle seconde d’une fonction de plusieurs variables est une forme quadratique importante qui mérite d’être présentée dans cette leçon.

La définition des coniques affines non dégénérées doit être connue, et les propriétés classiques des coniques doivent être données. On pourra présenter les liens entre la classification des formes quadratiques et celles des coniques ; de même il est intéressant d’évoquer le lien entre le discriminant de l’équation $ax^2 + bx + c = 0$ et la signature de la forme quadratique $ax^2 + bxy + cy^2$.

S’ils le désirent, les candidats peuvent aussi aller vers la théorie des représentation et présenter l’indicatrice de Schur-Frobenius qui permet de réaliser une représentation donnée sur le corps des réels.”

Introduction

Plan

I Généralités sur les formes quadratiques réelles

(a) Formes bilinéaires symétriques et formes quadratiques

Définition forme quadratique, forme polaire, identité de polarisation. Exemples [GOU]. Matrice d’une forme quadratique, rang, noyau. Relation d’équivalence, congruence.

(b) Classification des formes quadratiques réelles en dimension finie

Définition base orthogonale. Existence d’une base orthogonale. Méthode de Gauss. Signature d’une forme quadratique. Théorème de Sylvester. Lien entre la signature et le rang. Exemples de calcul de signature.

II Formes quadratiques positives

Définition formes quadratiques positives, négatives, définies positives, définies négatives. Stabilité par restriction. Caractérisation de Sylvester [FGN] ou [GOU]. Inégalité de Cauchy-Schwarz, inégalité de Minkowski. Norme euclidienne. Propriétés de l’orthogonal d’un sous-espace. Théorème de diagonalisation simultanée. Application : log concavité du déterminant. Application : ellipsoïde de John-Loewner. Application : sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.

III Application au calcul différentiel et à l’optimisation

(a) Différentielle seconde

Définition différentielle seconde, matrice hessienne. Lemme de Schwarz. Lemme de Morse. Étude locale en un point critique d’une courbe. Caractérisation de la convexité par la différentielle seconde.

(b) Optimisation

Conditions nécessaires et conditions suffisantes d’extrema. Exemples et contre-exemples. Principe du maximum faible.

IV Coniques affines

(a) Généralités sur les coniques affines

Définition d'une conique affine par les zéros d'un polynôme du second degré puis par la classe d'équivalence de ce polynôme. Définition conique propre (par la forme homogénéisée). Définion conique à centre. CNS pour qu'une conique ait un centre.

(b) Coniques dans le plan euclidien

Classification des coniques propre à centre. Quelques propriétés des coniques (compacité d'une ellipse, branches infinies d'une hyperbole). Coniques propre sans centre (parabole). Définition grand (petit) axe. Définition âr foyer et directrice. Définition métrique. Problème du jardinier. Ellipse de Steiner.

Développements

(i) Ellipsoïde de John-Loewner

(ii) Composantes connexes de l'ensemble des formes quadratiques non dégénérées

Références

(i) Gourdon, Algèbre

(ii) FGN, Algèbre 3

(iii) Rouvière, PGCD

(iv) Audin, Géométrie

190 : Méthodes combinatoires, problèmes de dénombrement.

Rapport du jury

"Il est nécessaire de dégager clairement différentes méthodes de dénombrement et les illustrer d'exemples significatifs. De nombreux domaines de mathématiques sont concernés par des problèmes de dénombrement, cet aspect varié du thème de la leçon doit être mis en avant. L'utilisation de séries génératrices est un outil puissant pour le calcul de certains cardinaux. De plus il est naturel de calculer des cardinaux classiques et certaines probabilités. Il est important de connaître l'interprétation ensembliste de la somme des coefficients binomiaux, et ne pas se contenter d'une justification par le binôme de Newton. L'introduction des corps finis (même en se limitant aux cardinaux premiers) permet de créer un lien avec l'algèbre linéaire. Les actions de groupes peuvent également conduire à des résultats remarquables.

S'ils le désirent, les candidats peuvent aussi présenter des applications de la formule d'inversion de Möebius ou de la formule de Burnside."

Introduction

Plan

I Ensembles finis

(a) Cardinal d'un ensemble fini

Il existe une bijection entre $[\![1\,,n]\!]$ et $[\![1\,,m]\!]$ si et seulement si $n = m$. Définition d'un ensemble fini. Exemples pipeaux. Sous-ensembles d'un ensemble fini. Pleins d'égalités ensemblistes.

(b) Dénombrement, coefficients binomiaux

Dénombrement du nombre de p-liste, du nombre d'applications, du nombre d'injections, du nombre de permutation.

Nombre de parties à p éléments, coefficients binomiaux. Propriétés. FOrmule fondamentales. Binôme de Newton.

Formule du crible.

(c) Nombre de dérangements

Exo 2 [FGN].

II Séries génératrices

III Applications

(a) En théorie des groupes

(b) Dans $\mathbb{Z}/n\mathbb{Z}$

(c) En théorie des corps finis

(d) Partition d'un entier

(e) En probabilité discrète

Développements

(i) Dénombrement des polynômes irréductibles sur un corps fini

(ii) Loi de réciprocité quadratique

Références

(i) Deschamps & Warusfel, Mathémétiques tout-en-un, MPSI-PCSI
(ii) Arnaudies & Fraysse
(iii)
(iv)

Deuxième partie

Analyse et probabilités

201 : Espaces de fonctions ; exemples et applications.

Rapport du jury

"C'est une leçon riche où le candidat devra choisir soigneusement le niveau auquel il souhaite se placer. Les espaces de fonctions continues sur un compact (par exemple l'intervalle $[0,1]$) offrent des exemples élémentaires et pertinents. Dans ce domaine, le jury attend une maîtrise du fait qu'une limite uniforme de fonctions continues est continue. Les candidats peuvent se concentrer dans un premier temps sur les espaces de fonctions continues et les bases de la convergence uniforme. Les espaces de Hilbert de fonctions comme l'espace des fonctions L^2 constituent ensuite une ouverture déjà significative.

Pour aller plus loin, d'autres espaces usuels tels que les espaces L^p ont tout à fait leur place dans cette leçon. Le théorème de Riesz-Fischer est alors un très bon développement pour autant que ses difficultés soient maîtrisées. Les espaces de fonctions holomorphes sur un ouvert de $\mathbb{C}$ constituent aussi une ouverture de très bon niveau."

Introduction

Plan

I Espaces de fonctions continues sur un compact

Soit (X, d) un espace métrique compact (non vide). On note $\mathcal{C}_{\mathbb{K}}(X)$ le $\mathbb{K}$-espace des fonctions continues sur X à valeurs dans $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

(a) Généralités

$\mathcal{C}(X)$ est un espace de Banach séparable. Théorème de Heine. Module de continuité et propriétés. Théorèmes de Dini et application à une suite de polynômes [GOU,HL].

(b) Parties compacts

Définition équicontinuité. Théorème d'Arzela-Ascoli. Application : théorème de Cauchy-Peano.

(c) Parties denses

Théorème des Stone-Weierstrass (réel et complexe). Exemples : les polynômes de Bernstein. Applications, contre-exemples. Théorèmes de Weierstrass. Application : théorème des moments de Hausdorff. $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2(\mathbb{T})$.

Rappel : théorème de Baire. Application : densité des fonctions continues nulle part dérivables.

II Les espaces L^p

(a) Généralités

Définition de L^p, de $\|\cdot\|_p$. Inégalité de Hölder, de Minkowski. Application : $L^p \subset L^q$ si $p \geq q$ si la mesure est de masse finie. Application : inégalité d'interpolation [BRE]. Théorème de Riesz-Fischer.

(b) Dual des espaces L^p

Théorème de Riesz-Fréchet. Contre-exemple pour $p = 1$ ou $+\infty$. Critère de densité. Densité de $\mathcal{C}_c$ dans L^p. L^p est séparable.

(c) Convolution et régularisation

Produit de convolution $L^1 - L^1$, $L^1 - L^p$ et $L^p - L^q$. Approximation de l'unité, suites régularisantes. Exemple de suite régularisante. Densité de $\mathcal{C}_c^\infty(\Omega)$ dans $L^p(\Omega)$.

III Espace des opérateurs linéaires continus

Soit E et F deux espaces de Banach (réels).

(a) Généralités

Définition de $\mathcal{L}(E, F)$. Théorème de Banach. Structure des inversibles. Définition de l'adjoint d'un opérateur. Opérateurs auto-adjoints. Spectre et valeurs propres d'un opérateur.

(b) Opérateurs compacts

Définition opérateurs compacts. Exemples : opérateurs à noyaux. $\mathcal{K}(E, F)$ est un sous-espace vectoriel fermé de $\mathcal{L}(E, F)$ et est stable par composition à droite et à gauche. Théorème de Schauder. Alternative de Fredholm.

(c) Diagonalisation des opérateurs auto-adjoints compacts

Spectre des opérateurs compacts. Spectre des opérateurs auto-adjoints. Diagonalisation.

Développements

(i) Alternative de Fredholm
(ii) Densité des fonctions continues nulle part dérivables

Références

(i) Hirsch-Lacombe, Éléments d'analyse fonctionnelle
(ii) Brézis, Analyse fonctionnelle
(iii) Briane & Pages, Théorie de l'intégration
(iv) Zuily & Quéffélec, Analyse pour l'agrégation

202 : Exemples de parties denses et applications.

Rapport du jury

"Il ne faut pas négliger les exemples élémentaires comme les sous-groupes additifs de $\mathbb{R}$ et leurs applications, ou encore les critères de densité dans un espace de Hilbert. Le théorème de Weierstrass via les polynômes de Bernstein peut être abordé à des niveaux divers suivant que l'on précise ou pas la vitesse de convergence voire son optimalité. Pour aller plus loin, la version plus abstraite du théorème de Weierstrass (le théorème de Stone-Weierstrass) est aussi intéressante et a de multiples applications. Cette leçon permet aussi d'explorer les questions d'approximation de fonctions par des polynômes et des polynômes trigonométriques, ou plus généralement la densité de certains espaces remarquables de fonctions dans les espaces de fonctions continues, ou dans les espaces L^p. Il est également possible de parler de l'équirépartition."

Introduction

Plan

Soit (E, d) un espace métrique. Définition densité, séparabilité.

I Exemples en dimension finie

(a) Dans $\mathbb{R}$

Exemples de parties dense dans $\mathbb{R}$ ($\mathbb{Q}$, $\mathbb{R} \setminus \mathbb{Q}$, dyadiques). Sous-groupes additifs et applications [FGN 1 p.29]. Critère de Weyl et application aux suites équiréparties module 1 [FGN 2, p.50].

(b) Dans $\mathcal{M}_n(\mathbb{K})$

Densité dans l'espace des matrices (densité de $\mathrm{GL}_n(\mathbb{K})$. Application : différentielle du déterminant. Densité des matrices à n valeurs propres (donc diagonalisables) dans $\mathcal{M}_n(\mathbb{C})$). Application : Théorème de Cayley-Hamilton par densité dans $\mathcal{M}_n(\mathbb{C})$.

II Densité en dimension quelconque

(a) Prolongement par densité

Prolongement des applications UC. Application : intégrale de Riemann des fonctions réglées.

(b) Théorème de Baire et applications

Théorème de Baire et applications (densité des fonctions continues nulle part dérivables). Théorème de Banach-Steinhauss et applications (fonctions continues dont la SDF diverge en 0 (ou même presque partout), fonctions continues dont l'interpolée diverge).

(c) Densité dans $\mathcal{C}(X, \mathbb{K})$

(X, d) est un espace métrique compact. $\mathcal{C}(X, \mathbb{K})$ est un espace de Banach séparable [HL]. Théorèmes de Stone-Weierstrass et applications (Weierstrass et Weierstrass trigonométrique). (module de continuité) et polynômes de Bernstein. Théorème de Fejér et conséquences, noyau de Jackson et conséquences (théorème de Jackson).

(d) Densité dans les espaces L^p

Exemples de parties denses (fonction étagées, $\mathcal{C}_c$). Application : L^p est séparable. Convolution et applications (densité de $\mathcal{C}_c^{+\infty}$). Prolongement de Plancherel. Application (de la convolution) : théorème de Riesz-Fréchet-Kolmogorov (nécessite Ascoli).

III Espaces de Hilbert et bases hilbertiennes

Définition bases hilbertiennes. Caractérisation de la densité par l'orthogonal et application aux polynômes orthogonaux. Diagonalisation des opérateurs auto-adjoints et compacts.

Développements

(i) $\mathcal{C}_c^\infty$ est dense dans L^p

(ii) Densité des fonctions continues nulle part dérivables

Références

(i) Gourdon, Analyse (prolongement par densité) et Algèbre (pour la topologie des matrices)

(ii) Rouvière, PGCD (différentielle déterminant)

(iii) FGN, Analyse 1

(iv) FGN, Analyse 2

(v) Briane & Pages, Théorie de l'intégration

(vi) Hirch & Lacombe, Éléments d'analyse fonctionnelle

(vii) Brézis, Analyse fonctionnelle

203 : Utilisation de la notion de compacité.

Rapport du jury

"Il est important de ne pas concentrer la leçon sur la compacité générale (confusion entre utilisation de la notion compacité et notion de compacité). Néanmoins, on attend des candidats d'avoir une vision synthétique de la compacité. Des exemples d'applications comme le théorème de Heine et le théorème de Rolle doivent y figurer et leur démonstration être connue. Par ailleurs, le candidat doit savoir quand la boule unité d'un espace vectoriel normé est compacte. Des exemples significatifs d'utilisation comme le théorème de Stone-Weierstrass, des théorèmes de point fixe, voire l'étude qualitative d'équations différentielles, sont tout-à fait envisageables. Le rôle de la compacité pour des problèmes d'existence d'extrema mériterait d'être davantage étudié. On peut penser comme application à la diagonalisation des matrices symétriques à coefficients réels.

Pour aller plus loin, les familles normales de fonctions holomorphes fournissent des exemples fondamentaux d'utilisation de la compacité. Les opérateurs auto-adjoints compacts sur l'espace de Hilbert relèvent également de cette leçon, et on pourra développer l'analyse de leurs propriétés spectrales."

Introduction

Lien entre compacité et existence. La propriété de Borel-Lebesgue permet des raisonnements du type local au global.

Plan

Soit (E, d) un espace métrique.

I Définition et premières propriétés

Propriété de Borel-Lebesgue. Précompacité. Critères de la compacité (Bolzano-Weierstrass). Propriétés basiques. Application : premier théorème de Dini. Compacts de $(\mathbb{R}^n, \|\cdot\|_\infty)$. Théorème Tychonoff dénombrable. Application : théorème de Banach-Alaoglu.

II Applications continues sur un compact

(a) Utilisation de la compacité dans la recherche d'extrema

Image d'un compact par une application continue. Bicontinuité automatique. Une application continue et réelle sur un compact est bornée et atteint ses bornes. Si f est propre et continue alors f admet un minimum. Application : la distance à un compact est atteinte. Application : théorème de Rolle. Application : diagonalisation des endomorphismes auto-adjoints. Compacts en dimension finie. Application : Équivalence des normes en dimension finie. Éllipsoïde de John-Loewner.

(b) Théorème de Heine

Théorème de Heine. Application : deuxième théorème de Dini. Approximation de l'unité.

(c) Théorème de Stone-Weierstrass

Théorème de Stone-Weierstrass et ses conséquences. Critère de Weyl.

III Compacité dans les espaces fonctionnels

Équicontinuité. Théorème d'Arzela-Ascoli. Application : théorème d'Arzela-Peano. Théorème de Montel.

IV Spectre des opérateurs compacts

(a) Opérateurs compacts

Définition opérateur compact. Exemples : opérateurs à noyaux, opérateurs de rang fini. Propriétés de $\mathcal{K}(E, F)$. Adjoint. Théorème de Schauder.

(b) Alternative de Fredholm et ses conséquences

Lemme et théorème de Riesz. Alternative de Fredholm. Spectre et valeurs propres d'un opérateur. Description du spectre d'un opérateur compact.

Développements

(i) Alternative de Fredholm
(ii) Ellipsoïde de John-Loewner

Références

(i) Brézis, Analyse fonctionnelle
(ii) Gourdon, Analyse
(iii) Hirsch-Lacombe, Éléments d'analyse fonctionnelle

204 : Connexité. Exemples et applications.

Rapport du jury

"Le rôle clef de la connexité dans le passage du local au global doit être mis en évidence dans cette leçon : en calcul différentiel, voire pour les fonctions holomorphes. Il est important de présenter des résultats naturels dont la démonstration utilise la connexité. La stabilité par image continue, l'identification des connexes de $\mathbb{R}$ sont des résultats incontournables. On distinguera bien connexité et connexité par arcs (avec des exemples compris par le candidat), mais il est pertinent de présenter des situations où ces deux notions coïncident. A contrario, on pourra distinguer leur comportement par passage à l'adhérence.

Des exemples issus d'autres champs (algèbre linéaire notamment) seront appréciés. Le choix des développements doit être pertinent, le préambule en fournit quelques exemples, même s'il fait aussi appel à des thèmes différents ; on peut ainsi suggérer le théorème de Runge."

Introduction

La connexité décrit les espaces d'un seul tenant. La notion, plus forte de connexité par arcs est plus maniable.

Plan

I Généralités

(a) Définitions de la connexité

Définitions équivalentes de la connexité. Remarque sur la définition fonctionnelle : on peut remplacer $\mathbb{Z}$ par n'importe quel espace discret à au moins 2 éléments. Première propriété [QUE]. Parties connexes d'un espace métrique, lemme du passage des douanes. Exemple de parties non connexes ($\mathbb{R} \setminus \{0\}$).

Théorèmes de stabilité : union de connexes non disjoints, union d'une chaîne de connexes non disjoints successivement. Image par d'un connexe par une application continue. Applications : les intervalles sont connexes. Adhérence d'un connexe. Produit de connexes.

(b) L'exemple de $\mathbb{R}$

Connexité de $[0\,,1]$. Application : description des connexes de $\mathbb{R}$. Théorème des valeurs intermédiaires. $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes. Théorème de Darboux. Théorème de Brouwer en dimension 1. Il existe deux points sur la terre où la température (pression, qualité de l'air,...) est égale (exo 4 [GOU]).

(c) Connexité par arcs

Définition chemin entre deux éléments. Relation d'équivalence. Définition connexité par arcs. Lien entre connexité par arcs et connexité. Cas des ouverts d'un evn. Quelques exemples du [QUE]. Partie connexe mais non connexe par arcs.

(d) Composantes connexes

Définition des composantes connexes. Propriétés. Condition suffisante. Un ouvert de $\mathbb{R}$ s'écrit de manière unique comme réunion disjointes d'intervalles ouverts. Théorème de Jordan.

Ensemble totalement discontinu. Exemple : ensemble triadique de Cantor.

II Applications

(a) en algèbre linéaire

Connexité par arcs de $\mathrm{GL}_n(\mathbb{C})$. Composantes connexes de $\mathrm{GL}_n(\mathbb{R})$ et application [FGN,QUE]. Composantes connexes de l'ensemble des formes quadratiques non dégénérées. Image de l'exponentielle.

(b) en analyse

Si la différentielle de f est nulle sur un connexe alors f est constante. Théorème d'Hadamard-Lévy.

Théorème de prolongement analytique. Principe d'incertitude de Heisenberg [QUE]. Prolongement de Γ.

(c) en topologie

Homotopie. Espace simplement connexe. Primitive holomorphe, logarithme holomorphe. Lemme de non-retractation, théorème de Brouwer en dimension 2.

Développements

(i) Théorème d'Hadamard-Lévy

(ii) Composantes connexes des formes quadratiques non dégénérées

Références

(i) Quéffélec, Topologie

(ii) Gourdon, Analyse

(iii) Zuily & Quéffélec, Analyse pour l'agrégation

(iv) FGN, Algèbre 3

(v) Rouvière, PGCD

205 : Espaces complets. Exemples et applications.

Rapport du jury

“Les candidats devraient faire apparaître que l’un des intérêts essentiels de la complétude est de fournir des théorèmes d’existence : que ce soit tout simplement dans $\mathbb{R}$ ou $\mathbb{C}$ mais aussi dans certains espaces de dimension infinie (par exemple dans certains espaces de fonctions). Il est important de présenter des exemples d’espaces usuels, dont on sait justifier la complétude. Rappelons ici que l’on attend des candidats une bonne maîtrise de la convergence uniforme. Les espaces L^p sont des exemples pertinents qui ne sont pas sans danger pour des candidats aux connaissances fragiles. On peut évoquer dans cette leçon des théorèmes classiques tels que le théorème de Cauchy-Lipschitz ou le théorème du point fixe des applications contractantes.

On ne s’aventurera pas à parler du théorème de Baire sans application pertinente et maîtrisée ; elles sont nombreuses. Rappelons à ce propos que la démonstration détaillée de l’existence d’une partie dense de fonctions continues dérivables en aucun point est délicate.”

Introduction

La complétude fournit des théorèmes d’existence et palie le manque de compacité de certains espaces (en dimension infinie par exemple).

Plan

Soit (E, d) un espace métrique.

I Généralités sur la complétude

(a) Définition et exemples

Définition suites de Cauchy,

Espace métrique complet. Espaces de Banach. Exemples : $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^d$. Contre-exemple : $\mathbb{Q}$.

Les espaces complets sont fermés, et les fermés dans un espace complet sont complets. Le produit fini d’espaces complets est complet.

(b) Prolongement d’applications uniformément continues

Critère de Cauchy. Théorème de prolongement. Prolongement de Fourier-Plancherel. Unicité du complété.

II Le théorème du point fixe de Picard et quelques applications

Théorème du point fixe. Exemple simple d’utilisation. Application : théorème de Cauchy-Lipschitz, théorème d’inversion locale.

III Les espaces de Banach

(a) Généralités

Exemples : $\mathcal{B}(X, \mathbb{R})$, $\mathcal{L}_c(E, F)$, L^p.

Caractérisation des espaces de Banach (séries absolument convergentes). Étude de l’inverse dans les espaces de Banach.

(b) Les espaces de Hilbert

Théorème de projection sur un convexe fermé. Théorème de Riesz-Fréchet. Application : définition de l’espérance conditionnelle sur L^2. Application : Théorèmes de Hahn-Banach dans un Hilbert.

IV Le théorème de Baire et quelques applications

Théorème de Baire. Application : $\mathbb{K}[X]$ n'est pas complet Densité des fonctions continues nulle part dérivables.

Théorème de Banach-Steinhauss. Densité des fonctions dont la série de Fourier diverge sur tout $\mathbb{Q}$.

Théorème de l'application ouverte, théorème de Banach. Application : la transformée de Fourier n'est pas surjective.

Développements

(i) Densité des fonctions continues nulle part dérivables

(ii) Théorème de projection sur convexe fermé et théorème de Riesz-Fréchet

Références

(i) Gourdon, Analyse

(ii) Brézis, Analyse fonctionnelle

208 : Espaces vectoriels normés, applications linéaires continues. Exemples.

Rapport du jury

“Une telle leçon doit bien sûr contenir beaucoup d’illustrations et d’exemples, notamment avec quelques calculs élémentaires de normes subordonnées. Lors du choix de ceux-ci (le jury n’attend pas une liste encyclopédique), le candidat veillera à ne pas mentionner des exemples pour lesquels il n’a aucune idée de leur pertinence et à ne pas se lancer dans des développements trop sophistiqués.

La justification de la compacité de la boule unité en dimension finie doit être maîtrisée. Il faut savoir énoncer le théorème de Riesz sur la compacité de la boule unité fermée d’un espace vectoriel normé. Le théorème d’équivalence des normes en dimension finie, ou le caractère fermé de tout sous-espace de dimension finie d’un espace normé, sont des résultats fondamentaux à propos desquels les candidats doivent se garder des cercles vicieux. A contrario, des exemples d’espaces vectoriels normés de dimension infinie ont leur place dans cette leçon et il faut connaître quelques exemples de normes usuelles non équivalentes, notamment sur des espaces de suites ou des espaces de fonctions.

Pour aller plus loin, on peut éventuellement considérer le cas d’espaces métrisables mais dont la métrique n’est pas issue d’une norme, par exemple dans le champ des espaces de fonctions analytiques (topologie de la convergence uniforme sur tout compact par exemple).”

Introduction

Plan

I Généralités

(a) Espaces vectoriels normés

Définition norme. Exemples : $\mathbb{K}^n, \ell^2(I), \ell^p(I), \ell^\infty, L^p$. Normes équivalentes. Exemples et contre exemples.

(b) Applications linéaires continues

Définition, continuité, normes d’opérateurs, struture d’algèbre normée. Exemples.

(c) Complétude

Espace de Banach. Exemples : L^p, $\mathcal{L}_c(E,F)$. Propriétés de l’inverse.

(d) Dualité

Définition dual, forme linéaire. Caractérisation de la continuité par le noyau des formes linéaires. Théorème de Hahn-Banach. Caractérisation de la densité.

II La dimension finie

Équivalence des normes en dimension finie. Conséquences : continuité des applications linéaires, complétude, compacts. Contre exemples. Lemme de Riesz. Théorème de Riesz.

(a) Lemme de Baire et applications

Théorème de Baire. Application aux polynômes. Théorème de Banach-Steinhauss. Théorèmes de l’application ouverte, de Banach. Applications aux séries de Fourier qui divergent, aux fonctions dont l’interpolée de Lagrange ne converge pas.

III Les espaces de Hilbert

Définition espaces de Hilbert. Exemples. Projection sur un convexe fermé et lemme de Riesz. Propriétés de l’opérateur de projection. Théorème de Parceval. Application pour les séries de Fourier.

IV Spectre des opérateurs linéaires continus

Opérateurs compacts. Exemple : opérateur à noyau. Opérateur adjoint. Opérateur auto-adjoint. Exemple : opérateur à noyau symétrique. Théorème de Schauder. Alternative de Fredholm. Spectre des opérateurs compacts. Spectre des opérateurs auto-adjoints. Diagonalisation des opérateurs auto-adjoint et compacts. Exemple : opérateur de Fourier-Plancherel.

Développements

(i) Alternative de Fredholm
(ii) Projection sur un convexe fermé et théorème de Riesz-Fréchet

Références

(i) Gourdon, Analyse
(ii) Brézis, Analyse fonctionnelle
(iii) Hirsch & Lacombe, Éléments d'analyse fonctionnelle
(iv) Zavidovique, Un max de maths

209 : Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications.

Rapport du jury

"Cette leçon comporte un certain nombre de classiques comme par exemple les polynômes de Bernstein, éventuellement agrémenté d'une estimation de la vitesse de convergence (avec le module de continuité). Il n'est pas absurde de voir la formule de Taylor comme une approximation locale d'une fonction par des polynômes. Les polynômes d'interpolation de Lagrange peuvent être mentionnés en mettant en évidence les problèmes qu'ils engendrent du point de vue de l'approximation.

Pour aller plus loin, le théorème de Fejér (versions L^1, L^p ou $\mathcal{C}(\mathbb{T})$) offre aussi la possibilité d'un joli développement, surtout s'il est agrémenté d'applications (polynômes trigonométriques lacunaires, injectivité de la transformée de Fourier sur L^1, ...), mais on peut aussi s'intéresser à la convolution avec d'autres noyaux."

Introduction

Plan

Simplicité des polynômes. Mettre en avant l'aspect pratique et l'aspect théorique de l'approximation polynomiale.

I Généralités sur l'approximation polynomiale

(a) Approximation locale

Formules de Taylors, inégalité de Taylor-Lagrange. Fonctions analytiques. Exemples simples (exp, sin).

(b) Théorèmes de Stone-Weierstrass et de Weierstrass

Soit (X, d) un espace métrique compact. Définition de $\mathcal{C}(X, \mathbb{K})$, de la norme uniforme. Théorèmes de Stone-Weierstrass (cas réel et complexe) et Weierstrass. Contre-exemple si X n'est pas compact (dans $\mathbb{R}$, dans $\mathbb{T}$). Application : $\mathcal{C}(X, \mathbb{K})$ est séparable.

(c) Approximation uniforme sur un segment

Polynôme de meilleure approximation uniforme. Application aux polynômes de Tchebychev. Module de continuité, propriétés, polynômes de Bernstein, théorème de Bernstein. Théorème de Jackson.

II Interpolation polynomiale

(a) Interpolation de Lagrange

Définition polynômes de Lagrange, interpolation de Lagrange. Formule d'erreur. Exemple : points équidistants et points de Tchebychev.

(b) Stabilité et convergence

Opérateur de Lagrange, constante de Lebesgue. Propriétés. Exemple : points équidistants et points de Tchebychev, cas général. Application du théorème de Banach-Steinhauss à l'opérateur de Lagrange. Phénomène de Runge. Cas des fonctions lipschitziennes.

(c) Application : méthode de quadrature de Newton-Cotes

Principe d'une méthode. Ordre d'une méthode. Méthode de Newtwon-Cotes. Exemples.

III Approximation quadratique

(a) Généralités

Défintiion d'un poids. Définition de $L^2(I,\rho)$. Définiton polynôme de meilleure approximation quadratique.

(b) Polynômes orthogonaux

Existence et unicité des polynômes orthogonaux. Formule de récurrence. Exemple : Laguerre, Hermite, Legendre, Tchebychev. Étude des racines. Conditions suffisantes de base hilbertienne.

(c) Application : méthode de Gauss

IV Approximation par des polynômes trigonométriques

(a) Généralités

Définition de $L^p(\mathbb{T})$, des polynômes trigonométriques, des coefficients de Fourier, de $S_N f$, de $\sigma_N f$.

(b) Convergence des séries de Fourier

Noyau de Dirichlet, de Fejér, théorème de Dirichlet, théorème de Fejér. Corollaires : Weierstrass trigonométrique, injectivité de l'opérateur de Fourier. Application de l'injectivité : si $(c_n(f))_{n\in\mathbb{Z}}$ est cva alors f a un unique représentant continu. Théorème de Parceval. Application aux fonctions $\mathcal{C}^1$ par morceaux. Application au calcul de $\zeta(2)$. Vitesse de convergence en fonction de la régularité.

(c) Noyau de Jackson et applications

Définition noyau de Jackson, propriétés. Théorème de Jackson. Application aux fonctions hölderiennes.

Développements

(i) Densité des polynômes orthogonaux
(ii) Théorème de Jackson

Références

(i) Demailly, Analyse numérique et équations différentielles
(ii) Hirch & Lacombe, Éléments d'analyse fonctionnelle
(iii) Zuily & Quéffélec, Analyse pour l'agrégation
(iv) BMP, Objectif Agrégation

213 : Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.

Rapport du jury

"Il est bon de connaître et savoir justifier le critère de densité des sous-espaces par passage à l'orthogonal. Il faut aussi illustrer la leçon par des exemples de bases hilbertiennes (polynômes orthogonaux, séries de Fourier, ...).

Il est important de faire la différence entre base algébrique et base hilbertienne. De plus, la formule de la projection orthogonale sur un sous espace de dimension finie d'un espace de Hilbert doit absolument être connue de même que l'interprétation géométrique de la méthode de Gramm-Schmidt. Le théorème de projection sur les convexes fermés (ou sur un sous-espace vectoriel fermé) d'un espace de Hilbert H est régulièrement mentionné. Les candidats doivent s'intéresser au sens des formules

$$x = \sum_{i \geq 0} \langle x, e_i \rangle \, e_i \quad \text{et} \quad \|x\|_2^2 = \sum_{i \geq 0} |\langle x, e_i \rangle|^2$$

en précisant les hypothèses sur la famille $(e_n)_{n \in \mathbb{N}}$ et en justifiant la convergence. La notion d'adjoint d'un opérateur continu peut illustrer agréablement cette leçon.

Pour aller plus loin, le programme permet d'aborder la résolution et l'approximation de problèmes aux limites en dimension 1 par des arguments exploitant la formulation variationnelle de ces équations. Plus généralement, l'optimisation de fonctionnelles convexes sur les espaces de Hilbert peut être explorée. Enfin, le difficile théorème spectral pour les opérateurs auto-adjoints compacts peut être abordé."

Introduction

$\ell^2(I)$ est l'exemple d'espace de Hilbert en dimension infinie le plus simple mais aussi le plus général (cf. théorème de Parceval) ! Mettre en valeur le théorème de projection sur un convexe fermé. Évoquer le cas complexe et notamment les différences. Expliquer le principe de la formulation variationnelle.

Plan

On se place uniquement sur $\mathbb{R}$.

I Généralités sur les espaces de Hilbert

(a) Définitions et propriétés élémentaires

Définition espace préhilbertien réel. Exemples : $\ell^2(\mathbb{N})$, L^2, $\mathcal{C}(X, \mathbb{K})$. Inégalité de Cauchy-Schwarz. Identité du parallélogramme. Orthogonalité de deux vecteurs, orthogonal d'une partie. Définition espace de Hilbert. Retour sur les exemples.

(b) Projection sur un convexe fermé

Théorème de projection sur un convexe fermé. Application : Théorème de Hahn-Banach géométrique dans un Hilbert [BMP]. Propriétés de l'application projection. Conséquence : critère de densité. Application : densité de $\mathcal{C}_c$ dans L^2.

II Dualité

(a) Le théorème de Riesz-Fréchet

Théorème de Riesz-Fréchet. Application Hahn-Banach analytique dans un espace de Hilbert [BMP]. Théorème de Lax-Milgram.

Convergence faible. Théorème de Banach-Alaoglu.

(b) Adjoint d'un opérateur

Adjoint d'un opérateur. Propriétés de l'adjoint. Exemples : Opérateurs à noyaux. Opérateurs auto-adjoints. Retour sur l'exemple, projection orthogonale.

III Bases hilbertiennes

(a) Définition

Définition famille orthonormale, base hilbertienne. Différence avec base algébrique. Exemple : $\ell^2(\mathbb{N})$. Procédé d'orthomalisation de Schmidt. Base hilbertienne dénombrable. Exemple : L^2 est séparable. Projection sur un sev de dimension finie (formule de la projection). Théorème de Parceval dans le cas séparable.

(b) Exemple 1 : les séries de Fourier

Définition de $L^p(\mathbb{T})$, des polynômes trigonométriques, des coefficients de Fourier, de $S_N f$, de $\sigma_N f$. Théorème de Féjer. Conséquences ($(e_n)_n$ est une base hilbertienne, formule de Parceval). Application : calcul de $\zeta(2)$.

(c) Exemple 2 : l'espace $L^2(I, \rho)$

Définition d'un poids. Définition de $L^2(I, \rho)$. Existence et unicité des polynômes orthogonaux. Formule de récurrence. Exemple : Laguerre, Hermite, Legendre, Tchebychev. Conditions suffisantes de base hilbertienne.

IV Théorie spectrale

Étude du spectre d'un opérateur auto-adjoint. Diagonalisation des opérateurs auto-adjoints et compacts. Remarque : limite d'opérateur de rang fini.

Développements

(i) Densité des polynômes orthogonaux
(ii) Projection sur un convexe fermé et théorème de Riesz-Fréchet

Références

(i) Brézis, Analyse fonctionnelle
(ii) Hirsch & Lacombe, Éléments d'analyse fonctionnelle
(iii) BMP, Objectif Agrégation

214 : Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.

Rapport du jury

"Il s'agit d'une belle leçon, formulée ici dans la version qui sera adoptée pour la session 2017, qui exige une bonne maîtrise du calcul différentiel. Même si le candidat ne propose pas ces thèmes en développement, on est en droit d'attendre de lui des idées de démonstration des deux théorèmes fondamentaux qui donnent son intitulé à la leçon. Il est indispensable de savoir mettre en pratique le théorème des fonctions implicites au moins dans le cas de deux variables réelles. On attend des applications en géométrie différentielle notamment dans la formulation des multiplicateurs de Lagrange. Plusieurs inégalités classiques de l'analyse peuvent se démontrer avec ce point de vue : arithmético- géométrique, Hölder, Carleman, Hadamard, . . . En ce qui concerne la preuve du théorème des extrema liés, la présentation de la preuve par raisonnement "sous-matriciel" est souvent obscure ; on priviligiera si possible une présentation géométrique s'appuyant sur l'espace tangent.

Pour aller plus loin, l'introduction des sous-variétés est naturelle dans cette leçon. Il s'agit aussi d'agrémenter cette leçon d'exemples et d'applications en géométrie, sur les courbes et les surfaces."

Introduction

Lire Rouvière. La grande idée du calcul différentiel est d'approcher localement une fonction quelconque par une application différentielle. Analogie avec le cas de la dimension 1. Illustrer heuristiquement le TIL. Illustrer géométriquement le TFI.

Plan

I Le théorème d'inversion locale

(a) Le théorème et ses variantes

Définition difféomorphisme local. Théorème d'inversion locale. Condition Nécessaire. Exemple simple [ROU]. Difféomorphisme local implique application ouverte. Généralisation du théorème de la base incomplète. Application à la résolution d'une EDP. Théorème d'inversion globale (version $\mathcal{C}^k$, version holomorphe). Théorème d'Hadamard-Lévy.

(b) Applications

$\lambda_1 < \cdots < \lambda_{n+1} \longmapsto \prod(X - \lambda_i) - X^n \in \mathbb{R}_n[X]$ induit un $\mathcal{C}^\infty$-difféomorphisme.

Changement de coordonnées, paramétrisation. Exemple : changement linéaire, polaire, hypersphérique. Racine carrée (voire k^{ieme}) [LAF]. exp est un difféo local en 0. Application : surjectivité de exp, sous-groupe à un paramêtre, paramétrisation de certains sous-groupes de $\mathrm{GL}_n(\mathbb{R})$.

Lemme formes quadratiques. Lemme de Morse.

II Le théorème des fonctions implicites

(a) Le théorème

Théorème des fonctions implicites. Exemple sur le cercle. Différentielle de la fonction implicite. Exemple sur le cercle. Équivalence avec le théorème d'inversion locale. Lien avec un cas particulier du théorème de Cauchy-Lipschitz.

(b) Applications

Variation du point fixe [ROU]. Asymptotique d'une équation du troisème degré.

III Sous-variétés de $\mathbb{R}^n$

(a) Définitions

Définitions équivalentes d'une sous-variété. Exemples : sphère, tore, groupe orthogonal, groupe spécial linéaire. Contre-exemple : point de rebroussement, cône de révolution.

(b) Espace tangent

Définition vecteur tangent, espace tangent. L'espace tangent est un espace vectoriel. Espace tangent dans le cadre de la définition par immersion, submersion, graphe. Interprétation géomtrique. Retour sur les exemples.

(c) Application : le théorème des extrema liés

Théorème des extrema liés. Application : existence d'une valeur propre réelle dans $\mathcal{S}_n(\mathbb{R})$. Application : inégalité arithmético-géométrique. Application : inégalité d'Hadamard. Application : entropie maximum.

Développements

(i) Théorème de Hadamard-Lévy
(ii) Lemme de Morse

Références

(i) Rouvière, PGCD
(ii) Lafontaine, Introduction aux variétés différentielles
(iii) Avez, Calcul différentiel
(iv) Gourdon, Analyse
(v) Testard, La maîtrise de l'implicite

215 : Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.

Rapport du jury

"Cette leçon requiert une bonne maîtrise de la notion de différentielle première et de son lien avec les dérivées partielles. On doit pouvoir mettre en pratique le théorème de différentiation composée pour calculer des dérivées partielles de fonctions composées dans des situations simples (par exemple le laplacien en coordonnées polaires). La différentiation à l'ordre 2 est attendue, notamment pour les applications classiques quant à l'existence d'extrema locaux. On peut aussi faire figurer dans cette leçon la différentielle d'applications issues de l'algèbre linéaire (ou multilinéaire).

Pour aller plus loin, l'exponentielle matricielle est une ouverture pertinente. D'autres thèmes issus de la leçon 214 trouvent aussi leur place ici."

Introduction

Lire Rouvière pour la motivation du calcul différentiel. Il faut faire des dessins.

Plan

$E \subset \mathbb{R}^n$ et $F \subset \mathbb{R}^m$ deux ouverts.

I Différentiabilité

(a) Différentielle

Définition de la différentielle en un point. Unicité. Exemples : fonction dérivables, applications constantes, linéaires, quadratiques. Différentielle de l'inverse dans $\mathrm{GL}_n(\mathbb{R})$, du déterminant. Définition gradient. Interprétation géométrique.

Linéarité de la différenciation. Lemme de composition. Exemple d'utilisation : différentielle de la réciproque + [ROU]. Définition de la classe $\mathcal{C}^1$, d'un difféomorphisme.

(b) Dérivées partielles

Dérivation selon une direction. Différentiable implique dérivée selon toutes les directions. Contre-exemples ! Définition dérivées partielles. Expression de la différentielle avec les dérivées partielles. Jacobienne & Jacobien.

Définition de la classe $\mathcal{C}^k$. Hessienne.

(c) Inégalité de la moyenne

Inégalité de la moyenne. Application : critère fonction constante. Application : fonction contractante + théorème du point fixe (exemple d'utilisation). Application : Caractérisation des fonctions $\mathcal{C}^1$. Exemple et contre-exemple. Application : Théorème de Schwarz. Conséquences sur la Hessienne.

II Applications

(a) Théorème d'inversion locale et théorème des fonctions implicites

Théorème d'inversion locale. Application : racine $k^{ième}$ d'une matrice [BMP]. Théorème d'inversion globale. Exemple ($f : (x_1, \ldots, x_n) \longmapsto (s_1, \ldots, s_n)$) et/ou contre-exemple ($f(x, y) = (x^2 - y^2, xy)$) [ROU]. Théorème d'Hadamard-Levy [ZQ].

Théorème des fonctions implicites. Faire un dessin. Application : Régularité d'une racine simple d'un polynôme [BMP].

(b) Théorème de changement de variables

Théorème de changement de variables. Exemples : coordonnées polaires. Application : calcul de l'intégrale de Gauss.

(c) Formules de Taylor

Formules de Taylors. Application : lemme d'Hadamard [BMP]. Lemme de Morse. Interprétation en termes de courbes de niveau.

(d) Recherche d'extrema

Conditions suffisantes, nécessaire du premier et second ordre. Application : principe du maximum faible ou moindres carrés [ROU] ou exemples et contre-exemples. Application : méthode du gradient à pas optimal.

Extrema liés. Application : existence d'une valeur propre d'une matrice symétrique réelle. Application : maximum de l'entropie [ROU].

Développements

(i) Théorème de Hadamard-Lévy

(ii) Lemme de Morse

Références

(i) Rouvière, PGCD

(ii) Gourdon, Analyse

(iii) BMP, Objectif Agrégation

(iv) Zuily & Quéffélec, Analyse pour l'agrégation ($3^{\text{ème}}$ édition)

218 : Applications des formules de Taylor.

Rapport du jury

"Il faut connaître les formules de Taylor et certains développements très classiques. En général, le développement de Taylor d'une fonction comprend un terme de reste qu'il est crucial de savoir analyser. Le candidat doit pouvoir justifier les différentes formules de Taylor proposées ainsi que leur intérêt. Le jury s'inquiète des trop nombreux candidats qui ne savent pas expliquer clairement ce que signifient les notations o ou O qu'ils utilisent. De plus la différence entre l'existence d'un développement limité à l'ordre deux et l'existence de dérivée seconde doit être connue. On peut aussi montrer comment les formules de Taylor permettent d'établir le caractère développable en série entière (ou analytique) d'une fonction dont on contrôle les dérivées successives.

Pour aller plus loin, on peut mentionner des applications en algèbre bilinéaire (lemme de Morse), en géométrie (étude locale au voisinage des points stationnaires pour les courbes et des points critiques pour la recherche d'extrema) et, même si c'est plus anecdotique, en probabilités (Théorème central limite). On peut aussi penser à la méthode de Laplace, du col, de la phase stationnaire ou aux inégalités contrôlant les dérivées intermédiaires lorsque f et sa dérivée n-ième sont bornées. On soignera particulièrement le choix des développements."

Introduction

Les formules de Taylor c'est génial c'est génial c'eset génial !

Plan

I Les formules de Taylor et quelques applications

Énoncer les formules de Taylor [GOU]. Formule de Taylor-Lagrange fausse en dimension supérieure à 2 [ROU]. Lemme d'Hadamard. Inégalité de Kolmogorov [FGN analyse 1]. Fonctionnelle elliptique implique coercive, strictement convexe.œ

II Du local au global

(a) Développement limité

Condition suffisante d'existence de développement limité [GOU]. Contre-exemple de la réciproque. Application des développements limité à l'étude d'une suite récurrente [FGN 1].

(b) Développement en série entière

Condition nécessaire et suffisante de DSE [GOU]. En pratique. Exemples et contre-exemples (application à l'exponentielle).

(c) La méthode de Laplace

III Étude d'extrema et application à la géométrie

(a) Conditions nécessaires et conditions suffisantes d'extrema

(b) Étude locale de courbe en des points critiques

Étude affine d'une courbe plane [ROU p.315].
Lemme sur les formes quadratiques. Lemme de Morse. Interprétation en termes de courbes de niveau.

IV Application en analyse numérique

Méthode d'Euler. Calcul de l'erreur de consistance [DEM].
Méthode de Newton. Application au calcul de la racine carrée.

V Applications en probabilités

Définition fonction caractéristique, développement limité et moments. TCL multi-dimensionnel. Application au test du χ^2.

Développements

(i) Méthode de Newton
(ii) Lemme de Morse

Références

(i) Gourdon, Analyse
(ii) Rouvière, PGCD
(iii) Demailly Analyse numérique et équations différentielles
(iv) BMP, Objectif Agrégation

219 : Extremums : existence, caractérisation, recherche. Exemples et applications.

Rapport du jury

“Comme souvent en analyse, il peut être opportun d’illustrer dans cette leçon un exemple ou un raisonnement à l’aide d’un dessin. Il faut savoir faire la distinction entre propriétés locales (caractérisation d’un extremum) et globales (existence par compacité, par exemple). Dans le cas important des fonctions convexes, un minimum local est également global. Les applications de la minimisation des fonctions convexes sont nombreuses et elles peuvent illustrer cette leçon.

L’étude des algorithmes de recherche d’extremums y a toute sa place : méthode de gradient, preuve de la convergence de la méthode de gradient à pas optimal, . . .Le cas particulier des fonctionnelles sur $\mathbb{R}^n$ de la forme $\frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$, où A est une matrice symétrique définie positive, ne devrait pas poser de difficultés. Les problèmes de minimisation sous contrainte amènent à faire le lien avec les extremums liés, la notion de multiplicateur de Lagrange et, là encore, des algorithmes peuvent être présentés et analysés. À ce sujet, une preuve géométrique des extrema liés sera fortement valorisée par rapport à une preuve algébrique, formelle et souvent mal maitrisée.

Les candidats pourraient aussi être amenés à évoquer les problèmes de type moindres carrés, ou, dans un autre registre, le principe du maximum et ses applications.”

Introduction

Mettre en avant l’utilisation de la compacité, sa facilité d’utilisation en dimension finie et au contraire, le problème que cela pose en dimension infinie (d’où la recherche de solution faible).

Plan

I Généralités

Soit X un espace métrique et $F : X \longrightarrow \mathbb{R}$ une application. Définition maximum, minimum, extremum, stric ou pas, local et global.

(a) Utilisation de la compacité

Existence par compacité. Coercivité et minimum. Distance entre deux parties [GOU]. Théorème du point fixe sur un compact [GOU]. Application à l’existence du polynôme de meilleure approximation uniforme (et unicité). Ellipsoïde de John-Loewner (existence).

(b) Dans un espace de Hilbert

Théorème de projection sur un convexe fermé non vide. Application à la définition de l’espérance conditionnelle dans L^2. Théorème de Lax-Milgram et application à la recherche d’une solution faible d’une EDP.

II Extrema de fonctions différentiables

(a) Extrema libres

Conditions nécessaires et conditions suffisantes d’extrema. Principe du maximum faible. Exemples [ROU].

(b) Extrema liés

Théorème des extrema liés. Application : existence d’une valeur propre réelle dans $\mathcal{S}_n(\mathbb{R})$. Application : inégalité arithmético-géométrique. Application : inégalité d’Hadamard. Application : entropie maximum.

(c) Utilisation de la convexité

Propositions [CIA]. Point de Fermat [BMP, ROU]. Méthode du gradient à pas optimal.

Développements

(i) Principe du maximum faible
(ii) Gradient à pas optimal

Références

(i) Gourdon, Analyse
(ii) Rouvière, PGCD
(iii) Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation
(iv) BMP, Objectif Agrégation

220 : Équations différentielles $X' = f(t, X)$. Exemples d'études des solutions en dimension 1 et 2.

Rapport du jury

“C'est l'occasion de rappeler une nouvelle fois que le jury s'alarme des nombreux défauts de maîtrise du théorème de Cauchy-Lipschitz. Il est regrettable de voir des candidats ne connaître qu'un énoncé pour les fonctions globalement lipschitziennes ou plus grave, mélanger les conditions sur la variable de temps et d'espace. La notion de solution maximale et le théorème de sortie de tout compact sont nécessaires. Bien évidemment, le jury attend des exemples d'équations différentielles non linéaires. Le lemme de Grönwall semble trouver toute sa place dans cette leçon mais est curieusement rarement énoncé. L'utilisation du théorème de Cauchy-Lipschitz doit pouvoir être mise en œuvre sur des exemples concrets. Les études qualitatives doivent être préparées et soignées.

Pour les équations autonomes, la notion de point d'équilibre permet des illustrations de bon goût comme par exemple les petites oscillations du pendule. Trop peu de candidats pensent à tracer et discuter des portraits de phase alors que le sujet y invite clairement.

Pour aller plus loin, il est possible d'évoquer les problématiques de l'approximation numérique dans cette leçon en présentant le point de vue du schéma d'Euler. On peut aller jusqu'à aborder la notion de problèmes raides et la conception de schémas implicites pour autant que le candidat ait une maîtrise convenable de ces questions.”

Introduction

Plan

I Généralités sur les équations différentielles

(a) Définitions et hypothèses

Hypothèses sur F (continuité, domaine de définition, ...). Définiton d'une solution, d'une solution maximale, de prolongement de solutions. Existence d'une solution maximale s'il existe une solution.

Régularité des solutions.

(b) Existence et unicité de solutions locales

Définition d'un problème de Cauchy. Application localement lipschitzienne. Cas des fonctions $\mathcal{C}^1$. Cylindre de sécurité. Théorème de Cauchy-Lipschitz local précisé. Exemple : $y' = 3\,|y|^{2/3}$.

Théorème de Cauchy-Peano-Arzela. Exemple : $y' = \sqrt{y}$.

(c) La théorie globale

Définition d'une solution globale. Contre exemple : globale $\neq$ maximale.

Théorème de Cauchy-Lipschitz global. Interprétation géométrique.

Critère de maximalité (théorème de sortie de tout compact). Une condition suffisante d'existence de solutions globales.

Exemples : système de Lokta-Voltera [FGN], pendule simple (équation du pendule simple, globalité de la solution, disjonction de cas en fonction de la vitesse initiale [FGN]).

(d) Continuité par rapport aux conditions initiales

Définition flot. Théorème de continuité. Application : théorème d'Hadamard-Lévy.

II Le cas des équations différentielles linéaires

Définition système linéaire. Théorème de Cauchy-Lipschitz linéaire. L'espace vectoriel (ou affine) des solutions. Résolution d'un système linéaire à coefficients constants. Application : résolution d'une équation matricielle. Application : étude des points critiques d'un champs de vecteurs dans $\mathbb{R}^2$.

Théorème de Liapunov.

III Approximation numérique : la méthode d'Euler

Définition des méthodes numériques à un pas. Erreur de consistance. Description de la méthode d'Euler. Majoration de l'erreur. Convergence (à l'aide du théorème d'Ascoli). Erreur de consistance.

Développements

(i) Théorème d'Hadamard-Lévy

(ii) Nombre de zéros d'une équation différentielle linéaire

Références

(i) Demailly, Analyse numérique et équations différentielles

(ii) Zuily & Quéffélec, Analyse pour l'agrégation

(iii) Gourdon, Analyse

(iv) FGN, Analyse 4

(v) Rouvière, PGCD

221 : Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.

Rapport du jury

"Le jury attend d'un candidat qu'il sache déterminer rigoureusement la dimension de l'espace vectoriel des solutions. Le cas des systèmes à cœfficients constants fait appel à la réduction des matrices qui doit être connue et pratiquée. Le jury attend qu'un candidat puisse mettre en œuvre la méthode de variation des constantes pour résoudre une équation différentielle linéaire d'ordre 2 simple (à coefficients constants par exemple) avec second membre.

L'utilisation des exponentielles de matrices a toute sa place ici. Les problématiques de stabilité des solutions et le lien avec l'analyse spectrale devraient être exploitées.

Le théorème de Cauchy-Lipschitz linéaire constitue un exemple de développement pertinent pour cette leçon. Les résultats autour du comportement des solutions, ou de leurs zéros, de certaines équations linéaires d'ordre 2 (Sturm, Hill-Mathieu, ...) sont aussi d'autres possibilités."

Introduction

Plan

I Étude générale

(a) Existence et unicité des solutions

Définition système différentiel linéaire d'ordre p. Remarque sur l'ordre (il suffit d'étudier l'ordre 1). Théorème de Cauchy-Lipschitz linéaire. Exemple : système linéaire à coefficients constants.

(b) L'espace des solutions

Espace des solutions, structure (espace vectoriel, affine, dimension). Résolvante et propriétés. Exemple [DEM]. Wronskien et propriétés. Calcul du Wronskien à l'aide de la résolvante.

II Résolution explicite

(a) Méthode de la variation de la constante

[GOU]. Méthode + exemple (exo 1).

(b) Système différentiel à coefficients constants : le cas général

Solution sans second membre. Application : équation matricielle. Solution avec second membre avec la variation de la constante. Exemples (exo 2 [GOU]).

(c) Système différentiel linéaires d'ordre p à coefficients constants sur $\mathbb{K}$

[DEM]

(d) Solution développable en série entière

[ZQ]

III Étude qualitative

(a) Stabilité et points singuliers

[DEM] + Liapunov.

(b) Équation du type : $y'' + py' + qy = r$

Nombre de zéros + lemme d'entrelacement de Sturm.

Développements

(i) Nombre de zéors d'une équation différententielle linéaire
(ii) Équation matricielle

Références

(i) Demailly, Analyse numérique et équations différentielles
(ii) Zuily & Quéffélec, Analyse pour l'agrégation
(iii) Gourdon, Analyse
(iv) FGN, Analyse 4

223 : Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications

Rapport

"Cette leçon permet souvent aux candidats de s'exprimer. Il ne faut pas négliger les suites de nombres complexes. Le théorème de Bolzano-Weierstrass doit être cité et le candidat doit être capable d'en donner une démonstration. On attend des candidats qu'ils parlent des limites inférieure et supérieure d'une suite réelle bornée, et qu'ils en maîtrisent le concept. Les procédés de sommation peuvent être éventuellement évoqués mais le théorème de Cesàro doit être mentionné et sa preuve maîtrisée par tout candidat à l'agrégation. Les résultats autour des sous-groupes additifs de $\mathbb{R}$ permettent d'exhiber des suites denses remarquables et l'ensemble constitue un joli thème.

Pour aller plus loin, un développement autour de l'équirépartition est tout à fait envisageable."

Introduction

Plan

I Généralités sur les suites numériques

(a) Limite d'une suite

Définiton d'une suite croissante, bornée, convergente vers une limite, vers $\pm\infty$. Propriétés de la limite (unicité, linéarité, stabilité par produit, par quotient, conservation des inégalités larges). Monotonie et limite. Exemple [FGN].

Théorème des gendarmes, suites adjacentes. Exemple : suite arithmético-géométrique [GOU]. Lemme de Césaro. Application : $(x_1 \cdots x_n)^{1/n}$.

Caractérisation séquentielle de la continuité.

(b) Valeurs d'adhérence

Définition et contre-exemple ($n(1+(-1)^n)$ ou $(-1)^n$ ne converge pas mais a une valeur d'adhérence). Lien avec la limite. Limite sup et inf, liens avec les valeurs d'adhérence.

Théorème de Bolzano-Weierstrass. Application : problème de convergence [FGN p.75]. Application : caractérisation des compacts en dimension finie. Application : fonction continue sur un compact à valeurs réelles.

(c) Suites de Cauchy

Définition. Propriétés. $\mathbb{R}$ et $\mathbb{C}$ sont complets. Application : caractérisation des espaces de Banach par les suites absolument convergentes [GOU].

II Quelques exemples de suites remarquables

(a) Suites arithmétiques et géométriques

(b) Suites homographiques

(c) Suites récurrentes

Définition générale + quelques exemples de systèmes dynamiques [FGN].

III Comparaison de suites

(a) Utilisation des notations de Landau

Domination, équivalents, négligiabilité. Passage au produit, au quotient (mais pas à la somme). Exemple : $\left(1+\frac{1}{n}\right)^n$. Application : intégrales de Wallis.

(b) Série numériques

Liens avec les séries numériques, comparaison série-intégrale. Sommation des équivalents, séries télé-scopiques. Application : formule de Stirling, développement asymptotique de la série harmonique.

