MASTER DE MATHÉMATIQUES M1
SUJETS T.E.R. - 2014/2015

FRÉDÉRIC KLOPP
RESPONSABLE DES T.E.R.

Date: 19 mars 2015.

Enseignant(e)	Bureau	Courriel	Intitulé du ou des sujets	Étudiant(e)(s)
Omer Adelman	16-26-2-10	Omer.Adelman@upmc.fr	Le théorème de densité de Lebesgue	Pedro Paez David Villamar
Cédric Boutillier	16-26-2-08	cedric.boutillier@upmc.fr	Processus semi-markoviens à temps discret	Yating Liu
Jean-Fran cois Babadjian	16-26-301	jean-francois.babadjian@upmc.fr	Mesures et dimension de Hausdorff	Clément Pagès
Xavier Claeys	15-25-310	claeys@ann.jussieu.fr	Élements finis multi-échelles pour un problème elliptique avec petite inclusion	
			Calcul d'intégrales quasi-singulières par méthode de quadrature adaptative	
Albert Cohen		cohen@ann.jussieu.fr	Classification automatique de signaux bioacoustiques	Randall Balestriero
Grégory Ginot	15-25-523	gregory.ginot@imj-prg.fr	Homologie des espaces cellulaires et applications	Yanis Amirou
Vincent Humilière	15-25-528	vincent.humiliere@imj-prg.fr	Quatrième problème de Hilbert et géométrie symplectique	
Pierre-Vincent Koseleff	15-25-506	pierre-vincent.koseleff@imj-prg.fr	Calcul des polynômes cyclotomiques	Antoine Rodriguez
Catherine Matias	15-16-216	catherine.matias@upmc.fr	État de l'art sur la loi de Poisson binomiale et application en assurance vie.	Yongshen Chen Hacine Yosni
Jan Nekovar		jan.nekovar@imj-prg.fr	Nombres p-adiques, distributions p-adiques et la fonction zêta p-adique	
Alexandru Oancea	15-25-509	alexandru.oancea@imj-prg.fr	le théorème d'Atiyah-Jänich	
Marie Postel	15-25-3-13	marie.postel@upmc.fr	Collision de satellites	
Tabea Rebafka	15-16-214	tabea.rebafka@upmc.fr	Estimation par algorithme EM d'un modèle de mélange. Application à la prédiction du nombre et de la taille des défauts d'un composant industriel.	Ibrahima Diallo Yanqing Liu
Etienne Roquain	15-16-2-13	etienne.roquain@upmc.fr	Confirmation de diagnostics	M. Bezza
Philippe Saint-Pierre	15-25-2-14	philippe.saint_pierre@upmc.fr	Validation de pronostics	Laura Chelfi
Damien Simon	16-26-2-07	damien.simon@upmc.fr	Entrelacements de semi-groupes markoviens: construction et applications	Thibaut Lemoine
Camille Tardif	16-26-1-18	camille.tardif@upmc.fr	Recuit simulé appliqué au problème du voyageur de commerce	Pierre Martin

Le théorème de densité de Lebesgue

(et deux autres théorèmes)

Thème proposé par Omer Adelman

Le théorème de densité de Lebesgue admet plusieurs déclinaisons. Une des plus simples dit ceci : si B est un borélien de $\mathbb{R}$, alors, pour Lebesgue-presque tout $x \in B$,

$$\liminf_{\substack{s \leqslant x \leqslant t \\ 0 < t-s \to 0}} \frac{\text{Lebesgue}(B \cap [s,t])}{t-s} = 1$$

(très facile à démontrer lorsque B est l'union d'un ensemble dénombrable d'intervalles). En d'autre termes, Lebesgue-presque tout $x \in B$ est **point de densité** de B (et, évidemment, Lebesgue-presque tout $x \in \mathbb{R} \setminus B$ est point de densité de $\mathbb{R} \setminus B$ – et il est aisé de convertir cela en un énoncé portant sur des partitions dénombrables de $\mathbb{R}$ en des ensembles boréliens).

Le théorème de densité de Lebesgue admet quelques versions légèrement plus corsées : le cadre unidimensionnel ne s'impose pas, même le contexte euclidien peut être délaissé au profit d'un contexte métrique un peu plus général, et on peut renoncer à la mesurabilité (et parler de mesure extérieure, par exemple).

Un espace de probabilité étant donné, **le théorème de Radon-Nikodym** permet de parler de la probabilité conditionnelle d'un évènement (ou de l'espérance conditionnelle de variables aléatoires réelles positives et de certaines autres variables aléatoires) relativement à une sous-tribu de la tribu de référence qui accompagne notre espace de probabilité.

En réfléchissant un peu, on peut observer que toute version de l'espérance conditionnelle (d'une variable aléatoire réelle bornée, par exemple) relativement à une sous-tribu $\mathscr{G}$ est essentiellement égale à une version de l'espérance conditionnelle relativement à une sous-tribu de $\mathscr{G}$ engendrée par un ensemble dénombrable (éventuellement fini).

