MASTER DE MATHÉMATIQUES M1 - SORBONNE UNIVERSITÉ
SUJETS T.E.R. - 2024/2025

FRÉDÉRIC KLOPP
RESPONSABLE DES T.E.R.

Date: 5 décembre 2024.

Encadrant(e)(s)	Bureau	Courriel	Intitulé du ou des sujets	Étudiant(e)(s)
Guillaume Garnier	15-16 22	guillaume.garnier@sorbonne-universite.fr	Introduction à la théorie du contrôle	Arslane Chaouach
Pierre-Vincent Koseleff	15-25-506	pierre-vincent.koseleff@imj-prg.fr	Calcul des polynômes cyclotomiques	
Jean-Pierre Marco	15-25-517	jean-pierre.marco@imj-prg.fr	Diffusion et percolation	
Frédéric Paugam	15-25-518	frederic.paugam@imj-prg.fr	Faisceaux et champs différentiables	
Etienne Roquain	15-16-213	etienne.roquain@upmc.fr	Autour de la prédiction conformelle	
			Interpolation et sur-apprentissage en estimation non-paramétrique	
			Autour de l' inférence sélective et des e-values	
			Autour de l' équité en apprentissage automatique	

Introduction à la théorie du contrôle

Encadrant : Guillaume Garnier
guillaume.garnier@sorbonne-universite.fr

La théorie du contrôle est un domaine de l'ingénierie et des mathématiques appliquées qui étudie les systèmes dynamiques afin de les faire fonctionner de manière optimale, stable et efficace. L'idée général consiste à faire passer un système d'un état itinial à un état voulu. Pour cela on dispose généralement d'une commande ou d'un terme de contrôle.

Par exemple, tout le monde sait instinctivement comment faire tenir un stylo sur son doigt. En revanche dès que l'on veut le refaire avec deux stylos l'un sur l'autre les choses se compliquent. Contrôler ce système revient à faire tenir les deux stylos dans les positions que l'on souhaite, et ce, à tout instant. Ici, le terme de contrôle est la position du doigt dans l'espace.

De manière analogue un automobiliste qui veut partir de Paris pour aller vers Marseille doit contrôler sa trajectoire. Ici le terme de contrôle est composé de la position des pédales, de la boîte de vitesse et de la rotation du volant.
Dans ce dernier exemple, on constate qu'il y a plusieurs trajets (et donc plusieurs termes de contrôle) possibles. L'automobiliste pourra choisir de passer par l'autoroute A6 s'il est pressé, il pourra faire un tour d'Europe s'il a plus de temps ou encore il peut passer par de petites nationales s'il veut économiser les frais de péages.

Le but de ce TER est de découvrir les principaux aspects de la théorie du contrôle afin de les appliquer sur différents exemples issus de la physique ou de la biologie. Pour cela on s'appuiera sur les livres classiques cités en références.

Références

[1] Jean-Michel Coron. *Control and nonlinearity.* Number 136. American Mathematical Soc., 2007.

[2] Emmanuel Trélat. *Contrôle optimal : théorie & applications*, volume 36. Vuibert Paris, 2005.

Calcul des polynômes cyclotomiques

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

Mots-clés : Arithmétique, calcul formel.

Dans [AM10], Arnold et Monagan 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.

Les ouvrages [GG13, AECF, SM] pourront utilement être consultés.

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

References

[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. [PDF]

[AECF] A. Bostan *etal.*, *Algorithmes Efficaces en Calcul Formel,* 2017, [PDF]

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

[SM] SageMath Reference Manual, 2023, [PDF]

Projet de TER : diffusion et percolation

Jean-Pierre Marco (jean-pierre.marco@imj-prg.fr)

La diffusion dans les systèmes hamiltoniens presque intégrables (comme par exemple le système solaire) consiste à étudier l'existence d'orbites dont certaines variables (par exemple les grands-axes des orbites des planètes) varient beaucoup pendant un temps éventuellement long, mais que l'on sait borner supérieurement. Ces études portent jusqu'à maintenant sur des orbites prises isolément, et ne prennent pas en compte la mesure de leur ensemble et les probabilités associées. Cependant on montre facilement l'existence de "réseaux" le long desquels les orbites de diffusion existent, avec la possibilité de changer de "direction" aux noeuds de ces réseaux. Ce phénomène global a de fortes analogies avec celui de la percolation (voir par exemple l'exposé de H. Duminil-Coppin : https://www.youtube.com/watch?v=vqoOjS0fgBc pour une introduction).