IV Applications

(a) Sous-groupes additifs de $\mathbb{R}$

(b) Critère d'équirépartition

Critère de Weyl. Application.

(c) Méthode de Newton

Méthode de Newton sur un segment

Développements

(i) Méthode de Newton sur un segment
(ii) Développement asymptotique de la série harmonique

Références

(i) Gourdon, Analyse
(ii) Rouvière, PGCD
(iii) FGN, Analyse 1 et 2

228 : Continuité et dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications.

Rapport du jury

"Cette leçon permet des exposés de niveaux très variés. Les théorèmes de base doivent être maîtrisés et illustrés par des exemples intéressants, par exemple le théorème des valeurs intermédiaires pour la dérivée. Le jury s'attend à ce que le candidat connaisse et puisse calculer la dérivée des fonctions usuelles. Les candidats doivent disposer d'un exemple de fonction dérivable de la variable réelle qui ne soit pas continûment dérivable. La stabilité par passage à la limite des notions de continuité et de dérivabilité doit être comprise par les candidats. De façon plus fine, on peut s'intéresser aux fonctions continues nulle part dérivables.

Pour aller plus loin, la dérivabilité presque partout des fonctions lipschitziennes ou des fonctions monotones relève de cette leçon. Les applications du théorème d'Ascoli (avec, par exemple, des exemples d'opérateurs à noyaux compacts), sont les bienvenues. L'étude de la dérivée au sens des distributions $x \in [a\,,b] \mapsto \int_a^x f(t)\,\mathrm{d}t$ pour une fonction intégrable $f \in L^1([a\,,b])$, est un résultat intéressant."

Introduction

Plan

Soit A une partie de $\mathbb{R}$.

I Généralités sur la continuité et la dérivabilité

(a) Définition et premières propriétés

Définiton continuité. Caractérisation séquentielle. Exemples : fonctions polynômes. Définition dérivabilité. Développement limité. Stabilité de la continuité, de la dérivabilité par opération élémentaires. Classe $\mathcal{C}^k$. Exemple de fonction dérivable mais de dérivée non continue ($x \longmapsto x\sin\left(\frac{1}{x}\right)$). Formule de Leibniz. Exemples et contre-exemples.

(b) Liens entre continuité et dérivabilité

La dérivabilité implique la continuité. L'ensemble des points de continuité de la dérivée est dense.

Exemple de fonctions continues mais pas dérivables en un point. Exemple de Weierstrass de fonction continue nulle part dérivable. L'ensemble des fonctions continues nulle part dérivable est dense.

II Théorèmes fondamentaux

(a) Théorème des valeurs intermédiaires

Théorème des valeurs intermédiaires. $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes. Théorème de Darboux.

(b) Théorème de Rolle et des accroissements finis

Théorème de Rolle. Théorème des accroissements finis. Contre-exemple. Formule d'erreur polynômes de Lagrange.

(c) Théorème de Heine

Définition de l'uniforme continuité. Théorème de Heine. Théorème de Dini. Approximation de l'unité. Module de continuité. Propriétés.

III Formules de Taylor

Formules de Taylor (toutes). Méthode de Newton. Méthode de Laplace.

IV L'espace $\mathcal{C}(K)$

(a) Théorèmes de transfert

Points de continuité d'une limite simple de fonctions continues [ZQ]. Limite uniforme d'une suite de fonction continue est continue. Théorème analogue pour la dérivabilité. Exemple. Polynômes de Bernstein, théorème de Bernstein. Théorème de Weierstrass. Théorème des moments d'Hausdorff.

(b) Le théorème d'Arzela-Ascoli

Définition équicontinuité. Théorème d'Arzela-Ascoli. Application opérateur à noyaux compact.

Développements

(i) Densité des fonctions continues nulle part dérivables

(ii) Polynômes de Bernstein

Références

(i) Gourdon, Analyse

(ii) Zuily Quéffélec, Analyse pour l'agrégation

(iii) Rouvière, PGCD

(iv) Hirsch-Lacombe, Éléments d'analyse fonctionnelle

229 : Fonctions monotones. Fonctions convexes. Exemples et applications.

Rapport du jury

"L'énoncé et la connaissance de la preuve de l'existence de limites à gauche et à droite pour les fonctions monotones sont attendues. Ainsi on doit parler des propriétés de continuité et de dérivabilité à gauche et à droite des fonctions convexes de la variable réelle. Il est souhaitable d'illustrer la présentation de la convexité par des dessins clairs. On notera que la monotonie concerne les fonctions réelles d'une seule variable réelle, mais que la convexité concerne également les fonctions définies sur une partie convexe de $\mathbb{R}^n$, qui fournissent de beaux exemples d'utilisation.

Pour aller plus loin, la dérivabilité presque partout des fonctions monotones est un résultat remarquable (dont la preuve peut être éventuellement admise). L'espace vectoriel engendré par les fonctions monotones (les fonctions à variation bornée) relève de cette leçon. Enfin, la dérivation au sens des distributions fournit les caractérisations les plus générales de la monotonie et de la convexité ; les candidats maîtrisant ces notions peuvent s'aventurer utilement dans cette direction."

Introduction

Plan

I Fonctions monotones

(a) Définition et premières propriétés

Définition fonction croissante, décroissante, monotones. Exemples simples (fonctions affines, fonctions de répartition d'une var, ...).

Théorème des accroissements finis. Critère de monotonie pour les fonctions dérivables. Exemples simples.

(b) L'espace des fonctions monotones

Propriétés de stabilité (somme, multiplication par un réel positif, composition). Théorèmes de Dini et application à une suite récurrente [HL]. Théorème de Helly [GOU].

Fonctions à variation bornées [GOU].

(c) Monotonie et régularité

Caractérisation des fonctions croissantes dérivables. Exemples simples. Les fonctions monotones admettent des limites à gauche et des limites à droites. Les fonctions monotones sont dérivables presque partout.

II Fonctions convexes

Définition fonction (strictement) convexe, concave. Épigraphe et caractérisation de la convexité. Exemple : norme sont convexes mais pas stritement. Caractérisation(s) de la convexité dans le cas différentiable. Exemple : exp, log, $\sin|_{[0,\frac{\pi}{2}]}$, $x \longmapsto x^p$ pour $p \geq 1$.

III Applications

(a) Processus de branchement

Processus de Galton-Watson.

(b) Quelques inégalités de convexité remarquables

Inégalité arithmético-géométrique. Inégalité de Hölder. Application : inégalité de Minkowski, $L^p \subset L^q$ si la mesure est finie et si $p \leq q$. Inégalité de Young [Brezis]. Inégalité de Jensen. Application à la convolution, application à l'entropie. log-convexité du déterminant. Application : inégalité de Brunn-Minkowski [FGN].

(c) Optimisation

Propositions [CIA]. Point de Fermat [BMP, ROU]. Méthode du gradient. Pas optimal, pas conjugué. Inéquation d'Euler pour la méthode des moindres carrés.

Application : unicité de l'ellipsoïde de John-Loewner.

Développements

(i) Processus de Galton-Watson

(ii) Gradient à pas optimal

Références

(i) Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation

(ii) Gourdon, Analyse

(iii) FGN, Algèbre 2

(iv) Brézis, Analyse fonctionnelle

(v) Hirsch & Lacombe, Éléménts d'analyse fonctionnelle

230 : Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

Rapport du jury

"De nombreux candidats commencent leur plan par une longue exposition des conditions classiques assurant la convergence ou la divergence des séries numériques. Sans être hors sujet, cette exposition ne doit pas former l'essentiel de la matière de la leçon. Le thème central de la leçon est en effet le comportement asymptotique des restes et sommes partielles (équivalents, développements asymptotiques — par exemple pour certaines suites récurrentes — cas des séries de Riemann, ...).

On peut aussi s'intéresser à certaines sommes particulières, que ce soit pour exhiber des nombres irrationnels (voire transcendants), ou mettre en valeur des techniques de calculs non triviales (par exemple en faisant appel aux séries de Fourier ou aux séries entières).

Enfin le jury apprécie que le théorème des séries alternées (avec sa version sur le contrôle du reste) soit maîtrisé, mais on rappelle aussi que la transformation d'Abel trouve toute sa place dans cette leçon."

Introduction

Plan

I Généralités sur les séries

(a) Définitions

Définition d'une série de terme général u_n, des sommes partielles, des restes. Définition de la convergence. Exemples : séries arthimétiques, séries géométriques. Si une suite converge alors son terme général tend vers 0. Contre exemple : série harmonique [HAU]. Critère de Cauchy et lien entre convergence d'une série et le reste qui tend vers 0.

Séries téléscopiques. Exemples : $\sum \frac{1}{(n(n+1)}$.

(b) Convergence des séries

Le cas des séries à termes positifs : la série converge si et seulement si la suite des sommes partielles est majorée, comparaison des termes généraux (contre exemples [HAU], exemple [FGN] avec le sinus). Séries de Riemann. Règles de d'Alembert, de Cauchy. Exemples, D'Alembert plus faible que Cauchy mais plus pratique, contre exemples lorsque la limite est 1, exemples d'utilisation (définition de e).

Séries absolument convergentes. Exemples : $\sum \frac{\cos(n)}{n^2}$. Contre exemple : $\sum \frac{(-1)^n}{n}$ [HAU]. Séries semi-convergente.

II Comportement des restes et des sommes partielles

(a) Séries à termes positifs

Comparaisons série-intégrale. Séries de Bertrand. Contre exemple [HAU]. Exo 5 [GOU].

Sommation des relations des comparaison. Application : développement asympototique de $u_n = \sin u_{n-1}$. Application : développement asymptotique de la série harmonique.

Comparaison des quotient successifs. Exemple : séries de Bertrand. Règle de Raabe-Duhamel. Contre-exemple avec ce qui précède. Application : formule de Stirling.

(b) Séries semi-convergentes

Séries alternée. Exemples : $\sum \frac{(-1)^n}{n^\alpha}$.

Transformation d'Abel. Exemple : $\sum \frac{e^{in\theta}}{n^\alpha}$. Application : théorème abélien radial, calcul de $\sum \frac{\sin(nt)}{n}$.

Comparaison série-intégrale et application (cas général [GOU]).

Développements

(i) Développement asymptotique de la série harmonique

(ii) Théorème abélien radial et application au calcul de $\sum \frac{\sin(nt)}{n}$

Références

(i) Gourdon, Analyse

(ii) FGN, Analyse 1

(iii) Hauchecorne, Contre exemples

(iv) Zuily & Quéffélec, Analyse pour l'agrégation

234 : Espaces L^p, $1 \leq p \leq +\infty$.

Rapport du jury

"Cette leçon nécessite d'avoir compris les notions de presque partout (comme par exemple les opérations sur les ensembles négligeables) et évidemment la définition des espaces L^p. Le jury a apprécié les candidats sachant montrer qu'avec une mesure finie $L^2 \subset L^1$ (ou même $L^p \subset L^q$ si $p \geq q$). Il est important de pouvoir justifier l'existence de produits de convolution comme par exemple le produit de convolution de deux fonctions de L^1). Par ailleurs, les espaces associés à la mesure de comptage sur $\mathbb{N}$ ou $\mathbb{Z}$ fournissent des exemples pertinents non triviaux à propos desquels des développements peuvent être proposés comme la description du dual. Par ailleurs, des exemples issus des probabilités peuvent tout à fait être mentionnés.

Pour aller plus loin, la complétude de L^p (p fini ou infini) offre aussi un bon développement. On peut aussi penser à certains résultats sur la dimension des sous-espaces fermés de L^p dont les éléments ont des propriétés particulières de régularité. Enfin, le cas particulier hilbertien $p = 2$ mérite attention mais il faut se concentrer sur les spécificités d'un espace de fonctions L^2 et éviter de faire un catalogue de propriétés vraies pour n'importe quel espace de Hilbert."

Introduction

Plan

Soit $(X, \mathcal{A}, \mu)$ un espace mesuré avec μ mesure σ-finie.

I Les espaces L^p : définition et premières propriétés

(a) L'espace L^1

Définition de $\mathcal{L}^1$, de L^1, de la norme $\|\cdot\|_1$. L^1 est un espace vectoriel normé. Exemples : mesure de comptage pour $\mathbb{N}$ ou $\mathbb{Z}$. Théorème de convergence dominée dans L^1.

(b) Les espaces de Banach L^p

Définition de L^p, de $\|\cdot\|_p$. Théorème de convergence dominée dans L^p. Inégalité de Hölder, de Minkowski. Application : $L^p \subset L^q$ si $p \geq q$ si la mesure est de masse finie. Application : inégalité d'interpolation [BRE]. Théorème de Riesz-Fischer. Réciproque partiel du théorème de convergence dominée. Conséquence : L^2 est un espace de Hilbert.

(c) Quelques théorèmes de densité

Densité des fonctions étagées, en escalier à support compact, de $\mathcal{C}_c$.

II Les espaces fonctionnels L^p

(a) Un critère de compacité forte [pas obligatoire]

Théorème de Riesz-Fréchet-Kolmogorov.

(b) Dual des espaces L^p

Théorème de Riesz-Fréchet. Contre-exemple pour $p = 1$ ou $+\infty$. Critère de densité. Densité de $\mathcal{C}_c$ dans L^p. L^p est séparable.

(c) L'espace de Hilbert L^2

Produit scalaire. Théorème de projection sur un convexe fermé. Espérance conditionnelle sur L^2. Théorème de Parceval.

Étude de $L^2(\rho)$. Polynômes orthogonaux.

III Produit de convolution et régularisation

Produit de convolution $L^1 - L^1$, $L^1 - L^p$ et $L^p - L^q$. Inégalité de Young. Approximation de l'unité, suites régularisantes. Exemple de suite régularisante. Support d'une fonction mesurable. Support de la convolution. Densité de $\mathcal{C}_c^\infty(\Omega)$ dans $L^p(\Omega)$.

IV Transformation de Fourier

Transformée de Fourier dans $L^1(\mathbb{R}^d)$. Propriétés, lemme de Riemann-Lebesgue. Lien entre moment et dérivation. Transformée de Fourier de la gaussienne. Formule d'inversion de Fourier. Application : transformée de Fourier d'une var de Cauchy. Injectivité.

Développements

(i) Densité de $\mathcal{C}_c^\infty(\Omega)$ dans $L^p(\Omega)$, $1 \leq p < +\infty$

(ii) Formule d'inversion de Fourier dans $\mathbb{R}^d$

Références

(i) Brézis, Analyse fonctionnelle

(ii) Hirsch-Lacombe, Éléments d'analyse fonctionnelle

(iii) Briane & Pages, Théorie de l'intégration

(iv) Ouvrard, Probabilités tome II

235 : Problèmes d'interversion de limites et d'intégrales.

Rapport du jury

"Cette leçon s'intéresse aux problèmes d'interversion limite-limite, limite-intégrale et intégrale-intégrale. Il ne s'agit pas de refaire un cours d'intégration. On pourra toutefois mettre en évidence le rôle important joué par des théorèmes cruciaux de ce cours. À un niveau élémentaire, on peut insister sur le rôle de la convergence uniforme, ou de la convergence normale (dans le cas de séries de fonctions).

À un niveau plus avancé, les théorèmes de convergence dominée, de convergence monotone et le théorème de Fubini (et Fubini-Tonelli) ont leur place dans cette leçon. On choisira des exemples pertinents pour illustrer l'intérêt de chacun de ces résultats, mais on pourra aussi exhiber des contre-exemples montrant que des hypothèses trop faibles ne permettent pas en général d'effectuer l'interversion tant désirée. Pour les candidats qui le souhaitent, on pourra parler de la transformée de Fourier et/ou de la transformée de Laplace."

Introduction

Plan

I Théorèmes d'interversion

(a) Intervertion limite-limite

(b) Intervertion limite-intégrale

(c) Théorèmes de Fubbini

II Théorèmes de régularités

(a) Régularité d'une suite de fonctions

(b) Régularité d'une intégrale à paramêtre

III Exemples

(a) Séries entières

(b) Convolution

(c) Transformée de Fourier

Développements

(i) Formule d'inversion de Fourier
(ii) Théorème abélien et application

Références

(i)
(ii)
(iii)
(iv)

236 : Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.

Rapport du jury

"Cette leçon doit être très riche en exemples simples, comme l'intégrale $\int_0^{+\infty} \frac{\sin(t)}{t}\,\mathrm{d}t$. Il est souhaitable de présenter des utilisations du théorème des résidus, ainsi que des exemples faisant intervenir les intégrales multiples comme le calcul de l'intégrale d'une gaussienne. Le calcul du volume de la boule unité de $\mathbb{R}^n$ ne doit pas poser de problèmes insurmontables. Le calcul de la transformation de Fourier d'une gaussienne a sa place dans cette leçon.

On peut aussi penser à l'utilisation du théorème d'inversion de Fourier ou du théorème de Plancherel. Certains éléments de la leçon précédente, comme par exemple l'utilisation des théorèmes de convergence monotone, de convergence dominée et/ou de Fubini, sont aussi des outils permettant le calcul de certaines intégrales."

Introduction

Plan

I Utilisation des théorèmes généraux de l'intégration

(a) Théorème fondamental du calcul intégral

(b) Intégration par parties

(c) Théorèmes de Fubini

(d) Théorème de changement de variables

II Intégrales à paramètres

III Intégrales le long d'un chemin

Théorème des résidus, plein de calculs.

IV Intégration numérique

Méthode de quadrature Newton-Cotes, de Gauss, de Monte-Carlo.

Développements

(i) Formule d'inversion de Fourier
(ii) Méthode de Gauss

Références

(i)
(ii)
(iii)
(iv)

239 : Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.

Rapport du jury

"Souvent les candidats incluent les théorèmes de régularité (version segment — a minima — mais aussi version "convergence dominée") ce qui est pertinent. Cette leçon peut être enrichie par des études et méthodes de comportements asymptotiques. Les propriétés de la fonction Γ d'Euler fournissent un développement standard (il sera de bon ton d'y inclure le comportement asymptotique). Les différentes transformations classiques (Fourier, Laplace, ...) relèvent aussi de cette leçon. On peut en donner des applications pour obtenir la valeur d'intégrales classiques (celle de l'intégrale de Dirichlet par exemple).

Pour aller plus loin, on peut par exemple développer les propriétés des transformations mentionnées (notamment Fourier), ainsi que de la convolution."

Introduction

Plan

I Régularité des intégrales à paramêtres

Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et E un espace métrique. Soit $f : X \times E \longrightarrow \mathbb{K}$ telle que pour tout $\lambda \in E$, $x \longmapsto f(x, \lambda)$ soit mesurable. Lorsqu'elle est définie, on note : $F(\lambda) = \int_X f(x, \lambda)\, \mathrm{d}\mu(x)$.

Théorème de continuité (version "convergence dominée"). Nécessité de l'hypothèse. Exemple : fonction Γ d'Euler, transformation de Fourier.

Théorème de dérivation. Nécessité de l'hypothèse. Exemple : fonction Γ d'Euler. Application : calcul de l'intégrale de la transformée de Fourier de l'intégrale de Dirichlet [FGN]. Application : calcul de l'intégrale de Fresnel [FGN].

Théorème d'holomorphie. Prolongement de la fonction Γ d'Euler en une fonction méromorphe sur $\mathbb{C} \setminus \mathbb{Z}^-$.

II Étude asympototique

Soient $-\infty < a < b \leq +\infty$, E un $\mathbb{R}$-espace de Banach, $f : [a\,, b[\, \longrightarrow \mathbb{R}$ et $g : [a\,, b[\, \longrightarrow \mathbb{R}^+$ deux applications continues par morceaux.

Théorème d'intégration des relations de comparaison [GOU]. Nécessité de la positivité de g [HAU]. Application : reste de l'intégrale de Gauss.

Méthode de Laplace. Application : formule de Stirling généralisée.

III Produit de convolution

Soient f et g deux fonctions mesurables de $\mathbb{R}^d$ dans $\mathbb{K}$.

Définition fonctions convoluables [BP]. Cas $L^1 - L^1$, cas $L^1 - L^p$, cas $L^p - L^q$ et conséquences sur la convolée. Commutativité de la convolution. Somme de deux variables aléatoires indépendantes à densité. $(L^1(\mathbb{R}^d), +, \cdot, \star)$ est une algèbre normée commutative non unitaire. Support d'une convolution. Application : théorème de Steinhauss. Théorème de dérivation. Approximation de l'unité. Exemple : gaussienne, Cauchy. Suite régularisante. Exemple. Application : lemme de Tietze-Urysohn. Application : densité de $\mathcal{C}_c^\infty$ dans L^p. Application : théorème de Weierstrass trigonométrique (via le théorème de Féjer).

IV Transformée de Fourier

Opérateur de translation, continuité. Lemme de Riemann-Lebesgue. Propriété de l'opérateur transformée de Fourier. Fonction caractéristique d'une variable aléatoire à densité. Exemple : transformée de fourier de la gaussienne. Lien avec la convolution. Formule d'inversion de Fourier. Formule d'inversion de Fourier. Application : transformée de Fourier d'une Cauchy. Injectivité de la transformée de Fourier.

Application : densité des polynômes orthogonaux.

Développements

(i) Formule d'inversion de Fourier

(ii) Densité des fonctions $\mathcal{C}_c^\infty(\Omega)$ dans $L^p(\Omega)$

Références

(i) Faraut, Calcul intégral

(ii) Gourdon, Analyse

(iii) Briane & Pages, Théorie de l'intégration

(iv) Zuily & Quéffélec, Analyse pour l'agrégation

(v) FGN, Analyse 3

241 : Suites et séries de fonctions. Exemples et contre-exemples.

Rapport du jury

"Une fois les résultats généraux énoncés, on attend du candidat qu'il évoque les séries de fonctions particulières classiques : séries entières, séries de Fourier. On pourra éventuellement s'intéresser aussi aux séries de Dirichlet. Il y a beaucoup de développements possibles et les candidats n'ont généralement aucun mal à trouver des idées que ce soit à un niveau élémentaire mais fourni en exemples pertinents ou plus avancé, voire nécessitant une certaine technicité. Par exemple, les théorèmes taubériens offrent une belle palette de développements.

Par ailleurs, la leçon n'exclut pas du tout de s'intéresser au comportement des suites et séries de fonctions dans les espaces de type L^p (notamment pour $p = 1$), ou encore aux séries de variables aléatoires indépendantes."

Introduction

Plan

I Généralités sur les suites et séries de fonctions

(a) Quelques modes de convergences

Convergence simple, uniforme, normale, presque partout, dans L^p. Des exemples. Critère de Cauchy uniforme (exemple des polynômes).

(b) Liens entre les modes de convergences

Les liens et pleins de contre-exemples.

II Passage à la limite

(a) Continuité

(b) Dérivabilité

(c) Intégration

Les théorèmes du Gourdon et ceux de théorie de l'intégration

III Séries entières, fonctions holomorphes

(a) Séries entières

Définition série entière. Lemme d'Abel. Définition rayon de convergence.

Définition disque de convergence. Convergence uniforme sur les compacts du disque de convergence.

Holomorphie dans le disque de convergence, dérivation termes à termes. Expression des coefficients en fonctions des dérivées de la fonction somme.

Théorème abélien radial et application au calcul de $\sum \frac{\sin(nt)}{n}$.

Fonction analytique. Principe des zéros isolés. Prolongement analytique. Formules de Cauchy, théorème de Liouville.

(b) Fonctions holomorphes

Fonction holomorphe. Holomorphe implique analytique. Suite de fonctions holomorphes (théorème de Weierstrass, de Montel). Application à l'holomorphie des séries de Dirichlet.

IV Séries de Fourier

(a) Généralités

Définition de $L^p(\mathbb{T})$, des polynômes trigonométriques, des coefficients de Fourier, de $S_N f$, de $\sigma_N f$.

(b) Convergence des séries de Fourier

Noyau de Dirichlet, de Fejér, théorème de Dirichlet, théorème de Fejér. Corollaires : Weierstrass trigonométrique, injectivité de l'opérateur de Fourier. Application de l'injectivité : si $(c_n(f))_{n\in\mathbb{Z}}$ est cva alors f a un unique représentant continu. Théorème de Parceval. Application aux fonctions $\mathcal{C}^1$ par morceaux. Application au calcul de $\zeta(2)$. Vitesse de convergence en fonction de la régularité.

Application : résolution de l'équation de la chaleur sur le tore.

Développements

(i) Théorème abélien radial et calcul de $\sum \frac{\sin(nt)}{n}$

(ii) Équation de la chaleur sur le tore

Références

(i) Gourdon, Analyse

(ii) Hauchecorne, Contre-exemples en mathématiques

(iii) FGN, Analyse 2

(iv) BMP, Objectif Agrégation

(v) Briane & Pages, Théorie de l'intégration

243 : Convergence des séries entières, propriétés de la somme. Exemples et applications.

Rapport du jury

"Les candidats évoquent souvent des critères (Cauchy, D'Alembert) permettant d'estimer le rayon de convergence mais oublient souvent la formule de Cauchy-Hadamard. Le jury attend bien sûr que le candidat puisse donner des arguments justifiant qu'une série entière en 0 dont le rayon de convergence est R est développable en série entière en un point z_0 intérieur au disque de convergence et de minorer le rayon de convergence de cette série. Sans tomber dans un catalogue excessif, on peut indiquer les formules de développement de fonctions usuelles importantes (exp, log, $1/(1-z)$, sin, ...). Le jury attend également que le candidat puisse les donner sans consulter ses notes. En ce qui concerne la fonction exponentielle, le candidat doit avoir réfléchi au point de vue adopté sur sa définition et donc sur l'articulation entre l'obtention du développement en série entière et les propriétés de la fonction. À ce propos, les résultats sur l'existence du développement en série entière pour les fonctions dont on contrôle toutes les dérivées successives sur un voisinage de 0 sont souvent méconnus.

Le théorème d'Abel (radial ou sectoriel) trouve toute sa place mais doit être agrémenté d'exercices pertinents. Réciproquement, les théorèmes taubériens offrent aussi de jolis développements. On pourra aller plus loin en abordant quelques propriétés importantes liées à l'analyticité de la somme d'une série entière."

Introduction

Plan

I Rayon de convergence d'une série entière

Définition série entière. Lemme d'Abel. Définition rayon de convergence. Outils pour calculer le rayon de convergence : critère de d'Alembert, formule d'Hadamard. Exemples : $\sum z^n$, $\sum n^\alpha z^n$, $\sum n! z^n$, exo 2 [Gou], exp.

Comparaison des coefficients et rayon de convergence. Rayon de convergence d'une somme, d'un produit, d'une composition de séries entières. Contre-exemples [Hauchecorne]. Application [BMP]. Application : l'exponentielle est un morphisme.

Analycité et estimation du rayon de convergence de la série entière translatée.

II Comportement à l'intérieur du disque de convergence

Définition disque de convergence. Convergence uniforme sur les compacts du disque de convergence.

Holomorphie dans le disque de convergence, dérivation termes à termes. Expression des coefficients en fonctions des dérivées de la fonction somme.

Principe des zéros isolés. Prolongement analytique. Formules de Cauchy, théorème de Liouville.

III Comportement sur le disque de convergence

Théorème d'Abel radial. Application : calcul de certaines séries ($\frac{\pi}{2}$, $\log(2)$), $\sum \frac{\sin(nt)}{n}$. Théorème taubérien faible (et fort).

Points réguliers, singuliers. Exemple puis existence d'un point singulier. Théorème des lacunes d'Hadamard

IV Fonctions développables en série entière

Définition fonction DSE. Conditions nécessaires (et suffisantes). Contre-exemples [GOU]. DSE de fonctions usuelles. DSE d'une fraction rationnelle.

V Applications

(a) Fonction génératrice d'une variable aléatoire discrète

Fonctions génératrice. Propriétés et application aux processus de Galton-Watson.

(b) Résolution d'équations différentielles linéaires

Résolution d'une EDO [ZQ].

Développements

(i) Théorème abélien radial et application
(ii) Processus de Galton-Watson

Références

(i) Gourdon, Analyse
(ii) Zuily & Quéffélec, Analyse pour l'agrégation
(iii) BMP, Objectif Agrégation
(iv) Hauchecorne, Les contre-exemples en mathématiques

246 : Série de Fourier. Exemples et applications.

Rapport du jury

"Les différents résultats autour de la convergence (L^2, Fejér, Dirichlet, ...) doivent être connus. Il faut avoir les idées claires sur la notion de fonctions de classe $\mathcal{C}^1$ par morceaux (elles ne sont pas forcément continues). Dans le cas d'une fonction continue et $\mathcal{C}^1$ par morceaux on peut conclure sur la convergence normale de la série Fourier sans utiliser le théorème de Dirichlet. Il est classique d'obtenir des sommes de séries remarquables comme conséquence de ces théorèmes.

On peut aussi s'intéresser à la formule de Poisson et à ses conséquences. L'existence d'exemples de séries de Fourier divergentes, associées à des fonctions continues (qu'ils soient explicites ou obtenus par des techniques d'analyse fonctionnelle) peuvent aussi compléter le contenu.

Mais il est souhaitable que cette leçon ne se réduise pas à un cours abstrait sur les coefficients de Fourier. La résolution d'équations aux dérivées partielles (par exemple l'équation de la chaleur) peuvent illustrer de manière pertinente cette leçon, mais on peut penser à bien d'autres applications (inégalité isopérimétrique, comportements remarquables des fonctions à spectre lacunaire, ...)."

Introduction

Historiquement, les séries de Fourier ont été introduites par Joseph Fourier pour résoudre l'équation de la chaleur. Elles permettent de décomposer un signal périodique en ses harmoniques et donc de passer du continu au discret.

Plan

On note $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. On munit $L^p(\mathbb{T})$ de $\|f\|_p = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^p \,\mathrm{d}t\right)^{1/p}$ $(1 \leq p < +\infty)$ et $L^\infty(\mathbb{T})$ de $\|F\|_\infty$. Soit $f \in L^1(\mathbb{T}, \mathbb{C})$. On note $e_n(t) = e^{int}$ et $\mathcal{P}_N = \mathrm{vect}(e_n, -N \leq n \leq N)$

I Définitions et premières propriétés

(a) Les coefficients de Fourier

Définition des coefficients de Fourier réels et complexes. Exemple sur un polynôme trigonométrique. Exemple sur les fonctions plateau, tente, boobs. Définition convolution sur $\mathbb{T}$. Propriétés des coefficients de Fourier (notamment lien entre régularité de f et décroissance des coefficients). Lemme de Riemann-Lebesgue. Définiton de l'application linéaire qui à f associe sa suite de coefficients de Fourier et ses propriétés.

Comparaison comportement – régularité.

(b) Séries trigonométriques

Définition du noyau de Dirichlet, de Féjer, de Jackson. Propriétés. Définition de $S_N f$, de $\sigma_N f$.

II Convergence des séries de Fourier

(a) Convergence de $S_N f$

Convergence lorsque $c_n(f)$ est sommable. Convergence dans le cas $\mathcal{C}^2$. Majoration de la vitesse de convergence.

Théorème de Dirichlet, exemple (explicite et par Banach-Steinhauss) de fonction continue dont la série de Fourier diverge en 0. Application : $\frac{\pi - x}{2} = \sum \frac{\sin(nx)}{n}$.

(b) Convergence de $\sigma_N f$

Théorème de Féjer et conséquences (injectivité mais pas surjective, théorème de Weierstrass et de Weierstrass trigonométrique).

(c) Convergence de $J_n f$

Propriétés du noyau de Jackson. Théorème de Jackson. Application : théorème de Jackson (cas non périodique).

III Le cas L^2

Structure d'espace de Hilbert de $L^2(\mathbb{T})$. $(e_n)_n$ est une base hilbertienne. Théorème de Parceval. Convergence normale dans le cas $\mathcal{C}^1$ par morceaux. Cas d'égalité dans les inégalités de Cauchy. Quelques belles formules.

IV Applications

(a) Inégalité Isopérimétrique

(b) Formule sommatoire de Poisson

(c) Résolution de l'équation de la chaleur

Développements

(i) Théorème de Jackson
(ii) Résolution de l'équation de la chaleur sur le tore

Références

(i) BMP, Objectif Agrégation
(ii) Zuily & Quéffélec, Analyse pour l'agrégation
(iii) Gourdon, Analyse
(iv) Amar & Matheron, Analyse complexe

250 : Transformation de Fourier. Applications.

Rapport du jury

"Cette leçon, reformulée pour la session 2017, offre de multiples facettes. Les candidats peuvent adopter différents points de vue : L^1, L^2 et/ou distributions. L'aspect "séries de Fourier" n'est toutefois pas dans l'esprit de cette leçon ; précisons aussi qu'il ne s'agit pas de faire de l'analyse de Fourier sur n'importe quel groupe localement compact mais bien sur $\mathbb{R}$ ou $\mathbb{R}^d$.

La leçon nécessite une bonne maîtrise de questions de base telle que la définition du produit de convolution de deux fonctions de L^1. En ce qui concerne la transformation de Fourier, elle ne doit pas se limiter à une analyse algébrique de la transformation de Fourier. C'est bien une leçon d'analyse, qui nécessite une étude soigneuse des hypothèses, des définitions et de la nature des objets manipulés. Le lien entre la régularité de la fonction et la décroissance de sa transformée de Fourier doit être fait, même sous des hypothèses qui ne sont pas minimales.

La formule d'inversion de Fourier pour une fonction L^1 dont la transformée de Fourier est aussi L^1 sont attendues ainsi que l'extension de la transformée de Fourier à l'espace L^2 par Fourier-Plancherel. Des exemples explicites de calcul de transformations de Fourier, classiques comme la gaussienne ou $(1+x^2)^{-1}$, paraissent nécessaires.

Pour aller plus loin, la transformation de Fourier des distributions tempérées ainsi que la convolution dans le cadre des distributions tempérées peuvent être abordées. Rappelons une fois de plus que les attentes du jury sur ces questions restent modestes, au niveau de ce qu'un cours de première année de master sur le sujet peut contenir. Le fait que la transformée de Fourier envoie $\mathcal{S}(\mathbb{R}^d)$ dans lui même avec de bonnes estimations des semi-normes doit alors être compris et la formule d'inversion de Fourier maîtrisée dans ce cadre. Des exemples de calcul de transformée de Fourier peuvent être données dans des contextes liés à la théorie des distributions comme par exemple la transformée de Fourier de la valeur principale.

La résolution de certaines équations aux dérivées partielles telle que, par exemple, l'équation de la chaleur, peut être abordée, avec une discussion sur les propriétés qualitatives des solutions."

Introduction

Introduction historique, équation de la chaleur. Le cadre L^1 est le plus naturel mais, malheureusement est trop restrictif, comme le montre la formule d'inversion de Fourier (il faut que la fonction soit égale presque partout à une fonction continue tendant vers 0). Pour palier cela, on se restreint à l'espace de Schwarz puis par dualité, on étend la transformée de Fourier aux distributions tempérées. On peut aussi l'étendre à L^2 par prolongement, puis à L^p ($p \in [1,2]$) par interpolation.

Plan

I La transformée de Fourier sur $L^1(\mathbb{R}^d)$

(a) Définition et premières propriétés

Transformée de Fourier dans $L^1(\mathbb{R}^d)$. Exemples simples. Propriétés, lemme de Riemann-Lebesgue.

Lien entre moment et dérivation. Transformée de Fourier de la gaussienne.

Convolution $L^1 - L^1$. Lien avec la convolution. Application : $(L^1, +, \cdot, \star)$ est une algèbre de Banach non unitaire.

(b) Formule d'inversion

Formule d'inversion de Fourier. Application : transformée de Fourier d'une var de Cauchy. Injectivité.

II La transformée de Fourier sur $\mathcal{S}(\mathbb{R}^d)$

Définition de l'espace de Schwarz. Exemple : $e^{-\|x\|^2}$. Propriétés de l'espace de Schwarz (stabilité par multiplication par un polynôme, par dérivation, etc). Estimation des semi-normes. La transformée de Fourier est un isomorphisme de l'espace de Schwarz sur lui-même.

III La transformée de Fourier sur $L^2(\mathbb{R}^d)$

Extension de la transformée de Fourier à L^2. Transformation de Fourier-Plancherel. Lien avec la convolution. Étude des vecteurs propres. Interpolation de Riesz-Thorin.

IV Applications

(a) Formule sommatoire de Poisson

Formule sommatoire de Poisson et application à la fonction de Jacobi [FGN Analyse 2].

(b) Résolution de l'équation de la chaleur sur $\mathbb{R}^d$

(c) Polynômes orthogonaux

Étude de $L^2(\rho)$. Polynômes orthogonaux.

(d) La fonction caractéristique

Définition de la fonction caractéristique. Caractérisation de la loi. Théorème central limite.

Développements

(i) Formule d'inversion de Fourier
(ii) Densité des polynômes orthogonaux

Références

(i) Briane & Pages, Théorie de l'intégration
(ii) Bony, Cours d'analyse
(iii) Zuily & Quéffélec, Analyse pour l'agrégation
(iv) FGN, Analyse 2
(v) Stein & Shakarchi, Fourier analysis

253 : Utilisation de la notion de convexité en analyse.

Rapport du jury

“Il s’agit d’une leçon de synthèse, très riche, qui mérite une préparation soigneuse. Même si localement (notamment lors de la phase de présentation orale) des rappels sur la convexité peuvent être énoncés, ceci n’est pas attendu dans le plan. Il s’agit d’aborder différents champs des mathématiques où la convexité intervient. On pensera bien sûr, sans que ce soit exhaustif, aux problèmes d’optimisation, au théorème de projection sur un convexe fermé, au rôle joué par la convexité dans les espaces vectoriels normés (convexité de la norme, jauge d’un convexe,...). Les fonctions convexes élémentaires permettent aussi d’obtenir des inégalités célèbres. On retrouve aussi ce type d’argument pour justifier des inégalités de type Brunn-Minkowski ou Hadamard. Par ailleurs, l’inégalité de Jensen a aussi des applications en intégration et en probabilités.

Pour aller plus loin, on peut mettre en évidence le rôle joué par la convexité dans le théorème de séparation de Hahn-Banach. On peut aussi parler des propriétés d’uniforme convexité dans certains espaces, les espaces L^p pour $p > 1$, par exemple, et de leurs conséquences.”

Introduction

En optimisation, la convexité permets de rendre les extrema globaux au lieu de locaux.

Plan

I Ensemble et fonction convexe : définitions et premières applications

(a) Ensemble convexe

Définition d’un ensemble convexe [BMP]. Exemple : convexes de $\mathbb{R}$, sous-espaces vectoriels, demi-espaces. Intersection de convexes. Enveloppe convexe. Théorème de Carathéodory [GOU]. Application : $\mathrm{conv}(X)$ est compact si X est compact. Continuité des fonctions convexes [BMP].

(b) Fonction convexe [BMP,GOU]

Définition fonction (strictement) convexe, concave. Épigraphe et caractérisation de la convexité. Exemple : norme sont convexes mais pas stritement. Caractérisation(s) de la convexité dans le cas différentiable. Exemple : exp, log, $\sin|_{[0,\frac{\pi}{2}]}$, $x \longmapsto x^p$ pour $p \geq 1$.

(c) Quelques inégalités de convexité et applications [GOU]

Inégalité arithmético-géométrique. Inégalité de Hölder. Application : inégalité de Minkowski, $L^p \subset L^q$ si la mesure est finie et si $p \leq q$. Inégalité de Young [Brezis]. Inégalité de Jensen. Application à la convolution, application à l’entropie. log-convexité du déterminant. Application : inégalité de Brunn-Minkowski [FGN]. Application : ellipsoïde de John-Loewner.

II Applications en analyse fonctionnelle

(a) Les théorèmes de Hahn-Banach [Brezis]

Lemme de Zorn. Théorème de Hahn-Banach (forme analytique). Corollaires. Jauge d’un convexe, propriétés. Lien avec entre la boule unité et la norme associée. Théorème de Hahn-Banach (forme géométrique). Corollaire : caractérisation de la densité par dualité. Application : densité des polynômes orthogonaux, densité de dans L^p.

(b) Projection sur un convexe fermé

Théorème de projection sur un convexe fermé. Démonstration facile de Hahn-Banach géométrique. Propriétés de de l’opérateur de projection. Définition de l’espérance conditionnelle sur L^2. Les moindres carrés.

Théorème de représentation de Riesz-Fréchet. Théorème de Lax-Milgram et/ou théorème de Stampacchia. Application à la recherche d’une solution faible d’une EDP elliptique [Brezis].

III Utilisation de la convexité en optimisation

Propositions [CIA]. Point de Fermat [BMP, ROU]. Méthode du gradient. Pas optimal, pas conjugué. Inéquation d'Euler pour la méthode des moindres carrés.

Développements

(i) Projection sur un convexe fermé et théorème de Riesz-Fréchet
(ii) Gradient à pas optimal

Références

(i) BMP, Objectif Agrégation
(ii) Gourdon, Analyse
(iii) Ciarlet, INtroduction à l'analyse numérique matricielle et à l'optimisation
(iv) Brézis, Analyse fonctionnelle
(v) Rouvière, PGCD

260 : Espérance, variance et moments d'une variable aléatoire.

Rapport du jury

"Le jury attend des candidats qu'ils donnent la définition des moments centrés, qu'ils rappellent les implications d'existence de moments (décroissance des L^p). Le candidat peut citer — mais doit surtout savoir retrouver rapidement — les espérances et variances de lois usuelles, notamment Bernoulli, binômiale, géométrique, Poisson, exponentielle, normale. La variance de la somme de variables aléatoires indépendantes suscite souvent des hésitations. Les inégalités classiques (de Markov, de Bienaymé-Chebychev, de Jensen et de Cauchy-Schwarz) pourront être données, ainsi que les théorèmes de convergence (loi des grands nombres et théorème central limite). La notion de fonction génératrice des moments pourra être présentée ainsi que les liens entre moments et fonction caractéristique.

Pour aller plus loin, le comportement des moyennes pour une suite de variables aléatoires indépendantes et identiquement distribuées n'admettant pas d'espérance pourra être étudié. Pour les candidats suffisamment à l'aise avec ce sujet, l'espérance conditionnelle pourra aussi être abordée."

Introduction

Faire en annexe un tableau des espérances et des variances de lois classiques.

Plan

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Une variable aléatoire réelle (v.a.r) est une application mesurable de $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(\overline{\mathbb{R}}, \mathcal{B}(\overline{R})$. La loi de X est la mesure image de $\mathbb{P}$ par X.

I Définition et premières propriétés

(a) Espérance

Définition espérance d'une v.a.r. Variables centrée. Théorème de transfert. Exemple : variable discrète, variable à densité. Caractérisation de la loi de X par l'espérance de $f(X)$, $f \in \mathcal{C}_c^+$. Exemple : calcul de la densité de $h(X)$ si h difféomorphisme. Application : loi de X^2 où $X \sim \mathcal{N}(0,1)$. Propriétés de l'espérance (linéarité, monotonie, indépendance). Calcul de espérance d'une loi de Bernouilli, puis d'une binômiale. Inégalité de Jensen. Inégalité de Markov. Inégalité de grande déviation.

(b) Moments

Définition des espaces L^p $(1 \leq p \leq +\infty)$, des moments. Les moments ne caractérisent pas le loi de X. Inégalité de Hölder, de Minkowski. Décroissance des espaces L^p.

Définition variance. V.a.r. réduite. Définiton covariance. Propriétés de la variance (produit scalaire, indépendance). Calcul de variance d'une loi normale, puis de somme de loi normale. Inégalité de Bienaymé-Tchebychev.

Inégalité de Kintchine. Théorème de Bernstein.

II Fonctions génératrice et fonction caractéristique

(a) Fonction génératrice

Définition fonction génératrice. Caractérisation de la loi. Propriétés (lien avec les moments). Exemple de calcul. Processus de Galton-Watson.

(b) Fonction caractéristique

Définition de la fonction caractéristique. Caractérisation de la loi. Propriétés (lien avec les moments). Contre-exemple. Calcul de fonctions caractéristiques.

III Théorèmes limites

(a) Lois faible et forte des grands nombres

Loi faible des grands nombres.

Loi forte des grands nombres. Contre-exemple : Loi faible loi forte [OUV]. Application : théorème de Glivenko-Cantelli

(b) Théorème central limite

Théorème central limite. Application à un suite de Bernouilli. Application χ^2.

Développements

(i) Processus de Galton-Watson

(ii) Polynômes de Bernstein

Références

(i) Ouvrard, Probabilités 2

(ii) Barbe & Ledoux, Probabilités

(iii) CGDM, Exercices de probabilités

(iv) Rivoirard & Stoltz, Statistiques mathématiques en action

Troisième partie

Développements

1 La méthode de Monte-Carlo

Contexte

Développement

Théorème. *On note $P = [0\,,1]^d$. Soit $f \in L^1(P)$ et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.i.i.d de loi uniforme sur P. On pose :*

$$e_N = \frac{1}{N}\sum_{n=1}^{N} f(X_n) - \int_P f\,.$$

Alors :

(i) $e_N \xrightarrow[N\to infty]{p.s} 0$,

(ii) si de plus il existe A et B tels que $|f| \leq A$ presque partout et $\int_P f^2 \leq B$ alors : $\forall \beta \in [0\,,\frac{B}{A^2}]$,

$$\mathbb{P}\left(|e_N| \geq \beta A\right) \leq 2\exp\left(\frac{-N\beta^2A^2}{4B}\right)\,.$$

Démonstration.

(i) On calcul l'espérance de $f(X_n)$. D'après le théorème de transfert et puisque X_n suit une loi uniforme sur P :

$$\mathbb{E}\left[f(X_n)\right] = \int_P f(x)\,\mathrm{d}\,\mathbb{P}_{X_n}(x) = \int_P f(x)\,.$$

La suite de variables aléatoires $(f(X_n))_n$ est i.i.d et intégrable. D'après la loi forte des grands nombres : $e_N \longrightarrow 0$ presque sûrement.

(ii) Soit $\alpha > 0$. Soit $\beta \in [0\,,\frac{B}{A^2}]$. On utilisant la monotonie de la fonction exponentielle et l'inégalité de Markov :

$$\mathbb{P}\left(e_N \geq \beta A\right) = \mathbb{P}\left(e^{\alpha e_N} \geq e^{\alpha\beta A}\right) \leq \frac{\mathbb{E}\left[e^{\alpha e_N}\right]}{e^{\alpha\beta A}}\,.$$

Or, par indépendance :

$$\mathbb{E}\left[e^{\alpha e_N}\right] = \left(\mathbb{E}\left[e^{\frac{\alpha}{N}\left(f(X_1)-\int_P f\right)}\right]\right)^N\,.$$

On sait de plus que : $\forall\,|t| \leq 1,\ e^t \leq 1+t+t^2$. Pour le voir, on pose : $\psi(t) = e^{-t}(1+t+t^2)$. On a :

$$\forall\,|t| \leq 1,\ \psi'(t) = e^{-t}(-1-t-t^2+1+2t) = e^{-t}(1-t)t\,.$$

Par l'inégalité triangulaire, on a :

$$\left|f(x) - \int_P f\right| \leq 2A \quad \text{p.s}\,,$$

Donc si $\alpha \leq \frac{N}{2A}$:

$$\mathbb{E}\left[e^{\alpha e_N}\right] \leq \mathbb{E}\left[1 + \frac{\alpha}{N}\left(f(X_1) - \int_P f\right) + \frac{\alpha^2}{N^2}\left(f(X_1) - \int_P f\right)^2\right]^N \leq \left[1 + \frac{\alpha^2}{N^2}\operatorname{Var}\left(f(X_1)\right)\right]^N\,.$$

Or, d'après les hypothèses : $\operatorname{Var}\left(f(X_1)\right) \leq B$. On rappelle la classique inégalité : $\forall t \geq -1,\ \ln(1+t) \leq t$. On obtient :

$$\mathbb{E}\left[e^{\alpha e_N}\right] \leq e^{\frac{\alpha^2}{N}B}\,.$$

On a donc obtenu :

$$\mathbb{P}\left(e_N \geq \beta A\right) \leq e^{\frac{\alpha^2}{N}B}e^{-\alpha\beta A}\,.$$

(iii) Il reste plus qu'à optimiser en α. On cherche α qui minimise la fonction $\varphi : \alpha \longmapsto \frac{\alpha^2}{N}B - -\alpha\beta A$. Ce qui donne : $\alpha = \frac{\beta AN}{2B}$.
La condition $\alpha \leq \frac{N}{2A}$ se réécrit $\beta \leq \frac{B}{A^2}$. Sous ces conditions, on obtient :

$$\mathbb{P}\left(e_N \geq \beta A\right) \leq \exp\left(\frac{-N\beta^2A^2}{4B}\right)\,.$$

Il reste plus qu'à faire la même chose pour les négatifs et c'est bon ! □

Leçons concernées : 236, 263

2 Densité des fonctions continues nulle part dérivables dans $\mathcal{C}([0\,,1])$

Contexte

Soit (E,d) un espace métrique complet.

Théorème (Lemme de Baire)**.** *Soit $(O_n)_{n\in\mathbb{N}}$ (resp. $(F_n)_{n\in\mathbb{N}}$) une suite dénombrable d'ouverts denses (resp. de fermés d'intérieur vide) de E. On pose :*

$$O := \bigcap_{n\in\mathbb{N}} O_n \quad (\textit{resp. } F = \bigcup_{n\in\mathbb{N}} F_n)\,.$$

Alors O est dense (resp. F est d'intérieur vide) dans E.

Une partie A de E est dite *grasse* si elle contient une G_δ-dense E i.e une intersection dénombrable d'ouverts dense. Une partie A de E est dite *maigre* si elle est incluse dans une réunion dénombrable de fermés d'intérieur vide.

Développement

Théorème. *Soit A l'ensemble des fonctions continues nulle part dérivables dans $\mathcal{C}([0\,,1])$. Alors A contient un G_δ-dense de $\mathcal{C}([0\,,1])$.*

Remarque. On dira que $f \in \mathcal{C}([0\,,1])$ est dérivable en 0 (resp. en 1) si f admet une dérivée à droite (resp. à gauche) en 0 (resp. en 1). Pour les points intérieurs, on garde la notion de dérivabilité usuelle.

Démonstration. Tout d'abord, on rappelle que $\mathcal{C}([0\,,1])$ muni de la norme uniforme $\|\cdot\|_\infty$ est un espace de Banach, on peut utiliser le lemme de Baire.

Soit $B = A^c$ l'ensemble des fonctions de $\mathcal{C}([0\,,1])$ dérivables en au moins un point. Montrons que B est maigre. Si $f \in B$ alors il existe $x \in [0\,,1]$ tel que l'application

$$y \in [0\,,1] \setminus \{x\} \longmapsto \frac{f(y) - f(x)}{y - x}$$

se prolonge par continuité sur $[0\,,1]$. Ainsi, par compacité et continuité, si $f \in B$ alors :

$$\exists C > 0,\ \exists x \in [0\,,1],\ \forall y \in [0\,,1],\ |f(y) - f(x)| \leq C\,|y - x|\ .$$

Posons pour $n \in \mathbb{N}^*$

$$F_n = \{f \in \mathcal{C}([0\,,1]) \mid \exists x \in [0\,,1],\ \forall y \in [0\,,1],\ |f(y) - f(x)| \leq n\,|y - x|\}\,.$$

Ce qui précède montre que $B \subset \bigcup_{n\in\mathbb{N}^*} F_n$.

Montrons maintenant que pour tout $n \in \mathbb{N}^*$, F_n est fermé et d'intérieur vide. Le lemme de Baire conclura.

F_n est fermé.

En effet, soit $(f_k)_{k\in\mathbb{N}}$ une suite de F_n qui converge vers f dans $\mathcal{C}([0\,,1])$. Montrons que $f \in F_n$.

$$\forall k \in \mathbb{N},\ \exists x_k \in [0\,,1],\ \forall y \in [0\,,1],\ |f(y) - f(x_k)| \leq n\,|y - x_k|\ .$$

Comme $[0\,,1]$ est compact, d'après la propriété de Bolzano-Weierstrass et quitte à extraire, on peut supposer que $(x_k)_{k\in\mathbb{N}}$ converge vers $x \in [0\,,1]$. Pour tout $k \in \mathbb{N}$ et tout $y \in [0\,,1]$, on a :

$$\begin{aligned}|f(x) - f(y)| &\leq |f(x) - f(x_k)| + |f(x_k) - f_k(x_k)| + |f_k(x_k) - f_k(y)| + |f_k(y) - f(y)| \\ &\leq |f(x) - f(x_k)| + 2\,\|f - f_k\|_\infty + n\,|x_k - x| + n\,|x - y|\ .\end{aligned}$$

Par continuité de f et en passant à la limite lorsque $k \to +\infty$, on a bien $f \in F_n$.

F est d'intérieur vide.

Pour le vérifier, il suffit de montrer que toute boule centrée en un élément de F_n rencontre F_n^c. Soient $f \in F_n$ et $\varepsilon > 0$. On cherche $g \in \mathcal{C}([0\,,1])$ tel que : $\|f - g\|_\infty < \varepsilon$ et

$$\forall x \in [0\,,1],\ \exists y \in [0\,,1],\ |g(x) - g(y)| > n\,|x - y|\ .$$

Pour trouver un tel g, l'idée est d'approcher f par une fonctions assez régulière (ici, ce sera un polynôme) que l'on perturbera par une fonction qui oscille fortement.

D'après le théorème de Weierstrass, l'ensemble des fonctions polynômiales est dense dans $\mathcal{C}([0\,,1])$ donc il existe P un polynôme tel que : $\|f - P\|_\infty \leq \varepsilon/2$.

Soit $N \in \mathbb{N}^*$. On définit la fonction $\frac{1}{N}$-périodique $g_N \in \mathcal{C}([0\,,1])$ avec :

$$g_N : x \longmapsto \begin{cases} \frac{\varepsilon N}{2} x & \text{si} \quad x \in [0\,, \frac{1}{2N}] \\ \frac{\varepsilon}{2} - \frac{\varepsilon N}{2} x & \text{si} \quad x \in [\frac{1}{2N}\,, \frac{1}{N}] \end{cases} \quad ,$$

g_N est une fonction “tente” que l'on a périodisée. Elle vérifie : $\|g_N\| = \frac{\varepsilon}{4}$ et $|g_N'(x)| = \frac{\varepsilon N}{2}$, là où g_N est dérivable. Soit $M = \|P'\|_\infty$. On choisit $N \geq 2\frac{M+n+1}{\varepsilon}$ pour avoir $|g_N'(x)| \geq M + n + 1$. On pose $g = g_N + P$. On a bien :

$$\|f - g\|_\infty \leq \|f - P\|_\infty + \|g_N\|_\infty \leq \frac{3\varepsilon}{4} < \varepsilon\,.$$

De plus, soit $x \in [0\,,1]$. Soit $y \in [0\,,1]$ assez proche de x de sorte que :

$$|g_N(x) - g_N(y)| = \frac{\varepsilon N}{2}\,|x - y|\ .$$

En utilisant l'inégalité de la moyenne sur P :

$$|g(x) - g(y)| \geq |g_N(x) - g_N(y)| - |P(x) - P(y)| \geq (n+1)\,|x - y| > n\,|x - y|\ .$$

Ce qui montre que F_n est d'intérieur vide. □

Compléments

Voici maintenant un exemple explicite de fonction continue sur $[0\,,1]$ et dérivable nulle part. Soit t (pour “tente”) la fonction 1-périodique sur $\mathbb{R}$ définie par :

$$t : x \in [0\,,1] \longmapsto \begin{cases} x & \text{si} \quad x \in [0\,, \frac{1}{2}] \\ 1 - x & \text{si} \quad x \in [\frac{1}{2}\,, 1] \end{cases} \quad ,$$

En fait, $t(x)$ est la distance de x à l'entier le plus proche. Pour $x \in \mathbb{R}$, on pose :

$$f(x) = \sum_{n \in \mathbb{N}} \frac{t(2^n x)}{2^n}\,.$$

t est continue donc, pour tout $n \in \mathbb{N}$, $x \mapsto \frac{t(2^n x)}{2^n}$. La série de fonctions qui définit f est normalement convergente (t est bornée), donc f est une fonction continue sur $\mathbb{R}$.