Nous avons parfois affaire à l'espérance conditionnelle d'une variable aléatoire (réelle, par exemple) X non pas relativement à une seule sous-tribu, mais relativement au terme général d'une suite de sous-tribus, et parler du comportement asymptotique de la famille d'espérances conditionnelles peut avoir un sens. **Le théorème de convergence des martingales** parle justement de la presque sûre convergence de certaines suites $(E(X \mid \mathscr{G}_n))_{n\in\mathbb{N}}$ où X est une variable aléatoire d'un certain genre (réelle bornée, par exemple) et $(\mathscr{G}_n)_{n\in\mathbb{N}}$ est une suite croissante de sous-tribus (croissante relativement à l'inclusion, bien sûr).

Il arrive qu'une famille intéressante de sous-tribus soit indexée par un ensemble totalement ordonné autre que $\mathbb{N}$ ou même par un ensemble ordonné filtrant qui n'est pas totalement ordonné, et nous avons une assez vaste famille de théorèmes de convergence des martingales (ou des sous-martingales)... Mais envoyer un étudiant en M1 se promener essentiellement seul dans un terrain aussi accidenté serait sans doute passablement méchant.

Chacun des trois théorèmes est intéressant – et les démonstrations sont elles aussi intéressantes. Votre mission, si vous décidez de l'accepter, consistera à comprendre ces théorèmes (et quelques-unes de leurs démonstrations) et à étudier les *liens* entre ces trois théorèmes. (Par exemple, présenter le théorème de densité de Lebesgue comme un corollaire ou comme une instance particulière du théorème de convergence des martingales, ce pourrait être une assez bonne chose.)

Pour l'instant, je préfère ne pas en dire plus et ne pas donner de références. (Dans une certaine mesure, la recherche de documents utiles peut être considérée comme une partie du travail, et la débrouillardise dans ce domaine s'acquiert en pratiquant.)

Mesures et dimension de Hausdorff

Sujet proposé par

Jean-François Babadjian,
Bureau 16-26-301, LJLL, jean-francois.babadjian@upmc.fr

Les mesures de Hausdorff de dimension k dans $\mathbb{R}^n$ permettent de donner une définition intrinsèque des notions de longueur (pour $k = 1$) et d'aire (pour $k = 2$) de n'importe quel sous ensemble de $\mathbb{R}^n$, sans aucune référence à la paramétrisation. Il s'agit d'un outil extrêmement utile, en particulier dans les problèmes variationnels de théorie géométrique de la mesure.

L'objet de ce TER consistera à étudier la construction de ces mesures par la méthode de Carathéodory ainsi que leurs propriétés de base. Il s'agira ensuite de comprendre la notion de dimension de Hausdorff qui coïncide avec la notion usuelle de dimension pour les ensembles réguliers. On s'intéressera également à certains ensembles fractals comme par exemple l'ensemble de Cantor ou le flocon de neige de Von Koch qui sont des ensembles de dimension de Hausdorff réelle non entière. On calculera explicitement leur dimension de Hausdorff ainsi que leur mesure de Hausdorff correspondante. Si le temps le permet, on essaiera aussi de comprendre les formules de l'aire et de la coaire qui sont des généralisations de la formule usuelle de changement de variable et du théorème de Fubini.

Références

[1] L. C. Evans, R. G. Gariey : Measure theory and fine properties of functions, CRC Press, Boca Raton, (1992).

[2] F. Maggi : Sets of finite perimeter and geometric variational problems. An introduction to geometric measure theory, Cambridge University Press (2012).

Processus semi-markoviens à temps discret

Cédric Boutillier *

Les processus de Markov ont été beaucoup utilisés pour la modélisation dans divers domaines, grâce à la propriété de Markov, qui dit qu'étant donné l'état passé et présent d'un système, son évolution future ne dépend que de l'état présent.

Le problème est que cette hypothèse impose des contraintes fortes sur la distribution du temps que le système doit passer dans chaque état.

Les processus semi-markoviens sont une façon de relaxer ces contraintes, tout en conservant une forme, affaiblie certes, de propriété de Markov.

Nous proposons d'étudier ces processus en temps discret. Nous démontrerons les propriétés-clé de ces processus, telles que le théorème de renouvellement markovien, et verrons leur pertinence pour la modélisation des statistiques d'apparition de motifs dans des suites chaînes d'ADN, ou si le temps le permet en traduction automatique [2, 3].

Le point de départ de cette étude sera constitué des premiers chapitres de l'ouvrage [1].

Références

[1] Vlad Stefan Barbu et Nikolaos Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications* Lecture Notes in Statistics 191 (2008), Springer.

[2] Peter F. Brown, Vincent J. Della Pietra, Stephen A. Della Pietra et Robert L. Mercer *The mathematics of statistical machine translation : parameter estimation*. Computational Linguistics (1993) Vol. 19, n2. pp. 263-311.

[3] Philipp Koehn. *Statistical Machine Translation*. Cambridge University Press. (2009).