On se propose dans un premier temps de donner tous les éléments pour une étude élémentaire des systèmes hamiltoniens intégrables et leurs perturbations, puis plus spécifiquement des problèmes de diffusion, encore assez mal compris. En parallèle, une introduction à la percolation sera aussi proposée. Après ces étapes préliminaires, on construira un modèle simple sur lequel les deux approches peuvent être comparées, à la fois sur le plan mathématique et sur le plan numérique (suivant goûts et aptitudes). Il s'agit donc d'une introduction à un sujet encore en plein développement.

Références

1. V.I. Arnold. Mathematical method of classical mechanics. Springer (second edition).

2. V. Werner. Percolation et modèle d'Ising. SMF.

3. J-P Marco. A symplectic approach to Arnold diffusion (in MSRI Publications – Volume 72).

4. T. Lazaro, J-P. Marco. Anisotropy in the neighborhood of double resonances. Preprint.

Faisceaux et champs différentiables

Proposition de mémoire

Frédéric Paugam
frederic.paugam at imj-prg.fr

13 octobre 2024

On propose d'étudier les bases de la théorie des faisceaux et champs différentiables, qui sont une généralisation naturelle des variétés différentielles, permettant (en particulier) d'étudier de manière fine les espaces de fonctions et les espaces quotients. Ces structures sont utiles dans la formalization mathématique de la physique.

On commencera par rappeler comment définir les variétés différentielles comme des faisceaux différentiables et les avantages que ce point de vue, dit "du foncteur des points", introduit par Grothendieck, apporte (on abordera seulement les bases de la théorie, mais en cas de besoin, on pourra se référer à [BH11], Section 4). On évoquera ensuite la théorie des champs différentiables en utilisant la référence [Gin13].

Prérequis :

- Un peu de géométrie différentielle.
- Un peu de théorie des catégories (voir [ML98] ; à rappeler si besoin en cours de mémoire).

Références

[BH11] John C. Baez and Alexander E. Hoffnung. Convenient categories of smooth spaces. *Trans. Amer. Math. Soc.*, 363(11) :5789–5825, 2011.

[Gin13] Gregory Ginot. Differentiable stacks. *(lecture notes)*, 2013.

[ML98] Saunders Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.

Autour de la prédiction conformelle

Sujet proposé par Etienne Roquain, etienne.roquain@upmc.fr
Année 2024-2025

Les méthodes actuelles de machine learning peuvent avoir des performances impressionnantes, mais font des erreurs. Ainsi, un travail est de quantifier l'incertitude liée à ces méthodes, et ce malgré le côté "boîte noire" de ces procédures et la complexité des données observées.

L'objectif de ce TER est de comprendre comment la prédiction conformelle permet d'apporter une solution à ce problème. Plus précisément, en s'appuyant sur Angelopoulos and Bates (2022), une feuille de route pourra être la suivante :

- Comprendre et prouver les théorèmes fondamentaux sur laquelle cette approche repose ;
- L'appliquer sur des modèles simples en régression et classification non paramétrique ;
- Aller vers des extension "en ligne", en s'appuyant par exemple sur Susmann et al. (2023).

La rédaction du rapport se fera en langage LaTeX et une soutenance orale devra appuyer ce document.

Il sera important que l'étudiante ou l'étudiant en question soit à l'aise avec les statistiques de base. De plus, il est conseillé de suivre le cours de "statistiques avancées" en parallèle.

Références

Angelopoulos, A. N. and Bates, S. (2022). A gentle introduction to conformal prediction and distribution-free uncertainty quantification.

Susmann, H., Chambaz, A., and Josse, J. (2023). AdaptiveConformal : An R Package for Adaptive Conformal Inference. working paper or preprint.

Interpolation et sur-apprentissage en estimation non-paramétrique

Sujet proposé par Etienne Roquain, `etienne.roquain@upmc.fr`
Année 2024-2025

Un obstacle majeur des procédures d'apprentissage automatique est le sur-apprentissage ("over-fitting" en "machine learning") : si la procédure "colle" trop aux données d'apprentissage, alors elle aura de piètres performances sur d'autres jeux de données, pourtant générés selon la même loi.