Proposition. *f n'est dérivable nulle part.*

Démonstration. f est 1-périodique donc il suffit de montrer que f n'est dérivable en aucun point de $[0\,,1[$. Soit $x \in [0\,,1[$ et on note $0,a_1a_2\cdots = \sum_{n\geq 1} a_n 2^{-n}$ son développement propre en base 2 (les a_i ne sont pas tous égaux à 1 à partir d'un certain rang). Considérons la suite $(x_m := 0,a_1\ldots a_m)_m$ et $(x_m' := x_m + 2^{-m})_m$. On a : $\forall m \geq 1,\ x_m < x < x_m'$. De plus x_m (resp. x_m') converge en croissant (resp. en décroissant) vers x. Si f était dérivable en x, on aurait :

$$A_m := \frac{f(x_m') - f(x_m)}{x_m' - x_m} = \sum_{n\geq 0} \underbrace{\frac{t(2^n x_m + 2^{n-m}) - t(2^n x_m)}{2^{n-m}}}_{:= A_{mn}} \underset{m\to+\infty}{\longrightarrow} f'(x)\,.$$

Si $n > m$ alors 2^{n-m} est un entier et par 1-périodicité : $t(2^n x_m + 2^{n-m}) - t(2^n x_m) = 0$. Par conséquent : $A_m = \sum_{n=0}^{m-1} A_{mn}$. Fixons $n < m$, deux cas s'offrent à nous :

– si $a_{n+1} = 0$ alors $2^n x_m = k_n + r_n$ où k_n est un entier et où :

$$r_n = \frac{a_{n+1}}{2} + \cdots + \frac{a_m}{2^{m-n}} = \frac{a_{n+2}}{4} + \cdots + \frac{a_m}{2^{m-n}} \leq \frac{1}{4}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{m-n-2}}\right)$$
$$= \frac{1}{4}\frac{1 - \left(\frac{1}{2}\right)^{m-n-1}}{1-\frac{1}{2}} = \frac{1}{2} - \frac{1}{2^{m-n}}$$

Donc $0 \leq r_n + \frac{1}{2^{m-n}} \leq \frac{1}{2}$. De plus, $2^n x'_m = k_n + r_n + \frac{1}{2^{m-n}}$ donc $(t(x) = d(x, \mathbb{Z}))$:

$$A_{mn} = \frac{r_n + \frac{1}{2^{m-n}} - r_n}{\frac{1}{2^{m-n}}} = 1\,.$$

– Si $a_{n+1} = 1$, avec le même genre de calcul et les mêmes notations : $\frac{1}{2} \leq r_n \leq r_n + \frac{1}{2^{m-n}} \leq 1$ et $A_{mn} = -1$.

Ainsi, par récurrence, on montre que A_m est un entier relatif pair si m est pair et impair si m est impair. Donc A_m ne converge pas et f n'est pas dérivable en x. □

Références

– Queffélec et Zuily, Analyse pour l'agrégation, p.269 - 270

Leçons concernées : 202, 205, 228

3 Alternative de Fredholm

Contexte

Soient E et F deux espaces de Banach réels. On note B_E la boule unité *fermée* de E. On rappelle qu'un opérateur (continu) $T \in \mathcal{L}(E, F)$ est dit *compact* si $T(B_E)$ est relativement compact dans F (pour la topologie forte). On note $\mathcal{K}(E, F)$ l'ensemble des opérateurs compacts de E vers F. On note aussi $\mathcal{K}(E) = \mathcal{K}(E, E)$. On notera I_E l'identité sur E (I s'il n'y a pas d'ambiguïté).

Exemple. Soient $a < b$ deux réels et $K \in \mathcal{C}([a\,,b]^2)$. Soient également α et β deux fonctions continues de $[a\,,b]$ dans $[a\,,b]$. On pose, pour $f \in \mathcal{C}([a\,,b])$:

$$Tf : x \in [a\,,b] \longmapsto \int_{\alpha(x)}^{\beta(x)} K(x, y) f(y)\,\mathrm{d}y\,.$$

L'application $T : f \mapsto Tf$ est un opérateur compact de $\mathcal{C}([a\,,b])$ (appliquer le théorème d'Arzela-Ascoli). En particulier, les opérateurs de convolution par une fonction continue et d'intégration sont compacts.

Théorème. *$\mathcal{K}(E, F)$ est un sous-espace vectoriel fermé de $\mathcal{L}(E, F)$ (pour la norme d'opérateur associée). De plus, la compacité est une notion stable par composition à droite et à gauche par des opérateurs (continus). En particulier, $\mathcal{K}(E)$ est un idéal bilatère de $\mathcal{L}(E)$.*

Remarque. Pour montrer que $\mathcal{K}(E, F)$ est fermé, on utilise la complétude de F.

On rappelle que E' désigne le dual topologique de E, c'est-à-dire l'ensemble des formes linéaires continues de E. On rappelle que E' est aussi un espace de Banach. On notera $\langle \cdot, \cdot \rangle_{E',E}$ le crochet de dualité.

Définition (adjoint d'un opérateur)**.** Soit $T \in \mathcal{L}(E, F)$. L'*adjoint* T^* de T est l'opérateur de $\mathcal{L}(F', E')$ défini par :

$$\forall v \in F',\ \forall u \in E,\ \langle A^* v, u \rangle_{E',E} = \langle v, Au \rangle_{F',F}\,.$$

Théorème (Schauder)**.** *Soit $T \in \mathcal{L}(E, F)$ alors : $T \in \mathcal{K}(E, F) \Longleftrightarrow T^* \in \mathcal{K}(F', E')$.*

Développement

Théorème (Alternative de Fredholm)**.** *Soit $T \in \mathcal{K}(E)$. Alors :*

(i) $\dim(\operatorname{Ker}(I - T)) < +\infty$,

(ii) $\operatorname{Im}(I - T)$ *est fermé et* $\operatorname{Im}(I - T) = \operatorname{Ker}(I - T^*)^{\perp}$,

(iii) $\operatorname{Ker}(I - T) = \{0\} \Longleftrightarrow \operatorname{Im}(I - T) = E$.

Lemme (de Riesz)**.** *Soit E un espace vectoriel normé. Soit $F \subset E$ un sous-espace fermé strict de E. Alors :*

$$\exists u \in E \mid \|u\| = 1 \quad \textit{et} \quad d(u, M) \geq \frac{1}{2}\,.$$

Démonstration du lemme de Riesz. Soit $v \in E \setminus F$. Comme F est fermé : $d = d(u, M) > 0$. Soit $m_0 \in F$ tel que :

$$d \leq \|v - m_0\| \leq 2d\,.$$

On pose : $u = \frac{u - m_0}{\|u - m_0\|}$. Comme pour tout $m \in F$, $m_0 + \|u - m_0\|\, m \in F$, on a :

$$\forall m \in F,\ \|u - m\| = \frac{1}{\|v - m_0\|} \|v - (m_0 + \|u - m_0\|\, m)\| \geq \frac{1}{2}\,.$$

$\square$

Lemme (Théorème de Riesz)**.** *Soit E un espace vectoriel normé. On note B_E sa boule unité fermée. Alors :*

$$B_E \textit{ est compacte} \Longleftrightarrow \dim(E) < +\infty\,.$$

Démonstration du théorème de Riesz. Le sens indirect provient de la caractérisation des compacts en dimension finie.

Supposons que $\dim(E) = +\infty$. Alors il existe une suite strictement croissante $(E_n)_n$ de sous-espaces vectoriels de dimension finie de E. Le lemme précedent montre que :

$$\forall n,\ \exists u_n \in E_n,\ \begin{cases} \|u_n\| = 1 \\ d(u_n, E_{n-1}) \geq 1/2 \end{cases}.$$

Ainsi :

$$\forall n \neq m,\ \|u_n - u_m\| \geq 1/2$$

et $(u_n)_n$ n'a pas de valeurs d'adhérence donc B_E n'est pas compacte. $\square$

Démonstration du théorème de Fredholm. Soit E un espace de Banach. Soit $T \in \mathcal{K}(E)$.

(i) Soit $N = \operatorname{Ker}(I - T)$, sous-espace T-stable. Alors $B_N \subset T_{|N}(B_N)$ or, par restriction, $T_{|N}$ est un opérateur compact de N et B_N est compacte (fermé inclus dans un ensemble relativement compact). D'après le théorème de Riesz, N est de dimension finie.

(ii) Montrons que $\operatorname{Im}(I - T)$ est séquentiellement fermée. Soit :

$$f_n = u_n - T(u_n) \underset{n \to +\infty}{\longrightarrow} f \in E\,.$$

Notons $d_n = d(u_n, \operatorname{Ker}(I - T))$. Comme $\operatorname{Ker}(I - T)$ est de dimension finie, la distance est atteinte :

$$\exists v_n \in \operatorname{Ker}(I - T),\ \|u_n - v_n\| = d_n\,.$$

Donc :

$$\forall n,\ f_n = (u_n - v_n) - T(u_n - v_n)\,.$$

Montrons que $(\|u_n - v_n\|)_n$ est bornée. Pour cela, raisonnons par l'absurde. Quitte à extraire, on peut supposer que $\|u_n - v_n\| \longrightarrow +\infty$ lorsque $n \mapsto +\infty$.
Comme $(\|f_n\|)_n$ est bornée car convergente, en posant $w_n = \frac{u_n - v_n}{\|u_n - v_n\|}$, on a :

$$\frac{f_n}{d_n} = w_n - T(w_n) \underset{n \to +\infty}{\longrightarrow} 0\,.$$

Comme T est compact, quitte à extraire, on peut supposer $(T(w_n))_n$ converge vers une élément z de E lorsque n tend vers $+\infty$. D'où :

$$w_n \underset{n \to +\infty}{\longrightarrow} z\,.$$

Ainsi : $z \in \operatorname{Ker}(I - T)$. Or, par continuité de l'application distance :

$$1 = d(w_n, \operatorname{Ker}(I - T)) \underset{n \to +\infty}{\longrightarrow} d(z, \operatorname{Ker}(I - T))\,,$$

ce qui est absurde. Donc $(\|u_n - v_n\|)_n$ est bornée.
Comme T est compact, quitte à extraire, on peut supposer que $(T(u_n - v_n))_n$ converge vers un élément ℓ de E. Donc :

$$u_n - v_n \underset{n \to +\infty}{\longrightarrow} f + \ell\,,$$

et en posant $g = f + \ell$, on a bien : $f = g - T(g) \in \operatorname{Im}(I - T)$. Donc $\operatorname{Im}(I - T)$ est fermée.
Puisque $\operatorname{Im}(I - T)$ est fermé, on a (proposition du cours) :

$$\operatorname{Im}(I - T) = \operatorname{Ker}(I - T^*)^{\perp} \quad \text{et} \quad \operatorname{Im}(I - T^*) = \operatorname{Ker}(I - T)^{\perp}\,.$$

(iii) Montrons d'abord le sens direct par l'absurde. Supposons que $E_1 := (I - T)(E) \neq E$. On construit par récurrence une suite décroissante $(E_n)_n$ de sous-espaces vectoriels en posant :

$$E_{n+1} = (I - T)(E_n) = (I - T)^{n+1}(E)\,.$$

En fait, cette suite est strictement décroissante. En effet, si ce n'était pas le cas alors $(E_n)_n$ serait stationnaire à partir d'un certain rang noté n_0. Si $x \in E \setminus E_1$ alors $(I-T)^{n_0}(x) \in E_{n_0}$ et comme $E_{n_0+1} = (I-T)^{n_0}(E_1) = E_{n_0}$, on aurait l'existence d'un élément $y \in E_1$ tel que :

$$(I-T)^{n_0}(x) = (I-T)^{n_0}(x)\,.$$

Par injectivité de $(I-T)$ on aurait $x = y$, ce qui est absurde. Donc $(E_n)_n$ est strictement décroissante. De plus, les E_n sont T-stables et $T_{|E_n}$ est un opérateur compact. D'après (ii), les E_n sont fermés.
On construit, par récurrence et à l'aide du lemme de Riesz, une suite $(u_n)_n$ de E telle que :

$$\forall n,\ \begin{cases} u_n \in E_n \\ \|u_n\| = 1 \\ d(u_n, E_{n+1}) \geq 1/2 \end{cases}.$$

On a :

$$\forall n > m,\ T(u_n) - T(u_m) = [-(u_n - T(u_n)) + (u_m - T(u_m)) + u_n] - u_m\,.$$

Comme $E_{n+1} \subset E_n \subset E_{m+1} \subset E_m$, on a : $-(u_n - T(u_n)) + (u_m - T(u_m)) + u_n \in E_{m+1}$. Donc :

$$\forall n > m,\ \|T(u_n) - T(u_m)\| \geq 1/2\,,$$

ce qui est absurde car T est compact.
Pour le sens indirect, on utilise le théorème de Schauder. Si $\mathrm{Im}(I-T) = E$ alors $\mathrm{Ker}(I-T^*) = \mathrm{Im}(I-T)^\perp = \{0\}$ et puisque $T^* \in \mathcal{K}(E')$, on peut appliquer ce qui précède à T^* : $\mathrm{Im}(I-T^*) = E'$. Or : $\mathrm{Ker}(I-T) = \mathrm{Im}(I-T^*)^\perp = \{0\}$.

□

Remarque. La proposition (iii) (injectivité $\iff$ surjectivité $\iff$ bijectivité) est familière en dimension finie mais est mise en défaut en dimension infinie (cf. *shift* à gauche ou à droite dans $\ell^1(\mathbb{N})$). C'est une propriété remarquable des opérateurs de la forme $I - T$ avec $T \in \mathcal{K}(E)$ qui se comportent presque comme des opérateurs de rang fini (en fait, tout opérateur de rang fini est compact et, dans un espace de Hilbert par exemple, tout opérateur compact est limite d'une suite d'opérateurs de rang fini).

Compléments

Une des principales applications de la théorie de Riesz-Fredholm est l'étude du spectre des opérateurs compacts qui mène à la diagonalisation des opérateurs auto-adjoints et compacts.

On rappelle d'abord quelques définitions. Soit $T \in \mathcal{L}(E)$ un opérateur. L'*ensemble résolvant* de T est défini par :

$$\rho(T) := \{\lambda \in \mathbb{R} \mid T - \lambda I \text{ est bijectif}\}\,.$$

En particulier, d'après le théorème de Banach, si $\lambda \in \rho(T)$ alors $(T - \lambda I)^{-1} \in \mathcal{L}(E)$. Le *spectre* $\sigma(T)$ de T est le complémentaire de $\rho(T)$ dans $\mathbb{R}$. L'ensemble des *valeurs propres* de T est :

$$\mathrm{Vp}(T) := \{\lambda \in \mathbb{R} \mid \mathrm{Ker}(T - \lambda I) \neq \{0\}\}\,.$$

On a $\mathrm{Vp}(T) \subset \sigma(T)$ et il y a égalité si E est de dimension finie (en général l'inclusion est stricte, cf. le *shift* à droite dans $\ell^1(\mathbb{N})$) Si $\lambda \in \mathrm{Vp}(T)$ alors $\mathrm{Ker}(T - \lambda I)$ est l'*espace propre* associé à λ.

Proposition. *Soit $T \in \mathcal{L}(E)$ alors $\sigma(T)$ est un ensemble compact inclus dans $[-\|T\|, +\|T\|]$ (où $\|T\|$ désigne la norme d'opérateur de T).*

Démonstration. Pour montrer l'inclusion, on raisonne par contraposée et on montre l'existence et l'unicité de la solution de l'équation $(T - \lambda I)(u) = f$ en appliquant le théorème de point fixe de Banach à la fonction $u \mapsto \frac{1}{\lambda}(T(u) - f)$. Pour montrer que $\sigma(T)$ est compact, on montre que $\rho(T)$ est ouvert de façon analogue à ce qui précède (en utilisant le point fixe de Banach). □

Théorème. *Supposons que* $\dim E = +\infty$. *Si $T \in \mathcal{K}(E)$ alors on a :*

(i) $0 \in \sigma(T)$,

(ii) $\mathrm{Vp}(T) \setminus \{0\} = \sigma(T) \setminus \{0\}$*,*

(iii) tous les points de $\sigma(T) \setminus \{0\}$ *sont isolés,*

(iv) l'une des situations suivantes :

- *ou bien* $\sigma(T) = \{0\}$*,*
- *ou bien* $\sigma(T) \setminus \{0\}$ *est fini,*
- *ou bien* $\sigma(T) \setminus \{0\}$ *est le support d'une suite injective qui tend vers 0.*

Démonstration. (i) Si $0 \notin \sigma(T)$ alors $I = T \circ T^{-1}$ est compact et $\dim E < +\infty$ d'après le théorème de Riesz, ce qui est absurde.

(ii) Soit $\lambda \in \sigma(T) \setminus \{0\}$. Supposons que $\ker(T - \lambda I) = \{0\}$, alors le théorème de Fredholm montre que $(T - \lambda I)$ est surjective et $\lambda \in \rho(T)$, ce qui est absurde.

(iii) Comme $\sigma(T)$ est compact, il suffit de vérifier que toute suite $(\lambda_n)_n$ convergente et injective de $\sigma(T) \setminus \{0\}$ converge vers 0. Soit une telle suite, pour n il existe $e_n \neq 0$ tel que : $(T - \lambda I)(e_n) = 0$. On pose $E_n := \mathrm{vect}(e_1, \ldots, e_n)$. On montre, par récurrence et par l'absurde, que $(E_n)_n$ est une suite strictement croissante. De plus, pour tout n, on a : $(T - \lambda_n I)(E_n) \subset E_{n-1}$. En utilisant le lemme de Riesz, on construit par récurrence une suite $(u_n)_n$ telle que :

$$\forall n \in \mathbb{N},\ u_n \in E_n, \quad \|u_n\| = 1 \quad \text{et} \quad d(u_n, E_{n-1}) \geq \frac{1}{2}\,.$$

Par le même type de procédé que pour la démonstration du (iii) du théorème de Fredholm, on montre que pour tout $2 \leq m < n$:

$$\left\| \frac{T(u_n)}{\lambda_n} - \frac{T(u_m)}{\lambda_m} \right\| \geq d(u_n, E_{n-1}) \geq \frac{1}{2}\,,$$

et si $\lambda_n \to \lambda \neq 0$, on aboutirait à une contradiction puisque $(T(u_n))_n$ admet une sous-suite convergeante.

(iv) Pour tout n, l'ensemble $\sigma(T) \cap \{\lambda \in \mathbb{R} \mid |\lambda| \geq \frac{1}{n}\}$ est vide ou fini (sinon, par compacité de $\sigma(T)$, il y aurait un point d'accumulation). Si $\sigma(T) \setminus \{0\}$ est infini, par récurrence, on peut ranger ses éléments par module décroissant.

□

Remarque. Réciproquement, si $(\alpha_n)_n$ est une suite qui tend vers 0 alors on peut construire un opérateur compact T tel que $\sigma(T) = \mathrm{supp}\,\alpha_n \cup \{0\}$ en considérant l'opérateur de $\ell^2(\mathbb{N})$ $T : u \mapsto t(u) = (\alpha_n u_n)_n$ qui est compact car limite l'opérateurs de rang fini.

On peut poursuivre le vice et discuter du spectre des opérateurs compacts auto-adjoints. On suppose à partir de maintenant que $E = H$ est un espace de Hilbert. Le théorème de représentation de Riesz-Fréchet permettant d'identifier H et H', on dit qu'un opérateur $T \in \mathcal{L}(H)$ est *auto-adjoint* si : $T = T^*$.

Proposition. *Soit* $T \in \mathcal{L}(H)$ *un opérateur auto-adjoint. On note :*

$$m = \inf_{\|u\| \leq 1} \langle Tu, u \rangle \quad \textit{et} \quad M = \sup_{\|u\| \leq 1} \langle Tu, u \rangle\ ,$$

alors $\sigma(T) \subset [m\,, M]$*. De plus,* m *et* M *sont dans* $\sigma(T)$*.*

Démonstration. Soit $\lambda > M$ alors : $\forall u \in H,\ \langle T(u), u \rangle \leq M \|u\|^2$. On en déduit :

$$\forall u \in H,\ \langle (\lambda I - T)(u), u \rangle \geq (\lambda - M) \|u\|^2\ ,$$

et en appliquant le théorème de Lax-Milgram, on conclut que $(T - \lambda I)$ est bijectif.

Pour montrer que $M \in \sigma(T)$, on applique l'inégalité de Cauchy-Schwarz à la forme bilinéaire symétrique et positive $a : (u, v) \in H^2 \mapsto \langle (MI - T)(u), v \rangle$:

$$\forall u, v \in H,\ |\langle (MI - T)(u), v \rangle| \leq \sqrt{\langle (MI - T)(u), u \rangle} \sqrt{\langle (MI - T)(v), v \rangle}\,.$$

En particulier, pour $v = \frac{(MI-T)(u)}{\|(MI-T)(u)\|}$ et $C = \sqrt{\|MI - T\|}$:

$$\forall u \in H,\ \forall u, v \in H,\ \|(MI - T)(u)\| \leq C \sqrt{\langle (MI - T)(u), u \rangle}\,.$$

Soit $(u_n)_n$ telle que pour tout n, $\|u_n\| = 1$ et : $\langle T(u_n), u_n\rangle \to M$. Ce qui précède montre que

$$\|(MI - T)(u)\| \to 0$$

et donc $M \in \sigma(T)$ car sinon :

$$u_n = (MI - T)^{-1}(MI - T)(u_n) \longrightarrow 0\,,$$

par continuité de $(MI - T)^{-1}$, ce qui est absurde. Pour m, il fait la même chose en remplaçant T par $-T$. □

Remarque. Ainsi, pour $T \in \mathcal{L}(H)$, si $\sigma(T) = \{0\}$ alors $T = 0$ (utiliser l'identité de polarisation des formes quadratiques).

Théorème (Diagonalisation des opérateurs compacts auto-adjoints)**.** *On suppose que H est séparable. Soit T un opérateur compact et auto-adjoint de H alors H admet une base hilbertienne formée de vecteurs propres de T.*

Démonstration. En dimension finie, c'est le théorème spectral de prépa. On suppose $\dim H = +\infty$. Soit $(\lambda_n)_{n\geq 1}$ la suite des valeurs propres de T exceptée 0 et on note $\lambda_0 = 0$. On note E_i l'espace propre de T associé à λ_i. On rappelle que, d'après les théorèmes précédents, on a :

$$0 \leq \dim E_0 \leq +\infty \quad \text{et} \quad \forall n \geq 1,\ 0 < \dim E_n < +\infty\,.$$

(i) Les E_n sont deux à deux orthogonaux car : $\forall n \neq m,\ \forall u \in E_n,\ \forall v \in E_m$,

$$\langle T(u), v\rangle = \lambda_n \langle u, v\rangle = \langle u, T(v)\rangle = \lambda_m \langle u, v\rangle \implies (\lambda_n - \lambda_m)\langle u, v\rangle = 0 \implies \langle u, v\rangle = 0\,,$$

(ii) L'espace F engendré par les espaces propres de T est dense dans H. En effet, clairement $T(F) \subset F$. T étant auto-adjoint, cela implique que $T(F^\perp) \subset F^\perp$. Par restriction, l'opérateur $T_0 = T_{|F^\perp}$ est auto-adjoint et compact. Or, $\sigma(T_0) \subset \{0\}$ (un vecteur propre de T_0 non associé à 0 est aussi un vecteur propre de T et est donc dans F, mais est aussi dans $F^\perp$ donc est forcément nul) et la proposition précédente montre que $\sigma(T_0) = \{0\}$. La remarque implique que $T_0 = 0$ et donc : $F^\perp \subset \ker(T) \subset F$ puis $F^\perp = \{0\}$. Donc F est dense dans H.

Les espaces E_n sont en somme hilbertienne dans H et comme H est séparable il suffit de choisir pour chaque E_n une base hilbertienne et de le recoller pour avoir une base hilbertienne qui diagonalise T. □

Remarque. L'hypothèse de compacité est indispensable ($Tf(x) \mapsto xf(x)$ pour $f \in L^2([0\,,1]$ est auto-adjoint et non compact, et n'a pas de valeurs propres).

Exemple. Soit H un espace de Hilbert séparable de dimension infinie. On note $(e_n)_n$ une base hilbertienne de H. On dit qu'un opérateur $T \in \mathcal{L}(H)$ est un *opérateur de Hilbert-Schmidt* si la série numérique $\sum_{n\geq 0} \|T(e_n)\|^2$ est convergente. On peut démontrer que cette définition est indépendante de la base hilbertienne choisie. Si on note P_n le projecteur sur l'espace $\mathrm{vect}(e_0, \ldots, e_n)$ alors, on montre que la suite d'opérateurs de rang fini $(P_n(T)$ converge vers T, ce qui montre que T est compact.

Dans le cas où $H = L^2(X, \mathcal{A}, \mu) = L^2(\mu)$ où μ est une mesure σ-finie sur un espace mesurable $(X, \mathcal{A})$ (et si H est séparable, ce qui est le cas si E est un ouvert de $\mathbb{R}^n$) alors, si T est un opérateur de Hilbert-Schmidt sur H, il existe $K \in L^2(\mu \otimes \mu)$ tel que pour tout $f \in H$ et pour μ-presque tout $x \in X$:

$$Tf(x) = \int_X K(x, y) f(y)\, \mathrm{d}\mu(y)\,.$$

Si de plus, K est symétrique i.e $K(x, y) = K(y, x)$ $\mu \otimes \mu$-presque partout, alors T est auto-adjoint.

Références

- Brézis, Analyse fonctionnelle, Chap. VI sur les opérateurs compacts, p.89
- Hirsch-Lacombe, Éléments d'analyse fonctionnelle, Chap. 6, p.186 (il y a plus d'exemples)

Leçons concernées : 203, 208

4 Formule d'inversion de Fourier dans L^1

Contexte

Si $f \in L^1(\mathbb{R}^d)$, on définit sa *transformée de Fourier* $\hat{f}$ par :

$$\forall \xi \in \mathbb{R}^d,\ \hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{i\langle x,\xi\rangle}\,\mathrm{d}x\,.$$

Pour f et g dans $L^1(\mathbb{R}^d)$, on définit la *convolution* de f par g, notée $f * g$ par :

$$\forall x \in \mathbb{R}^d,\ f * g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}y\,.$$

Si $x \in \mathbb{R}^d$, on note $|x|$ sa norme euclidienne. Pour $\sigma > 0$, on définit la gaussienne g_σ de variance σ par :

$$\forall x \in \mathbb{R}^d,\ g_\sigma = \frac{1}{(2\pi)^{d/2}\sigma^d} \exp(-\frac{|x|^2}{2\sigma^2})\,.$$

On remarque que $(g_\sigma)_\sigma$ est une approximation de l'unité (générée à partir de g_1) de $\mathbb{R}^d$ lorsque σ tend vers 0. On admet que :

Proposition.

$$\forall f \in L^1(\mathbb{R}^d),\ f * g_\sigma \underset{\sigma\to 0}{\longrightarrow} f \ \textit{dans}\ L^1(\mathbb{R}^d)\,.$$

Développement

Théorème (Formule d'inversion de Fourier)**.** *Soit $f \in L^1(\mathbb{R}^d)$ telle que $\hat{f} \in L^1(\mathbb{R}^d)$. Alors :*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi) e^{-i\langle x,\xi\rangle}\,\mathrm{d}\xi \quad p.p\,.$$

Lemme (Calcul de $\hat{g}_\sigma$)**.**

$$\forall \sigma > 0,\ \forall \xi \in \mathbb{R}^d,\ \hat{g}_\sigma(\xi) = \exp(-\frac{\sigma^2 |\xi|^2}{2})\,.$$

Démonstration. On calcule d'abord g_1 puis on se ramène au cas général par dilatation.

Soit $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$. Comme $g_1 \in L^1(\mathbb{R}^d)$, on peut utiliser le théorème de Fubini-Lebesgue :

$$\begin{aligned}(2\pi)^{d/2}\hat{g}_1(\xi) &= \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle - \frac{|x|^2}{2}}\,\mathrm{d}x \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{k=1}^d e^{ix_k\xi_k - \frac{x_k^2}{2}}\,\mathrm{d}x_k \\ &= \prod_{k=1}^d \hat{\varphi}(\xi_k)\end{aligned}$$

où on a posé : $\forall x \in \mathbb{R},\ \varphi(x) = e^{-\frac{x^2}{2}}$. On s'est ramené à la dimension 1. Comme $x \mapsto x\varphi(x)$ est intégrable, $\hat{\varphi}$ est $\mathcal{C}^1$ et :

$$\forall \xi \in \mathbb{R},\ \hat{\varphi}'(\xi) = \int_{\mathbb{R}} ixe^{ix\xi - \frac{x^2}{2}}\,\mathrm{d}x\,.$$

Par intégration par parties, on obtient une équation différentielle du premier ordre :

$$\forall \xi \in \mathbb{R},\ \hat{\varphi}'(\xi) + \xi\hat{\varphi}(\xi) = 0\,,$$

d'où :

$$\forall \xi \in \mathbb{R},\ \hat{\varphi}(\xi) = \hat{\varphi}(0)\exp(-\frac{\xi^2}{2})\,.$$

Et comme : $\hat{\varphi}(0) = \int_{\mathbb{R}} e^{-\frac{x^2}{2}} \,\mathrm{d}x = \sqrt{2}\int_{\mathbb{R}} e^{-x^2} \,\mathrm{d}x = \sqrt{2\pi}$, on obtient :

$$\forall \xi \in \mathbb{R}^d,\ \hat{g}_1(\xi) = \exp(-\frac{|\xi|^2}{2})\,.$$

Soit $\sigma > 0$. On a :

$$\forall \xi \in \mathbb{R}^d,\ \hat{g}_\sigma(\xi) = \int_{\mathbb{R}^d} g_\sigma(x) e^{i\langle x,\xi\rangle} \,\mathrm{d}x = \int_{\mathbb{R}^d} g_1(\frac{x}{\sigma}) e^{i\langle \frac{x}{\sigma},\sigma\xi\rangle} \frac{\mathrm{d}x}{\sigma^d} = \exp(-\frac{\sigma^2 |\xi|^2}{2})\,.$$

□

Lemme. *Soit $f \in L^1(\mathbb{R}^d)$. Alors :*

$$\forall \sigma > 0,\ \forall x \in \mathbb{R}^d,\ (f * g_\sigma)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi) e^{-i\langle x,\xi\rangle} e^{-\sigma^2 \frac{|\xi|^2}{2}} \,\mathrm{d}\xi\,.$$

Démonstration. Soient $\sigma > 0$ et $x \in \mathbb{R}^d$. On a, d'après le lemme précedent (et en utilisant la parité de g_σ) :

$$\begin{aligned}(f * g_\sigma)(x) &= \int_{\mathbb{R}^d} f(y) g_\sigma(x-y) \,\mathrm{d}y = \frac{1}{(2\pi)^{d/2}\sigma^d} \int_{\mathbb{R}^d} f(y) \hat{g}_{\sigma^{-1}}(y-x) \,\mathrm{d}y \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y) e^{i\langle \xi, y-x\rangle} e^{-\sigma^2 \frac{|\xi|^2}{2}} \,\mathrm{d}\xi \,\mathrm{d}y\,.\end{aligned}$$

Or l'intégrande est majoré par $(y,\xi) \mapsto |f(y)|\, e^{-\sigma^2 \frac{|\xi|^2}{2}}$ qui est dans $L^1(\mathbb{R}^d \times \mathbb{R}^d)$, d'après le théorème de Fubini-Tonelli. On peut donc appliquer le théorème de Fubini-Lebesgue et :

$$(f * g_\sigma)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \underbrace{\left[\int_{\mathbb{R}^d} f(y) e^{i\langle \xi,y\rangle} \,\mathrm{d}y\right]}_{\hat{f}(\xi)} e^{-i\langle \xi,x\rangle} e^{-\sigma^2 \frac{|\xi|^2}{2}} \,\mathrm{d}\xi\,.$$

□

Démonstration de la formule d'inversion de Fourier. Soit $f \in L^1(\mathbb{R}^d)$ telle que $\hat{f} \in L^1(\mathbb{R}^d)$.

— $(g_\sigma)_\sigma$ est une approximation de l'unité donc : $f * g_\sigma \underset{\sigma\to 0}{\longrightarrow} f$ dans $L^1(\mathbb{R}^d)$,

— le lemme précedent et théorème de convergence dominée de Lebesgue (applicable car $\hat{f} \in L^1(\mathbb{R}^d)$) montrent que :

$$f * g_\sigma \underset{\sigma\to 0}{\longrightarrow} \left[x \longmapsto \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi) e^{-i\langle x,\xi\rangle} \,\mathrm{d}\xi\right] \text{ dans } L^1(\mathbb{R}^d)\,.$$

Par unicité de la limite dans $L^1(\mathbb{R}^d)$, les deux expressions sont égales presque partout. □

Application. *Soit $f : x \mapsto \frac{1}{1+x^2}$, alors : $\forall \xi \in \mathbb{R},\ \hat{f}(\xi) = \pi e^{-|\xi|}$.*

Démonstration. On fait le calcul à "l'envers". Comme $x \mapsto e^{-|x|}$ est dans $L^1(\mathbb{R})$:

$$\begin{aligned}\forall \xi \in \mathbb{R},\ \int_{\mathbb{R}} e^{i\xi x - |x|} \,\mathrm{d}x &= \int_{-\infty}^0 e^{(i\xi+1)x} \,\mathrm{d}x + \int_0^{+\infty} e^{(i\xi-1)x} \,\mathrm{d}x \\ &= \frac{1}{i\xi+1} - \frac{1}{i\xi-1} = \frac{2}{1+x^2}\,.\end{aligned}$$

Or $x \mapsto \frac{2}{1+x^2}$ est dans $L^1(\mathbb{R})$ donc on peut appliquer la formule d'inversion de Fourier et :

$$e^{-|x|} = \frac{1}{\pi} \int_{\mathbb{R}} \frac{e^{-ix\xi}}{1+x^2} \,\mathrm{d}x = \frac{1}{\pi} \int_{\mathbb{R}} \frac{e^{ix\xi}}{1+x^2} \,\mathrm{d}x \quad \text{p.p}\,.$$

Comme les deux termes sont continus, on a égalité partout. □

Références

- Briane et Pagès, Théorie de l'intégration, p.321-323 (premier lemme), p.323-324 (deuxième lemme) et p.328 (application)
- Faraut, Calcul intégral, p.133 (conclusion)

Leçons concernées : 234, 235, 236, 239, 250

5 Densité de $\mathcal{C}_c^\infty(\Omega)$ dans $L^p(\Omega)$

Contexte

Dans ce qui suit, si Ω est un ouvert de $\mathbb{R}^d$, on notera $L^p(\Omega) = L^p(\Omega, \mathcal{B}(\Omega), \lambda_d)$ où $p \in [1\,, +\infty]$, où $\mathcal{B}(\Omega)$ désigne l'ensemble des boréliens de Ω et λ_d la mesure de Lebesgue sur $(\Omega, \mathcal{B}(\Omega))$. On suppose connu le théorème de densité suivant :

Théorème. *Soit Ω un ouvert de $\mathbb{R}^d$. Alors $\mathcal{C}_c(\Omega)$ est dense dans $L^p(\Omega)$ pour tout $1 \le p < +\infty$.*

Pour f et g dans $L^1(\mathbb{R}^d)$, on définit la *convolution* de f par g, notée $f * g$ par :

$$\forall x \in \mathbb{R}^d,\ (f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}y\,.$$

On dit qu'une suite $(\rho_n)_{n\in\mathbb{N}^*}$ de fonctions mesurables de $\mathbb{R}^d$ est une *approximation de l'unité* si :

$$\begin{aligned}
&\forall n \in \mathbb{N}^*,\ \rho_n \ge 0 \quad \text{p.p}\,,\\
&\forall n \in \mathbb{N}^*,\ \int_{\mathbb{R}^d} \rho_n = 1\,,\\
&\forall \varepsilon > 0,\ \int_{\{|x|\ge\varepsilon\}} \rho_n\,\mathrm{d}x \underset{n\to+\infty}{\longrightarrow} 0\,.
\end{aligned}$$

On dit que $(\rho_n)_n$ est une *suite régularisante* si de plus :

$$\begin{aligned}
&\forall n \in \mathbb{N}^*,\ \rho_n \in \mathcal{C}_c^\infty(\mathbb{R}^d)\,,\\
&\forall n \in \mathbb{N}^*,\ \operatorname{supp}\rho_n \subset \overline{B}(0, \frac{1}{n})\,.
\end{aligned}$$

Usuellement, pour construire une suite régularisante, on se donne une fonction $\rho \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ telle que $\operatorname{supp}\rho \subset \overline{B}(0,1)$ et $\int_{\mathbb{R}^d} \rho = 1$ et on pose, pour $n \in \mathbb{N}^*$: $\forall x \in \mathbb{R}^d,\ \rho_n(x) = n^d\rho(nx)$. Par exemple, on peut poser :

$$\rho : x \in \mathbb{R}^d \longmapsto \begin{cases} A\exp\left(-\frac{1}{1-|x|^2}\right) & \text{si} \quad |x| < 1 \\ 0 & \text{sinon}\end{cases}\,,$$

où A est une constante telle que $\int_{\mathbb{R}^d} \rho = 1$.

Enfin, on rappelle que le *support* d'une fonction continue f est défini par : $\operatorname{supp} f = \overline{\{f \neq 0\}}$ (attention, si f est juste mesurable alors le support de f est défini comme étant le complémentaire de l'union des ouverts où f est nulle presque partout).

Développement

Théorème. *Soit $\Omega \subset \mathbb{R}^d$ un ouvert. Alors, pour tout $1 \le p < +\infty$, $\mathcal{C}_c^\infty(\Omega)$ est dense dans $L^p(\Omega)$.*

Lemme (Régularisation par convolution)**.** $\forall \rho \in \mathcal{C}^\infty(\mathbb{R}^d),\ \forall g \in \mathcal{C}_c(\mathbb{R}^d),\ \rho * g \in \mathcal{C}^\infty(\mathbb{R}^d)\,.$

Démonstration. Comme $\mathcal{C}^\infty(\mathbb{R}^d)$ est stable par dérivation partielle, on se ramène par récurrence à montrer que :

$$\forall \rho \in \mathcal{C}^\infty(\mathbb{R}^d),\ \forall g \in \mathcal{C}_c(\mathbb{R}^d),\ \begin{cases} \rho * g \in \mathcal{C}^1(\mathbb{R}^d) \quad \text{et} \\ \forall i = 1, \dots, d,\ \forall x \in \mathbb{R}^d,\ \frac{\partial(\rho * g)}{\partial x_i}(x) = \frac{\partial \rho}{\partial x_i} * g(x)\end{cases}.$$

On pose : $\varphi : (x,y) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \rho(x-y)g(y)$ et on note $K_1 = \operatorname{supp} g$ compact :

- pour tout $x \in \mathbb{R}^d$, l'application $y \mapsto \varphi(x,y)$ est mesurable et intégrable,
- presque partout, l'application $x \mapsto \varphi(x,y)$ est $\mathcal{C}^1$,
- pour tout $i = 1, \dots, d$, on a, presque partout et pour tout compact K_2 :

$$\forall x \in K_2,\ \left|\frac{\partial \varphi}{\partial x_i}(x,y)\right| = \left|\frac{\partial \rho}{\partial x_i}(x-y)g(y)\right| \le \|g\|_\infty \sup_{x\in K_1, y \in K_2} \left\|\frac{\partial \rho}{\partial x_i}(x-y)\right\| \mathbb{1}_{K_1} \in L^1(\mathbb{R}^d)\,,$$

car $\lambda(K_1) < +\infty$ (λ désigne la mesure de Lebesgue).

On peut appliquer le théorème de dérivation sous le signe intégrale, ce qui montre le résultat. $\square$

Lemme (support d'une convolution)**.** $\forall f, g \in L^1(\mathbb{R}^d),\ \operatorname{supp}(f*g) \subset \overline{\operatorname{supp} f + \operatorname{supp} g}\,.$

Démonstration. Soit $x \notin \operatorname{supp} f + \operatorname{supp} g$ alors : $\forall y \in \operatorname{supp} g,\ x-y \notin \operatorname{supp} f$. Alors :

$$(f*g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}y = \int_{\operatorname{supp} g} f(x-y)g(y)\,\mathrm{d}y\,.$$

Ainsi, $f*g$ est en particulier nulle presque partout sur l'intérieur du complémentaire de $\operatorname{supp} f + \operatorname{supp} g$, d'où le lemme. $\square$

Ainsi, si $\rho \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ et $g \in \mathcal{C}_c(\mathbb{R}^d)$ alors $\rho * g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$.

Lemme. *Soit $(\rho_n)_n$ une suite régularisante. Alors :*

$$\forall g \in \mathcal{C}_c(\mathbb{R}^d),\ \rho_n * g \underset{n\to+\infty}{\longrightarrow} g \quad \textit{uniformément sur } \mathbb{R}^d\,.$$

Démonstration. Soit $g \in \mathcal{C}_c(\mathbb{R}^d)$. Soit $\varepsilon > 0$. On note $K = \operatorname{supp} g + \overline{B}(0,1)$ puis $\tilde{K} = K + \overline{B}(0,1)$. K et $\tilde{K}$ sont compacts. D'après le théorème de Heine appliqué à g sur $\tilde{K}$, on a :

$$\exists 0 < \delta < 1,\ \forall x \in K, \forall y \in B(0,\delta),\ |g(x-y) - g(x)| \le \varepsilon\,.$$

Ainsi, dès que $n > \frac{1}{\delta}$, on a :

$$\begin{aligned}
\forall x \in K,\ |(\rho_n * g)(x) - g(x)| &= \left|\int_{\mathbb{R}^d} \rho_n(y)(g(x-y)-g(x))\,\mathrm{d}y\right| = \left|\int_{B(0,\frac{1}{n})} \rho_n(y)(g(x-y)-g(x))\,\mathrm{d}y\right| \\
&\le \int_{B(0,\frac{1}{n})} \rho_n(y)\,|g(x-y)-g(x)|\,\mathrm{d}y \\
&\le \varepsilon \int_{B(0,\frac{1}{n})} \rho_n(y)\,\mathrm{d}y = \varepsilon
\end{aligned}$$

D'après le lemme précédent, pour tout $n \in \mathbb{N}^*$, on a : $\operatorname{supp}(\rho_n * g) \subset K$. Et comme $\operatorname{supp} g \subset K$, on a bien convergence uniforme sur $\mathbb{R}^d$. $\square$

Démonstration du théorème. Soient $f \in L^p(\Omega)$ et $\varepsilon > 0$. Soit $g \in \mathcal{C}_c(\Omega)$ tel que : $\|f-g\|_{L^p(\Omega)} \le \varepsilon/2$. On prolonge g à tout $\mathbb{R}^d$ en posant :

$$\tilde{g} : x \in \mathbb{R}^d \longmapsto \begin{cases} g(x) & \text{si} \quad x \in \Omega \\ 0 & \text{sinon} \end{cases}\,.$$

Comme g est continue à support compact sur l'ouvert Ω, $\tilde{g}$ est dans $\mathcal{C}_c(\mathbb{R}^d)$. Soit $(\rho_n)_n$ une suite régularisante. On pose $K = \operatorname{supp} \tilde{g} + \overline{B}(0,1)$ (ensemble compact) de sorte que $\operatorname{supp}(\rho_n * g) \subset K$ et $\operatorname{supp} g \subset K$. D'après le lemme précédent, il existe n_0 assez grand tel que :

$$\forall n \ge n_0,\ \forall x \in \mathbb{R}^d,\ |\tilde{g}(x) - (\rho_n * \tilde{g})(x)| \le \frac{\varepsilon/2}{\lambda(K)^{1/p}}\,.$$

Ainsi, pour tout $n \ge n_0$:

$$\|\tilde{g} - (\rho_n * \tilde{g})\|_{L^p(\mathbb{R}^d)}^p = \int_{\mathbb{R}^d} |\tilde{g}(x) - (\rho_n * \tilde{g})(x)|^p\, \mathbb{1}_K(x)\,\mathrm{d}x \le \left(\frac{\varepsilon}{2}\right)^p\,.$$

On choisit $n \ge n_0$ suffisament grand de sorte que $\operatorname{supp} \rho_n * \tilde{g} \subset \overline{\operatorname{supp} g + B(0,\frac{1}{n})} \subset \Omega$. On pose

$$h = (\rho_n * \tilde{g})_{|\Omega} \in \mathcal{C}_c^\infty(\Omega)\,,$$

alors :

$$\|f-h\|_{L^p(\Omega)} \le \|f-g\|_{L^p(\Omega)} + \|\tilde{g} - (\rho_n * \tilde{g})\|_{L^p(\mathbb{R}^d)} < \varepsilon\,,$$

d'où le théorème. $\square$

Compléments

Il faut au moins connaître la trame de la démonstration de la densité de $\mathcal{C}_c(\Omega)$ dans $L^p(\Omega)$. On suppose connus les deux résultats suivants.

Proposition. *Pour tout $1 \leq p < +\infty$, l'ensemble des fonctions étagées intégrables est dense dans $L^p(\mathbb{R}^d)$.*

Remarque. – C'est une conséquence de la construction des espaces L^p.

- On rappelle qu'une *fonction étagée* est une fonction mesurable ne prenant qu'un nombre fini de valeurs et qu'une telle fonction φ est intégrable si et seulement si $\lambda(\{\varphi \neq 0\}) < +\infty$.
- Ce résultat est valable pour des espaces mesurés quelconque.

Lemme (Tiezte-Urysohn)**.** *Soient O un ouvert de $\mathbb{R}^d$ et K un compact inclus dans O. La fonction*

$$\rho_{K,O} : x \in \mathbb{R}^d \longmapsto \frac{d(x, O^c)}{d(x, O^c) + d(x, K)}$$

est bien définie et continue. De plus, elle vérifie : $\mathbb{1}_K \leq \rho_{K,O} \leq \mathbb{1}_\Omega$.

On traite d'abord le cas $\Omega = \mathbb{R}^d$.

Démonstration de la densité de $\mathcal{C}_c(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$. Dans un premier temps, on montre que l'ensemble des fonctions combinaisons linéaires d'indicatrices de pavés compacts (*fonctions en escalier*) est dense dans $L^p(\mathbb{R}^d)$, puis on approche chaque indicatrice de pavés par une fonction continue à support compact.

(i) D'après la proposition précédente, il suffit d'approcher dans $L^p(\mathbb{R}^d)$ l'indicatrice d'un borélien $A \in \mathcal{B}(\mathbb{R}^d)$ par une fonction en escalier. Par régularité extérieure de la mesure de Lebesgue :

$$\lambda_d(A) = \inf\{\lambda_d(O) \mid A \subset O,\ O \text{ ouvert de } \mathbb{R}^d\}\,.$$

Il existe donc une suite d'ouverts $(O_n)_n$ telle que $A \subset O_n$ et $\lambda_d(O_n) \to \lambda_d(A)$. On pose alors $\tilde{O}_n = O_n \cap B(0,n)$ où $B(0,n)$ est la boule unité ouverte pour, par exemple, la norme euclidienne. La fonction $\mathbb{1}_{\tilde{O}_n}$ est à support compact et :

$$\begin{aligned}\left\|\mathbb{1}_A - \mathbb{1}_{\tilde{O}_n}\right\|_p &= \left\|\mathbb{1}_A - \mathbb{1}_{A\cap B(0,n)} - \mathbb{1}_{B(0,n)}(\mathbb{1}_A - \mathbb{1}_{O_n})\right\|_p \\ &\leq \left\|\mathbb{1}_A - \mathbb{1}_{O_n}\right\|_p + \left\|\mathbb{1}_{A\setminus B(0,n)}\right\|_p \\ &\leq (\lambda_d(A) - \lambda_d(O_n))^{1/p} + \lambda_d(A \setminus B(0,n))^{1/p} \underset{n\to+\infty}{\longrightarrow} 0\,.\end{aligned}$$

car $\lambda_d(A) < +\infty$ et $A \setminus B(0,n) \to\downarrow \emptyset$. Or, tout ouvert borné Ω de $\mathbb{R}^d$ est réunion dénombrable de pavés ouverts :

$$\Omega = \bigcup_{k\geq 0} P_k\,,$$

Et comme toute réunion finie de pavés ouverts est égale à une réunion finie de pavés ouverts disjoints, on peut écrire : $\cup_{k=0}^n P_k = \sqcup_{k=0}^{p_n} P_k^n$. On en déduit :

$$\mathbb{1}_\Omega = \lim_{n\to+\infty} \uparrow \mathbb{1}_{\cup_{k=0}^n P_k} = \lim_{n\to+\infty} \uparrow \mathbb{1}_{\sqcup_{k=0}^{p_n} P_k^n} = \lim_{n\to+\infty} \uparrow \sum_{k=0}^{p_n} \mathbb{1}_{P_k^n}\,.$$

Par convergence dominée, la convergence a lieu dans L^p, ce qui conclut la première partie de la preuve.

(ii) Il reste à approcher l'indicatrice d'un pavé ouvert (d'adhérence compacte) P par une fonction continue à support compact. On peut en fait supposer P fermé (et donc compact) car $\lambda_d(\partial P) = 0$. Pour cela, on utilise encore la régularité extérieure de la mesure de Lebesgue. Il existe

$$\lambda_d(P) = \inf\{\lambda_d(O) \mid P \subset O,\ O \text{ ouvert de } \mathbb{R}^d\}\,.$$

On choisit alors une suite d'ouverts $(O_n)_n$ telle que $P \subset O_n$ et $\lambda_d(O_n) \to \lambda_d(P)$. Alors, à l'aide du lemme de Tiezte-Urysohn :

$$\left\|\rho_{P,O_n} - \mathbb{1}_P\right\|_p \leq \left\|\mathbb{1}_{O_n} - \mathbb{1}_P\right\|_p \leq (\lambda_d(O_n) - \lambda_d(P))^{1/p} \underset{n\to+\infty}{\longrightarrow} 0\,.$$

$\square$

Démonstration du cas général. Soit Ω un ouvert de $\mathbb{R}^d$. Soient $f \in L^p(\Omega)$ et $\varepsilon > 0$. D'après le théorème précédent, il existe $g \in \mathcal{C}_c(\mathbb{R}^d)$ tel que : $\|f - g\|_{L^p(\mathbb{R}^d)} \leq \frac{\varepsilon}{2}$. On utilise cette fois la régularité intérieure de la mesure de Lebesgue :

$$\lambda_d(\Omega) = \inf\{\lambda_d(K) \mid K \subset \Omega,\ K \text{ compact de } \mathbb{R}^d\}\,.$$

On choisit K compact tel que $\lambda_d(\Omega) - \lambda_d(K) \leq \left(\frac{1}{2}\frac{\varepsilon}{\|g\|_\infty}\right)^p$. On choisit O ouvert borné inclus dans Ω et contenant K. On pose alors $\tilde{g} = (g \times \rho_{K,O})_{|\Omega}$, fonction continue à support compact car inclus dans O. On a alors :

$$\begin{aligned}
\|f - \tilde{g}\|_{L^p(\Omega)} &\leq \|(f-g)\mathbb{1}_\Omega\|_{L^p(\mathbb{R}^d)} + \|\mathbb{1}_\Omega - \rho_{K,O}\|_{L^p(\mathbb{R}^d)} \|g\|_\infty \\
&\leq \|f - g\|_{L^p(\mathbb{R}^d)} + (\lambda_d(\Omega) - \lambda_d(K))^{1/p} \|g\|_\infty \leq \varepsilon\,.
\end{aligned}$$

$\square$

Références

- Brézis, Analyse fonctionnelle, Chap. VI sur les opérateurs compacts, Chap. IV.4, p.66 (inspiration)
- Briane et Pagès, Théorie de l'intégration, p.170 (compléments)

Leçons concernées : 201, 202, 234, 239

6 Processus de branchement de Galton-Watson

Contexte

On s'intéresse à la propagation d'un nom de famille au fil des générations.

On suppose que le nom se transmet uniquement par l'homme. On note Z_n le nombre de membres de la n-ième génération et on fait l'hypothèse : $Z_0 = 1$. Soit $(p_k)_k$ une loi de probabilité sur $\mathbb{N}$ et on note T une variable aléatoire suivant cette loi. On suppose que les lois des variables aléatoires décrivants le nombre de fils des hommes sont constantes au fil des générations et indépendantes entre elles. Formellement, on se donne une suite $(T_{ij})_{ij}$ de variables aléatoires i.i.d de loi même loi que T et on suppose que :

$$\forall n \in \mathbb{N}^*,\ Z_n = \sum_{j=1}^{Z_{n-1}} T_{ni}\,.$$

On note G (resp. G_n) la fonction génératrice de T (resp. de Z_n). G est au moins définie sur $]-1\,,1[$:

$$\forall s \in\,]-1\,,1],\ G(s) = \sum_{k\geq 0} p_k s^k\,.$$

Comme $Z_0 = 1$, on remarque que : $G_1 = G$. On note $x_n = \mathbb{P}(Z_n = 0)$ la probabilité d'extinction du patronyme à la n-ième génération. On pose :

$$m := \mathbb{E}(T) = \lim_{s\to 1^-} G'(s) \in [0\,,+\infty]\,.$$

Développement

Théorème (Galton-Watson)**.** *On suppose que* $0 < p_0 < 1$*, alors :*

(i) $m \leq 1 \implies x_n \underset{n\to\infty}{\longrightarrow} 1$,

(ii) $m > 1 \implies x_n \underset{n\to\infty}{\longrightarrow} p \in\,]0\,,1[$,

Démonstration. Le but de la preuve est d'étudier les points fixes de G et de les relier à $(x_n)_n$.

(i) Dans un premier temps, on montre que G est strictement croissante et convexe sur $[0\,,1]$.
En effet, G est $\mathcal{C}^2$ sur $[0\,,1[$ et pour tout $s \in\,]0\,,1[$:

$$G'(s) = \sum_{k\geq 1} k p_k s^{k-1} > 0 \quad \text{car} \quad 0 < p_0 < 1\,,$$
$$G''(s) = \sum_{k\geq 2} k(k-1) p_k s^{k-2} \geq 0\,.$$

Si $p_0 + p_1 = 1$ alors G est affine.
Si $p_0 + p_1 < 1$ alors G est strictement convexe (il existe $k \geq 2$ tel que $p_k > 0$).

(ii) On cherche une relation de récurrence entre G_N et G_{N-1}.
Soit $s \in [0\,,1[$. Comme $\Omega = \bigsqcup_{n\geq 0}\{Z_{N-1} = n\}$ (union disjointe), on a :

$$\begin{aligned}
G_N(s) &= \sum_{k\geq 0} s^k\, \mathbb{P}(Z_N = k) = \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(Z_N = k, Z_{N-1} = n) \\
&= \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(\sum_{i=1}^{n} T_{Ni} = k, Z_{N-1} = n) \\
&= \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(\sum_{i=1}^{n} T_{Ni} = k)\, \mathbb{P}(Z_{N-1} = n) \\
&= \sum_{n\geq 0} \underbrace{\left[\sum_{k\geq 0} s^k\, \mathbb{P}(\sum_{i=1}^{n} T_{Ni} = k)\right]}_{=G(s)^n} \mathbb{P}(Z_{N-1} = n) \\
&= G_{N-1}(G(s)) = G^N(s)\,.
\end{aligned}$$

La troisième ligne est obtenue par indépendance et la quatrième grâce au théorème de Fubini-Tonelli. La dernière égalité est valide car, par croissance de G :

$$\forall s \in [0\,,1],\ 0 < p_0 \leq G(0) \leq G(s) \leq G(1) = 1\,.$$

(iii) On étudie la suite $(x_n)_n$.
Comme $\{Z_n = 0\} \subset \{Z_{n+1}\}$, $(x_n)_n$ est une suite croissante. De plus, elle est majorée par 1 donc :

$$\exists p \in]0\,,1],\ x_n \underset{n\to+\infty}{\longrightarrow} p\,.$$

($p \neq 0$ car $p \geq x_1 = p_0 > 0$)
De plus : $\forall n,\ x_{n+1} = G_{n+1}(0) = G \circ G^n(0) = G(x_n)$. Par continuité de G, on obtient : $p = G(p)$.

(iv) Soit F l'ensemble des points fixes de G. On sait déjà que $1 \in F$ (car $G(1) = \sum_{k\geq 0} p_k = 1$) et que $0 \notin F$ (car $G(0) = p_0 > 0$). Soit $u \in F$. Alors, par stricte croissance de G :

$$x_1 = p_0 = G(0) < G(u) = u\,,$$

et, par récurrence : $\forall n,\ x_{n+1} = G(x_n) < G(u) = u$. En passant à la limite : $p \leq u$ et $p = \min F$.
Si $p_0 + p_1 = 1$ alors $F = \{1\}$ (car $G(0) > 0$).
Si $p_0 + p_1 < 1$ alors $|F| \leq 2$. Si ce n'était pas le cas, en notant $u_1 < u_2 < 1$ deux points fixes de G, le théorème de Rolle appliqué à la fonction $s \mapsto G(s) - s$ entre u_1 et u_2 et entre u_2 et 1 donne l'existence de $v_1 < v_2$ tels que : $G'(v_1) = 1 = G'(v_2)$, ce qui contredit la stricte convexité de G.
Supposons que $|F| = 2$ alors $p \in]0\,,1[$ et d'après le théorème de Rolle, il existe $z \in]p\,,1[$ tel que : $G'(z) = 1$. Par stricte convexité : $m := \lim\limits_{n\to+\infty} G'(s) > 1$.
Supposons que $|F| = 1$ alors, puisque $p_0 > 0$ on a : $\forall s \in [0\,,1[,\ G(s) > s$, donc :

$$\int_s^1 G'(x)\,\mathrm{d}x = 1 - G(s) < 1 - s = \int_s^1 \mathrm{d}x \implies \int_s^1 (1 - G'(x))\,\mathrm{d}x \geq 0\,.$$

Nécessairement $m \leq 1$ car sinon il existerait $s < 1$ tel que : $\forall s < x \leq,\ G'(x) \geq 1$ et cela contredirait ce qui précède.