*cedric.boutillier@upmc.fr

Élements finis multi-échelles pour un problème elliptique avec petite inclusion

proposé par

Xavier Claeys, bureau 15-25-310, LJLL, claeys@ann.jussieu.fr

Les méthodes standard de résolution d'EDP de type éléments finis ou différences finies basées sur un maillage sont normalement incapables de restituer fidèlement un phénomène dont l'échelle de longueur caractéristique est plus petite que la taille moyenne d'une maille.

Pour pouvoir traiter des problèmes impliquant plusieurs échelles de longueur caractéristiques, Hou et Efendiev [1] ont proposé une approche de type éléments finis dans laquelle certaines mailles recoivent un traitement ad-hoc particulier adapté aux petites échelles. Cette méthode a initialement été conçu pour traiter des problèmes où les phénomènes à petites échelles interviennent sur une zone étalée. Mais qu'en est-il dans le cas où les phénomènes de ptite échelle sont concentré sur un point du domaine de calcul.

Le but de ce TER consistera dans un premier temps à comprendre les éléments finis multi-échelles pour un problème en 1D de petite inclusion à travers un travail bibliographique. Dans un deuxième temps, on cherchera à tester les performances de cette méthode sur pour le problème suivant:

$$-\frac{1}{x}\frac{d}{dx}\Big(x\frac{du_\delta}{dx}\Big)+u_\delta(x)=f(x),\quad x\in[\delta,1],\qquad u_\delta(\delta)=u_\delta(1)=0$$

où f est une fonction connue, et δ est un paramètre qui a vocation à tendre vers 0.

References

[1] Y. Efendiev and T.Y. Hou. *Multiscale finite element methods*, volume 4 of Surveys and Tutorials in the Applied Mathematical Sciences. Springer, New York, 2009. Theory and applications.

[2] T. Hughes, G. Feijoo, L. Mazzei, and J.-B. Quincy, *The variational multiscale method: A paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg., 166 (1998), pp. 3âĂŞ24.

Calcul d'intégrales quasi-singulières par méthode de quadrature adaptative

proposé par

Xavier Claeys, bureau 15-25-310, LJLL, claeys@ann.jussieu.fr

Les éléments finis de bord, méthode de résolution des équations d'ondes linéaires enmilieu homogène maintenant largement répandue, nécessite le calcul d'intégrales de la forme

$$\int_{T_1}\int_{T_2}\frac{d\boldsymbol{x}d\boldsymbol{y}}{|\boldsymbol{x}-\boldsymbol{y}|}$$

où $T_1, T_2 \subset \mathbb{R}^3$ sont deux triangles. Dans une situation standard, on peut supposer que ces deux triangles sont soit complètement disjoint, soit avec juste un noeud en commun, ou avec juste une arete en commun, ou encore ils coincident. Sous ces hypothèses, il existe des méthodes de quadrature relativement éfficace pour calculer ce type d'intégrale.

On souhaite s'intéresser au cas *non* standard où les deux triangles T_1, T_2 sont autorisés à s'intersecter d'une façon un peu quelconque. Le but de ce TER sera de tester, pour ce cas de figure, une méthode de quadrature adaptative.

References

[1] M. Gander, W. Gander and F. Kwok, *Scientific Computing: an Introduction Using Maple and MATLAB*. Texts in Computational Science and Engineering 11, Springer 2014.

[2] W. Gander and W. Gautschi, *Adaptive quadrature revisited*. BIT Numerical Mathematics, 40 (2000), no. 1, 84-101.

CLASSIFICATION AUTOMATIQUE DE SIGNAUX BIOACOUSTIQUES.

Sujet proposé par : Stéphane MALLAT et Hervé GLOTIN
Encadrant UPMC : Albert COHEN, LJLL
Étudiant : Randall Balestriero

L'analyse massive d'enregistrements environnementaux est susceptible de révéler des informations cruciales pour la préservation de la diversité écologique. C'est dans ce cadre que le challenge international BirdCLEF [1] propose de développer des méthodes automatiques d'identification de certaines espèces d'oiseaux à partir de leurs chants. L'édition 2015 de BirdCLEF comprend une base d'évaluation de 33000 enregistrements d'origines très diverses, issus d'un millier d'espèces différentes. Or, la grande variabilité intra-classe (à une seule espèce correspondent souvent plusieurs types de cris) et la faible discriminabilité inter-classes (des espèces distinctes peuvent présenter des chants très semblables) de cette base rendent peu efficaces les modèles classiques de description sonore. En revanche, la théorie de la transformée de scattering, développée à l'École normale supérieure, permet une typologie des sons portant sur de multiples échelles temporelles, d'où une performance accrue dans le cadre de la reconnaissance de la parole [2]. Ce TER consistera à étendre le champ d'application de la transformée de scattering aux signaux bioacoustiques, en mesurant l'efficacité de plusieurs de ses architectures et paramétrages. En particulier, on pourra envisager de recourir à une transformée de scattering multivariables, variante récente consistant à capturer la géométrie des signaux selon les axes mélodiques (chromas) et harmoniques (octaves) en plus de l'axe temporel. L'algorithme obtenu sera comparé à l'état de l'art dans le domaine [3]. Ce travail aboutira sur une participation au challenge BirdCLEF 2015, afin d'évaluer les systèmes proposés selon un protocole de grande envergure.

