MT5731: Advanced Bayesian Inference

Advanced Project 2021-22

The project chosen will constitute 40% of your final mark for this course. These projects all go beyond the scope of the lecture notes provided and will involve further independent research. An example project (not perfect!) is provided in Moodle.

Timing

- End of week 5: inform by email the module coordinator on the project title you chose by Friday **15th October 5pm**.

- Beginning of week 7: meet with module coordinator to discuss your ideas and initial work.

- Weeks 7-11: work on the project! You can have one more meeting with the module coordinator if you feel that this is something that will help you. You can also email if you have questions. However, it is important to know that these projects are also designed to teach you to work independently, and the amount of advice you receive will be in accordance with that.

- Week 11: **submit the project** by uploading it to MMS by the end of **Friday 26th November**.

Project titles

NOTE: **your work should not exceed 20 pages (references and appendix excluded) using 1.5 line spacing and font size 12pt**. Include only essential figures and tables in the main text. Additional figures and tables (if any) should be included in the appendix and linked to the main text. You are also required to submit an R script containing all your analyses (if relevant to your project).

Possible project titles that can be chosen are:

1. An essay exploring one computational approach alternative to MCMC, among the following: Variational Bayes, Approximate Bayesian Computation, Hamiltonian Monte Carlo. The project should include (but not be restricted to) the theoretical background for the method, description of its algorithmic implementation, and analysis of simulated and/or real data. If the analyses concern simple examples, it is expected that original code is written. If the analyses concern more complex statistical modelling, R packages can be used.

2. An essay on Bayesian spatial models which should explore Conditional Autoregressive (CAR) models and more recent extensions. A clear, well-detailed description of a few different methods (2-3) is required as well as practical examples with the analysis of simulated and/or real data. If the analyses include simple examples, it is expected that original code is written. If the analyses concern more complex statistical modelling, R packages can be used. In any case, code should be submitted with the project as an Appendix.

3. An essay on Bayesian nonparametric inference. This is a project with a stronger theoretical/mathematical background. One possible direction is to describe the theory behind the Dirichlet Process, both in its more general form and considering the Chinese Restaurant Process and the stick-breaking prior. Then, focus on mixture modelling with the Dirichlet process, either for clustering purposes or for density estimation. Simple examples will suffice if the theoretical component of the project is strong. Published manuscripts that can form the backbone of the project will be suggested by the module coordinator.

4. A complete Bayesian data analysis in R using NIMBLE. The analysis should be based on a real dataset to be selected among those available from the UCI repository (https://archive.ics.uci.edu/ml/datasets.php) or from other sources proposed by the student. In both cases, your choice needs to be discussed and approved by the module coordinator before you start working with the data. You should implement statistical models that are more complex than those considered in the lectures. The essay should include (a) the aim of the analysis; (b) a description of the modelling challenges (e.g., high dimensional dataset, spatial and/or temporal correlation, non-linear covariates…) and of the Bayesian methods employed to overcome them; (c) a description of the model implemented in NIMBLE, all diagnostic checks and goodness of fit measures employed; (d) interpretation of the results obtained and, (e) discussion of limitations and possible improvements. The code should be submitted with the project in the Appendix.

Marking scheme

The marking scheme will be the same as for MT5999 (final year project), although obviously not to the same level as this is just a component of a 15 credit module.

- Initiative, Originality and Effort – 10
- Presentation and Exposition – 10
- Understanding and Accuracy – 10
- Mathematical, Statistical or Historical Analysis – 10
- Overall Achievement – 10
- Total – 50

Good academic practice

Please note *that you must work independently of other students on your project*; you cannot submit joint work. For more information, see the university web pages on good academic practice, particularly

https://www.st-andrews.ac.uk/policy/academic-policies-assessment-examination-and-award-good-academic-practice/good-academic-practice-guidance-for-students-frequently-asked-questions.pdf

Plagiarism software will be used to check each submission, so make sure you don't fall foul of the rules – submit only your own work and cite anything you use from other sources in the appropriate way. Please contact the module coordinator if you have any uncertainty about this.