□

Références

– Cottrell, Genon-Catalot, Duhamel et Meyre, Exercices de probabilités, p.72

Leçons concernées : 229, 260, 261, 264

7 Principe du maximum faible

Contexte

Soient $A = (a_{ij})_{ij} \in \mathcal{S}_n^+(\mathbb{R}) \setminus \{0\}$ et $b = (b_i)_i \in \mathbb{R}^n$. On définit l'opérateur différentiel suivant :

$$L = \sum_{ij} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial}{\partial x_i}\,.$$

On cherche à étudier les solutions des problèmes de Dirichlet du type :

$$\left\{ \begin{array}{lll} Lu = 0 & \text{sur} & \Omega \\ u_{|\partial\Omega} = f & \text{sur} & \partial\Omega \end{array} \right. ,$$

où Ω est un ouvert borné de $\mathbb{R}^n$ et f une fonction continue sur $\partial\Omega$. L'équation de la chaleur ($L = \Delta - \frac{\partial}{\partial t}$) rentre dans ce cas de figure mais ce n'est pas le cas de l'équation d'onde ($L = \Delta - \frac{\partial^2}{\partial t^2}$).

Développement

Théorème (Principe du maximum faible)**.** *Soit $\Omega \subset \mathbb{R}^n$ un ouvert* borné. *On note $\partial\Omega$ la frontière de Ω. Soit u une fonction à valeurs réelles $\mathcal{C}^2$ sur Ω et continue sur $\overline{\Omega}$. Alors :*

$$\forall x \in \Omega,\ Lu(x) \geq 0 \implies \forall x \in \overline{\Omega},\ u(x) \leq \max_{x \in \partial\Omega} u(y)\,.$$

Démonstration. On décompose la démonstration en plusieurs étapes.

(i) Si A et B sont deux matrices symétriques positives alors : $\operatorname{tr}(AB) \geq 0$.
En effet, A étant symétrique réelle et positive, il existe $P \in \mathcal{O}_n(\mathbb{R})$ tel que $A' = {}^tPAP = (a'_{ij})_{ij}$ soit diagonale à coefficients diagonaux positifs. On note $B' = {}^tPBP = (b'_{ij})_{ij}$, alors :

$$\operatorname{tr}(AB) = \operatorname{tr}(A'B') = \sum_i a'_{ii} b'_{ii} \geq 0\,.$$

($b'_{ii} = \langle BPe_i, Pe_i\rangle \geq 0$)

(ii) Montrons que si $Lu(x) > 0$ sur Ω alors le maximum de u sur $\overline{\Omega}$ n'est pas atteint dans Ω.
Supposons que u atteigne son maximum en un point a de Ω alors, d'après les conditions nécessaires d'extrema :

$$Du(a) = 0 \quad \text{et} \quad -D^2u(a) \quad \text{est positive}\,.$$

D'où, en utilisant (i) :

$$Lu(a) = \operatorname{tr}(AD^2u(a)) + \langle b, Du(a)\rangle = \operatorname{tr}(AD^2u(a)) \leq 0$$

(iii) On suppose maintenant le théorème. Supposons que : $\forall x \in \Omega,\ Lu(x) \geq 0$. On va perturber u pour se ramener au cas (ii).
Soient $\varepsilon > 0$ et $\lambda \in \mathbb{R}$. On choisit $i \in [\![1\,, n]\!]$ tel que $a_{ii} > 0$. Un tel i existe car sinon :

$$\forall k \neq l,\ \langle A(e_k \pm e_l), (e_k \pm e_l)\rangle = \pm 2a_{kl} \geq 0\,,$$

d'où $A = 0$ ce qui n'est pas le cas par hypothèse. On pose $v(x) = u(x) + \varepsilon e^{\lambda x_i}$. Par linéarité de L, on a :

$$\forall x \in \Omega,\ Lv(x) = Lu(x) + \varepsilon e^{\lambda x_i}(a_{ii}\lambda^2 + b_i\lambda)\,.$$

Comme $a_{ii} > 0$, on peut choisir λ tel que $a_{ii}\lambda^2 + b_i\lambda > 0$. D'après (ii) :

$$\forall x \in \overline{\Omega},\ u(x) < v(x) \leq \max_{y \in \partial\Omega} v(y) \leq \max_{y \in \partial\Omega} u(y) + \varepsilon e^M\,,$$

où M est telle que $\overline{\Omega} \subset B(0, M/\lambda)$. En faisant tendre ε vers 0, on obtient le résultat.

$\square$

Remarque. Si $L = \Delta$, alors on peut sauter l'étape (i) et poser $v(x) = u(x) + \varepsilon x_i^2$.

Application (Unicité du problème de Dirichlet)**.** *Soit Ω est un ouvert borné de $\mathbb{R}^n$. Il y a au plus une solution u, de classe $\mathcal{C}^2$ sur Ω et continue sur $\overline{\Omega}$, au problème :*

$$\left\{ \begin{array}{lll} Lu = 0 & sur & \Omega \\ u_{|\partial\Omega} = f & sur & \partial\Omega \end{array} \right. ,$$

où f une fonction continue sur $\partial\Omega$.

Démonstration. Soient u_1 et u_2 deux solutions. Alors, par linéarité du problème, $v = u_1 - u_2$ vérifie :

$$\forall x \in \Omega,\ Lv(x) = 0 \quad \text{et} \quad v_{|\partial\Omega} = 0\,.$$

D'après le principe du maximum faible : $\forall x \in \overline{\Omega},\ v(x) \leq 0$. On peut faire de même avec $-v$ et : $v = 0$. □

Remarque.

- Il y a aussi un principe du maximum fort, on suppose de plus que A est définie et que Ω est connexe et on montre que si u atteint son maximum en un point intérieur, alors u est constante sur Ω. En particulier, les applications harmoniques (et donc les applications holomorphes) vérifient le principe du maximum.
- Ces résultats restent valables si on suppose que les a_{ij} et les b_i sont des fonctions de x et vérifient certaines hypothèses supplémentaires (pour le principe du maximum faible, on peut les supposer continues et bornées, pour le principe du maximum fort, on peut les supposer dans L^∞ et supposer que $A(x)$ est coercive, i.e $\langle Ax, x\rangle \geq \alpha \|x\|_2^2$ pour un certain $\alpha > 0$)

Références

- Rouvière, Petit Guide du Calcul Différentiel, p.382 (principe du maximum faible)
- Queffélec et Zuily, Analyse pour l'agrégation, p.464 (remarque)

Leçons concernées : 158, 219, 221, 222

8 Polynômes de Bernstein

Contexte

On note $\mathcal{C} = \mathcal{C}([0\,,1])$ l'ensemble des applications continues de $[0\,,1]$ à valeurs réelles. On munit $\mathcal{C}$ de la norme uniforme $\|\cdot\|_\infty$. Pour $f \in \mathcal{C}$, on définit le *module de continuité* de f par :

$$w_f : t \in \mathbb{R}^+ \longmapsto \sup\{|f(x) - f(y)|\,,\ |x-y| \leq t\}\,.$$

On rappelle que w_f est une application croissante, sous-additive et continue en 0 (conséquence du théorème de Heine). En particulier, on en déduit :

$$\forall \lambda \in \mathbb{R}^+,\ \forall t \in \mathbb{R}^+,\ w_f(\lambda t) \leq w_f(\lceil\lambda\rceil t) \leq \lceil\lambda\rceil w_f(t) \leq (\lambda+1) w_f(t)\,,$$

où $\lceil\lambda\rceil$ désigne la partie supérieure de λ.

On dit que les variables aléatoires $\varepsilon_1, \dots, \varepsilon_N$ forment une *suite de Rademacher* si les ε_i sont i.i.d, à valeurs dans $\{-1, 1\}$ et vérifient :

$$\mathbb{P}(\varepsilon_i = -1) = \frac{1}{2} = \mathbb{P}(\varepsilon_i = 1)\,.$$

Développement

Théorème (Bernstein)**.** *Pour $f \in \mathcal{C}$ et $n \in \mathbb{N}^*$, on note :*

$$B_n f : x \in [0\,,1] \longmapsto \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

le n-ième polynôme de Bernstein *associé à f. Alors :*

(i) $\|f - B_n f\|_\infty \leq C w_f(\frac{1}{\sqrt{n}})$ (en particulier, $B_n f \underset{n\to+\infty}{\longrightarrow} f$ uniformément sur $[0\,,1]$) et

(ii) la majoration est optimale : il existe $f \in \mathcal{C}$ lipchitzienne et une constante $\delta > 0$ tels que :

$$\|f - B_n f\|_\infty \geq \frac{\delta}{\sqrt{n}}\,.$$

Démonstration du (i). Soient $x \in [0\,,1]$ et $n \in \mathbb{N}^*$. Soit $(X_i)_{i=1,\dots,n}$ une suite i.i.d de variables aléatoires suivant une loi de Bernouilli de paramètre x. En notant $S_n = X_1 + \cdots + X_n$, on remarque que :

$$B_n f(x) = \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]\,.$$

On a :

$$\begin{aligned}
|f(x) - B_n f(x)| &= \left|\mathbb{E}\left[f(x) - f\left(\frac{S_n}{n}\right)\right]\right| \\
&\leq \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right] \\
&\leq \mathbb{E}\left[w_f\left(\left|x - \frac{S_n}{n}\right|\right)\right] = \mathbb{E}\left[w_f\left(\sqrt{n}\left|x - \frac{S_n}{n}\right| \frac{1}{\sqrt{n}}\right)\right] \\
&\leq w_f\left(\frac{1}{\sqrt{n}}\right) \mathbb{E}\left[1 + \sqrt{n}\left|x - \frac{S_n}{n}\right|\right] \\
&\leq w_f\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\left\|x - \frac{S_n}{n}\right\|_1\right)
\end{aligned}$$

or, par l'inégalité de Cauchy-Schwarz : $\left\|x - \frac{S_n}{n}\right\|_1 \leq \left\|x - \frac{S_n}{n}\right\|_2 = \sqrt{\mathrm{Var}\left(\frac{S_n}{n}\right)} = \sqrt{\frac{x(1-x)}{n}} \leq \frac{1}{2\sqrt{n}}$. On en déduit le résultat voulu avec $C = \frac{3}{2}$. $\square$

Pour montrer (ii), on va utiliser le :

Lemme (Inégalité de Kintchine)**.** *Soient* $a_1, \dots, a_N \in \mathbb{R}$*. Soient* $\varepsilon_1, \dots, \varepsilon_N$ *une suite de Rademacher. Alors, en notant* $X = \sum_j a_j \varepsilon_j$*, on a :* $\|X\|_1 \geq \frac{\|X\|_2}{\sqrt{e}}$*.*

Démonstration du lemme. Par homogénéité, on peut supposer $\|X\|_2 = 1$, i.e $\sum_j a_j^2 = 1$. Posons : $g = \prod_j (1 + i a_j \varepsilon_i)$. On a :

$$\forall \omega,\ |g(\omega| = \prod_j \sqrt{1 + a_j^2 \varepsilon_j^2} = \prod_j \sqrt{1 + a_j^2} \leq \prod_j \exp(a_j^2)^{1/2} = \exp(\sum_j a_j^2)^{1/2} = \sqrt{e}\,.$$

Or, par indépendance et puisque les variables sont centrées :

$$\begin{aligned}
\mathbb{E}(Xg) &= \sum_j a_j\, \mathbb{E}(\varepsilon_j g) = \sum_j a_j\, \mathbb{E}\left(\varepsilon_j (1 + i a_j \varepsilon_i) \prod_{k \neq j} (1 + i a_k \varepsilon_k)\right) \\
&= \sum_j a_j\, \mathbb{E}(\varepsilon_j (1 + i a_j \varepsilon_i)) \prod_{k \neq j} \underbrace{\mathbb{E}(1 + i a_k \varepsilon_k)}_{=1} \\
&= \sum_j i a_j^2\, \mathbb{E}(\varepsilon_j^2) = i
\end{aligned}$$

Ainsi :

$$1 = |i| = |\mathbb{E}(Xg)| \leq \mathbb{E}(|Xg|) \leq \|g\|_\infty \|X\|_1 \leq \sqrt{e}\, \|X\|_1\,.$$

□

Démonstration du (ii). On pose $f : x \mapsto |x - 1/2|$. On a :

$$\|f - B_n f\|_\infty \geq |f(1/2) - B_n f(1/2)| = |B_n(1/2)| = \mathbb{E}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right|\right) = \frac{1}{2n}\, \mathbb{E}(|\varepsilon_1 + \cdots + \varepsilon_n|)\,,$$

où $\varepsilon_1, \dots, \varepsilon_N$ est une suite de Rademacher. L'inégalité de Kintchine donne :

$$\|f - B_n f\|_\infty \geq \frac{1}{2n\sqrt{e}} \underbrace{\|\varepsilon_1 + \cdots + \varepsilon_n\|_2}_{= \sqrt{\mathrm{Var}(\varepsilon_1 + \cdots + \varepsilon_n)} = \sqrt{n}} = \frac{\delta}{\sqrt{n}}\,,$$

où $\delta = \frac{1}{2\sqrt{e}}$. □

Remarque. La vitesse de convergence des polynômes de Bernstein n'est pas optimale (le théorème de Jackson donne une convergence en $O\big(w_f(\frac{1}{n})\big)$) cependant, ils épousent bien les propriétés de f.

Références

– Queffélec et Zuily, Analyse pour l'agrégation, p.246 (inégalité de Kintchine) et p.519 (théorème de Bernstein)

Leçons concernées : 260, 264

9 Lemme de Morse

Contexte

On note $\mathcal{M} = \mathcal{M}_n(\mathbb{R})$ (resp. $\mathcal{S} = \mathcal{S}_n(\mathbb{R})$, resp. $\mathcal{A} = \mathcal{A}_n(\mathbb{R})$) l'ensemble des matrices carrées (resp. matrices symétriques, resp. matrices antisymétriques) à coefficients dans $\mathbb{R}$ et de taille n. On note aussi $\mathrm{GL}_n(\mathbb{R})$ l'ensemble des matrices inversibles de $\mathcal{M}$.

Développement

Théorème (Lemme de Morse)**.** *Soit $U \subset \mathbb{R}^n$ un ouvert contenant l'origine. Soit $f : U \to \mathbb{R}$ une application $\mathcal{C}^3$ telle que 0 soit un point critique quadratique non dégénéré de f, c'est-à-dire :*

$$Df(0) = 0 \quad \textit{et} \quad D^2f(0) \quad \textit{est non dégénérée.}$$

On note $(p, n-p)$ la signature de $D^2f(0)$, alors il existe $u : x \mapsto u(x) = (u_1(x), \dots, u_n(x))$ un difféomorphisme local et de classe $\mathcal{C}^1$ tel que :

$$u(0) = 0 \quad \textit{et} \quad f(x) - f(0) = u_1(x)^2 + \cdots + u_p(x)^2 - u_{p+1}(x)^2 - \cdots - u_n(x)^2 \,.$$

Lemme (Réduction des formes quadratiques, version différentielle)**.** *Soit $A_0 \in \mathcal{S} \cap \mathrm{GL}_n(\mathbb{R})$. Alors, il existe un voisinage V de A_0 dans $\mathcal{S}$ et une application $M : A \in V \mapsto M(A) \in \mathrm{GL}_n(\mathbb{R})$ de classe $\mathcal{C}^1$ tels que :*

$$\forall A \in V,\ A = {}^tM(A) \cdot A_0 \cdot M(A) \,.$$

Remarque. En fait, il est inutile pour la démonstration du lemme de Morse de prendre un voisinage dans $\mathrm{GL}_n(\mathbb{R})$.

Démonstration du lemme. Soit l'application $\varphi : M \in \mathcal{M} \mapsto {}^tMA_0M \in \mathcal{S}$. Le but de la preuve est "d'inverser" à droite φ.

Tout d'abord, φ est bilinéaire et $\mathcal{S}$ est de dimension finie, donc φ est de classe $\mathcal{C}^1$ et on a, pour tout $H \in \mathcal{M}$:

$$\varphi(I + H) - \varphi(I) = {}^t(A_0H) + (A_0H) + O\Big(\|H\|^2\Big) \,,$$

où $\|\cdot\|$ est, par exemple, une norme d'opérateur. En en déduit :

$$\forall H \in \mathcal{M},\ D\varphi(I)H = {}^t(A_0H) + (A_0H) \,.$$

Clairement : $\ker(D\varphi(I)) = A_0^{-1}\mathcal{A}$. Par ailleurs, $D\varphi(I)$ est surjective (pour le voir, on peut appliquer le théorème du rang ou, tout simplement, remarquer que $\frac{1}{2}A_0^{-1}A$ est un antécédant de A par $D\varphi(I)$). Comme $\mathcal{S}$ et $\mathcal{A}$ sont en somme directe et A_0 est inversible, on a : $\mathcal{M} = A_0^{-1}\mathcal{S} \oplus A_0^{-1}\mathcal{A}$. Ainsi, si on note ψ la restriction de φ à $F := A_0^{-1}\mathcal{S}$, on a $D\psi(I)$ inversible et on peut appliquer le théorème d'inversion locale. Il existent U voisinage de I dans F et V voisinage de $A_0 = \psi(I)$ dans $\mathcal{S}$ tels que $\psi : U \to V$ soit un $\mathcal{C}^1$-difféomorphisme. De plus, quitte à prendre U plus petit, on peut supposer U contenu dans l'ouvert $\mathrm{GL}_n(\mathbb{R})$. L'application $M = \psi_{|V}^{-1}$ est celle recherchée. □

Démonstration du lemme de Morse. On applique la formule de Taylor avec reste intégral à l'ordre 1. Pour x proche de 0 :

$$f(x) - f(0) = \int_0^1 (1-t)\,{}^txD^2f(tx)x \,\mathrm{d}t = {}^tx \underbrace{\left[\int_0^1 (1-t)D^2f(tx)\,\mathrm{d}t\right]}_{:=Q(x)} x \,.$$

Q est une application à valeurs dans $\mathcal{S}$ (intégration sur un sous-espace vectoriel) de classe $\mathcal{C}^1$ (f est $\mathcal{C}^3$ donc l'intégrande est $\mathcal{C}^1$, on utilise ensuite le théorème de dérivation sous le signe intégrale). Comme $D^2f(0) = 2Q(0)$ est non dégénérée, d'après le lemme, il existe un voisinage V de $Q(0)$ et $M : A \in V \mapsto M(A) \in \mathrm{GL}_n(\mathbb{R})$ de classe $\mathcal{C}^1$ tels que :

$$\forall A \in V,\ A = {}^tM(A) \cdot Q(0) \cdot M(A) \,.$$

Par continuité de Q, il existe un voisinage W de 0 dans $\mathbb{R}^n$ tel que $Q(W) \subset V$ et

$$\forall x \in W,\ Q(x) = {}^tM(Q(x))Q(0)M(Q(x)) \implies f(x) - f(0) = {}^tx{}^tM(Q(x))Q(0)\underbrace{M(Q(x))x}_{:=y(x)}.$$

y est une application $\mathcal{C}^1$ et sa différentielle à l'origine est $M(Q(0))$, matrice inversible. Par le théorème d'inversion locale, y est un $\mathcal{C}^1$-difféomorphisme local en 0. Enfin, $Q(0)$ étant de signature $(p, n-p)$, il existe $P \in \mathrm{GL}_n(\mathbb{R})$ tel que :

$$\forall u = (u_1, \ldots, u_n) \in \mathbb{R}^n,\ {}^tu{}^tPQ(0)Pu = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2\,.$$

$u = P^{-1}y$ (un $\mathcal{C}^1$-difféomorphisme local en 0 car P est inversible), répond au problème. $\square$

Remarque. Le lemme de Morse permet d'étudier le comportement local du graphe d'une application $f : U \subset \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^3$ en un point critique non dégénéré.

Références

– Rouvière, Petit Guide du Calcul Différentiel, p.209 (lemme) et p.344 (lemme de Morse)

Leçons concernées : 214, 215, 218

10 Densité des polynômes orthogonaux

Contexte

On confondra systématiquement polynôme et fonction polynômiale. Soit I un intervalle de $\mathbb{R}$ ($\mathring{I} \neq \emptyset$). On appelle *fonction de poids* toute fonction $\rho : I \to \mathbb{R}$ mesurable, strictement positive telle que :

$$\forall n \in \mathbb{N}, \ \int_I |x|^n \rho(x) \, \mathrm{d}x < +\infty \, .$$

On note $L^2(I, \rho)$ l'espace des fonctions de carré intégrable pour la mesure de densité ρ par rapport à la mesure de Lebesgue. On munit $L^2(I, \rho)$ du produit scalaire :

$$\langle f, g \rangle_\rho = \int_I f(x) g(x) \rho(x) \, \mathrm{d}x \, ,$$

(on note $\|\cdot\|_\rho$ la norme associée) ce qui en fait un espace de Hilbert qui contient, par hypothèse, l'ensemble des polynômes. On construit, à l'aide du procédé de Gram-Schmidt appliqué à la famille $(X^n)_{n\in\mathbb{N}}$ une famille $(p_n)_{n\in\mathbb{N}}$ échelonnée en degré de polynômes unitaires orthogonaux deux à deux pour $\langle \cdot, \cdot \rangle_\rho$ et telle que :

$$\forall N \in \mathbb{N}, \ \mathrm{vect}(X^n, n = 0, \ldots, N) = \mathrm{vect}(p_n, n = 0, \ldots, N) \, .$$

Cette famille est unique et ses éléments sont les *polynômes orthogonaux* associés au poids ρ.

Développement

Théorème (Densité des polynômes orthogonaux)**.** *On suppose qu'il existe $\alpha > 0$ tel que :*

$$\int_I e^{\alpha|x|} \rho(x) \, \mathrm{d}x < +\infty \, ,$$

alors $(p_n)_n$ est une base hilbertienne de $L^2(I, \rho)$.

Démonstration. On va utiliser la caractérisation de la densité par dualité. Soit $f \in L^2(I, \rho)$ telle que :

$$\forall n \in \mathbb{N}, \ \langle X^n, f \rangle_\rho = \int_I x^n f(x) \rho(x) \, \mathrm{d}x = 0 \, .$$

Posons :

$$\varphi : x \in \mathbb{R} \longmapsto \begin{cases} f(x)\rho(x) & \text{si } x \in I \\ 0 & \text{sinon} \end{cases} \, .$$

Par l'inégalité de Cauchy-Schwarz (dans $L^2(I, \rho)$) :

$$\int_\mathbb{R} |\varphi(x)| \, \mathrm{d}x = \int_I |f(x)| \, \rho(x) \, \mathrm{d}x \leq \|f\|_\rho \sqrt{\int_I \rho(x) \, \mathrm{d}x} < +\infty \, ,$$

et $\varphi \in L^1(\mathbb{R})$. On peut considérer sa transformée de Fourier $\hat{\varphi}$ définie par :

$$\forall \xi \in \mathbb{R}, \ \hat{\varphi}(\xi) = \int_\mathbb{R} \varphi(x) e^{i\xi x} \, \mathrm{d}x = \int_I f(x) e^{i\xi x} \rho(x) \, \mathrm{d}x \, .$$

Posons, pour $x \in I$ et $z \in \mathbb{C}$: $g(z, x) = f(x)\rho(x)e^{izx}$. Notons $B_\alpha := \{z \in \mathbb{C} \mid |\Im(z)| \leq \alpha/2\}$. On a :

- pour tout $z \in B_\alpha$, l'application $x \mapsto g(z, x)$ est mesurable sur I,
- pour (presque) tout $x \in I$, l'application $z \mapsto g(z, x)$ est holomorphe sur B_α,
- pour tout $z \in B_\alpha$ et (presque) tout $x \in I$:

$$|g(z, x)| \leq \underbrace{e^{\frac{\alpha|x|}{2}} \, |f(x)| \, \rho(x)}_{:=h(x)} \, ,$$

  or, par l'inégalité de Cauchy-Schwarz :

$$\int_I |h(x)| \, \mathrm{d}x \leq \|f\|_\rho^2 \times \int_I e^{\alpha|x|} \rho(x) \, \mathrm{d}x < +\infty \, ,$$

  donc $h \in L^1(I)$.

On peut appliquer le théorème d'holomorphie sous le signe intégrale et $\hat{\varphi}$ se prolonge en une fonction holomorphe F sur B_α qui vérifie :

$$\forall n \in \mathbb{N},\ \forall z \in B_\alpha,\ F^{(n)}(z) = \int_I f(x)(ix)^n e^{ixz} \rho(x)\,\mathrm{d}x\,.$$

En particulier : $\forall n \in \mathbb{N},\ F^{(n)}(0) = \int_I f(x)(ix)^n \rho(x)\,\mathrm{d}x = 0$. L'unicité du développement en série entière d'une fonction holomorphe montre que F est nulle sur un voisinage de 0 donc, par le principe du prolongement analytique, sur tout le connexe B_α. En particulier, F est nulle sur l'axe des réels. On en déduit que la transformée de Fourier de φ est nulle et par injectivité de la transformée de Fourier, $\varphi = 0$ presque partout. Comme $\rho > 0$, cela montre que $f = 0$ presque partout. □

Remarque. Pour $I =]0\,,+\infty[$ et $\rho(x) = x^{-\ln x}$, les polynômes orthogonaux ne forment pas une base hilbertienne de $L^2(I,\rho)$.

Démonstration. Tout d'abord, ρ est bien une fonction poids car, par le changement de variables $u = \ln x$:

$$\forall n \in \mathbb{N},\ \int_I |x|^n \rho(x)\,\mathrm{d}x = \int_{\mathbb{R}} e^{(n+1)u-u^2}\,\mathrm{d}u < +\infty\,.$$

On considère la fonction (non nulle) $f : x \in I \mapsto \sin(2\pi \ln x)$. Soit $n \in \mathbb{N}$. Par le changement de variables $u = \ln x$:

$$\langle f, X^n\rangle_\rho = \int_I \sin(2\pi \ln x) x^n x^{-\ln x}\,\mathrm{d}x = \int_{\mathbb{R}} \sin(2\pi u) e^{(n+1)u-u^2}\,\mathrm{d}u = e^{\frac{(n+1)^2}{4}} \int_{\mathbb{R}} \sin(2\pi u) e^{-\left(u-\frac{n+1}{2}\right)^2}\,\mathrm{d}u\,.$$

Par le changement de variables $v = u - \frac{n+1}{2}$ et par imparité, on trouve :

$$\langle f, X^n\rangle_\rho = (-1)^{n+1} e^{\frac{(n+1)^2}{4}} \int_{\mathbb{R}} \sin(2\pi v) e^{-v^2}\,\mathrm{d}v = 0\,.$$

□

Compléments

Remarque. Dans le cas où $I =]a\,,b[$ est borné, les polynômes orthogonaux forment toujours une base hilbertienne. En effet, on sait que $\mathcal{C}_c(I)$ est dense dans $L^2(I,\rho)$ or, d'après le théorème Weierstrass, l'espace des polynômes est dense dans $\mathcal{C}([a\,,b])$ (donc dans $\mathcal{C}_c(I)$) pour la convergence uniforme. Comme I est de mesure finie pour $\rho(x)\,\mathrm{d}x$, on en déduit que la norme $\|\cdot\|_\rho$ est dominée par la norme uniforme et l'espaces des polynômes est dense dans $L^2(I,\rho)$.

Exemple. Certains polynômes orthogonaux sont à connaître :

- Les polynômes de *Legendre* : $I =]-1\,,1[$, $\rho(x) = 1$ et $p_n(x) = \frac{n!}{(2n)!}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n (x^2-1)^n$,
- Les polynômes de *Laguerre* : $I =]0\,,+\infty[$, $\rho(x) = e^{-x}$ et $p_n(x) = (-1)^n e^x \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n (x^n e^{-x})$,
- Les polynômes de *Hermite* : $I = \mathbb{R}$, $\rho(x) = e^{-x^2}$ et $p_n(x) = (-1)^n 2^{-n} e^{x^2} \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n (e^{-x^2})$,
- Les polynômes de *Tchebychev* : $I =]-1\,,1[$, $\rho(x) = \frac{1}{\sqrt{1-x^2}}$ et $p_n(x) = \cos(n \arccos(x))$.

Proposition. *Les racines de p_n sont toutes simples et dans situées I.*

Démonstration. Soit $n \in \mathbb{N}$. Soient $x_1, \ldots, x_k$ les zéros distincts de p_n contenus dans I ($k \le n$) et $m_1, \ldots, m_k$ leurs multiplicités respectives. On pose $\varepsilon_i = 0$ si m_i est pair et $\varepsilon_i = 1$ si m_i est impair, ainsi que :

$$q(x) = \prod_{i=1}^{k} (x - x_i)^{\varepsilon_i}\,,$$

polynôme de degré au plus n. De cette façon, le polynôme qp_n est de signe constant et par conséquent :

$$\int_I q(x) p_n(x) \rho(x)\,\mathrm{d}x \neq 0\,,$$

or p_n est orthogonal à $\mathcal{P}_{n-1}$. Nécessairement $k = n$. □

On cherche à approcher numériquement l'intégrale d'une fonction f contre la mesure $\rho(x)\,\mathrm{d}x$. La méthode de Gauss est une méthode de quadrature du type :

$$\int_I f(x)\rho(x)\,\mathrm{d}x \simeq \sum_i \lambda_i f(x_i)\,,$$

où les x_i sont dans I et les λ_i des coefficients.

On rappelle qu'une méthode est d'*ordre* ℓ si elle est exacte pour tout polynôme de degré au plus ℓ et s'il existe un polynôme d'ordre $\ell+1$ tel que la méthode ne soit pas exacte.

Proposition. *Il existe un choix et un seul des points x_i et des coefficients λ_i de sorte que la méthode soit d'ordre $N = 2\ell+1$. Les points x_i appartiennent à I et sont les racines du $(\ell+1)$-ième polynôme orthogonal pour le poids ρ.*

Démonstration. On ne donne pas les détails de la preuve.

(i) Pour l'analyse, on pose $\pi_{\ell+1} = \prod_{i=0}^{\ell}(x-x_i)$ de sorte que : $\forall p \in \mathcal{P}_\ell,\ p\pi_{\ell+1} \in \mathcal{P}_{2\ell+1}$. L'hypothèse sur l'ordre de la méthode implique que $\pi_{\ell+1}$ est le $(\ell+1)$-ième polynôme orthogonal. Pour trouver l'expression des λ_i, on utilise les polynômes de Lagrange associés aux points x_i.

(ii) Pour la synthèse, on montre d'abord que la méthode est d'ordre au moins ℓ en remarquant qu'un polynôme est égal à son polynôme interpolateur de Lagrange. Pour montrer que la méthode est d'ordre au moins $2\ell+1$, on fait la division euclidienne par $\pi_{\ell+1}$. On admet que la méthode est exactement d'ordre $2\ell+1$ (utilise le noyau de Peano).

$\square$

Références

– Beck, Malick et Peyré, Objectif agrégation, p.141 (développement)

– Demailly, Analyse numérique et équations différentielles, Chap. II.5 et Chap.III-3 (compléments)

Leçons concernées : 201, 207, 209, 213, 245

11 Théorème de Jackson

Contexte

On note $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ et $\mathcal{C}_{2\pi} = \mathcal{C}(\mathbb{T}, \mathbb{C})$ l'ensemble des applications continues 2π-périodiques. Pour $n \in \mathbb{Z}$, on note $e_n : t \in \mathbb{R} \mapsto \exp(int)$ et pour $N \in \mathbb{N}$, on note $\mathcal{P}_N = \mathrm{vect}(e_n,\ n \in [\![-N\,,N]\!])$ l'ensemble des polynômes trigonométriques de degré au plus N. On rappelle les définitions des noyaux usuels :

Définition. Le noyau de *Dirichlet*, D_N, est défini par :

$$D_N = \sum_{|k| \leq N} e_k \in \mathcal{P}_N\,.$$

Le noyau de *Féjer*, K_N, est défini par :

$$K_N = \frac{1}{N} \sum_{k=0}^{N-1} D_k \in \mathcal{P}_{N-1} \subset \mathcal{P}_N\,.$$

Le noyau de *Jackson*, J_N, est défini par :

$$J_N = \frac{K_N^2}{\|K_N^2\|_1} \in \mathcal{P}_{2(N-1)} \subset \mathcal{P}_{2N}\,.$$

Proposition. *On a les propriétés suivantes :*

$$\forall t \in \mathbb{T},\ D_N(t) = \frac{\sin\big((N+\frac{1}{2})t\big)}{\sin(\frac{t}{2})}\,,$$

$$\forall t \in \mathbb{T},\ K_N(t) = \sum_{|n| \leq N-1} \left(1 - \frac{|n|}{N}\right) e_n(t) = \frac{1}{N}\left(\frac{\sin\frac{Nt}{2}}{\sin\frac{t}{2}}\right)^2 \geq 0\,,$$

$$\frac{1}{2\pi}\int_0^{2\pi} D_N(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi} K_N(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi} J_N(t)\,\mathrm{d}t = 1\,.$$

De plus, $(K_N)_N$ et $(J_N)_N$ sont des approximations de l'unité.

Soient f et g dans $L^1(\mathbb{T})$ (classes d'équivalences des fonctions 2π-périodiques et intégrables), on définit la *convolution* de f par g par :

$$\forall x \in \mathbb{T},\ (f * g)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(y)g(x-y)\,\mathrm{d}y\,.$$

Pour $f \in L^1(\mathbb{T})$, on note $S_N f = f * D_N \in \mathcal{P}_N$ sa série de Fourier à l'ordre N.

Pour $f \in \mathcal{C}_{2\pi}$, on définit le *module de continuité* de f par :

$$w_f : t \in \mathbb{R}^+ \longmapsto \sup\{|f(x) - f(y)|\,,\ |x-y| \leq t\}\,.$$

On rappelle que w_f est une application croissante, sous-additive et continue en 0 (conséquence du théorème de Heine). En particulier, on en déduit :

$$\forall \lambda \in \mathbb{R}^+,\ \forall t \in \mathbb{R}^+,\ w_f(\lambda t) \leq w_f(\lceil\lambda\rceil t) \leq \lceil\lambda\rceil w_f(t) \leq (\lambda+1)w_f(t)\,,$$

où $\lceil\lambda\rceil$ désigne la partie supérieure de λ.

Développement

Théorème (Jackson)**.** *Soit $f \in \mathcal{C}_{2\pi}$, alors : $\|f - S_n f\|_\infty \leq O\big(w_f(\frac{1}{n})\ln n\big)$.*

Lemme. $\forall k \in [0\,,3[,\ \int_{-\pi}^{\pi} |t|^k J_N(t)\,\mathrm{d}t = O\big(N^{-k}\big)$.

Démonstration du lemme. Signalons que l'inégalité suivante se montre par concavité (faire un putain de dessin !) :

$$\forall t \in [0\,,\frac{\pi}{2}],\ \frac{2}{\pi}t \leq \sin(t) \leq t\,.$$

On a :

$$\begin{aligned}
\left\|K_N^2\right\|_1 &= \frac{1}{2\pi N^2}\int_{-\pi}^{\pi}\left(\frac{\sin\frac{Nt}{2}}{\sin\frac{t}{2}}\right)^4 \mathrm{d}t \\
&= \frac{1}{\pi N^2}\int_0^{\pi}\left(\frac{\sin\frac{Nt}{2}}{\sin\frac{t}{2}}\right)^4 \mathrm{d}t \quad \text{(par parité)} \\
&= \frac{2}{\pi N^2}\int_0^{\pi/2}\left(\frac{\sin Nu}{\sin u}\right)^4 \mathrm{d}u \quad \text{(en posant} \quad u = \frac{t}{2}) \\
&\geq \frac{2}{\pi N^2}\int_0^{\pi/2}\left(\frac{\sin Nu}{u}\right)^4 \mathrm{d}u \quad \text{(car} \quad \sin u \leq u) \\
&\geq \frac{2N}{\pi}\int_0^{N\pi/2}\left(\frac{\sin v}{v}\right)^4 \mathrm{d}v \quad \text{(en posant} \quad v = Nu) \\
&\geq aN\,,
\end{aligned}$$

où $a = \frac{2}{\pi}\int_0^{\pi/2}\left(\frac{\sin v}{v}\right)^4 \mathrm{d}v > 0$.

Soit $k \in [0\,,3[$. On a, en utilisant des arguments similaires à précédement :

$$\begin{aligned}
\int_{-\pi}^{\pi}|t|^k J_N(t)\,\mathrm{d}t &= \frac{1}{2\pi\left\|K_N^2\right\|_1 N^2}\int_{-\pi}^{\pi}|t|^k\left(\frac{\sin\frac{Nt}{2}}{\sin\frac{t}{2}}\right)^4 \mathrm{d}t \\
&\leq \frac{2^{k+1}}{\pi a}N^{-3}\int_0^{\pi/2}u^k\left(\frac{\sin Nu}{u}\right)^4 \mathrm{d}u \\
&\leq \frac{2^{k+1}}{\pi a}N^{-k}\int_0^{N\pi/2}\frac{(\sin v)^4}{v^{4-k}}\,\mathrm{d}v \\
&\leq O\big(N^{-k}\big)\,,
\end{aligned}$$

puisque $\int_0^{+\infty}\frac{(\sin v)^4}{v^{4-k}}\,\mathrm{d}v < +\infty$. □

Démonstration du théorème de Jackson. Soit $f \in \mathcal{C}_{2\pi}$. Soit $\lambda_n = \lfloor\frac{n}{2}\rfloor$ de sorte que $J_{\lambda_n} \in \mathcal{P}_n$. On pose $g = f * J_{\lambda_n} \in \mathcal{P}_n$. Puisque J_{λ_n} est positive et d'intégrale égale à 1, on a :

$$\begin{aligned}
\forall x \in \mathbb{T},\ |g(x) - f(x)| &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}(f(x) - f(x-t))\,J_{\lambda_n}(t)\,\mathrm{d}t\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}w_f(|t|)J_{\lambda_n}(t)\,\mathrm{d}t \\
&\leq \frac{w_f(\frac{1}{n})}{2\pi}\int_{-\pi}^{\pi}(n\,|t| + 1)J_{\lambda_n}(t)\,\mathrm{d}t \\
&\leq w_f(\frac{1}{n})(1 + A) \quad \text{(en utilisant le lemme)}\,,
\end{aligned}$$

où A est une constante (ainsi, il existe une constante $B > 0$ telle que : $d(f, \mathcal{P}_n) \leq Bw_f(\frac{1}{n})$). Comme $g \in \mathcal{P}_n$, $S_n g = g$, ainsi :

$$\|f - S_n f\|_\infty \leq \|f - g\|_\infty + \|S_n(f-g)\|_\infty \leq (1 + \|D_n\|_1)\,\|f - g\|_\infty = O\bigg(w_f(\frac{1}{n})\ln n\bigg)\,,$$

car l'application linéaire $f \in \mathcal{C}_{2\pi} \mapsto S_n f$ est continue et sa norme est majorée par $\|D_n\|_1 = O(\ln n)$ (voir compléments). □

Application. *Si f est α-höldérienne ($\alpha \in]0\,,1[$) alors la série de Fourier de f converge uniformément sur $\mathbb{T}$ vers f.*

Remarque. En particulier, on a montré que $d(f, \mathcal{P}_n) \leq O\big(w_f(\frac{1}{n})\big)$, ce qui se généralise au cas des fonctions continues sur un intervalle compact de $\mathbb{R}$ et en remplaçant l'espace des polynômes trigonométriques par celui des polynômes. Par translation et dilatation, on se ramène à montrer le résultat pour $[-1\,,1]$. Si $f \in \mathcal{C}([-1\,,1])$ alors on pose $F(t) = f(\cos t)$, fonction de $\mathcal{C}_{2\pi}$. On applique le théorème de Jackson trigonométrique et pour conclure il suffit de remarquer que le noyau de Jackson est pair donc définie des polynômes uniquement en cosinus (utiliser les polynômes de Tchebychev), que cos est surjective de $\mathbb{T}$ dans $[-1\,,1]$ et 1-lipchitzienne.

Compléments

Proposition. *L'application linéaire $f \in \mathcal{C}_{2\pi} \mapsto S_n f = f * D_n$ est continu, de norme majorée par $\|D_n\|_1$. De plus :*

$$\|D_n\|_1 \underset{+\infty}{\sim} \frac{4}{\pi^2} \ln n\,.$$

Démonstration. Soit $f \in \mathcal{C}_{2\pi}$. On a :

$$\forall x \in \mathbb{T},\ |S_n f(x)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x-y)|\,|D_n(y)|\,\mathrm{d}y \leq \|f\|_\infty \|D_n\|_1\,.$$

et :

$$\|D_n\|_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{t}{2})} \right| \mathrm{d}t = \frac{1}{\pi} \int_0^{\pi} \frac{\left|\sin((n+\frac{1}{2})t)\right|}{\sin(\frac{t}{2})}\,\mathrm{d}t\,.$$

Or, la fonction $t \mapsto \frac{1}{\sin(\frac{t}{2})} - \frac{2}{t}$ est bornée sur $]0\,,\pi]$ donc, il existe $(A_n)_n$ une suite bornée telle que :

$$\|D_n\|_1 = \frac{2}{\pi} \int_0^{\pi} \frac{\left|\sin((n+\frac{1}{2})t)\right|}{t}\,\mathrm{d}t + A_n$$

et en faisant le changement de variables $u = (n+\frac{1}{2})t$, il existe un autre suite bornée $(A'_n)_n$ telle que :

$$\|D_n\|_1 = \frac{2}{\pi} \int_0^{n\pi} \frac{|\sin(u)|}{u}\,\mathrm{d}u + A'_n\,.$$

Comme $L_n := \frac{2}{\pi} \int_0^{n\pi} \frac{|\sin(u)|}{u}\,\mathrm{d}u \underset{n\to+\infty}{\longrightarrow} +\infty$, on a : $\|D_n\|_1 \underset{+\infty}{\sim} L_n$. Pour tout $k \geq 1$, on a :

$$\frac{1}{(k+1)\pi} \underbrace{\int_{k\pi}^{(k+1)\pi} |\sin(u)|\,\mathrm{d}u}_{=2} \leq \int_{k\pi}^{(k+1)\pi} \frac{|\sin(u)|}{u}\,\mathrm{d}u \leq \frac{1}{k\pi} \underbrace{\int_{k\pi}^{(k+1)\pi} |\sin(u)|\,\mathrm{d}u}_{=2},$$

et en sommant :

$$\frac{2}{\pi}\left(\frac{1}{2} + \cdots + \frac{1}{n}\right) \leq \int_{\pi}^{n\pi} \frac{|\sin(u)|}{u}\,\mathrm{d}u \leq \frac{2}{\pi}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right),$$

ce qui conclut. □

Références

– Queffélec et Zuily, Analyse pour l'agrégation, p.79 (propriétés du noyau de Jackson) et p.100 (théorème de Jackson)

Leçons concernées : 209, 246

12 Théorème central limite et statistique de Pearson

Contexte

Dans tout ce qui suit, $\mathbb{R}^d$ est muni du produit scalaire canonique $\langle\cdot,\cdot\rangle$ et de la norme euclidienne associée $\|\cdot\|_2$.

On note $\mathcal{C}_b(\mathbb{R}^d)$ (resp. $\mathcal{C}_0(\mathbb{R}^d)$, resp. $\mathcal{C}_c(\mathbb{R}^d)$) l'espace des applications continues et bornée (resp. et tendant vers 0 en l'infini, resp. et à support compact) sur $\mathbb{R}^d$. On dit qu'une suite de probabilités $(\mathbb{P}_n)_n$ sur $\mathbb{R}^d$ *converge étroitement* vers une probabilité $\mathbb{P}$ si :

$$\begin{aligned} & \forall f \in \mathcal{C}_b(\mathbb{R}^d),\ \int f \,\mathrm{d}\,\mathbb{P}_n \underset{n\to+\infty}{\longrightarrow} \int f \,\mathrm{d}\,\mathbb{P} && (✐) \\ \iff\ & \forall f \in \mathcal{C}_0(\mathbb{R}^d),\ \int f \,\mathrm{d}\,\mathbb{P}_n \underset{n\to+\infty}{\longrightarrow} \int f \,\mathrm{d}\,\mathbb{P} \\ \iff\ & \forall f \in \mathcal{C}_c(\mathbb{R}^d),\ \int f \,\mathrm{d}\,\mathbb{P}_n \underset{n\to+\infty}{\longrightarrow} \int f \,\mathrm{d}\,\mathbb{P} \end{aligned}$$

On se donne, pour tout $n \in \mathbb{N}$, une variable aléatoire X_n définie sur un espace probabilisé $(\Omega_n, \mathcal{A}_n, \mathbb{P}^n)$, à valeurs dans $\mathbb{R}^d$. On se donne aussi une autre variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$ ainsi que μ une probabilité sur $\mathbb{R}^d$. On dit que la suite $(X_n)_n$ *converge en loi* vers X (resp. vers μ) si la suite des $(P^n_{X_n})_n$ des lois des X_n converge étroitement vers la loi P_X de X (resp. vers μ) et on notera :

$$X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X \quad (\text{resp. } X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mu\ ).$$

Si X est une variable aléatoire à valeurs dans $\mathbb{R}^d$, on note ϕ_X sa *fonction caractéristique* définie par :

$$\phi_X : t \in \mathbb{R}^d \longmapsto \mathbb{E}\left[e^{i\langle X,t\rangle}\right].$$

En fait, ϕ_X est la transformée de Fourier de la loi de X et par injectivité elle caractérise la loi de X. Si $X_1, \dots, X_n$ sont des variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$ alors :

$$\forall t \in \mathbb{R}^d,\ \phi_{X_1+\dots+X_n}(t) = \mathbb{E}\left[e^{i\langle X_1+\dots+X_n,t\rangle}\right] = \mathbb{E}\left[\prod_{k=1}^n e^{i\langle X_k,t\rangle}\right] = \prod_{k=1}^n \mathbb{E}\left[e^{i\langle X_k,t\rangle}\right] = \prod_{k=1}^n \phi_{X_k}(t).$$

Le théorème de Lévy montre que pour qu'une suite de variables aléatoires converge en loi, il suffit de vérifier (✐) sur une classe de fonctions bien plus restreinte.

Théorème (Lévy)**.** *Soit $(X_n)_n$ une suite de variables aléatoires. Soit X une variable aléatoire. Alors :*

$$X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X \quad\iff\quad \forall t \in \mathbb{R}^d,\ \phi_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} \phi_X(t).$$

Remarque.
- On peut affaiblir le sens indirect en supposant seulement que la suite des fonctions caractéristiques converge simplement vers une fonction continue en 0, et le théorème donne l'existence d'une variable aléatoire X vers laquelle la suite $(X_n)_n$ converge en loi.
- En fait, dans tous les cas, la convergence des fonctions caractéristiques est uniforme sur tout compact.
- On peut généralisé le théorème à l'espace des mesures positives sur $\mathbb{R}^d$ (et même sur E un espace métrique localement compact et dénombrable à l'infini) dont la masse est bornée par une constante.

Enfin, si une variable aléatoire réelle X admet un moment d'ordre k alors ϕ_X est de classe $\mathcal{C}^k$ et les dérivées s'obtiennent en dérivant sous le signe $\mathbb{E}$. En particulier :

$$\forall 0 \leq j \leq k,\ \phi_x^{(j)}(t) = i^j\,\mathbb{E}\left[X^j\right],$$

et ϕ_X admet un développement limité au voisinage de 0 :

$$\phi_X(t) = 1 + it\,\mathbb{E}\,X - \frac{t^2}{2}\,\mathbb{E}\left[X^2\right] + \dots + \frac{(it)^k}{k!}\,\mathbb{E}\left[X^k\right] + o(t^k).$$

On introduit les notations relatives au théorème de Pearson qui est à la base du *test du* χ^2. On observe un n-échantillon $X = (X_1, \dots, X_n)$ dont chacune des variables est à valeurs dans un ensemble fini $\mathcal{X} = \{x_1, \dots, x_d\}$. On note $\mathbf{p} = (p_1, \dots, p_d)$ la loi commune des X_j et on se donne une loi de référence $\mathbf{p}^0 = (p_1^0, \dots, p_n^0)$ de support plein (i.e $p_j^0 > 0$ pour tout j). On veut tester si les X_j ont pour loi $\mathbf{p}^0$:

$$H_0 : \mathbf{p} = \mathbf{p}^0 \quad \text{contre} \quad H_1 : \mathbf{p} \neq \mathbf{p}^0 .$$

Pour estimer $\mathbf{p}$, on utilise la méthode des moments :

$$\widehat{p}_{i,n} = \frac{N_{i,n}}{n} \quad \text{où} \quad N_{i,n} = \sum_{j=1}^{n} \mathbb{1}_{X_j = x_i} .$$

$\widehat{\mathbf{p}}_n = (\widehat{p}_{1,n}, \dots, \widehat{p}_{d,n})$ est un estimateur fortement consistant de $\mathbf{p}$ (loi forte des grands nombres) et sans biais. On définit la *statistique de Pearson* par :

$$D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) = n \sum_{j=1}^{n} \frac{(\widehat{p}_{i,n} - p_i^0)^2}{p_i^0} = \sum_{j=1}^{n} \frac{(N_{i,n} - np_i^0)^2}{np_i^0} .$$

Pour démontrer le théorème de Pearson on aura besoin du théorème de Cochran. On rappelle d'abord la définition de la loi du χ^2.

Définition. Soit $X \sim \mathcal{N}(m, I_d)$ un vecteur gaussien de $\mathbb{R}^d$. La loi de $\|X\|_2^2$ ne dépend que de d et $\|m\|_2$: on dit que $\|X\|_2^2$ suit *une loi du* χ^2 à d degrés de liberté et de paramètre de décentrage $\|m\|_2^2$. On note :

$$\|X\|_2^2 \sim \chi^2(d, \|m\|_2^2) .$$

Si $m = 0$, on note plus simplement : $\|X\|_2^2 \sim \chi^2(d)$.

Théorème (Cochran)**.** *Soient $X \sim \mathcal{N}(m, I_d)$ un vecteur gaussien et $E_1 \oplus \cdots \oplus E_r$ une décomposition de $\mathbb{R}^d$ en sous-espaces orthogonaux de dimension respectives $d_1, \dots, d_r$. Les projections orthogonales $\Pi_{E_1} X, \dots, \Pi_{E_r} X$ forment des vecteurs gaussiens indépendants. De plus :*

$$\forall j \in [\![1, r]\!], \ \left\|\Pi_{E_j} X\right\|_2^2 \sim \chi^2\left(d_j, \left\|\Pi_{E_j} m\right\|_2^2\right) .$$

Remarque.
- Il est bon de savoir que la démonstration de ce théorème utilise efficacement l'équivalence entre décorrélation et indépendance des vecteurs gaussiens.
- Une des nombreuse applications de ce théorème en statistique est le test de Student dans le cadre du modèle linéaire gaussien.

Développement

Théorème (Théorème central limite)**.** *Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$, indépendantes, de même loi qu'une variable aléatoire X et admettant un moment d'ordre deux. Alors :*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mathbb{E}\, X_k) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0, C_X) ,$$

où C_X désigne la matrice de covariance des X_j.

Lemme. *Soit $(z_n)_n$ une suite de $\mathbb{C}$ qui converge vers $z \in \mathbb{C}$. Alors : $\left(1 + \frac{z_n}{n}\right)^n \xrightarrow[n \to +\infty]{} e^z$.*

Démonstration du lemme. On rappelle que le logarithme complexe est défini sur $D(1,1)$ par la série entière :

$$\forall z \in D(0,1), \ \log(1 - z) = -\sum_{k \geq 1} \frac{z^k}{k} ,$$

Pour tout n assez grand, $\left|\frac{z_n}{n}\right| < 1$ et :

$$\left(1 + \frac{z_n}{n}\right)^n = \exp\left(n \log\left(1 + \frac{z_n}{n}\right)\right) = \exp\left(z_n - \frac{1}{n} \sum_{k \geq 2} \frac{1}{k} \frac{(-z_n)^k}{n^{k-2}}\right) = \exp(z_n + o(1)) \xrightarrow[n \to +\infty]{} e^z .$$

□

Démonstration du théorème. On note $Y_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mathbb{E}\, X_k)$. Comme les variables X_k sont indépendantes et de même loi, on a pour tout $t \in \mathbb{R}^d$:

$$\phi_{Y_n}(t) = \prod_{k=1}^{n} \phi_{(X_k - \mathbb{E}\, X_k)}\left(\frac{t}{\sqrt{n}}\right) = \left[\phi_{\langle X - \mathbb{E}\, X, t\rangle}\left(\frac{1}{\sqrt{n}}\right)\right]^n .$$

La variable aléatoire réelle $\langle X - \mathbb{E}\, X, t\rangle$ est centrée et admet un moment d'ordre 2 (par l'inégalité de Cauchy-Schwarz), donc :

$$\phi_{Y_n}(t) = \left[1 - \frac{\mathbb{E}\left[\langle X - \mathbb{E}\, X, t\rangle^2\right]}{2n} + o\left(\frac{1}{n}\right)\right]^n \underset{n\to+\infty}{\longrightarrow} \exp\left(-\frac{1}{2}\,\mathbb{E}\left[\langle X - \mathbb{E}\, X, t\rangle^2\right]\right)$$

Dans le petit o se cache éventuellement des termes complexes, c'est pour cela qu'on a besoin du lemme. Comme, pour tout $t \in \mathbb{R}^d$ on a $\mathbb{E}\left[\langle X - \mathbb{E}\, X, t\rangle^2\right] = \langle C_X t, t\rangle$, le théorème de Lévy nous assure que :

$$Y_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0, C_X)\,.$$

□

Application (Théorème de Pearson)**.** *La statistique de Pearson $D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0)$ admet les comportements limites suivants :*

– *sous $H_0 : \mathbf{p} = \mathbf{p}^0$, on a : $D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \chi^2(d-1)$,*

– *sous $H_1 : \mathbf{p} \neq \mathbf{p}^0$, on a : $D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) \underset{n\to+\infty}{\longrightarrow} +\infty$ presque sûrement.*

Démonstration. On suppose H_0. Pour $j = 1, \ldots, n$, on définit la variable aléatoire :

$$Z_j := \left(\frac{1}{\sqrt{p_1^0}}(\mathbb{1}_{X_j = x_1} - p_1^0), \ldots, \frac{1}{\sqrt{p_d^0}}(\mathbb{1}_{X_j = x_d} - p_d^0)\right) .$$

Comme les X_j sont i.i.d, il en est de même pour les Z_j. Notons Z une variable aléatoire de même loi) que les Z_j. Z est centrée et admet un moment d'ordre 2. Calculons sa matrice de covariance $\Gamma = C_Z$. Pour tout $i \neq j$:

$$\Gamma_{i,i} = \mathrm{Var}\left(\frac{1}{\sqrt{p_i^0}}(\mathbb{1}_{X_1 = x_i} - p_i^0)\right) = \frac{p_i^0(1 - p_i^0)}{p_i^0} = 1 - p_i^0\,,$$

$$\Gamma_{i,j} = \mathrm{Cov}\left(\frac{1}{\sqrt{p_i^0}}(\mathbb{1}_{X_1 = x_i} - p_i^0), \frac{1}{\sqrt{p_j^0}}(\mathbb{1}_{X_1 = x_j} - p_j^0)\right) = -\sqrt{p_i^0 p_j^0}\,.$$

En notant $\sqrt{\mathbf{p}} = \sqrt{\mathbf{p^0}} = (\sqrt{p_1^0}, \ldots, \sqrt{p_n^0})$, on a : $\Gamma = I_d - {}^t\sqrt{\mathbf{p}}\sqrt{\mathbf{p}}$. La matrice ${}^t\sqrt{\mathbf{p}}\sqrt{\mathbf{p}}$ est la matrice de projection orthogonale sur le sous-espace de dimension 1 engendré par le vecteur ${}^t\sqrt{\mathbf{p}}$ (car $p_1^0 + \cdots + p_d^0 = 1$), donc Γ est la matrice de projection orthogonale sur l'orthogonal de ce sous-espace qui est de dimension $d-1$. On note V une variable aléatoire ayant pour loi $\mathcal{N}(0, I_d)$, alors : $\Gamma V \sim \mathcal{N}(0, \Gamma^t\Gamma) = \mathcal{N}(0, \Gamma)$. D'après le théorème central limite :

$$\sqrt{n}\left(\frac{1}{\sqrt{p_1^0}}(\widehat{p}_{1,n} - p_1^0), \ldots, \frac{1}{\sqrt{p_d^0}}(\widehat{p}_{d,n} - p_d^0)\right) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} Z_j \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \Gamma V\,.$$

puis :

$$D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \|\Gamma V\|_2^2\,.$$

Le théorème de Cochran montre alors que $\|\Gamma V\|_2^2$ suit une loi $\chi^2(d-1)$, ce qui conclut pour le premier cas.