BIBLIOGRAPHIE :

[1] Site officiel du challenge BirdCLEF. http ://www.imageclef.org/node/180.

[2] J. Bruna et S. Mallat. Deep Scattering Spectrum. http ://arxiv.org/abs/1304.6763

[3] D. Stowell et M. Plumbley. Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. http ://arxiv.org/abs/1405.6524

Quatrième problème de Hilbert et géométrie symplectique

Encadrant: Vincent Humilière, Bureau 15-25-528, humi@math.jussieu.fr

Enoncé lors du congrès international des mathématiciens en 1900, le quatrième des 23 problèmes de Hilbert peut se formuler en ces termes:

Caractériser toutes les géométries pour lesquelles la distance entre deux points est réalisée par un segment de droite de $\mathbb{R}^n$.

Bien sûr, le terme "géométrie" est vague et plusieurs interprétations sont possibles. La plupart des solutions existantes reposent sur la (très jolie) formule de Crofton, selon laquelle: *la longueur d'une courbe du plan est le nombre moyen de ses points d'intersection avec une droite quelconque.*

Dans le cadre des géométries Finsleriennes une caractérisation a été trouvée récemment par J.C. Alvarez Paiva. Cette caractérisation utilise la géométrie symplectique de l'espace des droites et fait donc interagir deux géométries (Finslerienne et symplectique) a priori sans lien.

Le but de ce TER est avant tout de se familiariser avec les objets apparaissant dans cette histoire (métriques, géodésiques, formes symplectiques, sous-variétés Lagrangienne). L'objectif sera ensuite de comprendre la solution au quatrième problème de Hilbert formulée dans le langage de la géométrie symplectique, à l'aide de la formule de Crofton. Comme prérequis, il sera préférable d'avoir suivi un cours de géométrie différentielle.

Références:

- S. Tabachnikov. Geometry and billiards.
AMS. 2005. http://www.math.psu.edu/tabachni/Books/billiardsgeometry.pdf

- J.C. Alvarez Paiva. Symplectic geometry and Hilbert's fourth problem.
To appear in JDG. http://math.univ-lille1.fr/ alvarez/pdfs/symp.pdf

Calcul des polynômes cyclotomiques

Sujet présenté par Pierre-Vincent Koseleff
pierre-vincent.koseleff@imj-prg.fr

Dans un article récent, Arnold et Monagan ([AM10]) proposent une méthode pour calculer efficacement et rapidement les polynômes cyclotomiques Φ_n.

Il s'agit de comprendre les méthodes utilisées et mises en œuvre, ainsi que d'expliquer, dans le cadre du programme du Master de Mathématiques, divers points restés obscurs ou imprécis : propriétés des polynômes cyclotomiques, irréductibilité, taille des coefficients; multiplications et divisions rapides, transformation de Fourier rapide.

L'ouvrage [GG13] pourra utilement être consulté.

Le travail peut mener à des calculs explicites et un premier contact avec des logiciels de calcul formel (Sage, Maple, etc).

Références

[AM10] A. Arnold, M. Monagan, *A high-performance algorithm for calculating cyclotomic polynomials. In* Proceedings of the 4th International Workshop on Parallel and Symbolic Computation, pp 112–120, 2010.

[GG13] J. von zur. Gathen, J. Gerhard, *Modern Computer Algebra*. Cambridge University Press, troisième édition, 2013.

ÉTAT DE L'ART SUR LA LOI DE POISSON BINOMIALE ET APPLICATION EN ASSURANCE VIE.

Sujet proposé par : Vincent LEPEZ, SCOR (vlepez@scor.com)
Encadrant UPMC : Catherine MATIAS, CNRS-LPMA (catherine.matias@upmc.fr)

Voir sujet page suivante

Etat de l'Art sur la loi de Poisson binomiale et application en assurance vie

Sujet de TER pour l'année 2014-2015

Contact : Vincent Lepez (vlepez@scor.com)

Si $\lambda_1, \ldots, \lambda_n \in [0\,;1]$ et, pour $i \in \{1, \ldots, n\}, X_i \sim \mathcal{B}(\lambda_i)$ sont des lois de Bernoulli indépendantes, alors la somme $Y = X_1 + \ldots + X_n$ suit une loi de Poisson binomiale. Généralisant la loi binomiale qui correspond au cas $\lambda_1 = \ldots = \lambda_n$, elle se comporte asymptotiquement comme une loi de Poisson de paramètre $\lambda_1 + \ldots + \lambda_n$ (on pourra se référer à Le Cam, 1960, pour un encadrement de la vitesse de convergence).

Cette loi joue un rôle fondamental en assurance vie. Elle intervient lorsque l'on s'intéresse à la distribution du nombre de décès dans un portefeuille d'assurés. On peut en effet associer à chaque assuré une probabilité de décès annuelle qui dépendra de l'âge ainsi que du sexe de l'assuré, et éventuellement d'autres variables. Ces décès pouvant de plus être considérés comme indépendants, le nombre total de décès dans le portefeuille chaque année suit ainsi une loi de Poisson binomiale.