Ces dernières années, les statisticiens portent un nouveau regard sur ce phénomène bien connu : il s'avère qu'il y a une façon de "coller" parfaitement aux données qui garantisse encore la validité de l'inférence statistique. Cela s'appelle l'"interpolation".

L'objectif de ce TER sera d'explorer ce nouveau phénomène dans le cadre de l'estimation non-paramétrique, en s'appuyant sur Belkin et al. (2018). Une feuille de route pourra être la suivante :

- Commencer par un travail bibliographique autour de l'interpolation dans les différents domaines de la statistique, cf par exemple les références dans Bartlett et al. (2021) ;
- Etudier profondément le travail de Belkin et al. (2018), par exemple avec des preuves détaillées en se restreignant à la classe des fonctions Lipschitziennes. Il faudra notamment percer le mystère de la Figure 1 ;
- Appliquer ce résultat théorique sur des données simulées et des données réelles ;
- Si le candidat est très à l'aise, regarder le cas de l'interpolation dans le cas de la régression linéaire, cf Bartlett et al. (2021).

La rédaction du rapport se fera en langage LaTeX et une soutenance orale devra appuyer ce document.

Il sera important que l'étudiante ou l'étudiant en question soit à l'aise avec les statistiques de base. De plus, il est conseillé de suivre le cours de "statistiques avancées" en parallèle.

Références

Bartlett, P. L., Montanari, A., and Rakhlin, A. (2021). Deep learning : a statistical viewpoint.

Belkin, M., Rakhlin, A., and Tsybakov, A. B. (2018). Does data interpolation contradict statistical optimality ?

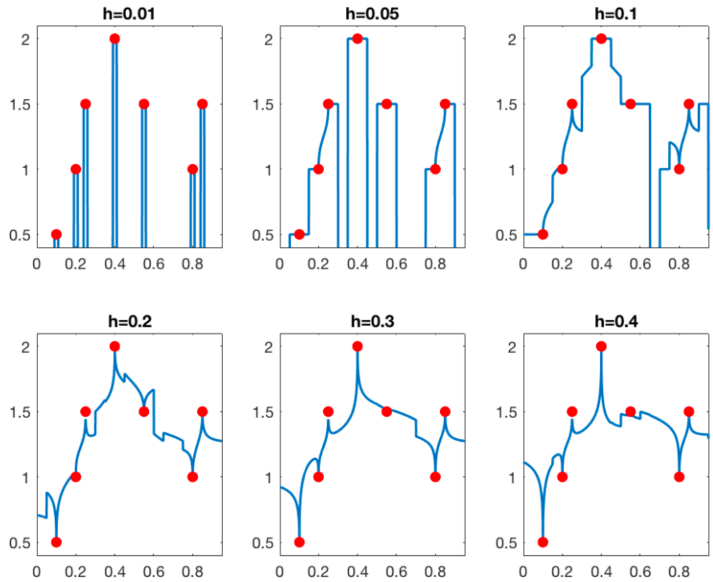

FIGURE 1 : Interpolation en régression non paramétrique (extraite de Belkin et al. (2018)).

Autour de l'inférence sélective et des e-values

Sujet proposé par Etienne Roquain, etienne.roquain@upmc.fr
Année 2024-2025

Une pratique courante dans les données actuelles est de reporter simultanément K intervalles de confiance $I_{1,\alpha}(X), \dots, I_{K,\alpha}(X)$ pour des vrais paramètres $\theta_1, \dots, \theta_K$ réels et des observations X. La propriété de couverture marginale est alors

$$\forall k \in \{1, \dots, K\}, \quad \mathbb{P}(\theta_k \in I_{k,\alpha}(X)) \geq 1 - \alpha.$$

Il est clair que cela assure un taux moyen global de fausse couverture plus petit que α :

$$\mathrm{FCR}(S) := \mathbb{E}\left(\frac{\sum_{k\in S} \mathbb{1}_{\theta_k \notin I_{k,\alpha}(X)}}{|S| \vee 1}\right) \leq \alpha,$$

pour une règle de selection $S \subset \{1, \dots, K\}$ basée sur des données indépendantes de X. Cependant, en pratique, il est très tentant de définir S à partir du même jeu de données X (i.e., $S = S(X)$). Il est clair que cela invalide complètement le contrôle donné plus haut. Par exemple, si tout les paramètres $\theta_1, \dots, \theta_K$ sont nuls et que l'on prend la règle de décision $S = \{k \in \{1, \dots, K\} : 0 \notin I_{k,\alpha}(X)\}$, le FCR vaut toujours 1. Cet *effet de sélection* est problématique car ce cas n'est pas singulier et est même typique lorsque l'on cherche à construire des intervalles de confiances sur les paramètres "intéressants".