Si on suppose H_1 alors il existe i tel que $p_i \neq p_i^0$ et, par la loi forte des grands nombres :

$$D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) \geq n\frac{(\widehat{p}_{i,n} - p_i^0)^2}{p_i^0} \underset{n\to+\infty}{\longrightarrow} +\infty \quad \text{presque sûrement}\,.$$

□

Compléments

Une démonstration du lemme qui n'utilise pas le logarithme complexe.

Autre démonstration du lemme. Soient $z \in \mathbb{N}$ et $n \in \mathbb{N}^*$. La formule du binôme de Newton donne :

$$\exp(z) - \left(1 + \frac{z}{n}\right)^n = \sum_{k \geq 0} \frac{z^k}{k!} - \sum_{k=0}^{n} \binom{n}{k} \frac{z^k}{n^k} = \sum_{k \geq n+1} \frac{z^k}{k!} + \sum_{k=0}^{n} \left[\frac{1}{k!} - \frac{1}{n^k}\binom{n}{k}\right] z^k ,$$

or, pour tout $0 < k \leq n$:

$$\frac{k!}{n^k}\binom{n}{k} = \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \leq 1 ,$$

et il y a égalité pour $k = 0$, donc :

$$\left|\exp(z) - \left(1 + \frac{z}{n}\right)^n\right| \leq \sum_{k \geq n+1} \frac{|z|^k}{k!} + \sum_{k=0}^{n} \underbrace{\left[\frac{1}{k!} - \frac{1}{n^k}\binom{n}{k}\right]}_{\geq 0} |z|^k = \exp(|z|) - \left(1 + \frac{|z|}{n}\right)^n \underset{n \to +\infty}{\longrightarrow} 0 .$$

Ensuite on découpe et on majore (c'est de l'analyse quoi) et c'est bon ! □

Proposition (Test du χ^2)**.** *Soit $\alpha \in]0\,,1[$. Le test défini par :*

$$\phi(X_1, \ldots, X_n) = \mathbb{1}_{D_n^2(\widehat{\mathbf{p}}_n, \mathbf{p}^0) > c_{d-1,1-\alpha}} ,$$

où $c_{d-1,1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de la loi du χ^2 à $d-1$ degrés de liberté, est un test de taille asymptotiquement égale à $1-\alpha$ lorsque n tend vers $+\infty$. En outre, il est consistant.

Remarque.
- La consistance signifie que la fonction puissance du test tend ponctuellement vers 1. On peut montrer, à l'aide de l'inégalité d'Hoeffding, que cette convergence est exponentielle.
- Les garanties de ce test sont asymptotiques. En pratique, on considère le test pertinent si $n \geq 30$ et si, pour tout $j = 1, \ldots, d$, $np_j^0 \geq 5$. Si ce n'est pas le cas, il faut regrouper certaines classes entre elles.

Références

- Ouvrard, Probabilités II, Chap. 14 (cours) et p.314 (TCL),
- Rivoirard et Stoltz, Statistiques mathématiques en action, p.56 (théorème de Pearson)
- Amar et Matheron, Analyse complexe, p.77 (logarithme complexe)

Leçons concernées : 218, 236, 250, 261, 262

13 Théorème d'Hadamard-Lévy

Contexte

Une application f de $\mathbb{R}^n$ dans $\mathbb{R}^n$ est dite *propre* si l'image réciproque par f de tout compact est un compact. Cette définition est équivalente à :

$$\|f(x)\| \underset{\|x\|\to+\infty}{\longrightarrow} +\infty\,.$$

On donne le contexte général des équations différentielles ordinaires ainsi que quelques résultats essentiels pour la suite.

Soient U un ouvert de $\mathbb{R}\times\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une application continue. On considère l'équation différentielle :

$$y' = f(t,y), \quad (t,y)\in U, \quad t\in\mathbb{R}, \quad y\in\mathbb{R}^n\,. \tag{E}$$

Une *solution* de (E) sur un intervalle I de $\mathbb{R}$ est une fonction dérivable $y : I \to \mathbb{R}^n$ telle que pour tout $t\in I$:

$$(t,y(t))\in U \quad \text{et} \quad y'(t) = f(t,y(t))\,.$$

Soient $y : I \to \mathbb{R}^n$ et $\tilde{y} : \tilde{I} \to \mathbb{R}^n$ des solutions de (E). On dit que $\tilde{y}$ est un *prolongement* de y si $I\subset\tilde{I}$ et si $\tilde{y}_{|I} = y$. On dit que y est une solution *maximale* si y n'admet pas de prolongement $\tilde{y}$ avec $I \subsetneq \tilde{I}$. On peut montrer que toute solution se prolonge en une solution maximale (pas nécessairement unique). Le théorème suivant est un critère de maximalité.

Théorème (Théorème de sortie de tout compact)**.** *Une solution $y :]a\,,b[\, \to \mathbb{R}^n$ de (E) (avec $-\infty \le a < b \le +\infty$) est maximale si et seulement si $t\mapsto (t,y(t))$ s'échappe de tout compact de U lorsque $t\to a^+$ ou $t\to b^-$.*

En particulier, si $U = I\times\mathbb{R}^n$ et si $b<+\infty$ (resp. $a>-\infty$) alors $y(t)$ n'est pas bornée au voisinage de b (resp de a). Si $(t_0,y_0)\in U$ alors le *problème de Cauchy* associé (E) avec conditions initiales (t_0,y_0) consiste à trouver une solution $y : I\to\mathbb{R}^n$ de (E) telle que $t_0\in I$ et $y(t_0)=y_0$.

Théorème (Cauchy-Lipschitz)**.** *Soient $(t_0,y_0)\in U$. On suppose de plus que f est localement lipchitzienne en la deuxième variable. Alors le problème de Cauchy associé à (E) avec conditions initiales (t_0,y_0) admet une unique solution maximale*

$$y : [t_0 - T\,, t_0+T] \to \overline{B}(y_0,r_0)\,.$$

Développement

Théorème (Hadamard-Lévy)**.** *Soit f une application de classe $\mathcal{C}^2$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Les deux conditions suivantes sont équivalentes :*

(a) f est un difféomorphisme de $\mathbb{R}^n$ sur $\mathbb{R}^n$,

(b) f est propre et sa différentielle est inversible en tout point de $\mathbb{R}^n$.

Remarque. En fait, ce théorème est vrai si on suppose f seulement de classe $\mathcal{C}^1$.

Démonstration de (a) ⇒ (b). Soit K un compact. L'image réciproque de K par f est égale à l'image de K par f^{-1} qui est continue donc f est propre. Ensuite, $f^{-1}\circ f = \mathrm{id}$ donc $d(f^{-1})(f(x))\circ df(x) = \mathrm{id}$ et la différentielle de f est inversible en tout point x, d'inverse $d(f^{-1})(f(x))$. □

Démonstration de (b) ⇒ (a). On va d'abord montrer la surjective de f puis son injectivité, le théorème d'inversion global conclura.

(i) **f est surjective.**

Nous allons montrer que $f(\mathbb{R}^n)$ est non vide (évident), ouvert et fermé dans $\mathbb{R}^n$. Par connexité, cela montrera la surjectivité. Soit $y\in f(\mathbb{R}^n)$ et $x\in\mathbb{R}^n$ tel que $f(x)=y$. $df(x)$ est inversible, le théorème d'inversion local montre l'existence d'un voisinage ouvert U de x tel que $f : U\to f(U)$ définisse un difféomorphisme. En particulier, $f(U)$ est ouvert et est un voisinage de y dans $f(\mathbb{R}^n)$, ce qui montre que $f(\mathbb{R}^n)$. Soit $(y_k)_k$ une suite de $f(\mathbb{R}^n)$ qui converge vers $y\in\mathbb{R}^n$. $K = \mathrm{supp}(y_k)_k$ est compact (fermé et borné) donc, comme f est propre, son image réciproque aussi. Pour tout k, il existe x_k dans $f^{-1}(K)$ tel que $f(x_k)=y_k$. Quitte à extraire, on peut supposer que x_k converge vers $x\in f^{-1}(K)\subset\mathbb{R}^n$ (propriété de Bolzao-Weierstrass) et par continuité $f(x)=y$ donc $y\in f(\mathbb{R}^n)$ et $f(\mathbb{R}^n)$ est fermé.

(ii) f **est injective.**

Soit $x_0 \in \mathbb{R}^n$ et posons $g(x) = f(x) - f(x_0)$ (g vérifie les deux conditions de (a)). On pose $S = g^{-1}(\{0\})$ et on va montrer que $S = \{x_0\}$.

- Tout d'abord, S **a un nombre fini d'éléments**. En effet, S est compact car f est propre et si S avait un nombre infini d'éléments alors S aurait un point d'accumulation, noté q. Or d'après le théorème d'inversion local appliqué à g en q, il existerait un voisinage U de q sur lequel g induit un difféomorphisme local. En particulier, g est injective sur U et : $S \cap U = \{q\}$, ce qui est absurde puisque q est un point d'accumulation de S.
- Notons $S = \{p_1, \dots, p_N\}$ et montrons que $N = 1$. Considérons la fonction

$$F : x \in \mathbb{R}^n \longmapsto [dg(x)]^{-1}(g(x)),$$

qui est de classe $\mathcal{C}^1$ car f est $\mathcal{C}^2$. Pour $\nu \in \mathbb{R}^n$, on introduit le système différentiel autonome :

$$\begin{cases} x' = -F(x) \\ x(0) = \nu \end{cases} . \tag{✠}$$

Puisque F est $\mathcal{C}^1$, le théorème de Cauchy-Lipschitz montre que ce système admet une unique solution maximale, que l'on note x_ν, définie sur $]T_*\,, T^*[$.

Montrons que $T^* = +\infty$. Pour cela, on démontre que x_ν reste borné sur $[0\,, T^*[$ et le théorème de sortie de tout compact conclura. Pour tout $t \in [0\,, T^*[$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[g(x_\nu(t))\right] = dg(x_\nu(t))\left[x_\nu'(t)\right] = -g(x_\nu(t)),$$

ce qui montre que :

$$\forall t \in [0\,, T^*[,\ g(x_\nu(t)) = g(\nu)e^{-t}, \tag{🍃}$$

Donc, pour tout $t \in [0\,, T^*[$, $x_\nu(t) \in g^{-1}\left(\overline{B}(0, \|g(\nu)\|)\right)$ qui est borné car compact, g étant propre. Donc $T^* = +\infty$.
- Soit $p \in S$. On montre maintenant que p **est asymptotiquement stable**, i.e :

$$F(p) = 0 \quad \text{et} \quad \exists \delta > 0,\ \forall \nu \in \mathbb{R}^n,\ |\nu - p| \leq \delta \implies \lim_{t \to +\infty} x_\nu(t) = p.$$

Puisque $g(p) = 0$, clairement $F(p) = 0$. Ensuite, d'après le théorème d'inversion local, il existe $\delta > 0$ et un voisinage V_δ de 0 tels que $g : B(p, \delta) \to V_\delta = g(B(p, \delta))$ soit un $\mathcal{C}^1$-difféomorphisme. Soit $\nu \in B(p, \delta)$. Puisque V_δ est ouvert et contient 0, pour t_0 assez grand et pour tout $t \geq t_0$, $e^{-t}g(\nu) \in V_\delta$, on peut donc poser $x(t) = g^{-1}_{|V_\delta}(e^{-t}g(\nu))$ pour tout $t \geq t_0$. x_ν et x coïncident au voisisnage de 0, par unicité, x est donc la restriction de la trajectoire issue de ν et donc : $x(t) \to g^{-1}_{|V_\delta}(0) = p$ lorsque $t \to +\infty$.

Pour $i \in [\![1\,, N]\!]$, on se donne $\delta_i > 0$ tel que $B(p_i, \delta_i)$ soit un voisinage d'asymptotique stabilité (terminaison non conventionnelle) pour p_i.
- **Montrons que** $\mathbb{R}^n = \cup_{i=1}^N W_i$ où W_i est l'ensemble des points ν de $\mathbb{R}^n$ tel que la trajectoire issue de ν converge vers p_i lorsque t tend vers $+\infty$. Soit $\nu \in \mathbb{R}^n$. On a vu que $x_\nu(t)$ restait dans un compact donc il existe $(t_k)_k$ telle que $t_k \to +\infty$ et $x_\nu(t_k) \to \ell$, lorsque $k \to +\infty$. Par continuité dans (🍃) : $g(\ell) = 0$ et $\ell = p_i$ pour un certain i. Soit k_0 tel que : $x_\nu(t_{k_0}) \in B(p_i, \delta_i)$. Alors, en notant $y_0 = x_\nu(t_{k_0})$, on a : $x_{y_0}(t) \to p_i$. Or, par unicité : $x_{y_0}(t) = x_\nu(t + t_{k_0})$ pour tout t. Ce qui montre que $\nu \in W_i$.
- Enfin, **chaque** W_i **est ouvert**. En effet, soit $\nu \in W_i$. Il existe $T > 0$ tel que : $x_\nu(T) \in B(p_i, \frac{\delta_i}{3})$. $x_\nu(t)$ reste dans un compact donc il existe donc $R > 0$ tel que $x_\nu(t)$ reste dans $\overline{B}(0, R)$ pour tout $t \geq 0$. F est $\mathcal{C}^1$ sur le connexe $\overline{B}(0, 2R)$ donc y est globalement lipchitzienne (inégalité des accroissements finis). Notons M une constante de Lipschitz associée. Soit $\mu \in \overline{B}(\nu, r)$ avec $0 < r < Re^{-MT}$. On pose $t_m = \sup\{t \geq 0,\ x_\mu(t) \in \overline{B}(0, 2R)\}$ ($t_m > 0$ car $\|\mu\| < 2R$). Montrons que $t_m \geq T$. Si ce n'est pas le cas alors :

$$\forall t \in [0\,, t_m],\ \left\|x_\nu'(t) - x_\mu'(t)\right\| = \|F(x_\nu(t)) - F(x_\mu(t))\| \leq M\,\|x_\nu(t) - x_\mu(t)\|\,.$$

Le lemme de Gronwall montre que :

$$\forall t \in [0\,,t_m],\ \|x_\nu(t) - x_\mu(t)\| \leq \|x_\nu(0) - x_\mu(0)\|\, e^{Mt} < \|\nu - \mu\|\, e^{MT} \leq R\,,$$

et par inégalité triangulaire :

$$\forall t \in [0\,,t_m],\ \|x_\mu(t)\| < 2R\,,$$

ce qui contredit la définition de t_m, donc $t_m \geq T$. Si on suppose de plus que $r < \frac{\delta_i}{3} e^{-MT}$ alors ce qui précède montre que :

$$\forall t \leq T,\ \|x_\nu(t) - x_\mu(t)\| \leq \frac{\delta_i}{3}\,,$$

et en particulier : $\|x_\nu(T) - x_\mu(T)\| \leq \frac{\delta_i}{3}$. Donc : $\|x_\mu(T) - p_i\| < \delta$. On a montré précédement que cela impliquait : $x_\mu(t) \longrightarrow p_i$, lorsque t tend $+\infty$. Donc $B(\nu, r) \subset W_i$ et W_i est ouvert.

Finalement, $\mathbb{R}^n$ est réunion disjointe d'ouverts non vides, si $N \geq 2$, cela contredit la connexité de $\mathbb{R}^n$. Donc $N = 1$.

□

Remarque. Faire un dessin n'est pas idiot.

Références

– Queffélec et Zuily, Analyse pour l'agrégation (*troisième édition*), p.392
– Demailly, Analyse numérique et équations différentielles, Chap.V, p.135 (cours)

Leçons concernées : 204, 214, 215, 220

14 Ellipsoïde de John-Loewner

Contexte

On munit $\mathbb{R}^n$ de sa structure euclidienne usuelle. On note Q (resp. Q_+, resp. Q_{++}) l'ensemble des formes quadratiques (resp. positives, resp. définies positives) sur $\mathbb{R}^n$. On munit l'espace vectoriel Q de la norme N définie par :

$$\forall q \in Q,\ N(q) = \sup_{\|x\| \leq 1} |q(x)|\ .$$

Un *ellipsoïde* plein centré en 0 est une partie de $\mathbb{R}^n$ définie par une équation du type $q(x) \leq 1$ où $q \in Q_{++}$ (c'est la boule unité de la norme associée à q). Pour $q \in Q_{++}$, on pose $\mathcal{E}_q = \{x \in \mathbb{R}^n,\ q(x) \leq 1\}$ et on note V_q le volume de $\mathcal{E}_q$.

Pour $q \in Q$, on définit l'*orthogonal* $\mathrm{O}(q)$ de q par :

$$\mathrm{O}(q) = \{g \in \mathrm{GL}_n(\mathbb{R}) \mid q = q \circ g\}\,.$$

Si de plus $q \in Q_{++}$ alors $\mathrm{O}(q)$ est isomorphe à $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{O}(q)$ est un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$.

Développement

Théorème (Ellipsoïde de John-Loewner)**.** *Soit K un compact d'intérieur non vide de $\mathbb{R}^n$. Il existe un unique ellipsoïde plein centré en 0 de volume minimal contenant K.*

Démonstration. On se ramène à un problème d'optimisation d'une fonctionnelle convexe sur un compact.

(i) Reformulons le problème. Pour cela explicitons V_q pour $q \in Q_{++}$.
D'après le théorème spectral, il existe une base orthonormée $\mathcal{B} = (e_1, \dots, e_n)$ telle que pour tout $x = \sum_i x_i e_i$, on ait :

$$q(x) = \sum_i a_i x_i^2\,,$$

où $a_1, \dots, a_n > 0$. Par le théorème de changement de variables et le théorème de Fubini-Tonelli :

$$V_q = \int_{\mathbb{R}^d} \mathbb{1}_{q(x)\leq 1}\,\mathrm{d}x = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \mathbb{1}_{a_1x_1^2+\dots+a_nx_n^2\leq 1}\,\mathrm{d}x_1 \dots \mathrm{d}x_n\,.$$

En posant $u_i = \frac{x_i}{\sqrt{a_i}}$ pour tout i, on obtient :

$$V_q = \frac{1}{\sqrt{a_1 \dots a_n}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \mathbb{1}_{u_1^2+\dots+u_n^2\leq 1}\,\mathrm{d}u_1 \dots \mathrm{d}u_n = \frac{V_0}{\sqrt{\det(q)}}\,,$$

où V_0 est le volume de la boule unité de la norme euclidienne canonique (qui ne dépend pas de q).
Ainsi, minimiser V_q revient à maximiser $\det(q)$. On pose :

$$\mathcal{A} = \{q \in Q_+,\ \forall x \in K,\ q(x) \leq 1\}\,,$$

et on cherche à maximiser la fonctionnelle continue $q \mapsto \det(q)$ sur $\mathcal{A}$.

(ii) **$\mathcal{A}$ est fermé.**
En effet, si $(q_n)_n \in \mathcal{A}^{\mathbb{N}}$ converge vers $q \in Q$, on a :

$$\forall x \in \mathbb{R}^n,\ |q(x) - q_n(x)| \leq N(q - q_n)\,\|x\|\ ,$$

et donc : $\forall x \in \mathbb{R}^n,\ q_n(x) \underset{n\to+\infty}{\longrightarrow} q(x)$. En particulier :

$$\forall x \in \mathbb{R}^n,\ q(x) = \lim_{n\to+\infty} q_n(x) \geq 0 \quad \text{et} \quad \forall x \in K,\ q(x) = \lim_{n\to+\infty} q_n(x) \leq 1\,,$$

donc $q \in \mathcal{A}$.

(iii) **$\mathcal{A}$ est borné.**
Comme K est d'intérieur non vide, il existe $a \in K$ et $r > 0$ tels que $B(a, 2r) \subset K$ (pour la norme euclidienne). Soit $q \in \mathcal{A}$. Par l'inégalité de Minkowski :

$$\forall \|x\| \leq r,\ \sqrt{q(x)} = \sqrt{q(x+a-a)} \leq \sqrt{q(x+a)} + \sqrt{q(-a)} = \sqrt{q(x+a)} + \sqrt{q(a)} \leq 1+1 = 2\,,$$

donc : $\forall \|x\| \leq 1,\ |q(x)| = \frac{q(rx)}{r^2} \leq \frac{4}{r^2}$. Donc $N(q) \leq \frac{4}{r^2}$ pour tout q dans $\mathcal{A}$ qui est donc borné.

(iv) **$\mathcal{A}$ est non vide.**
Puisque K est compact, il est borné, par exemple par $M > 0$. q défini par :

$$\forall x \in \mathbb{R}^n,\ q(x) = \frac{\|x\|^2}{M^2}$$

est bien un élément de $\mathcal{A}$ (et même de $\mathcal{A} \cap Q_{++}$) qui est donc non vide.

(v) **$\mathcal{A}$ est convexe.**
En effet, soient q_1 et q_2 dans $\mathcal{A}$ et soit $\lambda \in [0\,,1]$. On a :

$$\begin{aligned}
&\forall x \in \mathbb{R}^n,\ (\lambda q_1 + (1-\lambda)q_2)(x) = \lambda q_1(x) + (1-\lambda)q_2(x) \geq 0\,,\\
&\forall x \in K,\ (\lambda q_1 + (1-\lambda)q_2)(x) = \lambda q_1(x) + (1-\lambda)q_2(x) \leq \lambda + 1 - \lambda = 1\,,
\end{aligned}$$

donc $\lambda q_1 + (1-\lambda)q_2 \in \mathcal{A}$ et $\mathcal{A}$ est convexe.

(vi) L'application déterminant $q \mapsto \det(q)$ est continue (polynômiale en les coefficients) sur le compact non vide $\mathcal{A}$ donc y atteint sont maximum en un point q_0. De plus, comme $\mathcal{A}$ contient $x \mapsto \frac{\|x\|^2}{M^2}$, $\det(q_0) > 0$ et q_0 est définie positive. Pour montrer l'unicité, on utilise la stricte log-convexité du déterminant.
Supposons qu'il existe $q \neq q_0$ dans $\mathcal{A}$ tel que $\det(q) = \det(q_0)$ alors, par convexité de $\mathcal{A}$, on a $\frac{q+q_0}{2} \in \mathcal{A}$ et, en notant S (resp. S_0) la matrice de q (resp. q_0) dans la base canonique :

$$\det\left(\frac{q+q_0}{2}\right) = \det\left(\frac{S+S_0}{2}\right) > \sqrt{(\det S)(\det S_0)} = \sqrt{(\det q)(\det q_0)} = \det(q_0)\,,$$

ce qui contredit la maximalité de $\det(q_0)$.

□

Remarque. Si K est d'intérieur vide alors :

- ou bien $\operatorname{conv}(K \cup \{0\})$ est d'intérieur vide, i.e $\{0\} \in K$ alors il n'existe pas de forme quadratique q non dégénérée telle que $K \subset \mathcal{E}_q$ et $\mathcal{E}_q$ soit de volume minimal,
- ou bien $\operatorname{conv}(K \cup \{0\})$ est d'intérieur non vide et on peut appliquer le théorème au compact $\operatorname{conv}(K \cup \{0\})$.

Compléments

Lemme (log-convexité du déterminant)**.** *Soient A et B deux matrices symétriques réelles définies positives. Alors, pour tout $\lambda \in [0\,,1]$:*

$$\det(\lambda A + (1-\lambda)B) \geq (\det A)^\lambda (\det B)^{1-\lambda}\,,$$

avec inégalité stricte pour $\lambda \in \,]0\,,1[$ et $A \neq B$.

Démonstration. D'après le théorème de pseudo-réduction simultanée, il existe $P \in \mathrm{GL}_n(\mathbb{R})$, des réels strictement positifs $\alpha_1, \ldots, \alpha_n$ tels que, en notant $D = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$:

$$A = {}^tPP \quad \text{et} \quad B = {}^tPDP\,.$$

Pour tout $\lambda \in [0\,,1]$, on a :

$$\begin{aligned}
&\det(\lambda A + (1-\lambda)B) = \det(P)^2 \det(\lambda I_n + (1-\lambda)D) = \det(P)^2 \prod_{i=1}^n (\lambda + (1-\lambda)\alpha_i)\,,\\
&(\det A)^\lambda (\det B)^{1-\lambda} = \det(P)^{2\lambda} \det(I_n) \det(P)^{2(1-\lambda)} \det(D)^{1-\lambda} = \det(P)^2 \det(D)^{1-\lambda}\,.
\end{aligned}$$

En simplifiant par $\det(P)^2$ et en prenant le logarithme, on cherche à montrer :

$$\sum_{i=1}^{n} \ln(\lambda + (1-\lambda)\alpha_i) \geq (1-\lambda) \sum_{i=1}^{n} \ln(\alpha_i)\,.$$

On obtient le résultat par stricte concavité du log et en remarquant que si $A \neq B$ alors il existe i tel que $\alpha_i \neq 1$. □

Une application importante est la caractérisation des sous-groupes compacts maximaux de $\mathrm{GL}_n(\mathbb{R})$.

Application. *Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$, alors :*

(i) il existe une forme quadratique q définie positive telle que $G \subset \mathrm{O}(q)$,

(ii) il y a égalité si et seulement si G est maximal pour l'inclusion.

Démonstration. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$.

(i) Notons B la boule unité fermée pour la norme canonique. Soit :

$$K = \{g(x),\ x \in B,\ g \in G\} = \bigcup_{g \in G} g(B)\,.$$

C'est l'image du compact $G \times B$ par l'application continue $(g,x) \mapsto g(x)$ donc c'est un compact. Comme $I_n \in G$, K contient B qui est d'intérieur non vide. On considère l'ellipsoïde $\mathcal{E}_q$ associé à une forme quadratique $q \in Q_{++}$ de volume minimal contenant K. Soit $g \in G$. On définit la forme quadratique (définie positive) $q_g : x \mapsto q(g(x))$. Comme K est invariant par tous les éléments de G, on a $g(K) = K$ et donc $K \subset \mathcal{E}_{q_g}$. De plus, G contient la famille $\{g^n,\ n \in \mathbb{Z}\}$ or G est compact et l'application déterminant est continue donc $\det(\{g^n,\ n \in \mathbb{Z}\}) = \{\det(g)^n,\ \in \mathbb{Z}\}$ est bornée. Nécessairement, $\det(g) = 1$ et $\det(q_g) = \det(q)\det(g) = \det(q)$. Par unicité : $q_g = q$. Ce qui montre que $G \subset \mathrm{O}(q)$.

(ii) Si G est maximal pour l'inclusion, comme $\mathrm{O}(q)$ est un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$, nécessairement : $G = \mathrm{O}(q)$.

Réciproquement, supposons $G = \mathrm{O}(q)$. Soit G' un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ tel que $G \subset G'$. D'après (i), il existe une forme quadratique $q' \in Q_{++}$ telle que $G' \subset \mathrm{O}(q')$. Montrons alors que $\mathrm{O}(q) \subset \mathrm{O}(q')$. On note b (resp. b') la forme polaire associée à q (resp. q'). On se place dans $\mathbb{R}^n$ muni du produit scalaire b'. Le théorème de représentation nous fournit l'existence d'un endomorphisme u telle que :

$$\forall x, y \in \mathbb{R}^n,\ b'(u(x), y) = b(x,y)\,.$$

De plus, comme b et b' sont symétrique, u est auto-adjoint. Soit $v \in \mathrm{O}(q)$. Soient x et y dans $\mathbb{R}^n$. On a :

$$\forall x, y \in \mathbb{R}^n,\ b'((u \circ v)(x), v(y)) = b(v(x), v(y)) = b(x,y) = b'(u(x), y) = b'((v \circ u)(x), v(y))\,,$$

car $\mathrm{O}(q) \subset \mathrm{O}(q')$. Comme v surjective, on en déduit que u et v commutent. u commute avec tous les éléments de $\mathrm{O}(q')$ ce qui montre. Montrons que u est une homothétie. Pour cela, il suffit de montrer que pour tout $x \in \mathbb{R}^n$ la famille $(u(x), x)$ est liée. Soit $x \in \mathbb{R}^n \setminus \{0\}$ et notons H l'orthogonal de $\mathrm{vect}(x)$ pour q'. La symétrie orthogonale s par rapport à $\mathrm{vect}(x) = \mathrm{Ker}(s - \mathrm{id})$ et parallèlement à $H = \mathrm{Ker}(s + \mathrm{id})$ est un élément de $\mathrm{O}(q')$ donc commute avec u. Ainsi : $s(u(x)) = u(s(x)) = u(x)$. Donc $u(x) \in \mathrm{vect}(x)$, ce qui conclut.

□

Remarque. Tout les résultats exposés ici restent vrai dans un espace vectoriel réel E quelconque de dimension finie (il suffit de munir E d'une structure euclidienne).

Références

– Francinou, Gianella et Nicolas, Algèbre 3, p.222 (log-convexité du déterminant) et p.229 (développement et application)

Leçons concernées : 152, 171, 203

15 Méthode de Newton

Contexte

Soit $(x_n)_n$ une suite convergeant vers un réel x^*. S'il existe $0 < c < 1$ tel que :

$$\forall n \in \mathbb{N},\ |x_{n+1} - x^*| \leq c\,|x_n - x^*| \ ,$$

on dit que la convergence est *linéaire*. S'il existe $C > 0$ tel que :

$$\forall n \in \mathbb{N},\ |x_{n+1} - x^*| \leq C\,|x_n - x^*|^2 \ ,$$

on dit que la convergence est *quadratique*.

Soit $f : \mathbb{R} \to \mathbb{R}$ dérivable et soit x^* un zéro de f qu'on cherche à approcher. Si $f'(x^*) \neq 0$ alors, par continuité de f, il existe un voisinage $[c\,,d]$ de x^* tel que, pour tout $x \in [c\,,d]$, $f'(x) > 0$ (ou $f'(x) < 0$). Ainsi, les hypothèses du théorème ci-dessous ne sont pas restrictives. On pose $F(x) = x - \frac{f(x)}{f'(x)}$ (lorsque $f'(x) \neq 0$) et on appelle *méthode de Newton* pour f la suite récurrente $(x_n)_{n\in\mathbb{N}}$ définie par :

$$x_0 \in [c\,,d] \quad \text{et} \quad \forall n \geq 0,\ x_{n+1} = F(x_n)\,.$$

Développement

Théorème (Méthode de Newton)**.** *Soit $f : [c\,,d] \to \mathbb{R}$ de classe $\mathcal{C}^2$ telle que $f'(x) > 0$ pour $x \in \,]c\,,d[$ et il existe $x^* \in \,]c\,,d[$ tel que $f(x^*) = 0$ (x^* est unique par stricte croissance). Alors :*

(i) il existe un voisinage $I = [x^* - \alpha\,, x^* + \alpha]$ de x^* dans $[c\,,d]$, stable par F et tel que, pour $x_0 \in I$, $(x_n)_n$ converge quadratiquement vers x^*.

(ii) si, de plus, $f''(x) > 0$ pour tout $x \in \,]c\,,d[$ alors on peut choisir $I = [x^*\,,d]$ dans (i) et, pour $x_0 \in I \setminus \{x^*\}$, la suite $(x_n)_n$ est strictement décroissante et vérifie :

$$\exists C > 0,\ \forall n \in \mathbb{N},\ 0 \leq x_{n+1} - x^* \leq C(x_n - x^*)^2\,,$$
$$x_{n+1} - x^* \underset{+\infty}{\sim} \frac{1}{2}\frac{f''(x^*)}{f'(x^*)}(x_n - x^*)^2\,.$$

Démonstration. On remarque que, par stricte croissance de f : $f(c) < 0$ et $f(d) > 0$ et que, sous les hypothèses du théorème, F est bien définie sur $[c\,,d]$.

(i) Comme $f(x^*) = 0$, on a :

$$\forall x \in [c\,,d],\ F(x) - x^* = x - x^* - \frac{f(x) - f(x^*)}{f'(x)} = \frac{f(x^*) - f(x) - (x^* - x)f'(x)}{f'(x)}\,,$$

et en appliquant au numérateur la formule de Taylor d'ordre 2 d'origine x et d'extremité x^*, il existe $z_x \in \,]x^*\,,x[$ (ou $]x\,,x^*[$) tel que :

$$\forall x \in [c\,,d],\ F(x) - x^* = \frac{1}{2}\frac{f''(z_x)}{f'(x)}(x - x^*)^2\,.$$

En prenant : $C = \frac{1}{2}\frac{\max_{[c,d]}|f''|}{\min_{[c,d]}|f'|}$, on obtient :

$$\forall x \in [c\,,d],\ |F(x) - x^*| \leq C\,|x - x^*|^2 \ .$$

Soit $\alpha > 0$ assez petit pour que $C\alpha < 1$ et $I = [x^* - \alpha\,, x^* + \alpha] \subset [c\,,d]$, alors si $x \in I$, on a :

$$|F(x) - x^*| \leq C\alpha^2 < \alpha\,,$$

donc I est stable par F et pour $x_0 \in I$, la méthode de Newton est bien définie, $(x_n)_n$ reste dans I et converge quadratiquement vers x^*. En particulier :

$$\forall n \in \mathbb{N},\ |x_n - x^*| \leq (C\,|x_0 - \alpha|)^{2^n} \leq (C\alpha)^{2^n}\,.$$

(ii) Dans ces hypothèses, f' est strictement croissante et f est strictement convexe sur $]c\,,d[$. En particulier, pour tout $x \in \,]x^*\,,d]$, $f(x) > 0$ donc :

$$\forall x \in \;]x^*\,,d],\; F(x) = x - \frac{f(x)}{f'(x)} < x\,, \qquad (🍃)$$

de plus, d'après la démonstration du (i) :

$$\forall x \in \;]x^*\,,d],\; \exists z_x \in \,]x^*\,,x[,\; F(x) - x^* = \frac{1}{2}\frac{f''(z_x)}{f'(x)}(x-x^*)^2 > 0\,.$$

Ces deux inégalités (et le fait que x^* est un point fixe de F) montrent que l'intervalle $I = [c\,,d]$ est stable par F et que, pour $x_0 \in I \setminus \{x^*\}$, la méthode de Newton forme une suite $(x_n)_n$ strictement décroissante et minorée (par x^*) donc convergente vers une limite notée $\ell \in I$ qui vérifie, par continuité de F : $F(\ell) = \ell$. Nécessairement, d'après l'inégalité (🍃) : $\ell = x^*$. On a, de la même manière qu'en (i) :

$$\forall n \in \mathbb{N},\; 0 \leq x_{n+1} - x^* \leq C(x_n - x^*)^2\,.$$

ainsi :

$$\forall n \in \mathbb{N},\; \exists z_n \in \,]x^*\,,x_n[,\; \frac{x_{n+1}-x^*}{(x_n-x^*)^2} = \frac{1}{2}\frac{f''(z_n)}{f'(x_n)}\,.$$

Comme $x^* < z_n < x_n$ et $x_n \longrightarrow x^*$, on a par continuité :

$$x_{n+1} - x^* \underset{+\infty}{\sim} \frac{1}{2}\frac{f''(x^*)}{f'(x^*)}(x_n - x^*)^2\,.$$

□

Remarque. – Il est indispensable de savoir illustrer la méthode de Newton avec un dessin !

- Si x_0 est suffisament proche de x^*, on obtient une méthode qui converge quadratiquement. L'intérêt du cas où f est convexe est d'assouplir les conditions sur x_0.
- La méthode de Newton se généralise à des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On peut supposer, par exemple, que $f : \mathbb{R}^n \to \mathbb{R}^n$ est $\mathcal{C}^2$ et qu'il existe un point x^* tel que :

$$f(x^*) = 0 \quad \text{et} \quad Df(x^*) \text{ est inversible.}$$

Compléments

Un petit rappel sur les points fixes attractifs.

Proposition. *Soient $F : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^2$ et $x^* \in \mathbb{R}$ tel que $F(x^*) = x^*$. On considère la suite récurrente définie par :*

$$x_0 \in \mathbb{R}\,,$$
$$\forall n \geq 0,\; x_{n+1} = F(x_n)\,.$$

(i) Si $0 < |F(x^*)| < 1$ alors on dit que x^* est un point fixe attractif *et il existe un voisinage $I = [x^* - \alpha\,, x^* + \alpha]$ de x^* tel que si $x_0 \in I$ alors :*

$$\forall n \geq 1,\; x_n \in I\,, \quad x_n \underset{n\to+\infty}{\longrightarrow} x^* \quad \text{et} \quad x_{n+1} - x^* \underset{+\infty}{\sim} F'(x^*)(x_n - x^*)\,.$$

(ii) Si $|F'(x^*)| = 0$ et $|F''(x^*)| \neq 0$ alors on dit que x^* est un point fixe superattractif *il existe un voisinage $I = [x^* - \alpha\,, x^* + \alpha]$ de x^* tel que si $x_0 \in I$ alors :*

$$\forall n \geq 1,\; x_n \in I\,, \quad x_n \underset{n\to+\infty}{\longrightarrow} x^* \quad \text{et} \quad x_{n+1} - x^* \underset{+\infty}{\sim} \frac{F''(x^*)}{2}(x_n - x^*)^2\,.$$

Remarque. La démonstration de cette proposition utilise des outils du type (in)égalité de la moyenne ou formule de Taylor.

Références

- Rouvière, Petit Guide du Calcul Différentiel, p.152

Leçons concernées : 223, 226, 229

16 L'équation de la chaleur sur $\mathbb{T}$

Contexte

On note $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ et $\mathcal{C}_{2\pi} = \mathcal{C}(\mathbb{T}, \mathbb{C})$ l'ensemble des applications continues 2π-périodiques. Pour $n \in \mathbb{Z}$ et $f \in \mathcal{C}_{2\pi}$, on définit le n-ième *coefficient de Fourier* $c_n(f)$ de f par :

$$c_n(f) = \frac{1}{2\pi} = \int_0^{2\pi} f(x)e^{-inx}\,\mathrm{d}x\,.$$

Développement

Théorème (Équation de la chaleur sur $\mathbb{T}$)**.** *Soit $f \in \mathcal{C}_{2\pi}$ et $\mathcal{C}^1$ par morceaux. Il existe une unique application $u : \mathbb{R}_+ \times \mathbb{T} \to \mathbb{C}$ continue sur $\mathbb{R}_+ \times \mathbb{T}$ et $\mathcal{C}^\infty$ sur $\mathbb{R}_+^* \times \mathbb{T}$ vérifiant le problème*

$$\left\{ \begin{array}{ll} \frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x)\,, & \forall (t,x) \in \mathbb{R}_+^* \times \mathbb{T} \\ u(0,x) = f(x)\,, & \forall x \in \mathbb{T} \end{array} \right. .$$

Démonstration. On raisonne par analyse/synthèse.

(i) **Analyse.** Soit $u : \mathbb{R}_+ \times \mathbb{T} \to \mathbb{C}$ une solution. Comme u est $\mathcal{C}^\infty$, pour $t > 0$, l'application $u_t : x \in \mathbb{T} \mapsto u(t,x)$ l'est aussi. Sa série de Fourier de u_t converge donc normalement vers u_t. En notant $c_n(t)$ les coefficients de Fourier de u_t, on a :

$$\forall x \in \mathbb{T},\ u_t(x) = \sum_{n\in\mathbb{Z}} c_n(t)e^{inx}\,.$$

De plus, pour tout $n \in \mathbb{Z}$, on a :

$$\forall t > 0,\ c_n(t) = \frac{1}{2\pi}\int_0^{2\pi} u(t,x)e^{-inx}\,\mathrm{d}x\,,$$

et le théorème de dérivation (resp. de continuité) sous le signe intégrale montre que l'application $t > 0 \mapsto c_n(t)$ est $\mathcal{C}^1$ (resp. continue) sur $\mathbb{R}_+^*$ (resp. sur $\mathbb{R}_+$). De plus :

$$\forall t > 0,\ c_n'(t) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial u}{\partial t}(t,x)e^{-inx}\,\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial^2 u}{\partial x^2}(t,x)e^{-inx}\,\mathrm{d}x\,,$$

et par intégrations par parties successives :

$$\begin{aligned} \forall t > 0,\ c_n'(t) &= \frac{1}{2\pi}\underbrace{\left[\frac{\partial u}{\partial x}(t,x)e^{-inx}\right]_0^{2\pi}}_{=0} + \frac{in}{2\pi}\int_0^{2\pi} \frac{\partial u}{\partial x}(t,x)e^{-inx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\underbrace{\left[u(t,x)e^{-inx}\right]_0^{2\pi}}_{=0} - \frac{n^2}{2\pi}\int_0^{2\pi} u(t,x)e^{-inx}\,\mathrm{d}x \\ &= -n^2 c_n(t)\,. \end{aligned}$$

Or, par continuité :

$$\forall n \in \mathbb{Z},\ c_n(t) = \frac{1}{2\pi}\int_0^{2\pi} u(t,x)e^{-inx}\,\mathrm{d}x \underset{t\to 0^+}{\longrightarrow} \frac{1}{2\pi}\int_0^{2\pi} \underbrace{u(0,x)}_{=f(x)}e^{-inx}\,\mathrm{d}x = c_n(f)\,.$$

On en déduit que :

$$\forall t > 0, c_n(t) = c_n(f)\exp(-n^2 t)\,.$$

Nécessairement, pour tout $x \in \mathbb{T}$ et tout $t > 0$, u vérifie :

$$u(0,x) = f(x) \quad \text{et} \quad u(t,x) = \sum_{n\in\mathbb{Z}} c_n(f)e^{-n^2 t}e^{inx}\,. \qquad (✎)$$

(ii) **Synthèse.** Réciproquement, soit la fonction $u : \mathbb{R}_+ \times \mathbb{T} \to \mathbb{C}$ définie par (). Montrons d'abord que u est $\mathcal{C}^\infty$ sur $\mathbb{R}_+^* \times \mathbb{T}$ et vérifie l'équation de la chaleur. Procédons par récurrence et posons pour $m \in \mathbb{N}$:
$\mathcal{P}(m)$: "u est de classe $\mathcal{C}^m$ sur $\mathbb{R}_+^* \times \mathbb{T}$ et pour tout $(k, \ell) \in \mathbb{N}^2$ tels que $k + \ell \leq m$, on a :

$$\forall x \in \mathbb{T},\ \forall t > 0,\ \frac{\partial^{k+\ell} u}{\partial t^k \partial x^\ell}(t, x) = \sum_{n \in \mathbb{Z}} c_n(f)(-n^2)^k (in)^\ell e^{-n^2 t} e^{inx} . \quad \text{"}$$

Comme pour tout $t \in [a\,, b]$ (où $0 < a < b < +\infty$), on a :

$$\forall (k, \ell) \in \mathbb{N}^2,\ \forall n \in \mathbb{Z},\ \forall x \in \mathbb{T},\ \left| c_n(f)(-n^2)^k (in)^\ell e^{-n^2 t} e^{inx} \right| \leq \|f\|_\infty |n|^{2k+\ell} e^{-n^2 a} \in \ell^1(\mathbb{Z}) ,$$

le théorème de dérivation sous le signe somme montre que $\mathcal{P}_0$ est vraie puis que, si on suppose $\mathcal{P}(m)$ vraie pour m fixé dans $\mathbb{N}$, alors $\mathcal{P}(m+1)$ est vraie. De plus, u vérifie l'équation de la chaleur.
Pour montrer la continuité de u, on remarque que :

$$\forall t \geq 0,\ \forall n \in \mathbb{Z},\ \left| c_n(f) e^{-n^2 t} e^{inx} \right| \leq |c_n(f)| = \frac{|c_n(f')|}{|n|}$$

et, or la formule de Parseval montre que $(c_n(f'))_n \in \ell^2(\mathbb{Z})$ et l'inégalité de Cauchy-Schwarz que $\left(\frac{c_n(f')}{n} \right)_n \in \ell^1(\mathbb{Z})$. Ainsi, on a convergence normale sur $\mathbb{R}_+$ donc u est continue, ce qui conclut la démonstration du théorème.

□

Compléments

En fait, le théorème reste vrai si on suppose f seulement continue.

Démonstration. Jusqu'à la démonstration de la continuité, tout se passe comme dans le cas f $\mathcal{C}^1$ par morceaux.

Il reste plus qu'à montrer la continuité de u lorsque $t \to 0$. Soit $t > 0$. On définit le *noyau de la chaleur* (ou *noyau de Gauss*) :

$$\forall x \in \mathbb{T},\ p_t(x) = \sum_{n \in \mathbb{Z}} e^{-n^2 t} e^{inx} .$$

C'est une série de Fourier qui converge normalement donc : $\forall n \in \mathbb{Z},\ c_n(p_t) = e^{-n^2 t}$, et pour les mêmes raisons : $\forall n \in \mathbb{Z},\ c_n(u(t, \cdot)) = c_n(f) e^{-n^2 t}$. Or :

$$\forall n \in \mathbb{Z},\ c_n(f * p_t) = c_n(f) c_n(p_t) = c_n(f) e^{-n^2 t} = c_n(u(t, \cdot)) ,$$

et par injectivité de $f \in \mathcal{C}_{2\pi} \mapsto (c_n(f))_{n \in \mathbb{Z}}$, on a : $u(t, \cdot) = f * p_t$. On veut montrer que $f * p_t \longrightarrow f$ uniformément.

On va montrer que $(p_t)_t$ est une approximation de l'unité. Pour $t > 0$, on applique la formule de Poisson (voir plus bas) à l'application $f_t : x \in \mathbb{R} \mapsto \sqrt{\frac{\pi}{t}} \exp\left(-\frac{x^2}{4t} \right) \in \mathcal{S}(\mathbb{R})$ (une petite gaussienne des familles) :

$$\forall x \in \mathbb{T},\ \sqrt{\frac{\pi}{t}} \sum_{n \in \mathbb{Z}} \exp\left(-\frac{(x - 2n\pi)^2}{4t} \right) = \sum_{n \in \mathbb{Z}} e^{-n^2 t} e^{inx}$$

On a :

$$\begin{aligned}
\forall t > 0,\ \forall x \in \mathbb{T},\ (f * p_t)(x) &= \frac{1}{\sqrt{4\pi t}} \int_0^{2\pi} \sum_{n \in \mathbb{Z}} \exp\left(-\frac{(x - (y + 2n\pi))^2}{4t} \right) f(y)\, \mathrm{d}y \\
&= \frac{1}{\sqrt{4\pi t}} \sum_{n \in \mathbb{Z}} \int_0^{2\pi} \exp\left(-\frac{(x - (y + 2n\pi))^2}{4t} \right) f(y)\, \mathrm{d}y \\
&= \frac{1}{\sqrt{4\pi t}} \sum_{n \in \mathbb{Z}} \int_{2n\pi}^{2(n+1)\pi} \exp\left(-\frac{(x - y)^2}{4t} \right) f(y)\, \mathrm{d}y \\
&= \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} \exp\left(-\frac{(x - y)^2}{4t} \right) f(y)\, \mathrm{d}y .
\end{aligned}$$

On en déduit : $f * p_t \underset{t \to 0}{\longrightarrow} f$, uniformément sur tout compact de $\mathbb{R}$ donc sur $\mathbb{T}$. Donc u est continue. $\square$

Théorème (Formule sommatoire de Poisson)**.** *Soit $f \in \mathcal{S}(\mathbb{R})$. Alors :*

$$\forall x \in \mathbb{T}, \ \sum_{n \in \mathbb{Z}} f(x + 2n\pi) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx} \, .$$

Démonstration. Soit $f \in \mathcal{S}(\mathbb{R})$. On pose pour $x \in \mathbb{T}$:

$$g(x) = \sum_{n \in \mathbb{Z}} f(x + 2n\pi) \, .$$

Soit $R > 0$. Pour tout $k \in \mathbb{N}$ et puisque $\left| x^2 f^{(k)}(x) \right|$ est borné, on a pour tout $|n| > \frac{R}{2\pi}$:

$$\forall x \in B(0, R), \ \left| f^{(k)}(x + 2n\pi) \right| = O\left(\frac{1}{(x + 2n\pi)^2} \right) = O\left(\frac{1}{(2n\pi - R)^2} \right)$$

il y a donc convergence normale sur tout compact et g est de classe $\mathcal{C}^\infty$. g est clairement 2π-périodique, on peut calculer ses coefficients de Fourier. Pour $n \in \mathbb{Z}$, on a :

$$\begin{aligned} c_n(g) &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{k \in \mathbb{Z}} f(x + 2k\pi) e^{-inx} \, \mathrm{d}x = \sum_{k \in \mathbb{Z}} \frac{1}{2\pi} \int_0^{2\pi} f(x + 2k\pi) e^{-inx} \, \mathrm{d}x \\ &= \sum_{k \in \mathbb{Z}} \frac{1}{2\pi} \int_{2k\pi}^{2(k+1)\pi} f(x) e^{-in(x - 2k\pi)} \, \mathrm{d}x = \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-inx} \, \mathrm{d}x = \frac{\hat{f}(n)}{2\pi} \, . \end{aligned}$$

(on peut intervertir les signes $\sum$ et $\int$ car $f \in \mathcal{S}(\mathbb{R})$) Puisque $g \in \mathcal{C}^1$, sa série de Fourier converge normalement :

$$g(x) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx} \, .$$

$\square$

Références

– Francinou, Gianella et Nicolas, Analyse 4 , p.49

Leçons concernées : 222, 241, 246

17 Développement asymptotique de la série harmonique

Développement

Théorème. *Pour $n \in \mathbb{N}^*$, on pose $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{k=1}^{n} \frac{1}{k}$. Il existe une constante $\gamma > 0$ (constante d'Euler)*

$$H_n = \ln(n) + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + o\left(\frac{1}{n^2}\right).$$

Démonstration. On pose, pour $n \in \mathbb{N}^*$:

$$u_n = H_n - \ln n \quad \text{et} \quad v_n = u_n - \frac{1}{n}.$$

Montrons que u_n et v_n sont adjacentes.

- $u_n - v_n = \frac{1}{n}$ est positive et tend vers 0,
- $u_n - u_{n+1} = -\frac{1}{n+1} + \ln(n+1) - \ln(n) = -\frac{1}{n+1} - \ln\left(1 - \frac{1}{n+1}\right) \geq 0$ car : $\forall x > -1,\ \ln(1+x) \leq x$. Donc u_n est décroissante.
- $v_{n+1} - v_n = \frac{1}{n} - \ln(n+1) + \ln(n) = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) \geq 0$ pour les mêmes raisons. Donc v_n est croissante.

Ainsi u_n et v_n convergent vers un réel γ et $\gamma > 0$ car $v_2 = 1 - \ln 2 > 0$. Donc : $H_n = \ln n + \gamma + o(1)$.

Posons $t_n = u_n - \gamma$. Pour tout $n \geq 2$:

$$t_n - t_{n-1} = \frac{1}{n} + \ln\left(1 - \frac{1}{n}\right) \underset{+\infty}{\sim} -\frac{1}{2n^2},$$

Ainsi la série $\sum t_{k+1} - t_k$ converge et le théorème de sommation des équivalents donne ($t_n - t_{n-1}$ est de signe constant) :

$$-t_n = \sum_{k \geq n+1} t_k - t_{k-1} \underset{+\infty}{\sim} -\frac{1}{2} \sum_{k \geq n+1} \frac{1}{k^2}.$$

L'équivalent du reste de la série de Riemann s'obtient par comparaison série-intégrale : si $\alpha > 1$, la fonction $x \mapsto \frac{1}{x^\alpha}$ est décroissante sur $[1\,, +\infty[$, donc :

$$\forall k \geq 2,\ \int_k^{k+1} \frac{\mathrm{d}t}{t^\alpha} \leq \frac{1}{k^\alpha} \leq \int_{k-1}^{k} \frac{\mathrm{d}t}{t^\alpha},$$

et en sommant entre n et N puis en faisant N vers $+\infty$:

$$\int_n^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \leq \sum_{k \geq n} \frac{1}{k^\alpha} \leq \int_{n-1}^{+\infty} \frac{\mathrm{d}t}{t^\alpha}.$$

Les membres de gauche et de droite sont équivalents à $\frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}$ donc celui du milieu aussi et :

$$H_n = \ln n + \gamma + \frac{1}{2n} + o\left(\frac{1}{n}\right).$$

Posons $w_n = u_n - \gamma - \frac{1}{2n}$. Lorsque n tend vers $+\infty$:

$$\begin{aligned}
w_n - w_{n-1} &= \ln\left(1 - \frac{1}{n}\right) + \frac{1}{n} - \frac{1}{2n} + \frac{1}{2n-2} = -\frac{1}{n} - \frac{1}{2n^2} - \frac{1}{3n^3} + \frac{1}{n} - \frac{1}{2n} + \frac{1}{2n}\frac{1}{1-\frac{1}{n}} + o\left(\frac{1}{n^3}\right) \\
&= -\frac{1}{2n^2} - \frac{1}{3n^3} - \frac{1}{2n} + \frac{1}{2n}\left(1 + \frac{1}{n} + \frac{1}{n^2}\right) + o\left(\frac{1}{n^3}\right) \\
&= \frac{1}{6n^3} + o\left(\frac{1}{n^3}\right) \underset{+\infty}{\sim} \frac{1}{6n^3},
\end{aligned}$$

Ainsi n assez grand $w_n - w_{n-1}$ est positive et on peut appliquer le théorème de sommation des équivalents :

$$-w_n = \sum_{k \geq n+1} w_k - w_{k-1} \underset{+\infty}{\sim} \frac{1}{6} \sum_{k \geq n+1} \frac{1}{k^3} \underset{+\infty}{\sim} \frac{1}{12n^2}.$$

Ce qui montre que :

$$H_n = \ln(n) + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + o\left(\frac{1}{n^2}\right).$$

$\square$

Remarque. On sait peu de choses sur γ, notamment si γ est rationnel ou non. On a la valeur approchée par défaut : $\gamma \approx 0.577215$. Une manière d'approcher γ est d'utiliser les suites adjacentes u_n et v_n. En effet :

$$\forall n \in \mathbb{N},\ u_n - \gamma \leq u_n - v_n = \frac{1}{n}.$$

et même, en posant $s_n = \frac{u_n+v_n}{2} = u_n - \frac{1}{2n}$ (faire un dessin pour comprendre pourquoi on fait ça) et puisque $0 \leq u_n - \gamma \leq \frac{1}{n}$, on a :

$$|s_n - \gamma| = \left|u_n - \gamma - \frac{1}{2n}\right| = \begin{cases} u_n - \gamma - \frac{1}{2n} \leq \frac{1}{2n} & \text{si } u_n - \gamma \geq \frac{1}{2n} \\ \frac{1}{2n} - (u_n - \gamma) \leq \frac{1}{2n} & \text{si } u_n - \gamma \leq \frac{1}{2n} \end{cases}.$$

Références

– Francinou, Gianella et Nicolas, Analyse 1, p.156

Leçons concernées : 223, 224, 230

18 Théorème abélien et application à la série $\sum \frac{\sin(nt)}{n}$

Développement

Théorème (Abel)**.** *Soit* $\sum_{n\geq 0} a_n z^n$ *une série entière de rayon de convergence 1. On suppose que* $\sum_{n\geq 0} a_n$ *converge. On note* f *la somme de cette série entière sur le disque unité. On fixe* $\theta_0 \in [0\,,\frac{\pi}{2}[$ *et on pose :*

$$\Delta_{\theta_0} = \{z \in \mathbb{C},\ |z| \leq 1 \text{ et } \exists \rho > 0,\ \exists \theta \in [-\theta_0\,,\theta_0],\ z = 1 - \rho e^{\theta}\}\,.$$

Alors :

$$\lim_{z\to 1,\ z\in\Delta_{\theta_0}} f(z) = \sum_{n\geq 0} a_n\,.$$

Démonstration. On note $S = \sum_{n\geq 0} a_n$, $S_N = \sum_{n=0}^{N} a_n$ et $R_N = S - S_N$ pour tout $N \in \mathbb{N}$. On va effectuer une transformation d'Abel (intégration par parties discrète) en écrivant $a_n = R_{n-1} - R_n$. Pour tout $|z| < 1$ et pour tout $N \in \mathbb{N}^*$:

$$\begin{aligned}\left(\sum_{n=0}^{N} a_n z^n\right) - S_N &= \sum_{n=1}^{N}(R_{n-1} - R_n)(z^n - 1) = \sum_{n=0}^{N-1} R_n(z^{n+1} - 1) - \sum_{n=1}^{N} R_n(z^n - 1)\\ &= R_0(z-1) + \sum_{n=1}^{N-1} R_n(z^{n+1} - z^n) - R_N(z^N - 1) = (z-1)\sum_{n=0}^{N-1} R_n z^n - R_N(z^N - 1)\,,\end{aligned}$$

R_n est borné donc la série entière $\sum_{n\geq 0} R_n z^n$ a un rayon de convergence au moins égal à 1 et en faisant tendre N vers $+\infty$:

$$f(z) - S = (z-1)\sum_{n\geq 0} R_n z^n\,.$$

Soit $\varepsilon > 0$. Comme R_n tend vers 0, il existe N tel que : $\forall n > N,\ |R_n| \leq \frac{\varepsilon \cos\theta_0}{4}$. Donc :

$$|f(z) - S| \leq |z-1|\left|\sum_{n=0}^{N} R_n z^n\right| + \frac{\varepsilon\cos\theta_0}{4}|z-1|\left(\sum_{n>N}|z^n|\right) \leq |z-1|\underbrace{\sum_{n=0}^{N}|R_n|}_{:=A} + \frac{\varepsilon\cos\theta_0}{4}\frac{|z-1|}{1-|z|}\,.$$

Lorsque $z \in \Delta_{\theta_0}$, il existe $\varphi \in [-\theta_0\,,\theta_0]$ et $\rho > 0$ tel que : $|z|^2 = 1 - 2\rho\cos\varphi + \rho^2$. Et pour $\rho \leq \cos\theta_0$:

$$\frac{|z-1|}{1-|z|} = \frac{|z-1|}{1-|z|^2}(1+|z|) \leq \frac{2\rho}{2\rho\cos\varphi - \rho^2} \leq \frac{2}{2\cos\theta_0 - \cos\theta_0} \leq \frac{2}{\cos\theta_0}\,,$$

et pour $|z-1| \leq \min\{\frac{\varepsilon}{2A}, \cos\theta_0\}$, on a bien : $|f(z) - S| \leq \varepsilon$. □

Remarque.
- Il faut faire un dessin du secteur angulaire Δ_{θ_0}.
- Ce résultat se généralise à des séries entières de rayon de convergence R quelconque dans $]0\,,+\infty[$ et des secteurs angulaires d'origine n'importe quel point sur le cercle $C(0,R)$.