Déterminer directement la distribution de cette loi en étudiant toutes les combinaisons possibles n'est pas envisageable car le nombre de combinaisons augmente de manière exponentielle avec le nombre de composantes, avec pour n lois de Bernoulli, 2^n combinaisons possibles. Néanmoins, différentes méthodes ont été développées afin de permettre son calcul en un temps plus raisonnable (on pourra consulter Chen, 1994; Shah, 1994; Chen, 1997; Fernandez, 2010, pour une liste non-exhaustive de références sur le sujet).

L'objectif de ce TER consiste dans un premier temps à établir un état de l'art des résultats théoriques établis concernant la loi de Poisson binomiale. On s'intéressera ensuite aux différentes méthodes permettant le calcul exact de sa distribution, et on proposera une implémentation d'au moins deux de ces méthodes, en comparant leurs performances respectives.

Dans un second temps, on utilisera les résultats obtenus afin d'étudier la convergence de la loi de Poisson binomiale dans un cas concret. A partir de données issues de portefeuilles d'assurance vie et de simulations, on déterminera la distribution du nombre total de décès et on étudiera pour différentes tailles de portefeuille si l'approximation par une loi de Poisson se trouve justifiée en pratique.

References

Chen, S. X. (1997). Statistical applications of the poisson-binomial and conditional bernoulli distributions. *Statistica Sinica 7*, 875–892.

Chen, X. H. (1994). Weighted finite population sampling to maximize entropy. *Biometrika 81*, 457.

Fernandez, M. (2010). Closed-form expression for the poisson-binomial probability density function. *IEEE Transactions on Aerospace Electronic Systems 46*, 803–817.

Le Cam, L. (1960). An approximation theorem for the poisson binomial distribution. *Pacific Journal of Mathematics 10*, 1181–1197.

Shah, B. K. (1994). On the distribution of the sum of independent integer valued random variables. *American Statistician 27*, 123–124.

Nombres p-adiques, distributions p-adiques et la fonction zêta p-adique

Jan Nekovář

On peut donner un sens à la fonction zêta de Riemann $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ pour tout $s \in \mathbf{C} \setminus \{1\}$. Les valeurs en les entiers négatifs $\zeta(1-2n) = -B_{2n}/2n$ sont des nombres rationnels remarquables, qui sont liées par les congruences de Kummer $\zeta(1-2n) \equiv \zeta(1-2m) \pmod{p}$ si $n \equiv m \pmod{p-1}$. Les congruences analogues modulo toutes les puissances d'un nombre premier p s'expliquent en termes de la fonction zêta p-adique de Kubota-Leopoldt.

Le but du TER sera de comprendre au moins une construction de cette fonction (par exemple, celle qui utilise les distributions p-adiques donnés par les polynômes de Bernoulli).

Référence : N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, Graduate Texts in Mathematics **58**, Springer, New York-Heidelberg, 1977.

Sujet de TER : le théorème d'Atiyah-Jänich

Sujet proposé par Alexandru Oancea

Le théorème d'Atiyah-Jänich établit un lien surprenant entre deux types d'objets : les objets de nature analytique que sont les opérateurs de Fredholm sur un espace de Hilbert séparable, et les objets de nature géométrique que sont les fibrés vectoriels localement triviaux.

Un opérateur de Fredholm $T : V \to V$ sur un espace de Hilbert V de dimension infinie est un opérateur linéaire et continu tel que la dimension de son noyau et la dimension de son conoyau soient finies. On note $\mathcal{F}$ l'espace de ces opérateurs.

Le théorème d'Atiyah-Jänich affirme que $\mathcal{F}$ est un espace classifiant pour la K-théorie : pour tout espace compact X l'on a un isomorphisme canonique

$$[X, \mathcal{F}] \cong K(X)$$

entre le groupe des classes d'homotopie d'applications continues de X dans $\mathcal{F}$, et le groupe de K-théorie de X, noté $K(X)$. Ce dernier est un groupe qui contient naturellement $Vect(X)$, l'ensemble des classes d'équivalence de fibrés vectoriels sur X.

L'une des conséquences philosophiques de ce théorème est que les opérateurs de Fredholm trouvent naturellement leur place en topologie. Ce point de vue se manifeste de manière éclatante dans le théorème de l'indice d'Atiyah-Singer. Le théorème d'Atiyah-Jänich constitue une porte d'entrée vers ce cercle d'idées.

Ce sujet de TER utilise uniquement des notions élémentaires d'analyse fonctionnelle (opérateurs linéaires et continus, opérateurs de Fredholm, opérateurs compacts), et des notions élémentaires concernant les fibrés vectoriels (trivialisation, opérations).

Bibliographie.