L'objectif de ce TER est de comprendre comment l'article Xu et al. (2024) permet de résoudre ce problème. Une feuille route est la suivante :

- prendre la mesure du problème avec l'article fondateur Benjamini and Yekutieli (2005) ;
- se familiariser avec la notion de e-values avec par exemple certains passage de la revue Ramdas and Wang (2024) ;
- Explorer en détail l'article Xu et al. (2024) en s'appropriant les résultats et en reproduisant les expériences numériques.

La rédaction du rapport se fera en langage LaTeX et une soutenance orale devra appuyer ce document.

Il sera important que l'étudiante ou l'étudiant en question soit à l'aise avec les statistiques de base. De plus, il est conseillé de suivre le cours de "statistiques avancées" en parallèle.

Références

Benjamini, Y. and Yekutieli, D. (2005). False discovery rate–adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association*, 100(469) :71–81.

Ramdas, A. and Wang, R. (2024). Hypothesis testing with e-values. *arXiv preprint arXiv :2410.23614*.

Xu, Z., Wang, R., and Ramdas, A. (2024). Post-selection inference for e-value based confidence intervals. *Electronic Journal of Statistics*, 18(1) :2292–2338.

Autour de l'équité en apprentissage automatique

Sujet proposé par Etienne Roquain, etienne.roquain@upmc.fr
Année 2024-2025

Les algorithmes d'apprentissage automatique (machine learning) sont performants mais reproduisent *de facto* les biais/stéréotypes présent dans les données d'apprentissage. Une solution récemment proposée est le fait d'introduire des algorithmes "équitables" dans le sens où ils ont un comportement "peu influencé" par la variable sensible (genre, groupe éthnique, etc).

Par exemple, dans le modèle de régression $Y_i = f(X_i, S_i) + \sigma\varepsilon_i$, $1 \leq i \leq n$, avec des triplets (X_i, S_i, Y_i), $1 \leq i \leq n$, qui sont iid, indépendant de ε_i, iid $\mathcal{N}(0, 1)$, la variable Y_i peut représenter le salaire, S_i le genre, et X_i le CV de l'individu, et l'objectif est de produire un estimateur $\hat{f}(X, S)$ ou $\hat{f}(X)$ qui soit indépendant de la variable S. Ce critère d'équité est appelé "parité démographique". [1]

L'objectif de ce TER est de :

- Comprendre les différents types d'équité évoqués dans Barocas et al. (2023); Caton and Haas (2024), ainsi que leurs liens ;
- Etudier le cas particulier d'un modèle linéaire dans le cas du critère d'équité mentionné plus haut. Pour cela, on pourra simplifier le travail Chzhen et al. (2020) ;
- Appliquer ce résultat théorique sur des données simulées et des données réelles ;
- Si possible, explorer la théorie pour les autres critères d'équité et faire le lien avec le transport optimal de Lara (2023).

La rédaction du rapport se fera en langage LaTeX et une soutenance orale devra appuyer ce document. Il sera important que l'étudiante ou l'étudiant en question soit à l'aise avec les statistiques de base. De plus, il est conseillé de suivre le cours de "statistiques avancées" en parallèle.

Références

Barocas, S., Hardt, M., and Narayanan, A. (2023). *Fairness and machine learning : Limitations and opportunities*. MIT press.

Caton, S. and Haas, C. (2024). Fairness in machine learning : A survey. *ACM Computing Surveys*, 56(7) :1–38.

Chzhen, E., Denis, C., Hebiri, M., Oneto, L., and Pontil, M. (2020). Fair regression with wasserstein barycenters. *Advances in Neural Information Processing Systems*, 33 :7321–7331.

de Lara, L. (2023). *Modèles contrefactuels pour un apprentissage machine explicable et juste : une approche par transport de masse*. PhD thesis, Université Paul Sabatier-Toulouse III.

[1] Pour une présentation orale du sujet, le candidat ou le candidate pourra consulter la vidéo postée ici https://www.youtube.com/watch?v=v6OYu3Pbojw.