Application.

$$\forall 0 < t < 2\pi,\ \sum_{n\geq 1}\frac{\sin(nt)}{n} = \frac{\pi - t}{2}\,.$$

Démonstration. On se fixe t dans $]0\,,2\pi[$ et on considère la série entière $\sum_{n\geq 1} z^n\frac{\sin(nt)}{n}$ dont le rayon de convergence est ≥ 1 et qui converge en $z = 1$ (par une transformation d'Abel). Pour $-1 \leq r < 1$ et $u \in \mathbb{R}$, on pose :

$$f_r(u) = \sum_{n\geq 1}\frac{r^n\sin(nu)}{n} \quad \text{et} \quad g_r(u) = \arctan\left(\frac{r\sin(u)}{1 - r\cos(u)}\right)\,.$$

Comme la série de fonctions qui définie f_r converge normalement, f_r est $\mathcal{C}^\infty$ et, pour tout $u \in \mathbb{R}$:

$$f_r'(u) = \sum_{n\geq 1} r^n\cos(nu) = \Re\left(\sum_{n\geq 1}(re^{iu})^n\right) = \Re\left(\frac{re^{iu}}{1 - re^{iu}}\right) = \frac{r(\cos(u) - r)}{1 - 2r\cos(u) + r^2}\,.$$

g_r aussi est $\mathcal{C}^\infty$ et on a pour tout $u \in \mathbb{R}$:

$$g_r'(u) = \frac{\frac{r\cos(u)}{1-r\cos(u)} - \frac{r^2\sin^2(u)}{(1-r\cos(u))^2}}{1+\left(\frac{r\sin(u)}{1-r\cos(u)}\right)^2} = \frac{r\cos(u)-r^2}{(1-r\cos(u))^2+(r\sin(u))^2} = \frac{r(\cos(u)-r)}{1-2r\cos(u)+r^2}\,.$$

et comme $f_r(0) = g_r(0) = 0$: $f_r = g_r$. Par passage à la limite, lorsque $r \longrightarrow 1^-$, on obtient :

$$\sum_{n\geq 1} \frac{\sin(nt)}{n} = \arctan\left(\frac{\sin(t)}{1-\cos(t)}\right)\,.$$

D'autre part :

$$\frac{\sin(t)}{1-\cos(t)} = \frac{2\sin\frac{t}{2}\cos\frac{t}{2}}{2\sin^2\frac{t}{2}} = \cot\left(\frac{t}{2}\right) = \tan\left(\frac{\pi}{2}-\frac{t}{2}\right)\,.$$

et comme $-\frac{\pi}{2} < \frac{\pi-t}{2} < \frac{\pi}{2}$, on a : $\arctan\left(\tan\left(\frac{\pi-t}{2}\right)\right) = \frac{\pi-t}{2}$. □

Compléments

Exemple. On donne un exemple simple d'utilisation du théorème d'Abel ainsi qu'un contre exemple.

- On applique le théorème à la série alternée $\sum_n \frac{(-1)^n}{2n+1}$:

$$\sum_{n\geq 0}\frac{(-1)^n}{2n+1} = \lim_{x\to 1^-}\sum_{n\geq 0}\frac{(-1)^n}{2n+1}x^n = \lim_{x\to 1^-}\arctan x = \arctan 1 = \frac{\pi}{4}\,.$$

  En fait, les cas intéressants sont les cas où la série $\sum a_n$ est semi-convergente car si la série est absolument convergente, alors la série entière $\sum a_n z^n$ converge normalement sur tout le disque unité et est donc continue.

- La réciproque du théorème est fausse car :

$$\lim_{x\to 1^-}\sum_{n\geq 0}(-1)^n x^n = \lim_{x\to 1^-}\frac{1}{1+x} = \frac{1}{2}\,,$$

  pourtant $\sum_{n\geq 0}(-1)^n$ diverge. Cependant, on connaît des réciproques partielles qui constituent les théorèmes taubériens faible et fort.

Théorème (Théorème taubérien faible)**.** *Soit $\sum a_n z^n$ une série entière de rayon de convergence 1 et f la somme de cette série sur le disque unité. On suppose que :*

$$\exists S \in \mathbb{C},\ \lim_{x\to 1^-} f(x) = S\,.$$

Si $a_n = o\left(\frac{1}{n}\right)$ alors $\sum a_n$ converge vers S.

Rapide démonstration. On pose $S_N = \sum_{n=0}^N a_n$. Puisque $(n\,|a_n|)_n$ tend vers 0, la suite est majorée par une constante $M > 0$ et en remarquant que pour $x \in]0\,,1[$, $(1-x^n) = (1-x)(1+\cdots+x^{n-1}) \leq n(1-x)$:

$$|S_N - f(x)| \leq \sum_{n=1}^N |a_n|\,(1-x^n) + \sum_{n\geq N+1}\frac{n\,|a_n|}{N}x^k \leq (1-x)MN + \frac{\sup_{n>N} n\,|a_n|}{N(1-x)}\,,$$

et en choisissant x de la forme $x = 1-\frac{\varepsilon}{N}$, on conclut facilement. □

Remarque. Pour le théorème taubérien fort, on remplace l'hypothèse $a_n = o\left(\frac{1}{n}\right)$ par $a_n = O\left(\frac{1}{n}\right)$.

Références

- Gourdon, Analyse, p.252 (théorème et compléments)
- Queffélec et Zuily, Analyse pour l'agrégation, p.42 (application)

Leçons concernées : 230, 235, 241, 243

19 Prolongement de la fonction Γ d'Euler

Contexte

On définit la fonction Γ d'Euler sur $\mathfrak{D} = \{z \in \mathbb{C} \mid \Re(z) > 0\}$ par :

$$\Gamma : z \in \mathfrak{D} \longmapsto \int_0^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t \,.$$

Développement

Théorème. *La fonction Γ se prolonge de manière unique en une fonction méromorphe sur $\mathbb{C}$ dont les pôles, tous simples, sont les entiers négatifs ou nuls.*

Démonstration. Soit $z \in \mathfrak{D}$. On découpe l'intégrale en deux puis on développe l'exponentielle en série entière :

$$\Gamma(z) = \int_0^1 e^{-t} t^{z-1} \, \mathrm{d}t + \int_1^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t = \int_0^1 \sum_{n \geq 0} \frac{(-1)^n}{n!} t^{n+z-1} \, \mathrm{d}t + \int_1^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t \,.$$

Pour tout $t > 0$, $|t^z| = t^{\Re(z)}$ et :

$$\sum_{n \geq 0} \left| \frac{(-1)^n}{n!} \right| |t|^{n+z-1} = \sum_{n \geq 0} \frac{t^n}{n!} t^{\Re(z)-1} = e^t t^{\Re(z)-1} \,,$$

et comme $z \in \mathfrak{D}$, la fonction $t \mapsto e^t t^{\Re(z)-1}$ est intégrable sur $]0\,,1]$, on peut utiliser le théorème de Fubini-Lebesgue (appliquer à la mesure produit de la mesure de Lebesgue sur $]0\,,1[$ et de la mesure de comptage) :

$$\Gamma(z) = \sum_{n \geq 0} \frac{(-1)^n}{n!} \int_0^1 t^{n+z-1} \, \mathrm{d}t + \int_1^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t = \sum_{n \geq 0} \frac{(-1)^n}{n!(z+n)} + \int_1^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t \,.$$

La fonction $z \in \mathbb{C} \mapsto \int_1^{+\infty} e^{-t} t^{z-1} \, \mathrm{d}t$ définit une fonction holomorphe sur $\mathbb{C}$. En effet, on peut appliquer le théorème d'holomorphie sous le signe intégral car :

- pour tout $z \in \mathbb{C}$, la fonction $t \mapsto e^{-t} t^{z-1}$ est mesurable,
- pour presque tout $t \geq 1$, la fonction $z \mapsto e^{-t} t^{z-1}$ holomorphe $\mathbb{C}$,
- soit $K \subset \mathbb{C}$ un compact (borné par $M > 0$ par exemple) alors, pour tout $z \in K$:

$$\left| e^{-t} t^{z-1} \right| \leq e^{-t} t^{M-1} \in L^1([1\,,+\infty[) \,.$$

La série $\sum_{n \geq 0} \frac{(-1)^n}{n!(z+n)}$ est converge normalement sur tout compact de $\mathbb{C} \setminus \mathbb{Z}^-$. En effet, soit $K \subset \mathbb{C} \setminus \mathbb{Z}^-$ un compact et notons d la distance de K à $\mathbb{Z}^-$ alors :

$$\forall z \in K, \ \forall n \geq 0, \ \left| \frac{(-1)^n}{n!(z+n)} \right| \leq \frac{1}{n! d} \,.$$

Le théorème de Weierstrass montre que $z \mapsto \sum_{n \geq 0} \frac{(-1)^n}{n!(z+n)}$ définit une fonction holomorphe de $\mathbb{C} \setminus \mathbb{Z}^-$. Soit $k \in \mathbb{N}$. De la même manière, $\sum_{n \neq k} \frac{(-1)^n}{n!(z+n)}$ converge normalement sur $B(-k, \frac{1}{2})$ et définit une fonction holomorphe sur $B(-k, \frac{1}{2})$. Ainsi, au voisinage de $-k$, on peut écrire :

$$\Gamma(z) = \frac{(-1)^k}{k!} \frac{1}{z+k} + g(z) \,,$$

où g est une fonction holomorphe au voisinage de $-k$. On a donc construit un prolongement de Γ, méromorphe sur $\mathbb{C}$ et dont les pôles, tous simples, sont les entiers négatifs ou nuls. Le principe du prolongement analytique montre que ce prolongement est unique ($\mathbb{C} \setminus \mathbb{Z}^-$ contient un point d'accumulation).

$\square$

Références

- Amar et Matheron, Analyse complexe, p.147

Leçons concernées : 207, 245

20 Nombre de zéros d'une équation différentielle linéaire

Développement

Théorème. *Soient $a \in \mathbb{R}$ et $q \in \mathcal{C}^1([a\,,+\infty[)$ strictement positive et telle que :*

$$\int_a^{+\infty} \sqrt{q(t)}\,\mathrm{d}t = +\infty \quad \text{et} \quad q'(x) = o\Big(q(x)^{3/2}\Big) \text{ lorsque } x \longrightarrow +\infty\,.$$

Soit y une solution réelle non nulle *de l'équation différentielle $y'' + qy = 0$ sur $[a\,,+\infty[$ alors, en notant $N(x)$ le nombre de zéros de y entre a et x :*

$$N(x) \underset{+\infty}{\sim} \frac{1}{\pi}\int_a^x \sqrt{q(t)}\,\mathrm{d}t\,.$$

Lemme (Principe d'entrelacement de Sturm)**.** *Soient $a \in \mathbb{R}$ et y_1, y_2 dans $\mathcal{C}^1([a\,,+\infty[,\mathbb{R})$, sans zéro commum. On note $w = y_1y_2' - y_1'y_2$. Si $y_1(a) + iy_2(a) = r_0e^{i\theta_0}$ alors, on peut écrire :*

$$y_1 = r\cos\theta \quad \text{et} \quad y_2 = r\sin\theta\,,$$

où r et θ sont dans $\mathcal{C}^1([a\,,+\infty[,\mathbb{R})$ et sont données par :

$$r = \sqrt{y_1^2 + y_2^2} \quad \text{et} \quad \theta = \theta_0 + \int_a^x \frac{w(t)}{r(t)^2}\,\mathrm{d}t\,.$$

Démonstration du lemme. On pose $\varphi = y_1 + iy_2$ et $\psi(x) = \int_a^x \frac{\varphi'(t)}{\varphi(t)}\,\mathrm{d}t + \ln(r_0) + i\theta_0$. Alors :

$$(\varphi e^{-\psi})' = (\varphi' - \varphi\psi')e^{-\psi} = 0 \implies \forall x \geq a,\ \varphi(x)e^{-\psi(x)} = \varphi(a)e^{-\psi(a)} = r_0e^{i\theta_0}r_0^{-1}e^{-i\theta_0} = 1\,.$$

Donc : $y_1 + iy_2 = e^{\psi} = re^{\Im(\psi)}$ où $r = \sqrt{y_1^2 + y_2^2}$. On calcule $\Im(\psi)$:

$$\begin{aligned}
\forall x \geq a,\ \psi(x) &= \ln(r_0) + i\theta_0 + \int_a^x \frac{y_1'(t) + iy_2'(t)}{y_1(t) + iy_2(t)}\,\mathrm{d}t \\
&= \ln(r_0) + i\theta_0 + \int_a^x \frac{(y_1'(t) + iy_2'(t))(y_1(t) - iy_2(t))}{r(t)^2}\,\mathrm{d}t \\
&= \ln(r_0) + \int_a^x \frac{y_1y_1'(t) + y_2y_2'(t)}{r(t)^2}\,\mathrm{d}t + i\underbrace{\left[\theta_0 + \int_a^x \frac{w(t)}{r(t)^2}\,\mathrm{d}t\right]}_{=\Im(\psi)}.
\end{aligned}$$

□

Démonstration du théorème. La stratégie de la preuve consiste à se ramener en changeant de variables à une équation "proche" de l'équation bien connue du pendule oscillant.

(i) On pose $\tau(x) = \int_a^x \sqrt{q(t)}\,\mathrm{d}t$ qui, d'après les hypothèses sur q est bien un $\mathcal{C}^2$-difféomorphisme croissant (par le théorème d'inversion global, par exemple même si c'est abusé) de $[a\,,+\infty[$ sur $[0\,,+\infty[$. τ^{-1} est aussi croissante et, en posant $Y = y \circ \tau^{-1}$, on a pour tout $x \geq a$:

$$\begin{aligned}
&y'(x) = \tau'(x)Y'(\tau(x)) = \sqrt{q(x)}Y'(\tau(x))\,, \\
&y''(x) = \frac{q'(x)}{2\sqrt{q(x)}}Y'(\tau(x)) + q(x)Y''(\tau(x))\,, \\
&y''(x) + q(x)y(x) = q(x)Y''(\tau(x)) + \frac{q'(x)}{2\sqrt{q(x)}}Y'(\tau(x)) + q(x)Y(\tau(x)) = 0\,,
\end{aligned}$$

et en divisant par $q(x) > 0$, on obtient l'équation :

$$Y'' + \varphi Y' + Y = 0 \quad \text{où} \quad \forall t \geq 0,\ \varphi(t) = \frac{q'(\tau^{-1}(t))}{2q(\tau^{-1}(t))^{3/2}}\,.$$

D'après les hypothèses, $\varphi \to 0$ lorsque $t \to +\infty$.

(ii) Par unicité du problème de Cauchy, Y et Y' ne s'annulent pas en même temps (et sont $\mathcal{C}^1$ par composition), on peut donc appliquer le lemme à Y et Y' et écrire :

$$Y = r\sin\theta \quad \text{et} \quad Y' = r\cos\theta\,.$$

En dérivant :

$$\begin{aligned} Y' &= r'\sin\theta + r\theta'\cos\theta = r\cos\theta\,, \\ Y'' &= r'\cos\theta - r\theta'\sin\theta = -\varphi r\cos\theta - r\sin\theta\,. \end{aligned}$$

En multipliant la première équation par $\cos\theta$, la seconde par $-\sin\theta$ et en additionnant et en divisant par $r > 0$:

$$\theta' = 1 + \varphi\sin\theta\cos\theta \implies \forall t \geq 0,\ |\theta'(t) - 1| \leq \frac{1}{2}\varphi(t) \underset{t\to+\infty}{\longrightarrow} 0$$

Donc : $\theta'(t) \to 1$ et $\theta \underset{+\infty}{\sim} t$ (intégration des équivalents). On note $M(t)$ le nombre de zéros de Y entre 0 et t. Soit t_0 tel que pour tout $t \geq t_0$, on ait $\theta'(t) > 0$. Ainsi, θ est strictement croissante sur $[t_0\,, +\infty[$. Comme les zéros de Y sont isolés, $M(t_0) < +\infty$ et donc :

$$\begin{aligned} M(t) \sim \operatorname{Card}\{u \in [t_0\,, t],\ \sin(\theta(u)) = 0\} &= \operatorname{Card}\{v \in [\theta(t_0)\,, \theta(t)],\ \sin(v) = 0\} \\ &= \operatorname{Card}\{k \in \mathbb{Z},\ \theta(t_0) \leq k\pi \leq \theta(t)\} \sim \frac{\theta(t)}{\pi} \sim \frac{t}{\pi}\,. \end{aligned}$$

Enfin $N(x) = M(\tau(x))$ puisque :

$$M(\tau(x)) = \operatorname{Card}\{t \in [0\,, \tau(x)],\ Y(t) = 0\} = \operatorname{Card}\{s \in [a\,, x],\ \underbrace{Y(\tau(s))}_{=y(x)} = 0\} = N(x)\,.$$

ce qui achève la démonstration.

□

Remarque. L'hypothèse $q'(x) = o\big(q(x)^{3/2}\big)$ n'est pas totalement inutile (cf. le contre-exemple $a = 1$ et $q(x) = \frac{1}{4x^2}$).

Références

– Queffélec et Zuily, Analyse pour l'agrégation, p.405

Leçons concernées : 220, 224

21 Méthode du gradient à pas optimal

Contexte

Une fonction $J : \mathbb{R}^n \to \mathbb{R}$ est :

(i) *coercive* si :

$$|J(u)| \underset{\|u\| \to +\infty}{\longrightarrow} +\infty \,,$$

(ii) *convexe* si :

$$\forall t \in \,]0\,,1[,\ \forall u, v \in \mathbb{R}^n,\ J(tu + (1-t)v) \leq tJ(u) + (1-t)J(v) \,.$$

et *strictement convexe* si l'inégalité est stricte.

On s'intéresse au problème d'optimisation sans contraintes :

$$u^\star = \underset{u \in \mathbb{R}^n}{\operatorname{argmin}}\ J(u) \,, \tag{¤}$$

où $J : \mathbb{R}^n \to \mathbb{R}$ est une *fonctionnelle elliptique*, i.e qui est de classe $\mathcal{C}^1$ et qui vérifie :

$$\exists \alpha > 0,\ \forall u, v \in \mathbb{R}^n,\ \langle \nabla J(v) - \nabla J(u), v - u \rangle \geq \alpha \|v - u\|^2 \,.$$

Une *méthode du gradient* consiste à construire une suite minimisante $(u_k)_k$ pour (¤) de la manière suivante :

- u_0 est choisi arbitrairement,
- pour tout $k \geq 0$, u_{k+1} est de la forme

$$\exists \rho(u_k) \in \mathbb{R},\ u_{k+1} = u_k - \rho(u_k) \nabla J(u_k) \,.$$

L'idée qui se cache derrière est que $-\nabla J(u_k)$ est la direction de plus grande descente locale. En effet, on peut écrire :

$$J(u_k + w) = J(u_k) + \langle \nabla J(u_k), w \rangle + o(\|w\|) \,,$$

et ainsi, l'accroissement de J est localement majoré par $\|\nabla J(u_k)\| \, \|w\|$ (inégalité de Cauchy-Schwarz) et avec égalité si et seulement si w et $\nabla J(u_k)$ sont proportionnels. La *méthode du gradient à pas optimal* est la méthode de gradient où, pour tout k, on choisit $\rho(u_k)$ tel que :

$$J(u_k - \rho(u_k) \nabla J(u_k)) = \inf_{\rho \in \mathbb{R}} J(u_k - \rho \nabla J(u_k)) \,.$$

Développement

Théorème. *La méthode du gradient à pas optimal converge.*

Lemme. *Soit $J : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle elliptique. Alors J est strictement convexe, coercive et vérifie :*

$$\forall u, v \in \mathbb{R}^n,\ J(v) - J(u) \geq \langle \nabla J(u), v - u \rangle + \frac{\alpha}{2} \|v - u\|^2 \,.$$

Démonstration du lemme. J est de classe $\mathcal{C}^1$. D'après la formule de Taylor avec reste intégral :

$$\begin{aligned}
\forall u, v \in \mathbb{R}^n,\ J(v) - J(u) &= \int_0^1 \langle \nabla J(u + t(v-u)), v - u \rangle \,\mathrm{d}t \\
&= \langle \nabla J(u), v - u \rangle + \int_0^1 \langle \nabla J(u + t(v-u)) - \nabla J(u), t(v-u) \rangle \,\frac{\mathrm{d}t}{t} \\
&\geq \langle \nabla J(u), v - u \rangle + \alpha \|v - u\|^2 \underbrace{\int_0^1 t \,\mathrm{d}t}_{=\frac{1}{2}} .
\end{aligned}$$

On obtient ainsi l'inégalité demandée. Il en résulte aussi que :

$$\forall u \neq v \in \mathbb{R}^n,\ J(v) - J(u) > \langle \nabla J(u), v - u \rangle \,,$$

ainsi J est strictement convexe (voir compléments). De plus, par l'inégalité de Cauchy-Schwarz :

$$\forall v \in \mathbb{R}^n, \; J(v) \geq J(0) + \langle \nabla J(0), v \rangle + \frac{\alpha}{2} \|v\|^2 \geq J(0) - \|\nabla J(0)\| \, \|v\| + \frac{\alpha}{2} \|v\|^2 \, ,$$

et J est coercive. $\square$

Démonstration du théorème. Soit $k \in \mathbb{N}^*$. On peut supposer $\nabla J(u_k) \neq 0$ et $\rho(u_k) \neq 0$ car sinon on a atteint le minimum. Comme

$$\varphi_k : \rho \in \mathbb{R} \longmapsto \varphi_k(\rho) = J(u_k - \rho \nabla J(u_k))$$

est coercive et strictement convexe (donc continue), donc admet un unique minimum en $\rho(u_k)$ caractérisé par la relation $\varphi_k'(\rho(u_k)) = 0$ i.e :

$$\begin{aligned} \varphi_k'(\rho(u_k)) &= 0 \\ &= -\langle J(u_k - \rho(u_k) \nabla J(u_k)), \nabla J(u_k) \rangle \\ &= -\langle \nabla J(u_{k+1}), \nabla J(u_k) \rangle = \rho(u_k)^{-1} \langle \nabla J(u_{k+1}), u_{k+1} - u_k \rangle \, . \end{aligned}$$

En particulier, les directions de descente successives sont orthogonales. De plus, $\langle \nabla J(u_{k+1}), u_{k+1} - u_k \rangle = 0$ et en appliquant le lemme :

$$J(u_k) - J(u_{k+1}) \geq \frac{\alpha}{2} \|u_k - u_{k+1}\|^2 \, .$$

La suite $(J(u_k))_k$ est décroissante et minorée par $J(u^\star)$ donc converge. L'inégalité précedente montre que :

$$\|u_k - u_{k+1}\| \underset{k \to +\infty}{\longrightarrow} 0 \, . \qquad (\checkmark)$$

Par orthogonalité et par l'inégalité de Cauchy-Schwarz :

$$\|\nabla J(u_k)\|^2 = \langle \nabla J(u_k), \nabla J(u_k) - \nabla J(u_{k+1}) \rangle \leq \|\nabla J(u_k)\| \, \|\nabla J(u_k) - \nabla J(u_{k+1})\| \, ,$$

donc :

$$\|\nabla J(u_k)\| \leq \|\nabla J(u_k) - \nabla J(u_{k+1})\| \, .$$

Si $(u_k)_k$ n'était pas bornée alors, puisque J est coercive, la suite $(J(u_k))_k$ ne serait pas non plus bornée ce qui est absurde car cette suite converge. Donc $(u_k)_k$ est bornée, par exemple par $r > 0$. D'après le théorème de Heine, puisque J est $\mathcal{C}^1$, l'application $u \mapsto \nabla J(u)$ est uniformément continue sur $\overline{B}(0, r)$. D'après ($\checkmark$) :

$$\|\nabla J(u_k) - \nabla J(u_{k+1})\| \underset{k \to +\infty}{\longrightarrow} 0 \, ,$$

et on en déduit :

$$\|\nabla J(u_k)\| \underset{k \to +\infty}{\longrightarrow} 0 \, .$$

Par hypothèses :

$$\alpha \|u_k - u^\star\|^2 \leq \langle \nabla J(u_k) - \nabla J(u^\star), u_k - u^\star \rangle = \langle \nabla J(u_k), u_k - u^\star \rangle \leq \|\nabla J(u_k)\| \, \|u_k - u^\star\| \, ,$$

donc :

$$\|u_k - u^\star\| \leq \frac{\|\nabla J(u_k)\|}{\alpha} \underset{k \to +\infty}{\longrightarrow} 0 \, .$$

$\square$

Compléments

On donne un critère de (stricte) convexité lorsque la fonctionnelle est $\mathcal{C}^1$.

Proposition. *Une fonction $J : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^1$ est convexe si et seulement si :*

$$\forall u \neq v \in \mathbb{R}^n, \; J(v) - J(u) \geq \langle \nabla J(u), v - u \rangle \, .$$

J est strictement convexe si et seulement si on a en plus inégalité stricte.

Démonstration. Soient $u \neq v$ dans $\mathbb{R}^n$ et $t \in]0\,,1[$. Alors, si J est convexe :

$$J(u+t(v-u)) \leq J(u)+t(J(v)-J(u)) \implies J(v)-J(u) \geq \frac{J(tu+(1-t)v)-J(u)}{t} \underset{t\to 0^+}{\longrightarrow} \langle \nabla J(u), v-u\rangle \ .$$

Si on suppose de plus J strictement convexe alors pour $r \in]t\,,1[$:

$$u+t(v-u) = \frac{r-t}{r}u + \frac{t}{r}(u+r(v-u)) \implies J(u+t(v-u)) < \frac{r-t}{r}J(u) + \frac{t}{r}J(u+r(v-u))\,,$$

on en déduit :

$$J(v)-J(u) > \frac{J(u+t(v-u))-J(u)}{r} > \frac{J(u+t(v-u))-J(u)}{t} \underset{t\to 0^+}{\longrightarrow} \langle \nabla J(u), v-u\rangle \ .$$

Réciproquement, on écrit :

$$\begin{aligned} J(v) &\geq J(v+t(u-v)) - t\,\langle \nabla J(v+t(u-v)), v-u\rangle \ , \\ J(u) &\geq J(v+t(u-v) + (1-t)\,\langle \nabla J(v+t(u-v)), u-v\rangle \ , \end{aligned}$$

On multiplie l'inégalité de haut par $(1-t)$ et celle du bas par t puis on additionne pour montrer que J est convexe. Si on avait inégalité stricte, on aurait montrer la stricte convexité. □

Évidemment, la méthode du gradient à pas optimal n'est pas la seule méthode du gradient. Les autres qu'il faut connaître sont :

- la *méthode du gradient à pas constant* ($\rho(u_k) =$ pour tout k) qui concerne toute fonction elliptique. Sous certaines conditions (notamment que ∇J soit lipchitzienne), la convergence est linéaire.
- la *méthode du gradient conjugué* qui ne concerne que les fonctionnelles quadratiques elliptiques, i.e de la forme $J(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ où A est une matrice symétrique réelle et définie positive et b un vecteur. On construit une suite minimisante dont les gradients sont orthogonaux 2 à 2 (contrairement au gradient à pas optimal où seuls les gradients successifs sont orthogonaux), donc elle converge en au plus n itérations. Minimiser une fonctionnelle quadratique elliptique est équivalent à résoudre le problème inverse associé à A et b. Cette méthode est cependant plus coûteuse en opérations élémentaires que la méthode de Choleski. Elle est par contre bien adaptée aux matrices creuses.

Convergence de la méthode à pas constant. On suppose ∇J lipchitzienne de constante de Lipschitz L. On note α la constante d'ellipticité (non conventionnel) de J. On note $u^\star$ l'unique minimum de J. Alors, en posant $v_n = u_n - u^\star$, on a :

$$\begin{aligned} \|v_{n+1}\|^2 &= \|v_n\|^2 + \rho^2\,\|\nabla J(u_n)\|^2 - 2\rho\,\langle v_n, \nabla J(u_n)\rangle \\ &= \|v_n\|^2 + \rho^2\,\|\nabla J(u_n) - \nabla J(u^\star)\|^2 - 2\rho\,\langle u_n - u^\star, \nabla J(u_n) - \nabla J(u^\star)\rangle \\ &\leq \|v_n\|^2 + \rho^2 L^2\,\|v_n\|^2 - 2\rho\alpha\,\|v_n\|^2 = (1+\rho^2L^2 - 2\rho\alpha)\,\|v_n\|^2 \ , \end{aligned}$$

en faisant un petit dessin du trinôme $1 + X^2L^2 - 2\alpha X$, on en déduit des conditions suffisantes pour que la convergence soit linéaire.

□

Références

- Ciarlet, Introduction à l'analyse numérique matriciellle et à l'optimisation, p.189

Leçons concernées : 219, 226, 229, 253

22 Projection sur un convexe fermé et théorème de Riesz-Fréchet

Contexte

Dans tout ce qui suit, H désigne un espace de Hilbert *réel*. On notera $\langle \cdot, \cdot \rangle$ son produit scalaire et $\|\cdot\|$ la norme associée. On rappelle que dans un tel espace, on a l'inégalité de *Cauchy-Schwarz* :

$$\forall u, v \in H, \ |\langle u, v \rangle| \leq \|u\| \, \|v\| \ ,$$

ainsi que l'*identité du parallélogramme* :

$$\forall u, v \in H, \ \left\| \frac{u-v}{2} \right\| + \left\| \frac{u+v}{2} \right\| = \frac{1}{2}(\|u\|^2 + \|v\|^2) \, .$$

On note H' le dual (topologique) de H, i.e l'ensemble des formes linéaires continues sur H. On munit H' de la norme

$$\forall \varphi \in H', \ \|\varphi\| = \sup_{v \in H, \|v\| \leq 1} |\varphi(v)|$$

Développement

Théorème (Projection sur un convexe fermé)**.** *Soit $K \subset H$ un convexe fermé non vide. Alors, pour tout $f \in H$, il existe $u \in K$ unique tel que*

$$\|f - u\| = \min_{v \in K} \|f - v\| \ . \qquad (🍃)$$

De plus, u est caractérisé par la propriété :

$$u \in K \quad \textit{et} \quad \forall v \in K, \ \langle f - u, v - u \rangle \leq 0 \, . \qquad (🍃🍃)$$

On note $u = P_K f$ et on dit que $P_K f$ est la projection de f sur K.

Remarque. Il faut faire un dessin pour expliquer (🍃🍃).

Démonstration. Soit $f \in H$.

(i) Montrons l'existence de u.
Soit $(v_n)_n$ une suite de K $(K \neq \emptyset)$ telle que :

$$d_n := \|f - v_n\| \underset{n \to +\infty}{\longrightarrow} d = \min_{v \in K} \|f - v\| \ .$$

On montre que $(v_n)_n$ est de Cauchy en utilisant l'identité du parallélogramme. Pour tout m, n :

$$\left\| f - \frac{v_n + v_m}{2} \right\|^2 + \left\| \frac{v_n - v_m}{2} \right\|^2 = \frac{1}{2} \left(d_n^2 + d_m^2 \right) \ .$$

Puisque K est convexe, $\frac{v_n + v_m}{2} \in K$ et donc $\left\| f - \frac{v_n + v_m}{2} \right\|^2 \geq d^2$. On en déduit :

$$\|v_n - v_m\|^2 \leq 2(d_n^2 - d^2 + d_m^2 - d^2) \underset{n \to +\infty}{\longrightarrow} 0 \, .$$

Ainsi $(v_n)_n$ est de Cauchy dans un espace complet donc converge vers u et, K étant fermé, $u \in K$. Par continuité de l'application $v \mapsto \|f - v\|$, on a bien : $d = \|f - u\|$.

(ii) Montrons l'équivalence entre (🍃) et (🍃🍃).
Soit $u \in K$ vérifiant (🍃) et soit $w \in K$. Par convexité de K, on a pour tout $t \in \,]0\,,1]$:

$$v = (1-t) + tw \in K \, ,$$

donc :

$$\|f - u\| \leq \|f - v\| = \|(f - u) - t(w - u)\| \ ,$$

En élevant au carré :

$$\|f-u\|^2 \leq \|f-u\|^2 - 2t\langle f-u, w-u\rangle + t^2\|w-u\|^2 \implies 2\langle f-u, w-u\rangle \leq t\|w-u\|\ ,$$

et lorsque $t \to 0$, on obtient (🍃🍃).
Réciproquement, si u vérifie (🍃🍃), alors pour tout $v \in K$:

$$\|f-v\|^2 = \|(f-u)-(v-u)\|^2 = \|f-u\|^2 - 2\langle f-u, v-u\rangle + \|v-u\|^2\ ,$$

d'où :

$$\forall v \in K,\ \|f-v\| - \|f-u\| \geq 0\,,$$

ce qui montre (🍃)

(iii) Montrons enfin l'unicité de u à l'aide de (🍃🍃).
Si u_1 et u_2 vérifient (🍃🍃) alors, en particulier :

$$\langle f-u_1, u_2-u_1\rangle \leq 0\,,$$
$$\langle f-u_2, u_1-u_2\rangle \leq 0\,,$$

et en additionnant, on trouve $\|u_1-u_2\|^2 \leq 0$ donc $u_1 = u_2$.

□

Application (Théorème de Riesz-Fréchet)**.** *L'application*

$$\begin{array}{cccll} \varphi & : & H & \longrightarrow & H' \\ & & f & \longmapsto & [\varphi_f : v \in H \longmapsto \varphi_f(v) = \langle f, v\rangle] \end{array}$$

est un isomorphisme isométrique.

Démonstration. D'après l'inégalité de Cauchy-Schwarz :

$$\forall f \in H,\ \forall v \in H,\ |\varphi_f(v)| \leq \|f\|\,\|v\|\ .$$

On en déduit que $\|\varphi_f\| \leq \|f\|$ et φ est bien définie. De plus, en prenant $v = f$, on a égalité. Ensuite φ est clairement une application linéaire. On a donc montré que φ est une isométrie (ce qui implique l'injectivité).

Il reste à montrer la surjectivité. Soit $\psi \in H'$. On note $M := \operatorname{Ker}\psi$ qui est un sous-espace fermé (donc aussi convexe et non vide) de H puisque ψ est continue. Si $M = H$ alors ψ est l'application nulle et est donc dans l'image de φ. Supposons que $M \neq H$ et soit $g \in H \setminus M$. On pose :

$$h = \frac{g - P_M g}{\|g - P_M g\|}\,,$$

(h est bien défini car $g \neq P_M g$) et on a (voir compléments) :

$$h \notin M, \quad \|h\| = 1 \quad \text{et} \quad \forall w \in M,\ \langle h, w\rangle = 0\,.$$

Comme M est un hyperplan, on a une décomposition de la forme : $H = \langle h\rangle \oplus M$. Alors, pour tout $\lambda \in \mathbb{R}$ et $w \in M$, on a : $\psi(\lambda h + w) = \lambda\psi(h)$. On cherche $\lambda_0 \in \mathbb{R}$ et $w_0 \in M$ tel que pour $f = \lambda_0 h + w_0$ on ait :

$$\forall \lambda \in \mathbb{R},\ \forall w \in M,\ \langle f, \lambda h + w\rangle = \lambda\lambda_0\|h\| + \langle w_0, w\rangle = \lambda\psi(h) = \psi(\lambda h + w)\,.$$

Il suffit de prendre $\lambda_0 = \psi(h)$ et $w_0 = 0$. □

Remarque. On a des théorèmes analogues dans le cas où H est un espace de Hilbert complexe (i.e muni d'un produit scalaire hermitien). Il faut cependant faire quelques changements, notamment prendre la partie réelle du produit scalaire dans (🍃🍃) et identifier H à $\overline{H}'$ où $\overline{H}$ désigne le conjugué de H (la multiplication externe par λ devient la multiplication par $\bar{\lambda}$).

Compléments

Il faut connaître certaines propriétés de l'opérateur de projection.

Proposition. *Sous les hypothèses du théorème :*

$$\forall f_1, f_2 \in H,\ \|P_K f_1 - P_K f_2\| \leq \|f_1 - f_2\| \ .$$

De plus, si $K = M$ est un sous-espace vectoriel fermé de H alors, pour tout $f \in H$, $u = P_M f$ est caractérisé par

$$u \in M \quad \textit{et} \quad \forall v \in M,\ \langle f - u, v\rangle = 0 \,.$$

De plus, P_M est un opérateur linéaire.

Démonstration. Pour le premier point,on pose $u_1 = P_K f_1$ et $u_2 = P_k f_2$, alors :

$$\begin{aligned} \langle f_1 - u_1, u_2 - u_1\rangle &\leq 0 \,, \\ \langle f_2 - u_2, u_1 - u_2\rangle &\leq 0 \,, \end{aligned}$$

et en additionnant : $\|u_1 - u_2\|^2 \leq \langle f_1 - f_2, u_1 - u_2\rangle$. On conclut avec l'inégalité de Cauchy-Schwarz.

Pour le second point, on remarque que :

$$\forall t \in \mathbb{R},\ \forall v \in H,\ \langle f - u, tv - u\rangle \leq 0 \,,$$

et en faisant tendre t vers $\pm\infty$, on obtient $\langle f - u, v\rangle = 0$. Le reste est évident. $\square$

Il peut être intéressant de savoir caractérisé la continuité d'une forme linéaire en fonction de son noyau. Dans ce qui suit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Proposition. *Soient E un $\mathbb{K}$-espace vectoriel et $\varphi : E \to \mathbb{K}$ une forme linéaire. Alors φ est continue si et seulement si* $\operatorname{Ker}\varphi$ *est fermé.*

Démonstration. Le sens direct est évident puisque $\{0\}$ est fermé. Supposons que $H := \operatorname{Ker}\varphi$ soit fermé.

Si $H = E$ alors $\varphi = 0$ est continue.

Si $H \neq E$ alors $E \setminus H$ est un ouvert non vide. Soit $x_0 \in E \setminus H$ alors $E = \langle x_0\rangle \oplus H$ et il existe $\alpha_0 > 0$ tel que $B(x_0, \alpha_0) \subset E \setminus H$. Soit $x = y_x + \lambda_x x_0 \in E$ (où $y_x \in H$). Il vient :

$$|\varphi(x)| = |\lambda_x|\,|\varphi(x_0)| \ ,$$

où $|\varphi(x_0)| \neq 0$. Si $x \in H$ si et seulement si $\lambda_x = 0$. Pour tout $x \in E \setminus H$, $x_0 - \frac{x}{\lambda_x} \in H$ donc : $\frac{\|x\|}{\lambda_x} = \left\|x_0 - \frac{x}{\lambda_x}\right\| \geq \alpha_0$ et donc : $|\lambda_x| \leq \frac{\|x\|}{\alpha_0}$ ce qui montre la continuité de φ. $\square$

Remarque. On rappelle qu'un hyperplan est soit fermé soit dense.

Références

– Brézis, Analyse fonctionnelle, p.79

Leçons concernées : 205, 208, 213, 253

23 Lemme de Scheffé

Contexte

On notera λ_d la mesure de Lebesgue sur $\mathbb{R}^d$.

Proposition. *Soient, pour tout $n \in \mathbb{N}$, μ_n et μ des mesures de probabilités sur $\mathbb{R}^d$. Les assertions suivantes sont équivalentes :*

(i) *μ_n converge étroitement vers μ,*
(ii) *pour tout fermé F de $\mathbb{R}^d$, on a : $\limsup_n \mu_n(F) \leq \mu(F)$,*
(iii) *pour tout ouvert O de $\mathbb{R}^d$, on a : $\liminf_n \mu_n(O) \geq \mu(O)$,*
(iv) *pour tout borélien A de $\mathbb{R}^d$ tel que $\mu(\partial A) = 0$, on a : $\lim_n \mu_n(A) = \mu(A)$.*

En particulier, pour vérifier que une suite de variables aléatoires $(X_n)_n$ converge en loi vers une mesure μ, il suffit d'avoir :

$$\forall A \in \mathcal{B}(\mathbb{R}^d),\ \mathbb{P}_{X_n}(A) \underset{n\to+\infty}{\longrightarrow} \mu(A)\,.$$

Un critère de convergence en loi des variables aléatoires réelles concerne leurs fonctions de répartition.

Proposition. *Pour tout $n \in \mathbb{N}$, on se donne une variable aléatoire réelle X_n définie sur un espace probabilisé $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ à valeurs dans $\mathbb{R}^d$, de fonction de répartition F_n et soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ de fonction de répartition F. La suite $(X_n)_n$ converge en loi vers X si et seulement si la suite $(F_n(x))_n$ converge vers $F_X(x)$ en tout point de continuité de F.*

Remarque. On ne peut pas s'attendre à la convergence simple sur tout $\mathbb{R}$ de la suite de fonctions de répartition F_n (cf. $X_n = \frac{1}{n}$ et $X = 0$).

Développement

Théorème (Lemme de Scheffé)**.** *Pour tout $n \in \mathbb{N}$, on se donne une variable aléatoire X_n définie sur un espace probabilisé $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ à valeurs dans $\mathbb{R}^d$ et admettant une densité f_n. Si la suite $(f_n)_n$ converge λ_d-presque partout vers une fonction f telle que $\int_{\mathbb{R}^d} f \, \mathrm{d}\lambda_d = 1$, alors la suite $(X_n)_n$ converge en loi vers la loi $f \cdot \lambda_d$. De plus :*

$$\sup_{A \in \mathcal{B}(\mathbb{R}^d)} \left| P_{X_n}(A) - \int_A f \, \mathrm{d}\lambda_d \right| \underset{n\to+\infty}{\longrightarrow} 0\,.$$

Démonstration. Soit A un borélien de $\mathbb{R}^d$. Alors :

$$|P_{X_n}(A) - f \cdot \lambda_d(A)| = \left| \int_A (f_n - f) \, \mathrm{d}\lambda_d \right| \leq \int_{\mathbb{R}^d} |f_n - f| \, \mathrm{d}\lambda_d\,.$$

d'où :

$$\sup_{A \in \mathcal{B}(\mathbb{R}^d} |P_{X_n}(A) - f \cdot \lambda_d(A)| \leq \int_{\mathbb{R}^d} |f_n - f| \, \mathrm{d}\lambda_d\,.$$

Puisque pour tout a, b réel on a : $|a - b| = a + b - 2\min(a, b)$, il s'ensuit :

$$\int_{\mathbb{R}^d} |f_n - f| \, \mathrm{d}\lambda_d = \underbrace{\int_{\mathbb{R}^d} f_n \, \mathrm{d}\lambda_d + \int_{\mathbb{R}^d} f \, \mathrm{d}\lambda_d}_{=2} - 2 \int_{\mathbb{R}^d} \min(f_n, f) \, \mathrm{d}\lambda_d\,.$$

Cependant, $\min(f_n, f)$ λ_d-presque partout vers f et :

$$0 \leq \min(f_n, f) \leq f\,,$$

puisque $\int_{\mathbb{R}^d} f \, \mathrm{d}\lambda_d = 1$ et $f \geq 0$ λ_d-presque partout, on a $f \in L^1(\mathbb{R}^d)$ et on peut appliquer le théorème de convergence dominée :

$$\int_{\mathbb{R}^d} |f_n - f| \, \mathrm{d}\lambda_d \underset{n\to+\infty}{\longrightarrow} 2 - 2 \int_{\mathbb{R}^d} f \, \mathrm{d}\lambda_d = 0\,.$$

□

Application. *Soit pour $n \geq 1$ une variable aléatoire réelle suivant une loi de Student à n degrés de liberté. Alors :*

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)\,.$$

Démonstration. On rappelle que X_n a pour densité :

$$f_n : x \in \mathbb{R} \longmapsto \frac{1}{\sqrt{n\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}\,.$$

Soient $t \in \mathbb{R}$ et $n \geq 1$. D'une part :

$$\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \underbrace{\left(\left(1+\frac{t^2}{n}\right)^n\right)^{-\frac{1}{2}}}_{\underset{n\to+\infty}{\longrightarrow} e^{t^2}} \underbrace{\left(1+\frac{t^2}{n}\right)}_{\underset{n\to+\infty}{\longrightarrow} 1} \underset{n\to+\infty}{\longrightarrow} e^{-\frac{t^2}{2}}\,.$$

et d'autre part, d'après la formule de Stirling généralisée (voir compléments) :

$$\Gamma\left(t+1\right) = \int_0^{+\infty} e^{-x}x^t\,\mathrm{d}x \underset{t\to+\infty}{\sim} t^t e^{-t}\sqrt{2\pi t}\,,$$

donc, pour $\alpha > 0$ fixé et lorsque t tend vers $+\infty$

$$\begin{aligned}
\Gamma(t+1+\alpha) &\sim (t+\alpha)^{t+\alpha}e^{-t-\alpha}\sqrt{2\pi(t+\alpha)} \\
&\sim t^\alpha t^t e^{-t}\sqrt{2\pi t}\left(1+\frac{\alpha}{t}\right)^{t+\alpha}e^{-\alpha}\sqrt{\frac{t+\alpha}{t}} \\
&\sim t^\alpha\Gamma(t)\underbrace{\left(1+\frac{\alpha}{t}\right)^t}_{\underset{t\to+\infty}{\longrightarrow} e^\alpha}\underbrace{\left(1+\frac{\alpha}{t}\right)^\alpha}_{\underset{t\to+\infty}{\longrightarrow} 1}e^{-\alpha} \\
&\sim t^\alpha\Gamma(t+1)\,.
\end{aligned}$$

Pour $\alpha = \frac{1}{n}$, $t+1 = \frac{n}{2}$ et en remarquant que $t^\alpha \sim (t+1)^\alpha$, on a lorsque n tend vers $+\infty$:

$$\frac{1}{\sqrt{n\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sim \frac{1}{\sqrt{2\pi}}\,.$$

On a donc convergence presque partout de f_n vers $f : x \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)$ qui est la densité d'une loi normale centrée réduite. Le lemme de Scheffé conclut. □

Remarque. Cette convergence en loi peut aussi se démontrer à l'aide de la loi faible des grands nombres et le lemme de Slutsky (il faut utiliser l'autre définition de la loi de Student).

Remarque. La réciproque du lemme de Scheffé est fausse. En effet, si on se donne pour tout $n \in \mathbb{N}^*$ une variable aléatoire réelle X_n définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et admettant pour densité f_n où

$$f_n : x \in \mathbb{R} \longmapsto \mathbb{1}_{]0,1]}(x)(1-\cos(2\pi nx))\,,$$

alors la suite $(f_n)_n$ ne converge pas λ-presque partout (elle diverge sur $]0\,,1[$ et converge ailleurs) et $(X_n)_n$ converge en loi vers $\mathbb{1}_{]0,1]}\cdot\lambda$ (loi uniforme sur $]0\,,1]$). En effet, la fonction de répartition de X_n vérifie :

$$F_{X_n} : x \in \mathbb{R} \longmapsto \begin{cases} 0 & \text{si } x < 0\,, \\ x - \frac{\sin(2\pi nx}{2\pi n} & \text{si } 0 < x \leq 1\,, \\ 1 & \text{si } x > 1\,, \end{cases}$$

et clairement, $(F_{X_n})_n$ converge simplement (en tout point de continuité, ici tout $\mathbb{R}$) vers la fonction de répartition de $\mathbb{1}_{]0,1]}\cdot\lambda$, ce qui montre la convergence en loi.

Compléments

La formule de Stirling généralisée se montre en appliquant la méthode de Laplace.

Proposition (Méthode de Laplace)**.** *Soient $I = [a\,, b[$ un intervalle de $\mathbb{R}$ ($a < b \leq +\infty$) et $\varphi : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. Soit $f : I \to \mathbb{C}$ mesurable telle que $e^{-t_0\varphi} f$ soit Lebesgue intégrable pour un certain réel t_0. On suppose*

(i) f continue en a et $f(a) \neq 0$,

(ii) $\varphi'(a) = 0$, $\varphi''(a) > 0$ et $\varphi' > 0$ sur $]a\,, b[$,

On pose pour $t \geq t_0$:

$$F(t) = \int_a^b e^{-t\varphi(x)} f(x)\,\mathrm{d}x\,.$$

alors :

$$F(t) \underset{t\to+\infty}{\sim} \sqrt{\frac{\pi}{2\varphi''(a)}} \cdot \frac{e^{-t\varphi(a)} f(a)}{\sqrt{t}}\,.$$

Application (Formule de Stirling généralisée)**.** *On a :*

$$\Gamma\,(t+1) = \int_0^{+\infty} e^{-x} x^t\,\mathrm{d}x \underset{t\to+\infty}{\sim} t^t e^{-t}\sqrt{2\pi t}\,,$$

Démonstration de la formule de Stirling. On pose $x = t(u+1)$, ce qui donne : $\forall t > 0$,

$$\Gamma(t+1) = \int_0^{+\infty} e^{-x} x^t\,\mathrm{d}x = t^{t+1}\int_{-1}^{+\infty} e^{-t(u+1)}(u+1)^t\,\mathrm{d}u = t^{t+1}\int_{-1}^{+\infty} e^{-t\varphi(u)}\,\mathrm{d}u\,,$$

où $\varphi(u) = 1+u-\ln(1+u)$ de classe $\mathcal{C}^2$, sur $\mathbb{R}_+$ notamment. On a : $\varphi'(0) = 0$, $\varphi''(0) = 1$ et $\varphi'(u) = \frac{u}{u+1} > 0$ pour tout $u > 0$. On applique ce qui la méthode de Laplace avec $f = 1$:

$$\int_0^{+\infty} e^{-t\varphi(u)}\,\mathrm{d}u \underset{t\to+\infty}{\sim} \sqrt{\frac{\pi}{2t}} e^{-t}\,,$$

Il vient ensuite, en posant $u = -u$:

$$\int_{-1}^0 e^{-t\varphi(u)}\,\mathrm{d}u = \int_0^1 e^{-t\psi(u)}\,\mathrm{d}u$$

où $\psi(u) = 1-u-\ln(1-u)$ est $\mathcal{C}^2$ sur $[0\,, 1[$ et vérifie : $\psi'(0) = 0$, $\psi''(0) = 1$ et $\psi'(u) = \frac{u}{1-u} > 0$ pour tout $0 < u < 1$. D'où :

$$\int_{-1}^0 e^{-t\varphi(u)}\,\mathrm{d}u \underset{t\to+\infty}{\sim} \sqrt{\frac{\pi}{2t}} e^{-t}\,.$$

On obtient la bonne formule en additionnant. □

Références

- Ouvrard, Probabilités 2, p.307 (lemme de Scheffé et contre-exemple de la réciproque)
- Rouvière, Petit guide du calcul différentiel, p.353 (formule de Stirling généralisée)

Leçons concernées : 262, 263

24 Les théorèmes de Sylow

Contexte

Soit G un groupe d'ordre $n = p^\alpha m$ où p est premier et $m \wedge p = 1$. Un *p-Sylow* de G est un sous-groupe de G de cardinal p^α. Ainsi, un sous-groupe S de G est un p-Sylow si et seulement si $|G/S| \wedge p = 1$.

On rappelle que si G agit sur un ensemble E alors pour tout $x \in E$ le *stabilisateur* de x est un sous-groupe de G et est défini par :

$$H_x = \{g \in G \mid g \cdot x = x\} .$$

De plus, il y a clairement une bijection entre G/H_x et $w(x)$ où $w(x)$ désigne l'*orbite* de x sous l'action de G. Les orbites forment une partition de E et si E est de cardinal fini alors on a la formule des classes :

$$|E| = \sum_{x \in \mathcal{O}} |w(x)| ,$$

où $\mathcal{O}$ est un système de représentant des orbites de E sous l'action de G.

Développement

Théorème (Sylow)**.** *Soit G un groupe d'ordre $n = p^\alpha m$ où p est premier et $m \wedge p = 1$. Alors :*

(i) G admet un p-Sylow,

(ii) si H est un p-sous-groupe de G alors il existe un p-Sylow qui contient H,

(iii) si n_p est le nombre de p-Sylow de G alors : $n_p \equiv 1 \pmod p$.

Remarque. L'assertion (ii) implique en particulier que les p-Sylow de G sont conjugués entre eux et les p-Sylow forment une orbite sous l'action de G par conjugaison sur l'ensemble de ses sous-groupes. En particulier, on a n_p divise n. Si on rajoute l'assertion (iii), alors on obtient que n_p divise m.

Lemme. *Soit G un groupe d'ordre $n = p^\alpha m$ où p est premier et $m \wedge p = 1$. Soit H un sous-groupe de G et S un p-Sylow de G. Alors il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H.*

Démonstration du lemme. G opère par translation à gauche sur G/S (qui n'est pas forcément un groupe). Pour $a \in G$, le stabilisateur de aS est aSa^{-1}. Par restriction, H opère aussi sur G/S et le stabiliteur de aS par cette action est $aSa^{-1} \cap H$. Comme c'est un sous-groupe de aSa^{-1} qui est un p-groupe, c'est aussi un p-groupe. Il reste à montrer qu'au moins un est un p-Sylow de H. Si ce n'était pas le cas alors pour tout $a \in G$, p diviserait $\left|H/(aSa^{-1} \cap H)\right| = |w(aS)|$ où $w(aS)$ désigne l'orbite de aS dans G/S sous l'action de H. D'après la formule des classes p diviserait $|G/S|$, ce qui contredit le fait que S est un p-Sylow de G. □

Démonstration des théorèmes de Sylow.

(i) On plonge G dans un groupe que l'on connaît un peu mieux. G agit fidèlement sur lui-même (par translation), on a donc la série d'injection :

$$G \hookrightarrow \mathfrak{S}(|G|) \hookrightarrow \mathfrak{S}_n \hookrightarrow \mathrm{GL}_n(\mathbb{F}_p) .$$

La dernière injection envoie une permutation σ sur la matrice de permutation P_σ définie par : $\forall i,\ P_\sigma(e_i) = e_{\sigma(i)}$. Pour utiliser le lemme, on veut expliciter un p-Sylow de $p^{\frac{n(n-1)}{2}}$. Le cardinal de $\mathrm{GL}_n(\mathbb{F}_p)$ se calcule en dénombrant le nombre de bases de $\mathbb{F}_p^n$. On initialise en choisissant un vecteur non nul ($p^n - 1$ choix) pour construire le $k+1$-ième vecteur, on choisit un vecteur qui n'appartient pas au $\mathbb{F}_p$-espace vectoriel engendré par les k autres ($p^n - p^k$ choix). Donc :

$$|\mathrm{GL}_n(\mathbb{F}_p)| = (p^n - 1)(p^n - p) \cdots (p^n - p^{n-1}) = p^{0+1+\cdots+(n-1)}(p^n - 1)(p^{n-1} - 1) \cdots (p - 1) .$$

Le deuxième terme n'est pas divisible par p. Un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$ est donc de cardinal $p^{\frac{n(n-1)}{2}}$. Ça tombe bien puisque que le cardinal du sous-groupe des matrices supérieures strictes (avec que des 1 sur la diagonale) est exactement $p^{\frac{n(n-1)}{2}}$. Le lemme montre que l'image de G dans $\mathrm{GL}_n(\mathbb{F}_p)$ possède un p-Sylow donc G aussi.

(ii) Soit S un p-Sylow de G. Le lemme montre qu'il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H mais comme H est un p-groupe, on a : $aSa^{-1} \cap H = H$. Donc $H \subset aSa^{-1}$ qui est un p-Sylow de G.