M. Atiyah, *K-theory*, W.A. Benjamin, Inc., New York, 1967. (Chapitre 1, §2.1, Appendice).

K. Jänich, Vektorraumbündel und der Raum der Fredholm-Operatoren. *Math. Ann.* 161 (1965) 129–142.

COLLISION DE SATELLITES

Sujet proposé par : Max CERF, EADS
Contact UPMC : Marie POSTEL, LJLL (marie.postel@upmc.fr)

Le nombre de débris spatiaux en orbite au voisinage de la Terre ne cesse d'augmenter depuis les débuts de la conquête spatiale. Ces débris de toutes tailles représentent un danger important pour les satellites en raison de leur vitesse, de l'ordre de 28000 km/h. Un débris d'un centimètre de diamètre peut suffire à rendre inopérant un satellite d'une valeur de plusieurs millions d'euros. Il est nécessaire de surveiller constamment les trajectoires de ces milliers de débris, afin d'évaluer les risques de collision avec les satellites opérationnels, et le cas échéant de faire réaliser au satellite menacé une manœuvre d'évitement.

L'objectif du TER est de réaliser un logiciel pour estimer la probabilité de collision entre deux objets en orbite. On suppose que les deux objets se déplacent sur des orbites képlériennes (ellipses parcourues suivant la loi des aires), ce qui constitue une approximation suffisamment précise sur une durée de quelques jours. L'estimation de la probabilité de collision entre deux objets se déroule en trois phases.

La première phase consiste à identifier les points de rapprochement géométrique, sans tenir compte de la loi horaire des objets sur leur orbite. Il faut trouver les points de distance minimale entre les 2 ellipses (voir [1], [2]). Cette phase permet de localiser les points à risque, et éventuellement d'arrêter les calculs si la distance minimale trouvée est suffisamment grande.

La deuxième phase consiste à raffiner l'évaluation de distance aux points de rapprochement, en prenant en compte le mouvement horaire des satellites sur une durée de quelques jours (typiquement 3 à 7 jours). Les deux satellites ayant des périodes différentes, il faut identifier parmi les points géométriques de rapprochement ceux où les satellites passeront effectivement à des dates voisines (voir [3]). Cette phase basée sur une simulation temporelle des mouvements conduit à sélectionner les dates critiques pour lesquelles on évaluera la probabilité de collision.

La troisième phase consiste à calculer la probabilité de collision aux dates critiques. Ce calcul repose sur un modèle simplifié : objets sphériques, mouvement instantané rectiligne à la date critique, incertitude gaussienne sur la position relative des deux satellites (voir [4], [5]).

On pourra appliquer à des cas réels le logiciel développé pour le TER et comparer aux résultats publiés sur le site Celestrak.com du NORAD (voir [6]).

RÉFÉRENCES

[1] Armellin, R., Di Lizia, P., Berz, M. et Makino K. (2010). Computing the critical points of the distance function between two Keplerian orbits via rigorous global optimization, *Celestial Mechanics and Dynamical Astronomy*, **107**, 377-395.

[2] Gronchi, G. F. (2005), An algebraic method to compute the critical points of the distance function between two Keplerian orbits, *Celestial Mechanics and Dynamical Astronomy*, **93**, 295-329.

[3] Alfano, S. (1994), Determining satellites close approaches, Part II, *The Journal of the Astronautical Science*, **2**, 143-152.

[4] Alfano, S. (2005), Relating position uncertainty to maximum conjunction probability, *The Journal of the Astronautical Science*, **2**, 193-205.

[5] Chan, F.K. (2008), Spacecraft collision probability, *The Aerospace Press* (Chapitres 2, 4, 5 et 13 pour les méthodes simplifiées d'évaluation de la probabilité de collision).

[6] Navabi, M. et Hamrah , R., Assessment of the probability of collision among space objects, 2009 Iridium-Cosmos collision case, (IAA-AAS-DyCoSS1-11-06)

ESTIMATION PAR ALGORITHME EM D'UN MODÈLE DE MÉLANGE. APPLICATION À LA PRÉDICTION DU NOMBRE ET DE LA TAILLE DES DÉFAUTS D'UN COMPOSANT INDUSTRIEL.

Sujet proposé par : Merlin KELLER, EDF R&D (merlin.keller@edf.fr)
Encadrant UPMC : Tabea REBAFKA, LPMA (tabea.rebafka@upmc.fr)

CONTEXTE : On s'intéresse à l'estimation de la distribution des tailles de défauts présents sur un composant industriel d'une installation de production électrique. En effet, cette distribution intervient dans l'évaluation de la fiabilité de l'installation, quantité faisant l'objet d'un suivi régulier. On dispose pour effectuer cette estimation d'une série de mesures, effectuées par expertise destructrice (affouillement) sur des éléments de même constitution que le composant visé. L'approche courante consiste donc à ajuster une loi paramétrique sur la série de mesures disponibles, vue comme un échantillon parfait de la distribution d'intérêt. Cependant, on sait par ailleurs que deux populations distinctes de défauts sont en fait présentes sur le composant, dont l'une est caractérisée par une hauteur maximale inférieure à une borne connue. Il serait donc plus juste de modéliser les mesures disponibles à l'aide d'un mélange de deux lois, dont une à support borné. L'approche actuelle ne tenant pas compte de ce fait, nous souhaitons tester des méthodes plus adaptées afin de déterminer si une information plus riche peut être extraite des données.

