

Optimisation stochastique

1 Le document

Le document à étudier est soit un article publié dans une revue scientifique, soit un chapitre de livre spécialisé. Vous pouvez choisir ce texte dans la liste de ceux que nous vous proposons ou suggérer vous-même un article que vous jugez intéressant et en rapport avec ce cours. Dans cette dernière éventualité, vous devez au préalable nous soumettre ce document pour que nous puissions juger rapidement de son adéquation avec le travail attendu.

2 Le travail

Il vous est demandé de préparer un petit cours écrit de niveau M2 portant sur la (ou l'une des) question(s) traitée(s) dans l'article. Le document principal sera un fichier pdf de **15 pages maximum** et sera accompagné d'un **notebook** html généré à partir de RMarkdown ou Python expliquant de manière claire et concise **les simulations effectuées**. Les figures issues de ces simulations seront à intégrer dans le pdf du document principal.

Idéalement, l'article étudié constitue un point de départ et non une fin en soi, vous pourrez donc le compléter par d'autres lectures. Aucune forme ou aucun plan particulier ne vous est imposé. Disons qu'un bon mémoire pourrait être articulé de la façon suivante :

1. Une introduction détaillant le sujet et la problématique, avec quelques références fondamentales.
2. Une partie présentant les notions essentielles ainsi que les notations.
3. Une partie contenant les résultats mathématiques, avec les preuves les plus importantes. Les résultats et preuves trop techniques, ou moins importants, pourront être admis : c'est à vous de voir, en fonction du message que vous voulez faire passer.
4. Des simulations pour illustrer les résultats sont **obligatoires**.
5. Une conclusion présentant quelques perspectives.

Ne perdez pas de vue qu'il s'agit d'un cours et que votre rédaction se doit d'être la plus pédagogique possible. Vous serez certainement amenés à faire des choix, à admettre (ou à passer sous silence) certains résultats ou, au contraire, à en détailler d'autres. Il vous appartient donc en particulier de décider où mettre le curseur : ni trop haut (cours trop facile), ni trop bas (cours de trop haut niveau).

Sur la forme, votre document devra bien entendu être synthétique, agréable à lire et bien présenté. Il pourra être rédigé en français ou en anglais, et il vous est vivement conseillé d'utiliser le traitement de texte Latex.

3 Notation

La notation prendra en compte aussi bien les aspects de fond (précision mathématique, qualité des démonstrations et **des simulations**, choix effectués, etc.) que les aspects de forme (orthographe, qualité de la rédaction, concision, précision, etc.). Il vous est en particulier demandé de **respecter la contrainte des 15 pages et d’apporter le plus grand soin à la cohérence générale** de votre document : ne perdez pas de vue qu’il s’agit d’un cours et que vous devez donc l’organiser comme vous auriez vous-même aimé l’avoir reçu !

4 Les écueils à éviter

1. Se limiter au document initial, sans aucune recherche complémentaire.
2. Traduire le texte ligne à ligne, sans faire de choix. Un cours est une synthèse, pas un roman !
3. Changer les notations pour le plaisir, sans argument solide.
4. Admettre tous les résultats mathématiques, sans aucune démonstration.
5. Choisir de présenter les preuves trop faciles et ignorer des résultats plus difficiles mais néanmoins essentiels.
6. Partir dans des digressions mal maîtrisées.
7. Faire un copier-coller d’un article de Wikipedia ou équivalent.
8. Se mettre à réfléchir au sujet quelques jours avant la date limite.

5 Articles

5.1 Algorithmes stochastiques

- [A stochastic Gauss-Newton algorithm for regularized semi-discrete optimal transport](#) B Bercu, J Bigo, S Gadat E Siviero 2020
- [Communication trade-offs for synchronized distributed SGD with large step size](#) K. Patel, A. Dieuleveut, 2019
- [Recursive ridge regression using second-order stochastic algorithms](#) A. Godichon-Baggioni, W Lu and B Portier 2022
- [A robust model-based clustering based on the geometric median and the Median Covariation Matrix](#) A. Godichon-Baggioni and S Robin 2022
- [SGD with Coordinate Sampling: Theory and Practice](#) R Leluc and F Portier 2020
- [Online estimation of the inverse of the Hessian for stochastic optimization with application to universal stochastic Newton algorithms](#) A. Godichon-Baggioni, W. Lu and B. Portier

5.2 PAC-Bayes et Inférence variationnelle

- [On the properties of variational approximations of Gibbs posteriors](#) P. Alquier, J. Ridgway, N. Chopin, 2016.
- [Tighter Variational Bounds are Not Necessarily Better](#) T. Rainforth, A.R. Kosiorek, T.A. Le, C.J. Maddison, M. Igl, F. Wood, Y.W. Teh, 2018.
- [Convergence rates of variational posterior distributions](#) F. Zhang, C. Gao, 2020.
- [Tighter Risk Certificates for Neural Networks](#) M. Pérez-Ortiz, O. Rivasplata, J. Shawe-Taylor, C. Szepesvári, 2021.

- [How Tight Can PAC-Bayes be in the Small Data Regime?](#) A.Y.K. Foong, W. Bruinsma, D.R. Burt, R.E. Turner, 2021.
- [Split-kl and PAC-Bayes-split-kl Inequalities for Ternary Random Variables](#) Y.-S. Wu, Y. Seldin, 2022.
- [The Bayesian Learning Rule](#) M.E. Khan, H. Rue, 2023.
- [Better-than-KL PAC-Bayes Bounds](#) I. Kuzborskij, K.-S. Jun, Y. Wu, K. Jang, F. Orabona, 2024.
- [A PAC-Bayesian Link Between Generalisation and Flat Minima](#) M Haddouche, P. Viillard, U. Simsekli, B. Guedj, 2024.