(iii) Soit E l'ensemble des p-Sylow de G. G opère transitivement (grâce à (ii)) par conjugaison sur E. Soit S un p-Sylow de G. S opère sur E par restriction. La formule des classes montre que :

$$|E| \equiv \left|E^S\right| \pmod{p},$$

où E^S est l'ensemble des p-Sylow fixes sous l'action de S. Montrons que $E^S = 1$. Clairement $S \in E^S$ et c'est le seul car si T est un autre p-Sylow normalisé par S alors on considère le sous-groupe N de G engendré par S et T. T et S sont des p-Sylow de N mais par hypothèse, T est distingué dans N donc c'est le seul p-Sylow de N (conséquence de (ii)) et $S = T$.

□

Remarque. Pour montrer que les groupes de cardinal n ne sont pas simples, on peut essayer de calculer leur nombre de p-Sylow pour p un diviseur premier de n. Par exemple, les groupes de cardinal 63 ou 255 ne sont pas simples.

Compléments

Proposition. *Soit G un groupe d'ordre $n = p^\alpha m$ où p est premier et $m \wedge p = 1$. Alors G possède des sous-groupes d'ordre p^i pour tout $1 \leq i \leq \alpha$.*

Démonstration. On se ramène au cas des p-groupes en considèrant un p-Sylow de G. On suppose donc que G est un p-groupe. G agit par conjugaison sur lui-même et la formule des classes montre que le centre de G n'est pas trivial (et même p divise le cardinal de $Z(G)$). Soit g un élément du centre et d'ordre p (considérer un élément non nul, son ordre divise p^α puis le mettre à un puissance adéquate). Le sous-groupe $\langle g \rangle$ engendré par g est distingué dans G. On peut raisonner par récurrence sur l'ordre de G. Il existe un sous-groupe d'ordre $p^{\beta-1}$ où $\beta \leq \alpha$ dans le groupe $G/\langle g \rangle$. Son image inverse par la surjection canonique est un sous-groupe d'ordre p^β. □

Références

– Perrin, Cours d'algèbre, p.18

Leçons concernées : 101, 103, 104, 121

25 Dénombrement des polynômes irréductibles de $\mathbb{F}_q[X]$

Contexte

Soit K un corps et $P \in K[X]$ un polynôme *irréductible*. Une extension $L \supset K$ est appelée *corps de rupture* de P sur K si L est une extension monogène $L = K(\alpha)$ avec $P(\alpha) = 0$. Il existe toujours un corps de rupture de P sur K et de plus, il est unique à isomorphisme près.

Remarque. Pour montrer l'existence, on peut considérer $K[X]/(P)$ qui est un corps car P est irréductible ($K[X]$ est principal donc (P) est maximal dans l'ensemble des idéaux de $K[X]$).

On ne suppose plus P irréductible. Une extension $L \supset K$ est appelée *corps de décomposition* de P sur K si :

(i) P est produit de facteurs de degré 1 dans $L[X]$ (P a toutes ses racines dans L),

(ii) le corps L est minimal pour cette propriété (les racines de P engendrent L).

De même que pour le corps de rupture, un tel corps existe toujours et est unique à isomorphisme près.

Soit K un corps fini alors pour un certain nombre premier p, $\mathbb{F}_p$ est le sous-corps premier de K. K peut-être vu comme un $\mathbb{F}_p$-espace vectoriel et il existe n tel que $|K| = p^n$. Réciproquement, pour p premier et $n \in \mathbb{N}$, il existe un corps à $q = p^n$ éléments (c'est le corps de décomposition de $X^q - X$ dans $\mathbb{F}_p$) et il est unique à isomorphisme près. On le note $\mathbb{F}_q$.

Théorème. *Le groupe multiplicatif $\mathbb{F}_q^\times$ est cyclique.*

Enfin, on rappelle que la *fonction de Möbius* $\mu : \mathbb{N}^* \to \{-1, 0, 1\}$ est définie par :

$$\forall n \in \mathbb{N}^*,\ \mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n \text{ contient un facteur carré} \\ (-1)^r & \text{si } n = p_1 \cdots p_r \text{ où les } p_i \text{ sont des nombres premiers distincts} \end{cases} .$$

On vérifie que μ est multiplicative : $\forall n, m \in \mathbb{N},\ n \wedge m = 1 \implies \mu(nm) = \mu(n)\mu(m)$.

Développement

Théorème. *On note $A(n, q)$ l'ensemble des polynômes irréductibles unitaires et de degré n de $\mathbb{F}_q[X]$ ainsi que $I(n, q)$ le cardinal de $A(n, q)$. Alors :*

$$I(n, q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \underset{n\to+\infty}{\sim} \frac{q^n}{n} .$$

Lemme (Formule d'inversion de Möbius)**.** *Soit $f : \mathbb{N}^* \to G$ où G est un groupe abélien. On pose pour tout $n \in \mathbb{N}^*$:*

$$g(n) = \sum_{d|n} f(d) ,$$

alors :

$$\forall n \in \mathbb{N}^*,\ f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) .$$

Démonstration du lemme. On a :

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{sinon} \end{cases} ,$$

puisque pour $n = p_1^{\alpha_i} \cdots p_r^{\alpha_r}$:

$$\sum_{d|n} \mu(d) = 1 + \sum_{i=1}^{r} \mu(p_i) + \sum_{1 \le i < j \le r} \mu(p_i p_j) + \cdots + p_1 \cdots p_r = \sum_{i=0}^{r} \binom{n}{i} (-1)^i .$$

Ensuite :

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \left(\sum_{d' | \frac{n}{d}} f(d') \right) = \sum_{dd'|n} \mu(d) f(d')$$
$$= \sum_{d'|n} f(d') \left(\sum_{d | \frac{n}{d'}} \mu(d) \right) = f(n)\,.$$

□

Démonstration du théorème. On montre d'abord que :

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in A(d,q)} P\,. \tag{🍃}$$

Soit $d \mid n$ et $P \in A(d,q)$. Montrons que $P \mid X^{q^n} - X$. Soit $K = \mathbb{F}_q(x)$ un corps de rupture de P (x est un racine de P). Puisque P est irréductible, c'est le polynôme minimal de x et $[K : \mathbb{F}_q] = \deg P = d$. Par unicité, K est isomorphe à $\mathbb{F}_{q^d}$. De plus, par le théorème de Lagrange, on a $x^{q^d} = x$. On en déduit, en itérant $\frac{n}{d}$ fois que $x^{q^n} = x$. Donc x est une racine de $X^{q^n} - X$ et puisque P est la polynôme minimal de x, on a $P \mid X^{q^n} - X$.

Réciproquement, si P est un facteur irréductible de $X^{q^n} - X$ alors $d = \deg P \mid n$. En effet, $X^{q^n} - X$ est scindé dans $\mathbb{F}_q[X]$. Soit x une racine de P alors $\mathbb{F}_q(x)$ est un corps intermédiaire entre $\mathbb{F}_q$ et $\mathbb{F}_{q^n}$. Donc :

$$n = [\mathbb{F}_{q^n} : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q(x)]\underbrace{[\mathbb{F}_q(x) : \mathbb{F}_q]}_{=d}\,.$$

Puisque les racines de $X^{q^n} - X$ sont simples, ses facteurs irréductibles n'apparaissent qu'une seule fois et on obtient (🍃).

En comparant les degrés dans (🍃), on a :

$$q^n = \sum_{d|n} d I(d,q)\,,$$

et on applique le lemme :

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d\,.$$

Ensuite, on écrit $I(n,q) = \frac{q^n + r_n}{n}$ où :

$$r_n = \sum_{d|n, d\neq n} \mu\left(\frac{n}{d}\right) q^d\,.$$

Brutalement, on a :

$$|r_n| \leq \sum_{d=1}^{\lfloor \frac{n}{2} \rfloor} q^d = q \frac{q^{\lfloor \frac{n}{2} \rfloor} - 1}{q-1} \leq \frac{q^{\lfloor \frac{n}{2} \rfloor + 1}}{q-1} = o(q^n)\ ,$$

d'où l'équivalent.

□

Références

– Francinou et Gianella, Exercices pour l'agrégation, Algèbre 1, p.189

Leçons concernées : 120, 122, 123, 125, 141, 190

26 Théorème des deux carrés

Contexte

On cherche à déterminer l'ensemble des entiers qui s'écrivent comme somme de deux carrés, c'est-à-dire :

$$\Sigma := \{a^2 + b^2 \mid a, b \in \mathbb{N}\}\,.$$

On définit l'ensemble des *entiers de Gauss* par :

$$\mathbb{Z}[i] := \{a + ib \mid a, b \in \mathbb{Z}\ \}\,.$$

Développement

Théorème (Deux carrés)**.** *Soit $n \in \mathbb{N} \setminus \{0, 1\}$. On écrit :*

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)}\,.$$

Alors : $n \in \Sigma \iff v_p(n)$ pair pour tout $p \equiv 3 \pmod 4$.

Démonstration. La clé de la démonstration est la stabilité de Σ par multiplication, ce qui permet de restreindre le problème aux nombres premiers.

(i) Dans un premier temps on étudie $\mathbb{Z}[i]$.
$\mathbb{Z}[i]$ est clairement un anneau intégre. De plus on peut le munir de la norme

$$N : z = a + ib \in \mathbb{Z}[i] \longmapsto z\overline{z} = a^2 + b^2\,,$$

qui est multiplicative. Cela montre que :

$$n \in \Sigma \iff \exists z \in \mathbb{Z}[i],\ N(z) = n\,,$$

et la multiplciativité de N montre que Σ est stable par multiplication.
L'ensemble des inversibles de $\mathbb{Z}[i]$ est $\{\pm 1, \pm i\}$. En effet, si $z \in \mathbb{Z}[i]^\times$ alors il existe z' tel que $zz' = 1$ donc $N(z)N(z') = 1$. Nécessairement, $N(z) = 1$. Réciproquement, tous les éléments de $\{\pm 1, \pm i\}$ sont inversibles.
Enfin $\mathbb{Z}[i]$ muni du stathme est euclidien. Soient $z, t \in \mathbb{Z}[i] \setminus \{0\}$. Il existe $x, y \in \mathbb{R}$ tels que : $z/t = x + iy$. Soient a (resp. b) l'entier le plus proche de x (resp. y). On pose $q = a + ib$ alors : $|z/t - q|^2 = (x - a)^2 + (y - b)^2 \leq 1/4 + 1/4 = 1/2 < 1$. On pose $r = z - qt \in \mathbb{Z}[i]$. On a $|r| < |t|\,|z/t - q| < |t|$ d'où $N(r) < N(t)$.

(ii) Soit p un nombre premier alors : $p \in \Sigma \iff p = 2$ ou $p \equiv 1 \pmod 4$.
On démontre d'abord que : $p \in \Sigma \iff p$ n'est pas irréductible dans $\mathbb{Z}[i]$ (en fait une seule implication est utile ici). Pour le sens direct, on décompose $p = a^2 + b^2 = (a + ib)(a - ib)$ où a et b sont différents de 0 (sinon p ne serait pas premier) de sorte que p n'est pas irréductible. Pour le sens indirect, si $p = zz'$ avec z et z' non inversibles dans $\mathbb{Z}[i]$ alors on aurait $p^2 = N(z)N(z')$. Comme $N(z) \neq 1$ et $N(z') \neq 1$, nécessairement $N(z) = N(z') = p$.
Comme $\mathbb{Z}[i]$ est euclidien donc principal donc factoriel, dire que p est non irréductible signifie que l'idéal engendré par p est non premier ou encore que $\mathbb{Z}[i]/(p)$ est non intègre. On a la série d'isomorphismes :

$$\mathbb{Z}[i]/(p) \simeq \mathbb{Z}[X]/(X^2 + 1, p) \simeq [\mathbb{Z}[X]/(p)]/(X^2 + 1) \simeq \mathbb{F}_p[X]/(X^2 + 1)\,, \qquad (✐)$$

puisqu'on peut identifier $\mathbb{Z}[i]$ à $\mathbb{Z}[X]/(X^2 + 1)$. Donc p non irréductible dans $\mathbb{Z}[i]$ équivaut à ce que $X^2 + 1$ soit non irréductible dans $\mathbb{F}_p[X]$, ou encore à ce que $X^2 + 1$ ait une racine dans $\mathbb{F}_p$. Il s'ensuit :

$$p \in \Sigma \iff -1 \in (\mathbb{F}_p^\times)^2\,.$$

Étudions les carrés de $\mathbb{F}_p$. Le morphisme surjectif

$$\begin{array}{ccc} \mathbb{F}_p^\times & \longrightarrow & (\mathbb{F}_p^\times)^2 \\ x & \longmapsto & x^2 \end{array}$$

a pour noyau $\{-1, 1\}$ si $p \neq 2$. On en déduit que $(\mathbb{F}_p^\times)^2$ a $\frac{p-1}{2}$ éléments. Comme $\mathbb{F}_p^\times$ est cyclique, si $x \in (\mathbb{F}_p^\times)^2$ alors $x^{\frac{p-1}{2}} = 1$ et par cardinalité :

$$(\mathbb{F}_p^\times)^2 = \{x \in \mathbb{F}_p^\times \mid x^{\frac{p-1}{2}} = 1\}.$$

Donc : $-1 \in (\mathbb{F}_p^\times)^2 \iff (-1)^{\frac{p-1}{2}} = 1 \iff \frac{p-1}{2}$ pair $\iff p \equiv 1 \pmod 4$ ou $p = 2$.

Remarque. – La description des carrés de $\mathbb{F}_p$ est encore vraie pour $\mathbb{F}_q$.

- On peut montrer facilement que la condition est nécessaire. En effet, si a est pair alors $a^2 \equiv 0 \pmod 4$ et si a impair alors $a^2 \equiv 1 \pmod 4$. Donc si $p = a^2 + b^2$ alors $p \equiv 0, 1, 2 \pmod 4$. Cela est en fait vérifier pour tout n.

(iii) Puisque un carré est toujours somme de deux carrés, par stabilité par multiplication, on a montré la sens indirect du théorème.

Démontrons le sens direct. Soient $n \in \Sigma$ et $p \equiv 3 \pmod 4$. Si $v_p(n) = 0$ alors c'est bon, sinon p divise $n = a^2 + b^2 = (a + ib)(a - ib)$ et comme p est irréductible dans $\mathbb{Z}[i]$, par le théorème de Gauss, p divise l'un des deux termes. Mais comme p est entier, nécessairement $p \mid a$ et $p \mid b$ donc $p^2 \mid a^2 + b^2 = n$. On conclut par récurrence sur $v_p(n)$ en remarquant que n/p^2 est toujours dans Σ. □

Compléments

On peut achever la détermination des irréductibles de $\mathbb{Z}[i]$.

Proposition. *Les irréductibles de $\mathbb{Z}[i]$ sont, aux inversibles près :*

(i) les entiers premiers $p \in \mathbb{N}$ avec $p \equiv 3 \pmod 4$,

(ii) les entiers de Gauss $a + ib$ tels que $a^2 + b^2$ est premier.

Démonstration. Ces éléments sont irréductibles (pour (ii) considérer la norme). Réciproquement, si $z \in \mathbb{Z}[i]$ est irréductible alors soit p premier divisant $N(z) = z\overline{z}$. Si $p \equiv 3 \pmod 4$ alors puisque p est irréductible p divise z ou $\overline{z}$. Donc p divise z et $p = z$ à un inversible près. Sinon $p \in \Sigma$ et il existe a et b tels que $p = a^2 + b^2$. On pose $t = a + ib$ qui est irréductible. Alors t divise z ou $\overline{z}$ et $z = t$ ou $\overline{t}$ à un inversible près. □

La série d'isomorphisme de (✐) est à savoir redémontrer. Il faut pas s'emméler les pinceaux !

Démonstration. □

Références

– Perrin, Cours d'algèbre, p.56 (théorème) et p.75 (des bricoles)

Leçons concernées : 120, 122, 126

27 Composantes connexes des formes quadratiques réelles

Contexte

Soit $(E, \|\cdot\|)$ un $\mathbb{R}$-espace vectoriel normé de dimension $n \geq 1$. On note $Q(E)$ (resp. $\Omega(E)$) l'ensemble des formes quadratiques (resp. des formes quadratiques non dégénérées) sur E. On munit $Q(E)$ de la norme

$$N : q \in Q(E) \longmapsto \sup_{\|x\|=1} |q(x)| \ .$$

Développement

Théorème. *Les composantes connexes de $\Omega(E)$ sont les sous-ensembles $\Omega_k(E)$ des formes quadratiques de signature $(k, n-k)$ où $k \in [\![0\,, n]\!]$.*

Lemme. $\mathrm{GL}_n(\mathbb{R})^+$ *est connexe par arcs.*

Démonstration du lemme. On peut utiliser les matrices du dilatation et de transvection (qui engendrent $\mathrm{GL}_n(\mathbb{R})$), cependant ce n'est pas très joli. Utilisons plutôt la décomposition polaire. Soit $A \in \mathrm{GL}_n(\mathbb{R})^+$, alors il existe $S \in \mathcal{S}_n^{++}(\mathbb{R})$ et $O \in \mathrm{SO}_n(\mathbb{R})$ tels que $A = OS$. $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe donc connexe par arcs. $\mathrm{SO}_n(\mathbb{R})$ est aussi connexe par arcs. Pour le montrer on peut utiliser la réduction des matrices de $\mathrm{SO}_n(\mathbb{R})$:

$$\forall O \in \mathrm{SO}_n(\mathbb{R}),\ \exists P \in \mathrm{O}_n(\mathbb{R}),\ \exists r \in \mathbb{N},\ \exists \theta_1, \ldots, \theta_r \in \mathbb{R} \mid O = {}^tP\,\mathrm{diag}(I_p, M(\theta_1), \ldots, M(\theta_r))P$$

où $p = n - 2r$, où les θ_i sont non nuls modulo 2π et où pour tout $\theta \in \mathbb{R}$ on a posé :

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

À partir de là, on peut facilement relier continûment A à I_n. □

Remarque. $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe car son image par l'application déterminant est égale à $\mathbb{R}^*$ qui n'est pas connexe.

Démonstration du theorème. Les $\Omega_k(E)$ forment une partition de $\Omega(E)$.

(i) Montrons d'abord que les $\Omega_k(E)$ sont ouverts dans $Q(E)$.
Soit $q \in \Omega_k(E)$. Soit $(e_1, \ldots, e_n)$ une base q-othogonale telle que pour tout $x = \sum x_i e_i$, on a :

$$q(x) = \sum_{i=1}^{n} \alpha_i x_i^2$$

où $\alpha_i > 0$ pour tout $i \in [\![1\,, k]\!]$ et $\alpha_i < 0$ pour tout $i \in [\![k+1\,, n]\!]$. On pose $F = \mathrm{vect}(e_1, \ldots, e_k)$ et $G = \mathrm{vect}(e_{k+1}, \ldots, e_n)$. La restriction de q à F est une forme quadratique définie positive. Par équivalence des normes en dimension finie, il existe $r_1 > 0$ tel que $\sqrt{q(x)} \geq r_1 \|x\|$ pour tout $x \in F$. De même, la restriction de $-q$ à G est définie positive et il existe r_2 tel que $\sqrt{-q(x)} \geq r_2 \|x\|$ pour tout $x \in G$. En prenant $r = \min(r_1^2, r_2^2)$, on a :

$$\forall x \in F,\ q(x) \geq r \|x\|^2 \quad \text{et} \quad \forall x \in G,\ q(x) \leq -r \|x\|^2 \ .$$

Soit $q' \in Q(E)$ tel que $N(q - q') < r$ alors : $\forall x \in E \setminus \{0\},\ |q(x) - q'(x)| \leq N(q - q') \|x\|^2 < r \|x\|^2$. Par inégalité triangulaire :

$$\forall x \in F \setminus \{0\},\ q'(x) > q(x) - r\|x\|^2 \geq 0 \quad \text{et} \quad \forall x \in G \setminus \{0\},\ q'(x) < q(x) + r\|x\|^2 \leq 0\,.$$

Le théorème d'inertie de Sylvester montre que la signature de q' est bien la même que celle de q.

(ii) On se donne une base de E et on note φ l'isomorphisme qui à une forme quadratique de E associe sa matrice dans cette base. Soient q (resp. q') dans $\Omega_k(E)$ et $A = \varphi(q)$ (resp. $B = \varphi(q')$). A et B sont symétriques et il existe P et Q dans $\mathrm{GL}_n(\mathbb{R})$ telles que :

$$A = {}^tPD_kP \quad \text{et} \quad B = {}^tQD_kQ$$

où D_k est une matrice diagonale contenant k coefficients 1 et $n-k$ coefficients -1. Comme

$$\begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} D_k \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = D_k \,,$$

on peut supposer que P et Q sont de déterminant positif. $\mathrm{GL}_n(\mathbb{R})^+$ étant connexe par arcs, il existe un chemin continu $R : t \in [0\,,1] \mapsto R(t) \in \mathrm{GL}_n(\mathbb{R})^+$ qui relie P à Q et le chemin continu $t \in [0\,,1] \mapsto \varphi^{-1}\left[{}^tR(t)D_kR(t)\right]$ relie q à q' tout en restant dans $\Omega_k(E)$.

$\square$

Références

– Francinou, Gianella et Nicolas, Algèbre 3, p.214 (théorème) et p.65 (SO_n est connexe par arcs)

Leçons concernées : 170, 171, 204

28 L'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$

Développement

Théorème. *L'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.*

Lemme. *Soient A et B deux matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ telles que $\exp A = \exp B$. Alors $A = B$.*

Démonstration du lemme. Si A et B sont diagonales alors le résultat est évident. Montrons que A et B sont codiagonalisables. Pour cela, il suffit de montrer que A et B commutent. On sait que B est un polynôme en $\exp B$ (et inversement d'ailleurs). En effet, on écrit $B = P \operatorname{diag}(d_1, \dots, d_n) P^{-1}$ avec P dans $\mathrm{GL}_n(\mathbb{R})$. Donc $\exp B = P \operatorname{diag}(e^{d_1}, \dots, e^{d_n}) P^{-1}$. Ensuite, on choisit le polynôme interpolateur de Lagrange $Q \in \mathbb{R}[X]$ qui envoie e^{d_k} sur d_k et on a bien $Q(\exp B) = B$. Comme A commute avec $\exp A = \exp B$, A commute avec B, ce qui conclut. □

Démonstration du théorème. On montre d'abord que l'exponentielle induit une bijection continue de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$ puis que la bijection réciproque est continue.

(i) Par continuité de la transposition, si $M \in \mathcal{M}_n(\mathbb{R})$ alors $\exp({}^tM) = {}^t(\exp M)$ donc l'exponentielle d'une matrice symétrique est symétrique. Si de plus M est symétrique, d'après la théorème spectral, il existe $P \in \mathrm{O}_n(\mathbb{R})$ et $D = \operatorname{diag}(d_1, \dots, d_n)$ tels que $M = {}^tPDP$. Par continuité de $A \mapsto {}^tPAP$, on a : $\exp M = {}^tP \operatorname{diag}(e^{d_1}, \dots, e^{d_n}) P \in \mathcal{S}_n^{++}(\mathbb{R})$.

Réciproquement, si $N \in \mathcal{S}_n^{++}(\mathbb{R})$ alors il existe $P \in \mathrm{O}_n(\mathbb{R})$ et $D = \operatorname{diag}(d'_1, \dots, d'_n)$ où $d'_1, \dots, d'_n > 0$ tels que : $N = {}^tPDP$. Ce qui précède montre que $N = \exp M$ pour $M = {}^tP \operatorname{diag}(\ln d'_1, \dots, \ln d'_n) P$. Et d'après le lemme, c'est la seule solution.

L'exponentielle étant définie par une série qui converge normalement, sa restriction à $\mathcal{S}_n(\mathbb{R})$ est bien continue.

(ii) Montrons que la récurrence est continue. On utilise la norme d'opérateur induite par la norme euclidienne. On rappelle que pour tout matrice symétrique M, on a :

$$\|M\| = \rho(M) = \max\{|\lambda|, \lambda \in \operatorname{Sp}(M)\}.$$

Soit $(N_k)_{k\geq 0}$ une suite de $\mathcal{S}_n^{++}(\mathbb{R})$ qui converge vers $N \in \mathcal{S}_n^{++}(\mathbb{R})$. Il existe une unique suite $(M_k)_{k\geq 0}$ et une unique matrice M dans $\mathcal{S}_n(\mathbb{R})$ telles que :

$$\forall k \geq 0,\ N_k = \exp M_k \quad \text{et} \quad N = \exp M.$$

Montrons que la suite $(M_k)_k$ est bornée. Comme la suite $(N_k)_k$ converge, elle est bornée par une constante $C > 0$. Donc, pour tout k les valeurs propres de N_k sont bornées par C et les valeurs propres de M_k sont majorées par $\ln C$. Il faut aussi minorer le spectre des M_k. La continuité de $A \in \mathrm{GL}_n(\mathbb{R}) \mapsto A^{-1}$ montre que $(N_k^{-1})_k$ converge vers N^{-1} et est donc bornée par une constante $c > 0$. On en déduit que toutes les valeurs propres des M_k sont minorées par $-\ln c$. Donc la suite $(M_k)_k$ est bornée par $\max(|\ln C|, |\ln c|)$.

La propriété de Bolzano-Weierstrass montre que $(M_k)_k$ admet des valeurs d'adhérences dans $\mathcal{S}_n(\mathbb{R})$ ($\mathcal{S}_n(\mathbb{R})$ est fermé). De plus, M est la seule. En effet, soit $\varphi : \mathbb{N} \to \mathbb{N}$ une extractrice et $M' \in \mathcal{S}_n(\mathbb{R})$ telle que $(M_{\varphi(k)})_k$ converge vers M'. Alors, par continuité de l'exponentielle, on a $\exp M' = N = \exp M$ et le lemme montre que $M = M'$. $(M_k)_k$ admet une unique valeur d'adhérence donc est convergente et converge vers M, ce qui montre que la réciproque est continue.

□

Compléménts

Il y a deux résultats intermédiaires et admis dans cette preuve, à savoir le théorème de codiagonalisation et la continuité de l'inverse dans une algèbre de Banach.

Théorème (Codiagonalisation)**.** *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{R}$). Soient u et v deux endomorphismes diagonalisables de E qui commutent alors u et v sont codiagonalisables.*

Démonstration. On remarque d'abord que la restriction d'un endomorphisme diagonalisable à un sous-espace stable est encore diagonalisable (caractérisation par un polynôme annulateur scindé à racines simples). Ensuite les sous-espaces propres de u sont stables par v. En effet, si $u(x) = \lambda x$ alors :

$$u(v(x)) = v(u(x)) = \lambda v(x)\,.$$

On conclut en recollant les bases obtenues. □

Remarque. Ce résultat est en fait vrai pour une famille quelconque d'endomorphismes commutants deux à deux. Pour le montrer, on peut raisonner par récurrence forte en excluant le cas où tous les endomorphismes sont des homothéties.

La continuité de l'inverse peut se montrer avec la formule explicite faisant intervenir la comatrice et le déterminant. Cependant, on a le résultat plus général suivant :

Théorème. *Soit A une $\mathbb{R}$-algèbre de Banach. On note G le groupe des inversibles. Alors :*

(i) si $\|x\| < 1$ alors $1 - x$ est inversible d'inverse $\sum_{k\geq 0} x^k$,

(ii) G est ouvert dans A,

(iii) l'application $x \in G \mapsto x^{-1}$ est continue.

Démonstration.

(i) La série $\sum_{k\geq 0} x^k$ converge absolument dans un espace de Banach donc converge. De plus, pour tout n on a :

$$(1-x)\sum_{k=0}^{n} x^k = \sum_{k=0}^{n}(1-x) = 1 - x^{n+1} \underset{n\to+\infty}{\longrightarrow} 1\,,$$

car $\left\|x^{n+1}\right\| \leq \|x\|^{n+1}$, d'où le résultat.

(ii) Soit $x \in G$. Soit $y \in G$ tel que $\|y\| < \left\|x^{-1}\right\|^{-1}$. Alors

$$\left\|x^{-1}y\right\| \leq \left\|x^{-1}\right\| \|y\| < 1\,.$$

On écrit ensuite $x - y = x(1 - x^{-1}y)$ et on utilise (i) pour voir que $x - y$ est dans G.

(iii) Soit $x \in G$. Soit $h \in G$ tel que $\|h\| < \left\|x^{-1}\right\|^{-1}$. Alors :

$$(x+h)^{-1} - x^{-1} = \left[(1 - (-x^{-1}h))^{-1} - 1\right] x^{-1} = \left(\sum_{k\geq 1}(-x^{-1}h)^k\right) x^{-1}\,.$$

On en déduit :

$$\left\|(x+h)^{-1} - x^{-1}\right\| \leq \|h\| \left\|x^{-1}\right\|^2 \sum_{k\geq 0}\left(\|h\| \left\|x^{-1}\right\|\right)^k \leq \|h\| \frac{\left\|x^{-1}\right\|^2}{1 - \|h\| \left\|x^{-1}\right\|} = O(\|h\|)\,.$$

□

Références

– Francinou, Gianella et Nicolas, Algèbre 2, p.243

Leçons concernées : 155, 156, 158, 160

29 Théorème de la base de Burnside

Contexte

Soit G un groupe. Un sous-groupe propre H de G est dit *maximal* si pour tout sous-groupe propre H' de G tel que $H \subset H'$ on a $H = H'$. On note $F(G)$ le *sous-groupe de Frattini* de G défini comme l'intersection de tous les sous-groupes maximaux de G.

On suppose G fini. On dit que $g \in G$ est *mou* si pour toute famille génératrice $(g, g_1, \ldots, g_r)$ contenant g de G, $(g_1, \ldots, d_r)$ est encore une famille génératrice de G.

Développement

Théorème (de la base de Burnside)**.** *Soit G un p-groupe. Toutes les parties génératrices minimales de G ont même cardinal.*

Lemme. *Soit G un groupe fini. Le groupe de Frattini de G est égal à l'ensemble des éléments mous de G.*

Démonstration du lemme. On note F l'ensemble des éléments mous de G.

Soit $g \in F(G)$ et $(g, g_1, \ldots, g_r)$ une famille génératrice de G. Supposons que le sous-groupe $H = \langle g_1, \ldots, g_r \rangle$ engendré par la famille $(g_1, \ldots, g_r)$ est propre. Il existe un sous-groupe maximal H' de G tel que $H \subset H'$. Or, par définition, $F(G)$ est inclus dans H' et $g \in H'$. Donc $H' = G$, ce qui est absurde donc $g \in F$.

Soit $g \in F$ et H un sous-groupe maximal de G. Supposons que $g \notin H$ et notons $(g_1, \ldots, g_r)$ une famille génératrice de H. Si $\langle g, g_1, \ldots, g_r \rangle$ était propre alors H ne serait pas maximal donc $\langle g, g_1, \ldots, g_r \rangle = G$ et comme g est mou on aurait $H = G$. Cela est absurde donc $g \in H$ et $g \in F(G)$. □

Lemme. *Le centre $Z(G)$ d'un p-groupe G est non trivial.*

Démonstration du lemme. On fait agir G sur lui-même par conjugaison. La formule des classes donne :

$$|G| = |Z(G)| + \sum_{g \in \mathcal{O}} |w(g)|$$

où $\mathcal{O}$ est un système de représentant des orbites non réduites à un seul élément. Si $g \in \mathcal{O}$ il y a une bijection entre $w(g)$ et G/H_g où H_g est le stabilisateur de g. Comme $H_g \neq G$, on a $p \mid |G/H_g| = |w(g)|$ et donc :

$$p \mid |Z(G)| \, .$$

Comme $0 \in \mathbb{Z}(G)$, on a $|Z(G)| \geq p$. En particulier, $Z(G)$ n'est pas trivial. □

Lemme. *Les sous-groupes maximaux d'un p-groupe sont d'indice p.*

Démonstration du lemme. On raisonne par récurrence sur le cardinal du groupe. Si $|G| = p$ alors le seul sous-groupe maximal est le sous-groupe trivial qui est bien d'indice p. Fixons $\alpha \geq 1$. Supposons que pour tout p-groupe G de cardinal inférieur à p^α, tous les sous-groupes maximaux de G soient d'indice p. Soit G un p-groupe de cardinal $p^{\alpha+1}$. Soit H un sous-groupe maximal de G. Considérons $H' = Z(G) \cap H$. H' est un sous-groupe distingué de G et de H.

Si H' est non trivial alors G/H' est un p-groupe de cardinal p^β avec $\beta \leq \alpha$. Soit $\pi : G \longmapsto G/H'$ la projection canonique. Si $\pi(H)$ n'était pas maximal, il existerait un sous-groupe propre $\widetilde{H}$ de G/H' tel que

$$\pi(H) \subsetneq \widetilde{H} \subsetneq G/H' \, .$$

Mézalor, par maximalité de H dans G, on aurait $\pi^{-1}(\widetilde{H}) = G$ et $\widetilde{H} = \pi(\pi^{-1}(\widetilde{H})) = G/H'$ ce qui est absurde. Donc $\pi(H)$ est maximal. Par hypothèse de récurrence, $\pi(H)$ est d'indice p dans G/H'. On a :

$$[\pi(G) : \pi(H)] = \frac{|G/H'|}{|H/H'|} = \frac{|G|}{|H|} = [G : H] \, .$$

Si H' est trivial alors, d'après le lemme précédant, il existe un élément g d'ordre p dans $Z(G)$ mais qui n'est pas dans H (prendre un élément non trivial du centre de G et le mettre à une puissance adéquate). On considère le groupe engendré par H et g qui est de cardinal $p\,|H|$ puisque $g \in Z(G)$ et qui contient strictement H. Par maximalité, c'est G et par cardinalité, H est d'indice p dans G. □

Lemme. *Tout sous-groupe maximal d'un p-groupe G est distingué. En particulier, $F(G)$ est distingué.*

Démonstration du lemme. Soit H un sous-groupe maximal de G. D'après le lemme, H est d'indice p. G agit par translation à gauche sur $G/H \simeq \mathbb{Z}/p\mathbb{Z}$. On note π la projection canonique de G sur G/H.

On a donc un morphisme de groupes, noté φ, de G dans $\mathfrak{S}_{G/H} \simeq \mathfrak{S}_p$. Soit S l'image de G. D'après le théorème de Lagrange, on a :

$$|S| \mid p! \quad \text{et} \quad |G| = |S| \times |\mathrm{Ker}(\varphi)| \ .$$

Comme φ est non trivial, nécessairement, $|S| = p$. De plus, $H \subset \mathrm{Ker}(\varphi)$ et par cardinalité : $H = \mathrm{Ker}\,\varphi$. Donc H est distingué dans G. Une intersection de sous-groupes distingués étant distinguée, $F(G)$ est distingué. □

Démonstration du théorème. Soit H un sous-groupe maximal de G.

Soit $g \in G$. Comme $G/H \simeq \mathbb{Z}/p\mathbb{Z}$, l'image de g^p par la projection canonique est H. Donc $g^p \in H$. On en déduit que : $\forall g \in G,\ g^p \in F(G)$.

Comme $\mathbb{Z}/p\mathbb{Z}$ est commutatif, π envoie le groupe dérivé $D(G)$ de G sur H. Et de même : $D(G) \subset F(G)$. Cela prouve que $G/F(G)$ est abélien.

Ces deux points et le fait que $F(G)$ soit distingué montrent qu'on peut munir $G/F(G)$ d'une structure de $\mathbb{Z}/p\mathbb{Z}$-espace vectoriel. Montrons que sa dimension s est égale à la taille de toute famille génératrice minimale de G. Soit $(g_1, \ldots, g_r)$ une telle famille. L'image $(\overline{g}_1, \ldots, \overline{g}_r)$ de cette famille par la projection canonique de G sur $G/F(G)$ engendre $G/F(G)$, donc $s \leq r$. Maintenant, si on extrait de $(\overline{g}_1, \ldots, \overline{g}_r)$ une base $(\overline{g}'_1, \ldots, \overline{g}'_s)$ de $G/F(G)$, alors on peut compléter $(g'_1, \ldots, g'_r)$ en une famille génératrice $(g'_1, \ldots, g'_r, a_1, \ldots, a_t)$ de G où les a_i sont dans $F(G)$. D'après le premier lemme, les a_i sont mous et $(g'_1, \ldots, g'_r)$ est une famille génératrice de G. Nécessairement $r = s$ et l'unicité de la dimension montre le théorème. □

Remarque. Ce théorème n'est plus vrai si G n'est pas un p-groupe. Par exemple, si $G = \mathfrak{S}_n$ alors $\{(1\ 2),\ (1\ 2\ \cdots\ n)\}$ et $\{(1\ i),\ i = 2, \ldots, n\}$ sont des familles génératrices minimales de cardinal différent dès que $n \geq 4$.

Compléments

Dans la même veine, on a le résultat suivant (qui se généralise facilement aux sous-groupe de Sylow) :

Proposition. *Soit G un p-groupe de cardinal p^α. Alors, pour tout $0 \leq \beta \leq \alpha$, G possède un groupe d'ordre p^β.*

Démonstration du lemme. Montrons qu'il existe toujours un élément du centre d'ordre p. Soit $g_0 \neq 1$ un élément du centre. g_0 est d'ordre p^r pour un certain $r \geq 1$ et si on pose $g = g_0^{p^{r-1}}$ alors g est un élément du centre d'ordre p.

Maintenant, on raisonne par récurrence sur le cardinal de G. Si $|G| = p$ alors c'est évident. Supposons que tout groupe de cardinal p^α possède des sous-groupes de cardinal p^β pour tout $0 \leq \beta \leq \alpha$. Soit G un groupe de cardinal $p^{\alpha+1}$ et $g \in Z(G)$ d'ordre p. Le sous-groupe engendré par g est distingué et de cardinal p. Donc $G/\langle g \rangle$ est un groupe de cardinal p^α. Pour $0 \leq \beta \leq \alpha$ il existe un sous-groupe H_β de $G/\langle g \rangle$ de cardinal p^β. Soit π la projection canonique de G sur $G/\langle g \rangle$. Alors, $\pi^{-1}(H_\beta)$ est un sous-groupe de G (évident) de cardinal $p^{\beta+1}$. En effet, si $(h_1, \ldots, h_{p^\beta})$ est un système de représentants de H_β alors $\pi^{-1}(H_\beta) = \{h_i g^j,\ i = 1, \ldots, p^\beta,\ j = 1, \ldots, p\}$. □

Références

– Zavidovique, Un max de maths, p.13

Leçons concernées : 101, 104, 108, 151

30 Théorème de Frobenius-Zolotarev

Contexte

Soit p un nombre premier ≥ 3. Pour $x \in \mathbb{F}_p^\times$, on définit le *symbole de Legendre* de x dans $\mathbb{F}_p^\times$ par :

$$\left(\frac{x}{p}\right) = \begin{cases} 1 & \text{si } x \in (\mathbb{F}_p^\times)^2 \\ -1 & \text{si } x \notin (\mathbb{F}_p^\times)^2 \end{cases} .$$

Pour tout $x \in \mathbb{F}_p^\times$, on a : $\left(\frac{x}{p}\right) \equiv x^{\frac{p-1}{2}} \pmod p$. On en déduit que l'application $x \mapsto \left(\frac{x}{p}\right)$ est un morphisme de $\mathbb{F}_p^\times$ vers $\{-1, 1\}$.

Développement

Théorème (Frobenius-Zolotarev)**.** *Soit p un nombre premier ≥ 3. Soit V un $\mathbb{F}_p$-espace vectoriel de dimension finie ≥ 3. Soit $u \in \mathrm{GL}(V)$. Alors :*

$$\varepsilon(u) = \left(\frac{\det(u)}{p}\right) ,$$

où $\varepsilon(u)$ est la signature de u vu comme élément de $\mathfrak{S}(V)$.

Lemme. *Soit k un corps. Si $n \geq 3$, alors : $D(\mathrm{GL}_n(k)) = \mathrm{SL}_n(k)$.*

Démonstration du lemme. Tout commutateur est dans $\mathrm{SL}_n(k)$. Réciproquement, $\mathrm{SL}_n(k)$ est engendré par les transvections. Soit $i \neq j$ et $\lambda \in k$. Alors, si $\ell \notin \{i, j\}$ (ce qui est possible car $n \geq 3$) :

$$I_n + \lambda E_{ij} = (I_n + \lambda E_{i\ell})(I_n + \lambda E_{\ell j})(I_n - \lambda E_{i\ell})(I_n - \lambda E_{\ell j})$$

(pour le vérifier, on peut faire les calculs ou interpréter en termes d'opérations sur les lignes ou les colonnes). Et comme $(I_n - \lambda E_{i\ell}) = (I_n + \lambda E_{i\ell})^{-1}$ et $(I_n - E_{\ell j}) = (I_n + E_{\ell j})^{-1}$, le lemme est vérifié. $\square$

Lemme. *Soit k un corps et A un groupe abélien. Si $n \geq 3$, alors pour tout morphisme de groupes $\varphi : \mathrm{GL}_n(k) \to A$, il existe un unique morphisme de groupes $\delta : k^\times \to A$ tel que :*

$$\varphi = \delta \circ \det .$$

Démonstration du lemme. Comme $D(\mathrm{GL}_n(k))$ est dans le noyau de φ, on peut factoriser φ :

$$\begin{tikzcd} \mathrm{GL}_n(k) \arrow[r, "\varphi"] \arrow[d, two heads, "\pi"'] & A \\ \mathrm{GL}_n(k)/\mathrm{SL}_n(k) \arrow[ur, "\overline{\varphi}"'] & \end{tikzcd}$$

et comme $\mathrm{SL}_n(k) = \mathrm{Ker}(\det)$, on a aussi la factorisation :

$$\begin{tikzcd} \mathrm{GL}_n(k) \arrow[r, two heads, "\det"] \arrow[d, two heads, "\pi"'] & k^\times \\ \mathrm{GL}_n(k)/\mathrm{SL}_n(k) \arrow[ur, hook, two heads, "\overline{\det}"'] & \end{tikzcd}$$

On en déduit : $\varphi = \overline{\varphi} \circ \left(\overline{\det}\right)^{-1} \circ \overline{\det} \circ \pi = \delta \circ \det$ où $\delta = \overline{\varphi} \circ \left(\overline{\det}\right)^{-1}$. La surjectivité du déterminant assure l'unicité de δ. $\square$

Lemme. *Soit p un nombre premier ≥ 3. Le symbole de Legendre est l'unique morphisme non trivial de $\mathbb{F}_p^\times$ dans $\{-1, 1\}$.*

Démonstration du lemme. Le symbole de Legendre est non trivial puisqu'il existe des éléments de $\mathbb{F}_p^\times$ qui ne sont pas des carrés (il y en a $\frac{p-1}{2}$). Réciproquement, soit $\alpha : \mathbb{F}_p^\times \to \{-1, 1\}$ un morphisme non trivial. Alors $\mathrm{Ker}(\alpha)$ est un sous-groupe d'indice 2 de $\mathbb{F}_p^\times$.

$\mathbb{F}_p^\times$ est cyclique et de cardinal pair donc il existe un unique sous-groupe H d'indice 2. En effet, si $x \in \mathbb{F}_p^\times$ est un générateur alors le sous-groupe généré par x^2 est d'indice 2. Réciproquement, soit H un sous-groupe d'indice 2 de $\mathbb{F}_p^\times$. Soit $r = \min\{k \in \mathbb{N}^\times \mid x^k \in H\}$. Par division euclidienne on peut montrer que g^r engendre H et si $r \neq 2$ alors H n'est pas d'indice 2.

Si $x \notin H$ alors $\mathbb{F}_p^\times = H \cup xH$. Donc α est entièrement déterminé par H, il y a au plus un morphisme non trivial de $\mathbb{F}_p^\times$ dans $\{-1, 1\}$. □

Démonstration du théorème. Soit $\varepsilon : \mathrm{GL}(V) \to \{-1, 1\}$ obtenu en composant l'inclusion $\mathrm{GL}(V) \subset \mathfrak{S}(V)$ et la signature. D'après le deuxième lemme, il existe un unique morphisme $\delta : \mathbb{F}_p^\times \to \{-1, 1\}$ tel que : $\varepsilon = \delta \circ \det$. Montrons que δ est non trivial, ce qui conclura. Soit $d = \dim V$ et posons $q = p^d$. Vus comme $\mathbb{F}_p$-espaces vectoriels, V et $\mathbb{F}_q$ sont isomorphes. Il faut donc trouver une bijection $\mathbb{F}_p$-linéaire de $\mathbb{F}_q$ de signature -1. $\mathbb{F}_q^\times$ est cyclique d'ordre $q-1$. Soit g un générateur. La bijection $x \mapsto gx$ de $\mathbb{F}_q^\times$ est $\mathbb{F}_p$-linéaire, fixe 0 et agit comme le cycle $(g, g^2, \ldots, g^{q-1})$ de longueur $q-1$ donc de signature $(-1)^q = -1$. □

Références

- Beck, Malick et Peyré, Objectif Agrégation, p.251 (théorème)
- Gourdon, Algèbre, p.156 (groupe dérivé de $\mathrm{GL}_n(\mathbb{F}_p)$)

Leçons concernées : 103, 105, 106, 152

31 Théorème de Chevalley-Warning

Contexte

Pour du blabla sur les corps finis, voir 26.

Développement

Théorème (Chevalley-Warning)**.** *Soit $q = p^n$. Soit $(f_a)_{a\in A}$ une famille de polynômes de $\mathbb{F}_q[X_1,\dots,X_m]$ (A ensemble fini). On suppose que :*

$$\sum_{a\in A} \deg(f_a) < m\,.$$

On note $V \subset \mathbb{F}_q^m$ l'ensemble des points d'annulation des f_a. Alors :

$$|V| \equiv 0 \pmod p\,.$$

Démonstration. Pour $u \geq 0$, on pose : $S(X^u) = \sum_{x\in\mathbb{F}_q} x^u$. Si $u = 0$ alors $S(X^0) = q$. Si u n'est pas divisible par $q-1$ alors, comme $\mathbb{F}_q^\times$ est cyclique de cardinal $q-1$, il existe $y \in \mathbb{F}_q^\times$ tel que $y^u \neq 1$. D'où :

$$S(X^u) = \sum_{x\in\mathbb{F}_q^\times} x^u = \sum_{x\in\mathbb{F}_q^\times} (yx)^u = y^u S(X^u)\,.$$

On en déduit : $S(X^u) = 0$.

Soit $P = \prod_{a\in A}(1 - f_a^{q-1}) \in \mathbb{F}_q[X_1,\dots,X_m]$. Clairement : $\forall x \in V,\ P(x) = 1$. Réciproquement, s'il existe $a \in A$ tel que $f_a(x) \neq 0$ alors $f_a(x)^{q-1} = 1$ et $P(x) = 0$. Ainsi P est la fonction caractéristique de V. Pour tout polynôme f, on introduit la fonction $S(f) = \sum_{x\in\mathbb{F}_q^m} f(x)$. Alors :

$$|V| \equiv S(P) \pmod p\,.$$

Par hypothèse, $\sum_{a\in A}\deg(f_a) < m$ donc $\deg P < m(q-1)$. Donc P est une combinaison linéaire de monômes de la formes $X^u = X_1^{u_1}\cdots X_m^{u_m}$ où $|u| = \sum_i u_i < m(q-1)$. Il suffit de montrer que $S(X^u) \equiv 0 \pmod p$ pour conclure. D'après le principe des tiroirs, il existe i tel que $u_i < q-1$. D'après ce qui précède, on a $S(X^{u_i}) = 0$. Mézalor :

$$\begin{aligned} S(X^u) &= \sum_{(x_1,\dots,x_m)\in\mathbb{F}_q^m} x_1^{u_1}\cdots x_m^{u_m} = \left(\sum_{x_i\in\mathbb{F}_q} x_i^{u_i}\right)\left(\sum_{(x_1,\dots,\hat{x}_i,\dots,x_m)\in\mathbb{F}_q^{m-1}} x_1^{u_1}\cdots\hat{x}_i^{u_i}\cdots x_m^{u_m}\right) \\ &= S(X^{u_i})S(X_1^{u_1}\cdots\widehat{X}_i^{u_i}\cdots X_m^{u_m}) = 0\,, \end{aligned}$$

Le chapeau signifie que l'on omet le terme. □

Application. *Soit p un nombre premier. Si $a_1,\dots,a_{2p-1}$ sont $2p-1$ entiers alors on peut en trouver p dont la somme est divisible par p.*

Démonstration. Pour un entier relatif $a \in \mathbb{Z}$, on note $\overline{a}$ sa classe module $\mathbb{Z}/p\mathbb{Z}$. Considérons les polynômes de $\mathbb{F}_p[X_1,\dots,X_{2p-1}]$ suivants :

$$\begin{aligned} P_1(X_1,\dots,X_{2p-1}) &= \sum_{i=1}^{2p-1} X_i^{p-1}\,, \\ P_2(X_1,\dots,X_{2p-1}) &= \sum_{i=1}^{2p-1} \overline{a_i} X_i^{p-1}\,. \end{aligned}$$

P_1 et P_2 ont au moins une racine commune, à savoir $(0,\dots,0)$. D'après le théorème de Chevalley-Warning, ils en ont une autre que l'on note $(x_1,\dots,x_{2p-1}) \neq (0,\dots,0)$. Comme :

$$\forall x \in \mathbb{F}_p,\ x^{p-1} = \begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{sinon} \end{cases}\,,$$

on déduit de l'égalité $P_1(x_1,\dots,x_{2p-1}) = 0$ que parmi les x_i il y en a exactement p qui sont non nuls. Et l'égalité $P_2(x_1,\dots,x_{2p-1}) = 0$ montre le résultat annoncé. □

Compléments

En fait, l'application est vraie pour tout entier n.

Théorème (Erdös-Ginzburg-Ziv)**.** *Soit n un entier. Si $a_1, \ldots, a_{2n-1}$ sont $2n-1$ entiers alors on peut en trouver n dont la somme est divisible par n.*

Démonstration. On raisonne par récurrence forte sur l'entier n. L'initialisation est évidente et le cas n premier a été traité. Si $n = pm$ où p est premier. On construit alors par récurrence des ensembles $E_0, E_1, \ldots, E_{2m-1}$ disjoints de $[\![1\,, 2n-1]\!]$ de la manière suivante : $E_0 = \emptyset$ et E_{k+1} est l'ensemble donné par l'application du corollaire du théorème de Chevalley-Warning aux $2p-1$ premiers (pour l'ordre induit par l'indexation) éléments de $[\![1\,, 2n-1]\!] \setminus (\cup_{i=0}^{k} E_i)$. L'égalité $2n-1 = (2m-1)p + p - 1$ montre que l'on peut en construire au plus $2m-1$ (sans compter E_0). On note pour $k \in [\![1\,, 2m-1]\!]$:

$$S_k = \sum_{i \in E_k} a_i \,,$$

et il existe S'_k tel que $S_k = pS'_k$. On applique l'hypothèse de récurrence aux S'_k et il existe $k_1, \ldots, k_m$ indices tels que :

$$m \mid \sum_{i=1}^{m} S'_{k_i} \,.$$

Il suffit ensuite de considére la réunion des E_{k_i} pour conclure. □

Remarque. Ce résultat est optimal puisqu'il n'est plus vrai si on ne dispose que de $2n-2$ entiers. Pour le voir, il suffit de considérer la suite formée de $n-1$ fois le terme 0 et de $n-1$ fois le terme 1.

Références

– Zavidovique, Un max de maths, p.32

Leçons concernées : 121, 123, 126, 142

32 Décomposition de Dunford

Contexte

Soit k un corps commutatif et E un k-espace vectoriel de dimension finie. Si $f \in \mathcal{L}(E)$, on définit son *polynôme caractéristique* par $\chi_f = \det(A - XI_n) \in k[X]$ où A est la matrice de f dans une base quelconque (χ_f ne dépend pas du choix de la base). D'après le théorème de Cayley-Hamilton, χ_f est un polynôme annulateur de f.

Développement

Théorème (Décomposition de Dunford). *Soit $f \in \mathcal{L}(E)$ tel que χ_f soit scindé sur k. Il existe un unique couple (d, n) d'endomorphismes tel que :*

(i) d est diagonalisable, n est nilpotente,

(ii) $f = d + n$,

(iii) d et n commutent.

De plus, d et n sont des polynômes en f.

Lemme. *Soit $f \in \mathcal{L}(E)$ et $F \in k[X]$ un polynôme annulateur de f. Soit $F = \beta M_1^{\alpha_1} \cdots M_s^{\alpha_s}$ la décomposition en facteurs irréductibles de F sur $k[X]$. Pour tout i, on note $N_i = \operatorname{Ker} M_i^{\alpha_i}(f)$. Alors $E = N_1 \oplus \cdots \oplus N_s$ et, pour tout i, la projection sur N_i parallèlement à $\oplus_{j \neq i} N_j$ est un polynôme en f.*

Démonstration du lemme. Sans perte de généralité, on peut supposer $\beta = 1$. D'après le lemme des noyaux : $E = N_1 \oplus \cdots \oplus N_s$.

Pour tout i, on note $Q_i = \prod_{j \neq i} M_j^{\alpha_j}$. Comme aucun facteur n'est commum à tous les Q_i, ils sont premiers entre eux dans leur ensemble. L'égalité de Bézout montre qu'il existe $U_1, \ldots, U_s \in k[X]$ tels que :

$$U_1 Q_1 + \cdots + U_s Q_s = 1\,.$$

On pose $P_i = U_i Q_i$ et $p_i = P_i(f)$. On a :

$$\sum_i p_i = \mathrm{id}_E\ .$$

De plus, comme $F \mid Q_i Q_j$ si $i \neq j$, f annule $P_i P_j$ et : $\forall i \neq j,\ p_i \circ p_j = 0$. En combinant les deux égalités précédentes, on en déduit que :

$$\forall i,\ p_i = p_i \circ \sum_j p_j = p_i^2\,.$$

Les p_i sont donc des projecteurs et, par construction, des polynômes en f. Pour les caractériser, il ne reste plus qu'à déterminer leur image et leur noyau.

Pour tout i, $\operatorname{Im} p_i = N_i$. En effet, si $y \in \operatorname{Im} p_i$ alors il existe x tel que $y = U_i(f) \circ Q_i(f)(x)$ et :

$$M_i^{\alpha_i}(f)(y) = U_i(f) \circ F(f)(x) = 0\,.$$

Donc $\operatorname{Im} p_i \subset N_i$. Réciproquement, si $y \in N_i$ alors :

$$y = \sum_j p_j(y) = p_i(y) + \sum_{j \neq i} U_j(f) \circ Q_j(f)(x) = p_i(y)\,,$$

puisque $M_i^{\alpha_i}$ divise Q_j si $i \neq j$. Donc $\operatorname{Im} p_i = N_i$.

Pour tout i, $\operatorname{Ker} p_i = \oplus_{j \neq i} N_j$. En effet, si $x \in \operatorname{Ker} p_i$ alors (en utilisant ce qui précède) :

$$x = \sum_{j \neq i} p_j(x) \in \bigoplus_{j \neq i} N_j\,.$$

Réciproquement, si $x \in \oplus_{j \neq i} N_j$ alors $p_i(x) = U_i(f) \circ Q_i(f)(x) = 0$ puisque $M_j^{\alpha_j}$ divise Q_i pour tout $j \neq i$. Donc $\operatorname{Ker} p_i = \oplus_{j \neq i} N_j$ pour tout i. □

Démonstration du théorème.

Existence. On écrit $\chi_f = (-1)^n \prod_{i=1}^{s}(X-\lambda_i)^{\alpha_i}$ et on note $N_i = \mathrm{Ker}(f-\lambda_i \,\mathrm{id}_E)^{\alpha_i}$ pour tout i. On applique le lemme précédent à χ_f et on réutilise les mêmes notations. Posons :

$$d = \sum_i \lambda_i p_i \quad \text{et} \quad n = f - d = \sum_i (f-\lambda_i \,\mathrm{id}_E) p_i \,.$$

d est diagonalisable (dans une base adaptée à la décomposition $E = N_1 \oplus \cdots \oplus N_s$). On utilisant les propriétés des projecteurs, on montre par récurrence que :

$$\forall q \geq 1,\ n^q = \sum_i (f-\lambda_i \,\mathrm{id}_E)^q p_i \,.$$

Si on prend $q \geq \max(\alpha_1, \ldots, \alpha_s)$ alors $n^q = 0$. En effet, $(f-\lambda_i \,\mathrm{id}_E)^q p_i = [(X-\lambda_i)^q P_i](f) = 0$ pour tout i car χ_f divise $(X-\lambda_i)^q P_i$. De plus, d'après le lemme, d et n sont des polynômes en f donc commutent.