OBJECTIF : L'objectif de ce TER est de développer une méthodologie permettant d'estimer le mélange des distributions de tailles des deux populations de défauts présentes sur le composant, ainsi que la proportion moyenne de leurs effectifs, puis de l'appliquer à la série de mesures disponibles. Cette estimation se fera à l'aide de l'algorithme Expectation-Maximization (EM), particulièrement bien adapté à ce type de problèmes. Nous nous limiterons ici à la modélisation de la hauteur des défauts, la modélisation conjointe du couple hauteur/longueur n'étant pas un objectif premier.

DÉROULEMENT :

1. Dans un premier temps, les étudiants conduiront un travail bibliographique sur l'estimation de modèles de mélange, et plus particulièrement l'algorithme EM. Le but de cette étude est de proposer un ou plusieurs algorithmes détaillés, adaptés au cas présent, en tenant compte notamment du caractère positif des observations, et de la présence d'une loi à support borné.
2. Dans un second temps, la ou les méthodes retenue(s) sera(ont) implémentée(s), de préférence en R ou en Python, et validée(s) sur des données simulées, avant d'être appliquée(s) sur les données réelles.
3. De manière complémentaire, plusieurs aspects pourront également être abordés : détermination d'intervalles de confiances sur les paramètres estimés, sélection de modèles paramétriques pour chaque classe du mélange, modélisation bivariée du couple hauteur/longueur . . .

BIBLIOGRAPHIE SUCCINCTE :

Dempster, A.P. ; Laird, N.M. ; Rubin, D.B. (1977). *Maximum Likelihood from Incomplete Data via the EM Algorithm.* Journal of the Royal Statistical Society, Series B 39 (1) : 1-38.

McLaughlan, G.J. (2000), Finite Mixture Models, Wiley.

van der Vaart, A.W. (1998). Asymptotic statistics. Cambridge University Press.

CONFIRMATION DE DIAGNOSTICS

Sujet proposé par : Jérôme LACAILLE, SNECMA-SAFRAN (jerome.lacaille@snecma.fr)
Encadrant UPMC : Etienne ROQUAIN, LPMA (etienne.roquain@upmc.fr)

Des algorithmes actuellement embarqués sur les moteurs SNECMA recherchent des patterns vibratoires anormaux et émettent des alarmes qui sont envoyées pendant le vol à l'aide de communications satellites. Une grande difficulté rencontrée pour la mise en œuvre de ces algorithmes est d'éviter de générer trop de fausses alarmes. Une fausse alarme entraîne un démontage très coûteux du moteur suivi d'une analyse qui finalement ne mène à aucune conclusion. Pour limiter le taux de fausses alarmes, les diagnostics d'anomalies émis par l'algorithme sont d'abord capitalisés sur le moteur et seulement quand suffisamment de tels diagnostics sont générés dans un laps de temps limité, une alarme est envoyée. Ce processus assure un certain niveau de qualité de l'alarme, mais par contre dégrade rapidement le taux de détection de l'algorithme. Nous cherchons des stratégies optimales pour prendre en charge ces algorithmes de détection d'anomalies.

Le sujet que nous proposons consiste à :

- comprendre dans un premier temps la stratégie en exploitation
- étudier ce que donnerait son remplacement par une meilleure stratégie (par exemple k détections parmi n observations)
- étudier l'état de l'art, notamment ce qui est mis en œuvre dans d'autres domaines
- formaliser le problème et proposer une solution.

RÉFÉRENCES

[1] J. Massé, A. Humeau, P. Lalonde, et A. Alimardani, Placement of alert thresholds on abnormality scores, in *PHME*, 2014, pp. 1-8.

[2] J. Massé, O. Hmad, P. Beauseroy, et E. Grall-Maës, System PHM Algorithm Maturation, *Chem. Enginering Trans.*, vol. 33, pp. 283Ð288, 2013.

[3] O. Hmad, J. Massé, E. Grall-Maës, et P. Beauseroy, Maturation of Detection Functions by Performances Benchmark . Application to a PHM Algorithm, *Chem. Enginering Trans.*, vol. 33, pp. 109-114, 2013.

[4] K. Pipe et B. Culkin, Dynamic alert generation technology for condition monitoring systems, in *CM & MFPT*, 2011.

VALIDATION DE PRONOSTICS

Sujet proposé par : Jérôme LACAILLE, SNECMA-SAFRAN (jerome.lacaille@snecma.fr)
Encadrant UPMC : Philippe Saint-Pierre, LSTA (philippe.saint_pierre@upmc.fr)

Pour une compagnie aérienne, il est utile de savoir à l'avance de combien de temps elle dispose avant de devoir réaliser une opération de maintenance. Cette estimation peut être dictée par la mesure d'un indicateur d'anomalie dont l'évolution est aléatoire mais suit tout de même des règles physiques. Le motoriste aimerait mettre à disposition de son client un test statistique lui permettant d'affirmer qu'une panne va arriver dans 100 vols par exemple, avec un seuil de confiance assez important, mais qu'il y a au moins 80% de chances qu'elle ne se produise pas dans les 60 prochains vols.

De nombreux articles traitent de ce sujet. La plupart estiment des durées de vie de batteries, ce qui est assez facile car il est possible d'observer la dégradation. Bien entendu, ce n'est pas le cas dans l'aéronautique car on a tendance à changer les composants bien avant qu'ils soient usés. Nous proposons un sujet consistant à formaliser clairement le problème mathématique, puis à étudier l'état de l'art, en faire une synthèse et proposer une démarche acceptable basée sur des documents scientifiques pour appliquer cette méthode de décision aux pannes des moteurs aéronautiques.

RÉFÉRENCES

[1] L. Tang, M. E. Orchard, K. Goebel, et G. Vachtsevanos, Novel Metrics and Methodologies for the Verification and Validation of Prognostic Algorithms, in *IEEE Aerospace Conference*, 2011, pp. 11/0502.

[2] M. E. Orchard, L. Tang, et G. Vachtsevanos, A Combined Anomaly Detection and Failure Prognosis Approach for Estimation of Remaining Useful Life in Energy Storage Devices, in *PHM*, 2011, pp. 1-7.

[3] C. Byington, N. Mackos, G. Argenna, A. Palladino, J. Reimann, et J. Schmitigal, Application of Symbolic Regression to Electrochemical Impedance Spectroscopy Data for Lubricating Oil Health Evaluation, in *PHM*, 2012, pp. 1-11.

[4] M. Daigle, A. Bregon, et I. Roychoudhury, A Distributed Approach to System-Level Prognostics, in *PHM*, 2012, pp. 1-12.

[5] J. A. M. Penna, C. L. Nascimento, et L. R. Rodrigues, Health Monitoring and Remaining Useful Life Estimation of Lithium-Ion Aeronautical Batteries, in *IEEE Aerospace conference*, 2012, pp. 1-12.

Entrelacements de semi-groupes markoviens : construction et applications

Damien Simon
(Laboratoire de probabilités et modèles aléatoires)

Année 2014–2015, deuxième semestre

Le but de ce TER est d'étudier des chaînes de Markov de semi-groupes de transition $(P_t)_{t\in\mathbb{R}_+}$ et $(Q_t)_{t\in\mathbb{R}_+}$ sur des espaces X et Y tels qu'il existe un opérateur Λ tel qu'on ait l'entrelacement suivant :

$$\Lambda P_t = Q_t \Lambda$$

Au-delà des cas de couplage pour lesquels il existe une chaîne de Markov sur un espace Z plus gros avec des projections $\iota_1 : Z \to X$ et $\iota_2 : Z \to Y$ telles que les chaînes projetées soient encore markoviennes, Ph. Biane et d'autres mathématiciens ont construit des entrelacements à partir d'espaces de probabilité non-commutatifs sur des groupes qui satisfont les propriétés idoines, les paires de Gelfand.

Le but de ce TER est d'étudier cette construction d'entrelacements à partir de paires de Gelfand au moins dans des cas simples où tous les calculs sont explicites et de l'utiliser pour produire des identités en loi ou des mesures invariantes pour des modèles probabilistes concrets.

Ce TER permet d'aborder superficiellement les probabilités non-commutatives et de voir comment algèbre et probabilités peuvent interagir.

Il est basé sur l'article de revue [1] à partir duquel l'étudiant devra étudier quelques aspects particuliers et approfondir quelques-uns des modèles cités.

Références

[1] Philippe Biane (2011). Entrelacements de semi-groupes provenant de paires de Gelfand. ESAIM : Probability and Statistics, 15, pp S2-S10. doi :10.1051/ps/2010025.

Recuit simulé appliqué au problème du voyageur de commerce

Sujet proposé par Camille Tardif, LPMA, camille.tardif@upmc.fr

Le problème du voyageur de commerce consiste à trouver un parcours de longueur minimale passant par N points fixés du plan. Il existe un algorithme probabiliste, le recuit simulé, qui permet de résoudre ce problème avec un temps de calcul acceptable et ce même pour des grandes valeurs de N. Il s'agit de simuler une chaine de Markov, sur l'ensemble des parcours possibles, qui a tendance à se diriger vers des états d'énérgie minimale (i.e correspondant à des parcours de longueur minimale). Pour échapper à d'éventuels minima locaux la chaîne peut, à l'occasion, sauter vers des états de plus grande énergie. Ceci arrive avec une certaine probabilité dépendant d'un paramètre T (la température) et qui est d'autant plus grande que T l'est. En diminuant la température suffisamment lentement on parvient à piéger la chaîne sur les états correspondants aux parcours de longueur minimale.

L'objectif de ce TER est de présenter la théorie sous jacente du recuit simulé (mesures de Gibbs, algorithme de Métropolis-Hasting) et de montrer le théorème de convergence vers les états d'énergie minimale. Il est aussi demander d'implémenter cet algorithme probabiliste et de le comparer, sur des exemples, à d'autres algorithmes classiques pour ce problème.

Références

— *Recuit simulé et voyageur de commerce.* Texte de Florent Malrieu pour l'agrégation de mathématiques. `http://perso.univ-rennes1.fr/florent.malrieu/AGREG/TEXTES/recuit.pdf`

— *Modélisation stochastique et simulation.* B. Bercu et D. Chafaï.