Unicité. Soit (d', n') un autre couple vérifiant (i), (ii) et (iii). d' et n' commutent avec $d' + n' = f$ donc avec d et n qui sont des polynômes en f. d et d' sont donc co-diagonalisables, ce qui entraîne que $d - d'$ est diagonalisable. De plus, puisque que n et n' commutent, on peut utiliser la formule de Newton :

$$\forall q \geq 1,\ (n-n')^q = \sum_{i=0}^{q} \binom{q}{i} n^i (-n')^{q-i} \,,$$

et si q est plus grand que la somme des indices de nilpotence de n et n' alors $(n-n')^q = 0$. $d - d' = n' - n$ est nilpotent et diagonalisable, c'est donc l'endomorphisme nul. $\square$

Compléments

Les sous-espaces $N_i = \mathrm{Ker}(f-\lambda_i \,\mathrm{id}_E)^{\alpha_i}$ sont appelés *sous-espaces caractériques* de f. Pour tout i, N_i est de dimension α_i. En effet, la restriction f_i de f à N_i a pour polynôme caracteristique $(X-\lambda_i)^{\dim N_i}$ et comme χ_{f_i} divise χ_f, $\dim N_i \leq \alpha_i$. Et puisque $n = \sum_i \dim N_i$, on a égalité.

Calcul pratique. Il suffit de calculer les projecteurs p_i. Pour cela, on décompose $\frac{1}{\chi_f}$ en éléments simples :

$$\frac{1}{\chi_f} = \sum_{i=1}^{s} \left[\sum_{j=1}^{\alpha_i} \frac{x_{i,j}}{(X-\lambda_i)^j} \right] .$$

On pose ensuite $U_i = \sum_{j=1}^{\alpha_i} x_{i,j} (X-\lambda_i)^{\alpha_i - j}$ de sorte que

$$\frac{1}{\chi_f} = \sum_{i=1}^{s} \frac{U_i}{(X-\lambda_i)^{\alpha_i}} \quad \text{et} \quad 1 = \sum_{i=1}^{s} U_i Q_i \,.$$

Les démonstrations de la décomposition de Dunford et du calcul pratique sont en fait valables pour tout polynôme annulateur.

Applications. La décomposition de Dunford est très pratique pour certains calculs, de puissance d'endomorphisme (à l'aide du binôme de Newton) ou d'exponentielle d'endomorphisme par exemple. Si $f = d + n$ alors $\exp(f) = \exp(d)\exp(n)$. Dans une base $\mathcal{B}$ où d est diagonale, on a :

$$[d]_{\mathcal{B}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \quad \text{et} \quad [\exp(d)]_{\mathcal{B}} = \exp([d]_{\mathcal{B}}) = \mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}) \,.$$

et $\exp(n) = \sum_{p=0}^{q-1} \frac{n^p}{p!}$. Dans la pratique, on utilise les projecteurs p_i :

$$\exp(d) = \sum_{i=1}^{s} e^{\lambda_i} p_i \quad \text{et} \quad \exp(n) = \sum_{i=1}^{s} \left[\sum_{p=0}^{\alpha_i - 1} \frac{(f-\lambda_i \,\mathrm{id}_E)^p}{p!} \right] p_i \,,$$

d'où :

$$\exp(f) = \sum_{i=1}^{s} e^{\lambda_i} \left[\sum_{p=0}^{\alpha_i - 1} \frac{(f-\lambda_i \,\mathrm{id}_E)^p}{p!} \right] p_i \,,$$

Références

- Gourdon, Algèbre, p.194

Leçons concernées : 153, 154, 155, 157

33 Polygones réguliers constructibles à la règle et au compas

Contexte

On considère le plan euclidien $\mathbb{R}^2$ (qu'on identifie au plan complexe $\mathbb{C}$). On pose $O = (0,0)$ et $I = (1,0)$. Si A est une partie de $\mathbb{R}^2$, on considère deux types de figures construites à partir de A :

(i) les *droites affines* $\langle P, Q\rangle$ pour $P \neq Q$ dans A,

(ii) les *cercles* centrés en un point P de A et passant par un autre point $Q \neq P$ de A.

On dit qu'un point $M \in \mathbb{R}^2$ est constructible en un pas à partir de A s'il existe deux éléments distincts (droite ou cercle) dont M est l'intersection. Un point $M \in \mathbb{R}^2$ est dit *constructible* s'il existe une suite $A_0 \subset \cdots \subset A_n$ de partie de $\mathbb{R}^2$ telle que :

$$A_0 = \{O, I\}, \quad M \in A_n, \quad A_i = A_{i-1} \cup \{M_i\} \text{ où } M_i \text{ est constructible en un pas à partir de } A_i\,.$$

Un réel x est dit constructible si le point $(x, 0)$ l'est.

Théorème (Wantzel)**.** *Un point d'affixe z est constructible si et seulement si z est dans une extension L de $\mathbb{Q}$ telle qu'il existe une tour d'extension :*

$$\mathbb{Q} = L_0 \subsetneq L_1 \subsetneq \cdots \subsetneq L_r = L\,,$$

avec : $\forall i \in [\![1\,, r]\!]$, $[L_i : L_{i-1}] = 2$.

Pour $n \geq 3$, on note P_n le polygone régulier à n côtés dont les sommets sont les racines n-ième du cercle unité. On dira que P_n est constructible si toutes les racines n-ième de l'unité le sont. En fait, P_n est constructible si et seulement si $\omega_n = e^{\frac{2i\pi}{n}}$ l'est (car si $e^{i\alpha}$ et $e^{i\beta}$ sont constructibles $e^{i(\alpha+\beta)}$ et $e^{i(\alpha-\beta)}$ le sont aussi).

Un *nombre de Fermat* est un nombre de la forme $2^{2^n} + 1$ où $n \in \mathbb{N}$.

Développement

Théorème (Gauss-Wantzel)**.** *P_n est constructible si et seulement si n est de la forme $n = 2^r p_1 \ldots, p_s$ où $r \in \mathbb{N}$ et où les p_i sont des nombres premier de Fermat distincts deux à deux.*

Démonstration. Soit $n = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$ où les p_i sont des nombres premiers distincts deux à deux. Alors P_n **est constructible si et seulement si** $P_{p_i^{\alpha_i}}$ **est constructible pour tout** i**.** En effet, soient m et n deux entiers tels que m divise n. Comme l'ensemble des sommets de P_m est inclus dans l'ensemble des sommets de P_n (en prendre 1 sur $\frac{n}{m}$), P_n est constructible implique que P_m est constructible. Cela montre le sens direct de l'assertion. Soient m et n deux entiers tels que $n \wedge m = 1$. On suppose que P_n et P_m sont constructibles. Par l'égalité de Bézout, il existe des entiers a et b tels que : $an + bm = 1$. On en déduit que :

$$\frac{2\pi}{nm} = a\frac{2\pi}{m} + b\frac{2\pi}{n}$$

est constructible donc P_{mn} est constructible. Cela montre le sens indirect de l'assertion.

Si α est constructible, on sait construire $\frac{\alpha}{2}$ (tracer la bissectrice), on en déduit par récurrence que P_{2^k} **est constructible pour tout** $k \in \mathbb{N}$**.**

Soit p premier ≥ 3. Montrons que P_{p^α} **est constructible si et seulement si** $\alpha = 1$ **et** p **est un nombre de Fermat.** On pose $q = p^\alpha$. Si P_q est constructible alors ω_q l'est aussi. D'après le théorème de Wantzel, il existe $\ell \in \mathbb{N}$ tel que

$$[\mathbb{Q}(\omega_q) : \mathbb{Q}] = 2^\ell\,.$$

ω_q est une racine $q^{\text{ième}}$ de l'unité donc son polynome minimal dans $\mathbb{Q}$ est le $q^{\text{ième}}$ polynôme cyclotomique Φ_q. Ainsi :

$$\deg(\Phi_q) = \varphi(q) = [\mathbb{Q}(\omega_q) : \mathbb{Q}] = p^{\alpha-1}(p-1) = 2^\ell\,,$$

Comme p est premier différent de 2, on en déduit que $\alpha = 1$ et que $p = 1 + 2^\ell$. Il reste à montrer que ℓ est une puissance de 2. Pour cela, on écrit : $\ell = \lambda 2^\beta$ où λ est un entier impair. Ensuite :

$$2^\ell = 2^{\lambda 2^\beta} = \left(2^{2^\beta}\right)^\lambda\,.$$

Comme λ est impair, le polynôme $1 + X^\lambda$ est divisible par $1 + X$. Ainsi, p est divisible par $1 + 2^{2^\beta}$ et comme p est premier : $p = 1 + 2^{2^\beta}$.

Réciproquement, soit $p = 1+2^r$ est un nombre de Fermat premier (r est nécessairement une puissance de 2). On pose $\omega = \omega_p$. On considère l'ensemble G des automorphismes de $\mathbb{Q}(\omega)$ dont la restriction à $\mathbb{Q}$ est égale à l'identité (c'est le groupe de Galois de l'extension $\mathbb{Q}(\omega)$ de $\mathbb{Q}$). $\varphi \in G$ est entièrement déterminé par sa valeur en ω. De plus, $\varphi(\omega)$ est nécessairement une racine $p^{\text{ième}}$ de l'unité. On obtient l'isomorphisme de groupes :

$$\begin{array}{ccc} G & \longrightarrow & \mathbb{F}_p^\times \\ \varphi & \longmapsto & m_\varphi \end{array}$$

où m_φ est l'unique élément de $\mathbb{F}_p^\times$ tel que $\varphi(\omega) = \omega^{m_\varphi}$. Donc G est cyclique d'ordre $p - 1 = 2^r$. Soit σ un générateur de G. Pour tout $k \in [\![0\,, r]\!]$, on pose :

$$L_k = \{z \in \mathbb{Q}(\omega) \mid \sigma^{2^k}(z) = z\}\,.$$

Les L_k sont des sous-corps de $\mathbb{Q}(\omega)$ tels que :

$$\mathbb{Q} \subset L_0 \subsetneq L_1 \subsetneq \cdots \subsetneq L_r = \mathbb{Q}(\omega)\,.$$

On a $L_{k-1} \subsetneq L_k$ pour tout k puisque :

$$x = \sum_{m=0}^{2^{r-k}-1} \sigma^{2^k m}(\omega) = \omega + \sigma^{2^k}(\omega) + \cdots + \sigma^{2^r - 2^k}(\omega) \in L_k$$

mais $x \notin L_{k-1}$ car :

$$\sigma^{2^{k-1}}(x) = \sum_{m=0}^{2^{r-k}-1} \sigma^{2^{k-1}(2m-1)}(\omega) \neq \sum_{m=0}^{2^{r-k}-1} \sigma^{2^{k-1}(2m)}(\omega) = x\,,$$

en effet, $(1, \omega, \ldots, \omega^{p-1})$ étant une base de $\mathbb{Q}(\omega)$ et σ un automorphisme, $(\sigma^k(\omega) = \sigma(\omega^k), k = 0, \ldots, p-1)$ est aussi une base de $\mathbb{Q}(\omega)$ et les deux combinaisons linéaires ci-dessus sont bien différentes l'une de l'autre. Ainsi, par multiplicativité du degré :

$$2^r = [\mathbb{Q}(\omega) : \mathbb{Q}] = [L_r : L_{r-1}] \times [L_{r-1} : L_{r-2}] \times \cdots \times [L_1 : L_0] \times [L_0 : \mathbb{Q}]\,,$$

ce qui implique que pour tout k : $[L_k : L_{k-1}] = 2$ (et $L_0 = \mathbb{Q}$). On en déduit, par le théorème de Wantzel, que ω est constructible. □

Remarque. De la correspondance de Galois se cache derrière tout ce beau merdier.

Références

– Carréga, Théorie des corps, p.36
– Perrin, Cours d'algèbre, p.69 (cours)

Leçons concernées : 102, 125, 141, 144, 151, 182, 183

34 Simplicité de $\mathcal{A}_n$

Contexte

Soit G un groupe. Un sous-groupe H de G est *distingué* dans G s'il est stable par automorphisme intérieur. Un groupe est dit *simple* s'il ne possède pas de sous-groupe distingué autre que lui-même et le sous-groupe trivial.

Développement

Théorème. *$\mathcal{A}_n$ est simple si $n \geq 5$.*

Démonstration. On montre d'abord le résultat pour $n = 5$ et on généralise ensuite pour $n > 5$.

(i) $\mathcal{A}_5$ possède $60 = 2^2 \times 3 \times 5$ éléments :
  - le neutre $\rightarrow$ 1 élément,
  - les éléments d'ordre 2 (produit de 2 transpositions disjointes) $\rightarrow 5 \times 3 = 15$ éléments (on choisit d'abord l'élément qui n'est pas dans le support puis on se fixe un élément du support et on choisit parmi les trois restant à qui on l'associe),
  - les éléments d'ordre 3 (les 3-cycles) $\rightarrow$ $\binom{5}{3} \times 2 = 20$ éléments (on choisit d'abord le support, on se fixe ensuite un élément et il reste à choisir parmi les deux éléments restant lequel on met ensuite),
  - les éléments d'ordre 5 (les 5-cycles) $\rightarrow$ $4! = 24$ éléments (classique).

  Les 3-cycles sont conjugués dans $\mathcal{A}_5$. En effet, soient $(x_1\ x_2\ x_3)$ et $(y_1\ y_2\ y_3)$ deux 3-cycles. Ils sont conjugués dans $\mathfrak{S}_5$:
  $$\exists \sigma \in \mathfrak{S}_5,\ \sigma(x_1\ x_2\ x_3)\sigma^{-1} = (y_1\ y_2\ y_3)\,.$$
  Si $\sigma \in \mathcal{A}_5$ alors c'est bon. Sinon on choisit x_4 et x_5 différents de x_1, x_2 et x_3 et on peut réécrire ce qui précède en remplaçant σ par $\sigma(x_4\ x_5) \in \mathcal{A}_5$.

  Les éléments d'ordre 2 sont aussi conjugués dans $\mathcal{A}_5$. Si $\tau = (a\ b)(c\ d)(e)$ et $\tau' = (a'\ b')(c'\ d')(e')$ alors il existe $\sigma \in \mathcal{A}_5$ tel que $\sigma(a) = a'$, $\sigma(b) = b'$ et $\sigma(e) = e'$ (multiplier ou pas par $(c\ d)$) et $\sigma\tau\sigma^{-1} = \tau'$.

  Soit H un sous-groupe distingué non trivial de $\mathcal{A}_5$. Si H contient
  - un 3-cycle alors H les contient tous,
  - un 2-cycle alors H les contient tous,
  - un 5-cycle alors H contient le 5-Sylow engendré par cet élément et comme les 5-Sylow sont conjugués entre eux, H les contient tous et donc tous les éléments d'ordre 5.

  H ne peut contenir un seul des trois type d'éléments car ni $16 = 15 + 1$, ni $21 = 20 + 1$, ni $25 = 24 + 1$ ne divisent 60. Nécessairement, H contient au moins deux de ces types d'éléments mais alors $|H| \geq 31$. Comme le seul diviseur de 60 supérieur à 31 est 60, on a $H = \mathcal{A}_5$. Donc $\mathcal{A}_5$ est simple.

(ii) Soit H un sous-groupe distingué non trivial de $\mathcal{A}_n$. Soit $\sigma \in H$ tel que $\sigma \neq 1$. On va construire à partir de σ une permutation non trivial avec $n - 5$ points fixes.

  Pour tout $\tau \in \mathcal{A}_n$, $\rho = \tau(\sigma\tau^{-1}\sigma^{-1}) = (\tau\sigma\tau^{-1})\sigma^{-1}$ est dans H. Soit $a \in [\![1\,,n]\!]$ tel que $b = \sigma(a) \neq a$. Soit $c \in [\![1\,,n]\!] \setminus \{a, b\}$. On pose $\tau = (a\ c\ b)$ de sorte que $\tau^{-1} = (a\ b\ c)$. On a aussi $\rho = \tau(\sigma\tau^{-1}\sigma^{-1}) = (a\ c\ b)(\sigma(a)\ \sigma(b)\ \sigma(c))$. L'ensemble $F = \{a, b, c, \sigma(a), \sigma(b), \sigma(c)\}$ a au plus 5 éléments puisque $b = \sigma(a)$. De plus, $\rho(F) = F$ et $\rho_{|F^c} = \mathrm{id}_{F^c}$. ρ n'est pas l'identité car $\rho(b) = \tau(\sigma(b)) \neq b$ puisque $\sigma(b) \neq c$.

  Quitte à rajouter des éléments, on peut supposer que $|F| = 5$. $\mathcal{A}(F)$ est isomorphe à $\mathcal{A}_5$ et se plonge dans $\mathcal{A}_n$ par $u \mapsto \overline{u}$ où :
  $$\overline{u}_{|F} = u \quad \text{et} \quad \overline{u}_{|F^c} = \mathrm{id}_{F^c}\,.$$
  $H_0 = H \cap \mathcal{A}(F)$ est distingué dans $\mathcal{A}(F)$ et contient $\rho_{|F}$. D'après le cas $n = 5$, $H_0 = \mathcal{A}(F)$. Si u est un 3-cycle de $\mathcal{A}(F)$ alors $\overline{u}$ est dans H et est aussi un 3-cycle. H contient un 3-cycle donc les contient tous et comme les 3-cycles engendrent $\mathcal{A}_n$ on a : $H = \mathcal{A}_n$. Donc $\mathcal{A}_n$ est simple.

$\square$

Remarque. $\mathcal{A}_4$ n'est pas simple car $\{\mathrm{id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$ est un sous-groupe distingué de $\mathcal{A}_4$. Les cas de $\mathcal{A}_2$ et $\mathcal{A}_3$ se traitent facilement.

Compléments

Application. *Pour $n \geq 5$, les seuls sous-groupes distingués de $\mathfrak{S}_n$ sont $\{1\}$, $\mathcal{A}_n$ et $\mathfrak{S}_n$.*

Démonstration. Soit H un sous-groupe distingué non trivial. Alors, puisque $H \cap \mathcal{A}_n$ est distingué dans $\mathcal{A}_n$, l'un des deux cas suivants se présente :

- $H \cap \mathcal{A}_n = \mathcal{A}_n$ et donc $\mathcal{A}_n \subset H$ et pour des raisons de cardinal et de divisibilité, $H = \mathfrak{S}_n$,
- $H \cap \mathcal{A}_n = \{1\}$ et la signature induit un isomorphisme de H sur $\{-1, 1\}$. Donc $|H| = 2$. Notons $H = \{1, \sigma\}$ où $\sigma \notin \mathcal{A}_n$. On peut voir que puisque H est distingué dans $\mathfrak{S}_n$, cela implique que σ est dans le centre de $\mathfrak{S}_n$, ce qui est absurde puisqu'il est réduit au neutre dès que $n \geq 3$.

□

Références

- Perrin, Cours d'algèbre, p.28

Leçons concernées : 105, 108

35 Quaternions et $\mathrm{SO}_3(\mathbb{R})$

Contexte

On présente un petit topo sur les quaternions.

Il existe une $\mathbb{R}$-algèbre $\mathbb{H}$ de dimension 4, appelée algèbre des *quaternions*, munie d'une base $(1, i, j, k)$ telle que :

(i) 1 est l'élément neutre pour la multiplication,

(ii) on a les formules :

$$i^2 = j^2 = k^2 = ijk = -1,\ jk = -kj = i,\ ki = -ik = j,\ ij = -ji = k\,.$$

Si $q = a + bi + cj + dk \in \mathbb{H}$ alors on définit son *conjugué* par $\overline{q} = a - bi - cj - dk$. On note $P = \{bi + cj + dk \mid b, c, d \in \mathbb{R}\} = \{q \in \mathbb{H} \mid \overline{q} = -q\}$ l'ensemble des *quaternions purs*. On définit la *norme* de $q \in \mathbb{H}$ par $N(q) = q\overline{q} = \overline{q}q \in \mathbb{R}^+$. $q \mapsto N(q)$ est une forme quadratique euclidienne sur $\mathbb{H}$ et la base $(1, i, j, k)$ est orthonormée. En particulier, $\mathbb{R}$ et P sont orthogonaux. On a de plus :

(i) l'algèbre $\mathbb{H}$ est un corps non commutatif dont le centre est égal à $\mathbb{R}$,

(ii) $N : \mathbb{H}^\times \to \mathbb{R}_+^*$ est un morphisme de groupes surjectif dont le noyau, le groupe des quaternions de norme 1, est noté G.

Si q est dans G alors $q^{-1} = \overline{q}$. On peut écrire $G = \{(a, b, c, d) \in \mathbb{R}^4 \mid a^2 + b^2 + c^2 + d^2 = 1\}$, ce qui montre que G est homéomorphe à la sphère $\mathbb{S}_3$ de $\mathbb{R}^4$. En particulier, G est connexe.

Quant à $\mathrm{SO}_3(\mathbb{R})$, on ne le présente plus tellement il nous est devenu familier au cours de ces dernières années !

Développement

Théorème. *On a un isomorphisme :*

$$\overline{s} : G/\{-1, 1\} \xrightarrow{\sim} \mathrm{SO}_3(\mathbb{R})\,.$$

Remarque. On étudie ici les automorphismes intérieurs de $\mathbb{H}^\times$ sur $\mathbb{H}$ qui ne sont pas tous triviaux puisque $\mathbb{H}$ n'est pas commutatif. Comme $q^{-1} = \frac{1}{N(q)}\overline{q}$, il suffit de regarder l'action de G sur $\mathbb{H}^\times$.

Démonstration. Soit $q \in G$. On pose :

$$S_q : q' \in \mathbb{H} \longmapsto qq'q^{-1} = qq'\overline{q}\,.$$

S_q est clairement linéaire, elle est de plus bijective. On a, en effet, $(S_q)^{-1} = S_{\overline{q}}$. On obtient ainsi une application :

$$S : G \longrightarrow \mathrm{GL}_4(\mathbb{R})\,.$$

S est un morphisme de groupes de noyau égal à $Z(\mathbb{H}) \cap G = \mathbb{R} \cap G = \{-1, 1\}$.

S_q est un élément de $\mathrm{O}(N) \simeq \mathrm{O}_4(\mathbb{R})$. En effet : $\forall q' \in \mathbb{H},\ N(S_q(q')) = N(q)N(q')N(\overline{q}) = N(q')$ puisque $N(q) = 1$.

On a $S_{q|\mathbb{R}} = \mathrm{id}_\mathbb{R}$ car $\mathbb{R}$ est le centre de $\mathbb{H}$.

L'espace des quaternions purs P est stable par S_q. En effet, P est l'orthogonal de $\mathbb{R}$ et :

$$\forall q' \in P,\ \langle S_q(q'), 1\rangle = \langle S_q(q'), S_q(1)\rangle = \langle q', 1\rangle = 0\,,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire induit par N. On pose $s_q = S_{q|P}$ et, par restriction, on a encore $s_q \in \mathrm{O}(N_{|P}) \simeq \mathrm{O}_3(\mathbb{R})$ et $s : G \to \mathrm{O}_3(\mathbb{R})$ un morphisme de groupes de noyau $\{-1, 1\}$.

Il reste à montrer que l'image de s est $\mathrm{SO}_3(\mathbb{R})$.

s est une application continue. En effet, la matrice de S_q dans la base $(1, i, j, k)$ est clairement à coefficients polynomiaux en les coordonnées de q. Comme l'application déterminant $\det : \mathrm{O}_3(\mathbb{R}) \to \{-1, 1\}$ est elle aussi continue, l'application $\det \circ s$ est continue de G dans $\{-1, 1\}$. Or G est connexe et $s_1 = \mathrm{id}$. On en déduit que l'image de s est incluse dans $\mathrm{SO}_3(\mathbb{R})$. Réciproquement, si $p \in P \cap G$ alors $s_p(p) = pp\overline{p} = p$. p est un point fixe de $s_p \in \mathrm{SO}_3(\mathbb{R})$, donc s_p est une rotation d'axe (p). De plus, comme $p^2 = -p\overline{p} = -1$, on a $(s_p)^2 = s_{p^2} = s_{-1} = \mathrm{id}$. s_p est une involution, c'est donc le renversement d'axe (p) et commme les renversements engendrent $\mathrm{SO}_3(\mathbb{R})$, on en déduit que l'image de s est $\mathrm{SO}_3(\mathbb{R})$.

Pour conclure, il suffit de quotienter s par son noyau. $\square$

Remarque. Réciproquement, on peut montrer que tout les automorphismes de corps de $\mathbb{H}$ sont intérieurs.

Remarque. Les quaternions de norme 1 permettent de paramétrer $\mathrm{SO}_3(\mathbb{R})$ ce qui, dans certains cas, allège les calculs.

Références

– Perrin, Cours d'algèbre, p.163

Leçons concernées : 106, 160, 161, 182, 183

36 Réciprocité quadratique (via les formes quadratiques)

Contexte

Pour p premier impair et a un élément de $\mathbb{F}_p$, on définit le *symbole de Legendre* de a par :

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{si } a \text{ est un carré dans } \mathbb{F}_p^\times \\ -1 & \text{si } a \text{ n'est pas un carré dans } \mathbb{F}_p^\times \\ 0 & \text{si } a = 0 \end{cases}.$$

Le symbole de Legendre est multiplicatif et vérifie :

$$\forall a \in \mathbb{F}_p,\ \left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \pmod{p}.$$

Développement

Théorème (Réciprocité quadratique)**.** *Soit p et q deux nombres premiers impairs et distincts. Alors :*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Lemme. *Soit p premier impair et $a \in \mathbb{F}_p^\times$. Alors :*

$$\left|\{x \in \mathbb{F}_p,\ ax^2 = 1\}\right| = 1 + \left(\frac{a}{p}\right).$$

Démonstration du lemme. Si a n'est pas un carré dans $\mathbb{F}_p$ alors cet ensemble est vide. Si a est un carré alors cet ensemble a deux éléments, à savoir : b^{-1} et $-b^{-1}$ où $b^2 = a$. □

Démonstration du théorème. On calcule le cardinal modulo p de la sphère unité sur $\mathbb{F}_q$ de deux façons différentes :

$$X := \{(x_1, \ldots, x_p) \in \mathbb{F}_q^p \mid \sum_{i=1}^p x_i^2 = 1\}.$$

(i) **Première façon.** On fait agir le groupe $\mathbb{Z}/p\mathbb{Z}$ sur X de la manière suivante :

$$\forall k \in \mathbb{Z}/p\mathbb{Z},\ \forall (x_1, \ldots, x_p) \in \mathbb{F}_p^q,\ k \cdot (x_1, \ldots, x_p) = (x_{\overline{1+k}}, \ldots, x_{\overline{n+k}}),$$

où $\overline{x}$ désigne la classe de x module $\mathbb{F}_p$. Les orbites sont de deux types : d'un côté les singletons $\{(x, \ldots, x)\}$ où $x \in \mathbb{F}_q$ dont le stabilisateur est $\mathbb{Z}/p\mathbb{Z}$ tout entier, et de l'autre côté les orbites non réduites à un singleton, dont le stabilisateur est nécessairement trivial (c'est un sous-groupe de $\mathbb{Z}/p\mathbb{Z}$), sont de cardinal p. La réunion des orbites réduites à un singleton est égale à :

$$\{x \in \mathbb{F}_q \mid px^2 = 1\}.$$

D'après le lemme et la formule des classes, on a :

$$|X| \equiv \left|\{x \in \mathbb{F}_q \mid px^2 = 1\}\right| \equiv 1 + \left(\frac{p}{q}\right) \pmod{p}.$$

(ii) **Deuxième façon.** On pose $d = \frac{p-1}{2}$. Les matrices de taille $p \times p$ et à coefficients dans $\mathbb{F}_q$

$$I_p = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \quad \text{et} \quad A = \begin{pmatrix} 0 & 1 & & & & \\ 1 & 0 & & & & \\ & & \ddots & & & \\ & & & 0 & 1 & \\ & & & 1 & 0 & \\ & & & & & a \end{pmatrix},$$

où $a = (-1)^d$, sont symétriques, ont même rang p et ont même déterminant 1. Elles sont donc congruentes sur $\mathbb{F}_q$ et il existe un changement de variables linéaires qui identifie X à :

$$X' := \{(y_1, \dots, y_d, z_1, \dots, z_d, t) \in \mathbb{F}_q^p,\ 2(y_1 z_1 + \cdots + y_d z_d) + at^2 = 1\}\,.$$

Il y a deux types de points $(y_1, \dots, y_d, z_1, \dots, z_d, t)$ dans X' :

- les points tels que $y_1 = \cdots = y_d = 0$ et le lemme montre que chaque valeur de t telle que $at^2 = 1$ détermine q^d points : il y a donc $q^d(1 + a^{\frac{q-1}{2}})$ tels points,
- les points pour lesquels au moins l'un des y_i est non nuls : une fois fixé $(y_1, \dots, y_d, t)$, il reste à choisir $(z_1, \dots, z_d)$ dans un hyperplan affine de $\mathbb{F}_q^d$. : il y a $(q^d - 1)qq^{d-1}$ tels points.

Finalement :

$$|X| = |X'| \equiv q^d \left(q^d + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\right) \pmod p\,.$$

En comparant :

$$1 + \left(\frac{p}{q}\right) \equiv q^{\frac{p-1}{2}} \left(q^{\frac{p-1}{2}} + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\right) \pmod p\,,$$

puis :

$$1 + \left(\frac{p}{q}\right) \equiv \left(\frac{q}{p}\right) \left[\left(\frac{q}{p}\right) + (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\right] \pmod p\,,$$

ce qui donne le résultat attendu puisque $\left(\frac{q}{p}\right)^2 = 1$ et qu'on travaille modulo $p \geq 3$. $\square$

Compléments

La loi de réciprocité quadratique ainsi que la multiplicativité du symbole de Legendre permettent de savoir si un nombre est un carré :

$$\left(\frac{58}{23}\right) = \left(\frac{12}{23}\right) = \left(\frac{3}{23}\right)\left(\frac{4}{23}\right) = (-1)^{1\times 11}\left(\frac{23}{3}\right) = -\left(\frac{2}{3}\right) = 1\,.$$

Références

- Caldero et Germoni, Histoires hédonistes de groupes et de géométries, p.185

Leçons concernées : 150, 170, 190

37 Enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$

Contexte

Théorème (Hahn-Banach - deuxième forme géométrique)**.** *Soit E un $\mathbb{R}$-espace vectoriel normé. Soient A et B deux convexes non vides et disjoints de E tels que A est compact et B est fermé. Alors il existe un hyperplan fermé qui sépare strictement A et B, i.e :*

$$\exists \varphi \in E^*, \ \sup_{x\in A} \varphi(x) < \inf_{x \in B} \varphi(x) \, .$$

On munit $\mathcal{M}_n(\mathbb{R})$ de la norme d'opérateur associée à la norme euclidienne canonique sur $\mathbb{R}^n$:

$$\forall M \in \mathcal{M}_n(\mathbb{R}), \ \|M\| = \sup_{x \neq 0} \frac{\|Mx\|_2}{\|x\|_2} \, .$$

Développement

Théorème. *L'enveloppe convexe de $\mathrm{O}_n(\mathbb{R})$ est la boule unité de $\mathcal{M}_n(\mathbb{R})$ pour la norme $\|\cdot\|$.*

Lemme (Théorème de représentation de Riesz)**.** *On munit $\mathbb{R}^n$ d'une structure euclidienne induite par un produit scalaire $\langle \cdot, \cdot \rangle$. Alors, l'application*

$$\begin{array}{rccl} \varphi & : & \mathbb{R}^n & \longrightarrow & (\mathbb{R}^n)^* \\ & & f & \longmapsto & [\varphi_f : v \in \mathbb{R}^n \longmapsto \varphi_f(v) = \langle f, v \rangle] \end{array}$$

est un isomorphisme (isométrique).

Démonstration du lemme. A FINIR

D'après l'inégalité de Cauchy-Schwarz :

$$\forall f \in \mathbb{R}^n, \ \forall v \in \mathbb{R}^n, \ |\varphi_f(v)| \leq \|f\| \, \|v\| \, .$$

On en déduit que $\|\varphi_f\| \leq \|f\|$ et φ est bien définie. De plus, en prenant $v = f$, on a égalité. Ensuite φ est clairement une application linéaire. On a donc montré que φ est une isométrie (ce qui implique l'injectivité). □

Démonstration du théorème. Notons B la boule unité de $\mathcal{M}_n(\mathbb{R})$ pour $\|\cdot\|$. Puisque $\mathrm{O}_n(\mathbb{R}) \subset B$, on a : $\mathrm{conv}(\mathrm{O}_n(\mathbb{R})) \subset B$.

Réciproquement, soit $M \in \mathcal{M}_n(\mathbb{R})$ tel que $\|M\| \leq 1$. Pour montrer que $M \in \mathrm{conv}(\mathrm{O}_n(\mathbb{R}))$, il suffit de montrer que :

$$\forall \varphi \in \mathcal{M}_n(\mathbb{R})^*, \ \varphi(M) \leq \sup_{O \in \mathrm{O}_n(\mathbb{R})} \varphi(O) \left(\leq \sup_{O \in \mathrm{conv}(\mathrm{O}_n(\mathbb{R}))} \varphi(O) \right) \, .$$

En effet, c'est la contraposée du théorème de Hahn-Banach appliqué à $A = \mathrm{conv}(\mathrm{O}_n(\mathbb{R}))$ qui est convexe et compact (d'après le théorème de Carathéodory) et $B = \{M\}$ qui est convexe et fermé. Puisque $(A, B) \mapsto \mathrm{tr}({}^tAB)$ est un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$, le théorème de représentation de Riesz permet de représenter chaque forme linéaire de $\mathcal{M}_n(\mathbb{R})$ par un vecteur :

$$\forall \varphi \in \mathcal{M}_n(\mathbb{R})^*, \exists A \in \mathcal{M}_n(\mathbb{R}), \ \forall M \in \mathcal{M}_n(\mathbb{R}), \ \varphi(M) = \mathrm{tr}\left({}^tAM\right) \, .$$

Il suffit donc de montrer :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \ \mathrm{tr}\left({}^tAM\right) \leq \sup_{O \in \mathrm{O}_n(\mathbb{R})} \mathrm{tr}\left({}^tAO\right) \, .$$

Soit $A \in \mathcal{M}_n(\mathbb{R})$. D'après le théorème de décomposition polaire, il existe $S \in \mathcal{S}_n^+(\mathbb{R})$ et $O \in \mathrm{O}_n(\mathbb{R})$ tels que $A = OS$. Le théorème spectral montre l'existence de $(f_i)_i$ une base orthonormée de vecteurs propres associés à S telle que :

$$\mathrm{tr}(S) = \mathrm{tr}\left({}^tS\right) = \mathrm{tr}({}^tAO) = \sum_{i=1}^n \langle Sf_i, f_i \rangle = \sum_{i=1}^n \|Sf_i\|_2 \, .$$

Ensuite, en utilisant l'inégalité de Cauchy-Schwarz :

$$\operatorname{tr}\left({}^t\!AM\right) = \operatorname{tr}\left({}^t\!MA\right) = \sum_{i=1}^{n} \left\langle {}^t\!MAf_i, f_i \right\rangle = \sum_{i=1}^{n} \langle Af_i, Mf_i \rangle \leq \sum_{i=1}^{n} \|Af_i\|_2 \|Mf_i\|_2 = \sum_{i=1}^{n} \|OSf_i\|_2 \|Mf_i\|_2 ,$$

or puisque $O \in \mathrm{O}_n(\mathbb{R})$ et $M \in B$:

$$\operatorname{tr}\left({}^t\!AM\right) \leq \sum_{i=1}^{n} \|Sf_i\|_2 \|Mf_i\|_2 \leq \sum_{i=1}^{n} \|Sf_i\|_2 = \operatorname{tr}({}^t\!AO) \leq \sup_{O' \in \mathrm{O}_n(\mathbb{R})} \operatorname{tr}\left({}^t\!AO'\right) ,$$

ce qui conclut. □

Références

– Szpirglas, Algèbre L3, p.344

Leçons concernées : 159, 161

38 Théorème de Kronecker

Développement

Théorème. *Soit P un polynôme unitaire de $\mathbb{Z}[X]$ dont les racines complexes sont toutes de module inférieur ou égal à 1. Si $P(0) \neq 0$ alors les racines de P sont des racines de l'unité.*

Démonstration. Soit $n = \deg P$. On note $z_1, \ldots, z_n$ les racines complexes de P comptées avec leur multiplicité et $\sigma_1, \ldots, \sigma_n$ les fonctions symétriques élémentaires des z_i. On a donc :

$$P(X) = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^n \sigma_n \,,$$

de plus les σ_i sont dans $\mathbb{Z}$. On note $\mathcal{P}_j$ l'ensemble des parties à j éléments de $[\![1\,,n]\!]$, il y en a $\binom{n}{j}$. Comme les z_i sont de modules inférieur à 1, on a :

$$|\sigma_j| \leq \left| \sum_{I \in \mathcal{P}_j} \prod_{i \in I} z_i \right| \leq |\mathcal{P}_j| = \binom{n}{j} .$$

Cela montre que l'ensemble des polynômes unitaires et de degré n de $\mathbb{Z}[X]$ dont les racines complexes sont toutes de module inférieur à 1 (ensemble noté Ω_n) est de cardinal fini.

Pour tout $k \geq 0$, on pose $P_k(X) = (X - z_1^k) \cdots (X - z_n^k)$, polynôme unitaire de degré n dont les racines sont toutes de module inférieur ou égal à 1. Montrons que pour tout k, P_k est dans $\mathbb{Z}[X]$. Pour tout r, le coefficient de X^{n-r} dans P_k est $(-1)^r \sigma(z_1^k, \ldots, \sigma_n^k)$. Il s'agit d'un polynôme symétrique en $z_1, \ldots, z_n$ à coefficients dans $\mathbb{Z}$. Il est donc égal à un polynôme à coefficients entiers en les fonctions symétriques élémentaires des z_i, ce qui montre que $P_k \in \Omega_n$.

L'ensemble des racines des éléments de Ω_n est fini car majoré par $n\,|\Omega_n|$ (un polynôme de degré n a au plus n racines différentes). L'application $k \mapsto z_i^k$ ne peut donc pas être injective et il existe $p \neq q$ tels que : $z_i^p = z_i^q$. Comme z_i est supposée non nul, z_i est un racine de l'unité. $\square$

Application. *Soit P un polynôme unitaire de $\mathbb{Z}[X]$ dont les racines complexes sont toutes de module inférieur ou égal à 1. Si P est irréductible alors $P = X$ ou P est un polynôme cyclotomique.*

Démonstration. Si $P(0) = 0$ alors nécessairement $P = X$ (P est irréductible). Si $P(0) \neq 0$ alors, d'après le théorème précédent, les racines de P sont des racines de l'unité. En notant $z_1, \ldots, z_n$ les racines de P avec leur multiplicité, alors :

$$\forall i,\ \exists k_i \geq 1 \mid z_i^{k_i} = 1 \,,$$

et pour $N = k_1 \vee \cdots \vee k_n$, on a : $\forall i,\ z_i^N = 1$. Or P est à racines simples car sinon $P \wedge P'$ serait un polynôme non constant divisant P, ce qui contredirait l'irréductibilité de P. On en déduit :

$$P \mid X^N - 1 = \prod_{d \mid n} \Phi_d \,,$$

où Φ_d est d-ième polynôme cyclotomique. Puisque les polynômes cyclotomiques sont irréductibles et unitaires, P est un polynôme cyclotomique. $\square$

Références

- Francinou, Gianella et Nicolas, Algèbre 1, p.213
- Szpirglas, Algèbre L3, p.573

Leçons concernées : 102, 142, 144

39 Invariants de similitudes (réduction de Frobenius)

Contexte

Soient $\mathbb{K}$ un corps (commutatif) et E un $\mathbb{K}$-espace vectoriel de dimension finie. On notera Π_f le polynôme minimal de f. On pose $\mathcal{L}_f := \{P(f) \mid P \in \mathbb{K}[X]\}$. Si $x \in E$, on note P_x le polynôme unitaire qui engendre l'idéal $\{P \in \mathbb{K}[X] \mid P(f)(x) = 0\}$ ainsi que E_x l'idéal $\{P(f)(x) \mid P \in \mathbb{K}[X]\}$. Alors :

Proposition. *Si $k = \deg(\Pi_f)$ (resp. $\ell = \deg(P_x)$) alors l'ensemble $\mathcal{L}_f$ (resp. E_x) est une sous-espace vectoriel de $\mathcal{L}(E)$ (resp. de E) de dimension k (resp. ℓ) dont une base est $(\mathrm{id}_E, f, \dots, f^{k-1})$ (resp. $(x, f(x), \dots, f^{\ell-1}(x))$).*

Cette proposition se montre en identifiant $\mathcal{L}_f$ à $\mathbb{K}[X]/(\Pi_f)$ (resp. E_x à $\mathbb{K}[X]/(P_x)$). On peut montrer la proposition suivante (voir compléménts) :

Proposition. $\exists x \in E,\ P_x = \Pi_f$.

On dira qu'un endomorphisme f de E est *cyclique* s'il existe $x \in E$ tel que $E_x = E$, ce qui est équivalent à dire que $\deg \Pi_f = n$ ou encore que $\Pi_f = (-1)^n P_f$ où P_f désigne le polynôme caractéristique de f.

Soit $P = X^p + a_{p-1}X^{p-1} + \cdots + a_0 \in \mathbb{K}[X]$. On appelle *matrice compagnon* de P la matrice :

$$\mathcal{C}(P) = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{p-2} \\ 0 & \cdots & 0 & 1 & -a_{p-1} \end{pmatrix} \in \mathcal{M}_p(\mathbb{K}).$$

Le polynôme caractéristique de $\mathcal{C}(P)$ est égal à P. De plus, si f est un endomorphisme cyclique alors il existe une base de E dans laquelle la matrice de f soit égale à $\mathcal{C}(\Pi_f)$. Il existe $x \in E$ tel que $E_x = E$ et il suffit de prendre la base $(x, f(x), \dots, f^{n-1}(x))$.

Développement

Théorème (Invariants de similitudes)**.** *Soit $f \in \mathcal{L}(E)$. Il existe une suite $F_1, \dots, F_r$ de sous-espaces vectoriels de E, tous stables par f, telle que :*

(i) $E = F_1 \oplus \cdots \oplus F_r$,

(ii) pour tout $i \in [\![1, r]\!]$, $f_i = f_{|F_i}$ est un endomorphisme cyclique de F_i,

(iii) en notant $P_i = \Pi_{f_i}$, $P_{i+1} \mid P_i$ pour tout $i \in [\![1, r-1]\!]$.

La suite de polynômes $P_1, \dots, P_r$ de dépend que de f et non du choix de la décomposition. On l'appelle suite des invariants de similitude *de f.*

Démonstration de l'existence. Pour montrer l'existence d'une telle décomposition, on raisonne par récurrence forte sur la dimension de E.

Soit $x \in E$ tel que $P_x = \Pi_f$. On pose $F = E_x$, F est f-stable et est de dimension $k = \deg(P_x)$. De plus, la famille de vecteurs $(e_1, \dots, e_k) := (x, f(x), \dots, f^{k-1}(x))$ forme une base de F. On complète cette famille libre en une base $(e_1, \dots, e_n)$ de E et on note où $(e_1^*, \dots, e_n^*)$ la base duale associée. On cherche G sous-espace f-stable tel que $E = F \oplus G$. Pour cela, on pose :

$$G = \Gamma' \quad \text{où} \quad \Gamma = \{e_k^* \circ f^i, i \in \mathbb{N}\}.$$

G est l'ensemble des $x \in E$ tel que la k-ième coordonnée de $f^i(x)$ est nulle pour tout $i \in \mathbb{N}$. Montrons que $E = F \oplus G$.

Soit $y \in F \cap G$. Si $y \neq 0$, on peut écrire $y = a_1 e_1 + \cdots + a_p e_p$ avec $p \leq k$ et $a_p \neq 0$. On compose par $e_k^* \circ f^{k-p}$:

$$0 = e_k^*(a_1 e_{k-p+1} + \cdots + a_p e_k) = a_p\,,$$

ce qui est absurde, donc $F \cap G = \{0\}$.

Puisque $\dim(G) = n - \dim(\Gamma)$, pour montrer que $\dim(F) + \dim(G) = n$, il suffit de montrer que $\dim(\Gamma) = k$. L'application

$$\begin{array}{rccc} \varphi & : & \mathcal{L}_f & \longrightarrow & \mathrm{vect}(\Gamma) \\ & & g & \longmapsto & e_k^* \circ g \end{array},$$

est linéaire et surjective par définition. Soit $g \in \operatorname{Ker}\varphi$. Si $g \neq 0$, on peut écrire $g = a_1 \,\mathrm{id}_E + \cdots + a_p f^p$ où $p \leq k$ et $a_p \neq 0$. On aurait :

$$0 = e_k^* \circ g(f^{k-p}(x)) = e_k^*(a_1 f^{k-p}(x) + \cdots + a_p f^k(x)) = a_p\,,$$

ce qui est absurde. Donc φ est injective et c'est donc un isomorphisme. On en déduit que $\dim \Gamma = \dim \mathcal{L}_f = k$.

Il reste à montrer que $\Pi_{f_{|G}} \mid \Pi_{f_{|F}}$ pour pouvoir appliquer la récurrence, ce qui est la cas. En effet, G est f-stable donc le polynôme minimal de $f_{|G}$ divise $\Pi_f = P_x = \Pi_{f_{|F}}$. □

Démonstration de l'unicité. Soient $F_1, \ldots, F_r$ et $G_1, \ldots, G_s$ deux suites de sous-espaces f-stables vérifiant les conditions (i), (ii) et (iii). On note $P_i = \Pi_{f_{|F_i}}$ et $Q_i = \Pi_{f_{|G_i}}$.

Supposons que les listes $(P_1, \ldots, P_r)$ et $(Q_1, \ldots, Q_s)$ soient différentes et notons j le premier indice tel que $P_j \neq Q_j$ (un tel indice existe toujours, si $r = s$ c'est l'hypothèse et si $r \neq s$ c'est toujours le cas car $\sum_i \deg(P_i) = n = \sum_i \deg(Q_i)$). $j \neq 1$ car on a $P_1 = \Pi_f = Q_1$. On a : $E = F_1 \oplus \cdots \oplus F_r$. Pour tout $k \geq j$, on a $P_j(f)(F_k) = \{0\}$. On en déduit, les F_i étant f-stables :

$$P_j(f)(E) = P_j(f)(F_1) \oplus \cdots \oplus P_j(f)(F_{j-1})\,.$$

Pour des raisons similaires :

$$P_j(f)(E) = P_j(f)(G_1) \oplus \cdots \oplus P_j(f)(G_{j-1}) \oplus P_j(f)(G_j) \oplus \cdots \oplus P_j(f)(G_s)\,.$$

Pour tout $1 \leq i \leq j-1$, il existe des bases telles que les matrices de $f_{|F_i}$ et $f_{|G_i}$ dans ces bases soient égales à $\mathcal{C}(P_i) = \mathcal{C}(Q_i)$. On en déduit que $\dim P_j(f)(F_i) = \dim P_j(f)(G_i)$. Ainsi :

$$0 = \dim P_j(f)(G_j) = \cdots = \dim P_j(f)(G_s)\,,$$

En particulier, P_j annule $f_{|G_j}$ donc $Q_j \mid P_j$. Comme P_j et Q_j jouent des rôles symétriques, on en déduit que $P_j = Q_j$, ce qui est absurde. Nécessairement, $r = s$ et $(P_1, \ldots, P_r) = (Q_1, \ldots, Q_s)$. □

Application (Réduction de Frobenius)**.** *Si $P_1, \ldots, P_r$ désigne la suite des invariants de similitudes de $f \in \mathcal{L}(E)$, il existe une base de E dans laquelle la matrice de f est égal à :*

$$\begin{pmatrix} \mathcal{C}(P_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(P_r) \end{pmatrix}.$$

Application. *Deux endomorphismes sont semblables si et seulement si ils ont mêmes invariants de similitudes.*

Compléments

On démontre l'assertion : $\exists x \in E,\ P_x = \Pi_f$.

Lemme. *Soient $x_1, \ldots, x_p \in E$ tels que $E_{x_1}, \ldots, E_{x_p}$ soient en somme directe. Alors : $P_{x_1 + \cdots + x_p} = \mathrm{ppcm}(P_{x_1}, \ldots, P_{x_p})$.*

Démonstration. Le lemme se démontre par récurrence sur p. Soient x et y dans E tels que $E_x \cap E_y = \{0\}$. Il suffit de montrer que $P_{x+y} = \mathrm{ppcm}(P_x, P_y)$.

Comme $P_{x+y}(x+y) = 0$, on a par linéarité $P_{x+y}(x) = -P_{x+y}(y) \in E_x \cap E_y = \{0\}$ donc $P_x \mid P_{x+y}$ et $P_y \mid P_{x+y}$, ce qui montre que $\mathrm{ppcm}(P_x, P_y) \mid P_{x+y}$.

Puisque $P_x \mid \mathrm{ppcm}(P_x, P_y)$ et $P_y \mid \mathrm{ppcm}(P_x, P_y)$, on a $\mathrm{ppcm}(P_x, P_y)(x) = \mathrm{ppcm}(P_x, P_y)(y) = 0$ et par linéarité $\mathrm{ppcm}(P_x, P_y)(x+y) = 0$, ce qui montre que $P_{x+y} \mid \mathrm{ppcm}(P_x, P_y)$. □

Lemme. *Soient $x_1, \dots, x_p \in E$ tels que $P_{x_1}, \dots, P_{x_p}$ soient premiers entre eux deux à deux. Alors : $E_{x_1+\cdots+x_p} = E_{x_1} \oplus \cdots \oplus E_{x_p}$.*

Démonstration. On montre le cas $p = 2$ et un raisonnement par récurrence permet de conclure. Soient x et y dans E tels que P_x et P_y soient premiers entre eux.

Soit $z \in E_x \cap E_y$. Il existe P et Q dans $\mathbb{K}[X]$ tel que : $z = P(f)(x) = Q(f)(y)$. On a :

$$0 = (PP_x)(f)(x) = P_x(f)(z) = (P_xQ)(f)(y)$$

donc $P_y \mid P_xQ$ et par le théorème de Gauss $P_y \mid Q$. Donc $z = 0$ et $E_x \cap E_y = \{0\}$. D'après le lemme précédent, on a : $P_{x+y} = \text{ppcm}(P_x, P_y) = P_xP_y$. Donc :

$$\dim(E_{x+y}) = \deg(P_{x+y}) = \deg(P_x) + \deg(P_y) = \dim(E_x) + \dim(E_y)\,.$$

Puisque : $E_{x+y} \subset E_x + E_y$, on a bien : $E_{x+y} = E_x \oplus E_y$. □

Démonstration de l'assertion. On montre d'abord que si M est un facteur irréductible de Π_f de multiplicité α alors il existe $x \in \operatorname{Ker} M^\alpha(f)$ tel que $P_x = M^\alpha$. On écrit $\Pi_f = M^\alpha N$ où N est premier avec M. D'après le lemme des noyaux :

$$E = \operatorname{Ker} M^\alpha(f) \oplus N(f)\,.$$

Pour tout $x \in \operatorname{Ker} M^\alpha(f)$, on a $M^\alpha(f)(x) = 0$ donc $P_x \mid M^\alpha$ et comme M est irréductible, il existe $\beta_x \leq \alpha$ tel que $P_x = M^{\beta_x}$. Supposons, par l'absurde, que pour tout x dans $\operatorname{Ker} M^\alpha(f)$, on ait $\beta_x < \alpha$. On aurait : $\forall x \in \operatorname{Ker} M^\alpha(f),\ P_x \mid M^{\alpha-1}$. Cela impliquerait que $\operatorname{Ker} M^\alpha(f) = \operatorname{Ker} M^{\alpha-1}(f)$. D'après le lemme des noyaux, on aurait : $E = \operatorname{Ker}(M^{\alpha-1}N)(f)$ et $M^{\alpha-1}N$ serait un polynôme annulateur de f, ce qui contredit la minimalité de Π_f.

Soit maintenant $\Pi_f = \prod_{i=1}^p M_i^{\alpha_i}$ la décomposition en facteurs irréductibles de Π_f. D'après ce qui précède, pour tout i il existe $x_i \in \operatorname{Ker} M_i^{\alpha_i}(f)$ tel que $P_{x_i} = M_i^{\alpha_i}$. Le deuxième lemme montre que : $E_{x_1+\cdots+x_p} = E_{x_1} \oplus \cdots \oplus E_{x_p}$. Le premier lemme montre que :

$$P_{x_1+\cdots+x_p} = \text{ppcm}(P_{x_1}, \dots, P_{x_p}) = \prod_{i=1}^p P_{x_i} = \Pi_f\,.$$

□

Références

– Gourdon, Algèbre, p.289 (théorème + compléments) et p.178 (compléments)

Leçons concernées : 150, 153, 154, 159

40 Équation matricielle

Contexte

Développement

Théorème. *Soit $A, B \in \mathcal{M}_n(\mathbb{R})$. On suppose que les valeurs propres de A et de B sont strictement négatives. Alors, pour tout $C \in \mathcal{M}_n(\mathbb{R})$, l'équation :*

$$AX + XB = C$$

a pour unique solution :

$$X = -\int_0^{+\infty} e^{tA} C e^{tB} \,\mathrm{d}t\,.$$

Lemme. *Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice dont les valeurs propres $\lambda_1, \ldots, \lambda_n$ (comptées avec leur multiplicité) soient strictement négatives. Si $\|\cdot\|$ désigne une norme d'algèbre alors :*

$$\exists \alpha, K > 0,\ \forall t \geq 0,\ \left\|e^{tA}\right\| \leq K e^{-\alpha t}\,.$$

Démonstration du lemme. Soit $A = D + N$ la décomposition de Dunford de A, avec D diagonalisable, N nilpotente et $DN = ND$. Il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $D = PD_1P^{-1}$ où $D_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Par linéarité et continuité de l'application $B \mapsto PBP^{-1}$, on a : $\forall t \geq 0,\ e^{tD} = P\,\mathrm{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})P^{-1}$.

On note $\|(a_{ij})_{ij}\|_\infty = \sup_{ij} |a_{ij}|$ la norme ℓ^∞ sur $\mathcal{M}_n(\mathbb{C})$. Alors :

$$\forall t \geq 0,\ \left\|e^{tD_1}\right\|_\infty = \left\|\mathrm{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})\right\|_\infty \leq e^{-2\alpha t}\,,$$

où $-2\alpha = \sup_i \lambda_i < 0$. Puisqu'on est en dimension finie, les normes sont équivalentes et il existe $K_1 > 0$ tel que : $\|\cdot\|_\infty \leq K_1 \|\cdot\|$. On en déduit :

$$\left\|e^{tD}\right\| \leq K_2 e^{-2\alpha t}\,,$$

avec $K_2 = \|P\|\, K_1 \left\|P^{-1}\right\|$. Comme N est nilpotente, on a :

$$\forall t \geq 0,\ e^{tN} = I_n + tN + \cdots + \frac{(tN)^{n-1}}{(n-1)!}\,,$$

donc $\left\|e^{tN}\right\| = o(t^n)$ lorsque t tend vers $+\infty$. Puisque D et N commutent :

$$\forall t \geq 0,\ \left\|e^{tA}\right\| = \left\|e^{tD}e^{tN}\right\| \leq \left\|e^{tN}\right\| e^{-\alpha t} K_2 e^{-\alpha t}\,,$$

et puisque $\left\|e^{tN}\right\| e^{-\alpha t}$ est borné, on en déduit le lemme. □

Démonstration du théorème. Le problème de Cauchy suivant

$$\begin{cases} Y' = AY + YB \\ Y(0) = C \end{cases}$$

est linéaire à coefficients constants, d'après le théorème de Cauchy-Lipschitz linéaire, il existe une unique solution globale, c'est l'application :

$$Y : t \in \mathbb{R} \longmapsto e^{tA} C e^{tB}\,.$$

On intègre entre 0 et t :

$$\forall t \geq 0,\ Y(t) - C = A\left(\int_0^t Y(s)\,\mathrm{d}s\right) + \left(\int_0^t Y(s)\,\mathrm{d}s\right) B\,.$$

Le lemme montre l'existence de $\alpha > 0$ et $M > 0$ tels que :

$$\forall t \geq 0,\ \left\|e^{tA}\right\| \leq M e^{-\alpha t} \quad \text{et} \quad \left\|e^{tB}\right\| \leq M e^{-\alpha t}\,.$$

On en déduit :

$$\forall t \geq 0,\ \|Y(t)\| \leq \left\|e^{tA}\right\| \|C\| \left\|e^{tB}\right\| \leq Ke^{-2\alpha t},$$

avec $K = M^2 \|C\|$. Ainsi, l'intégrale $\int_0^{+\infty} Y(s)\,\mathrm{d}s$ converge absolument donc converge et $Y(t)$ tend vers 0 lorsque t tend vers $+\infty$. On en déduit :

$$C = AX + XB \quad \text{avec} \quad X = -\int_0^{+\infty} e^{tA}Ce^{tB}\,\mathrm{d}t\,.$$

Pour montrer l'unicité de X, il suffit de montrer que l'endomorphisme

$$\begin{array}{cccc} \Phi & : & \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ & & X & \longmapsto & AX + XB \end{array}$$

est injective. Or on vient de montrer sa surjectivité et puisqu'on est en dimension finie, cela implique son injectivité. □

Références

– Gourdon, Analyse, p.390

Leçons concernées : 156, 157, 221