Course Outline

Summer Reading

Please go through the R package programming notes and practicals on Robin Evans' webpage before the course starts:

`http://www.stats.ox.ac.uk/~evans/teaching.htm#R`

We assume that you are either happy and comfortable with the material in the notes or have a list of things you would like better explained when you arrive.

Abstract

Computational statistics is one of the most dynamic areas of research in modern statistics and is key to such fields as Bayesian inference, model choice, Graphical models, big data and others. In this course students will be exposed to main concepts of computational statistics and to milestone ideas, such as Monte Carlo, Markov chain Monte Carlo, Sequential Monte Carlo, Reversible Jump, Approximate Bayesian Computation, Intractable Likelihood, Monte Carlo Expectation Maximisation, Simulated Annealing, Belief Propagation, LASSO, LARS, to name a few, and will study two of them in depth based on original research papers. The study will involve analysing real data and coding one of the methodologies. The course is designed not only to cover computational statistics and coding in R, but also to support acquisition of transferable skills such as collaborative work, giving and receiving constructive feedback and good practice in producing code.

Timetable

- **Mon:** Review of R; including calling C code, debugging; some practicals
- **Tue:** Presentation of $6-8$ papers from computational classics and related projects, about 30 minutes per topic.
- **Wed:** Morning: breaking into pairs; assigning projects to 6 pairs (each pair will have one Oxford and one Warwick student)
- **rest of week 1:** work on the project: produce code and draft reports
- **Mon:** 9am: hand in code and report to another team (not reciprocal)
  2pm: review second teams work, give feedback (round 1)
  4pm: review second teams work, give feedback (round 2)
- **Tue:** work on revising your project
  research presentations from Oxford and Warwick faculty members
- **Wed:** work on revising your project
  pizza-lunch + lunch time speaker (internal or external)
- **Thu:** 9am: hand in code and reports for marking
  work on presentation

- **Fri:** Symposium at Warwick
  11am-12.30pm: presentations
  12.30pm: lunch
  individual feedback over lunch
  2-4pm: double seminar: James Scott (Texas) + TBA
  4pm: wine
  6pm: bus to Oxford

Objectives

After the course the students will acquire the following skills:

- Familiarity with R package and ability to implement mainstream computational algorithms
- Reading and discussing code; understanding the benefits of producing clear well written code
- Collaborate in understanding a research problem and write code
- Critically assessing someone else's code and work, giving constructive feedback
- Receiving critical feedback and using it constructively to improve work
- Exposure to several mainstream topics in computational statistics
- Thorough familiarity with two of these topics (project done + project reviewed)
- Ability to present work and research

Assessment

Each student will produce an individual report of maximum 4 pages including graphs and references. The report will be marked.

Each pair will produce code and do a presentation. There will be feedback on the code and presentation.

List of Papers

The list is subject to (minor) changes.

1. Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. Hybrid monte carlo. Physics Letters B, 195(2): 216-222, 1987.
2. Tibshirani, R. Regression shrinkage and selection via the LASSO. J. Roy. Statist. Soc. Ser. B., 58(1): 267-288, 1996.
3. Propp, J.G. and Wilson, D.B. Exact sampling with coupled Markov chains and applications to statistical mechanics. Random structures and Algorithms, 9: 223-252, 1996.
4. Del Moral, P., Doucet, A & Jasra, A. Sequential Monte Carlo samplers. J. Roy. Statist. Soc. Ser. B., 68(3): 411-436, 2006.

5. Lauritzen, S.L. and Spiegelhalter, D.J. Local computations with probabilities on graphical structures and their application to expert systems, Journal of the Royal Statistical Society. Series B. 50(2): 157–224. 1988.

6. Green, P. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination Biometrika, Biometrika Trust, 1995, 82, 711-732

7. Andrieu, C. & Roberts, G. O. The pseudo-marginal approach for efficient Monte Carlo computations The Annals of Statistics, JSTOR, 2009, 697-725

8. Haario, H.; Saksman, E. & Tamminen, J. An adaptive Metropolis algorithm Bernoulli, JSTOR, 2001, 7, 223-242

9. Roberts, G. O. & Stramer, O. On inference for partially observed nonlinear diffusion models using the Metropolis–Hastings algorithm Biometrika, Biometrika Trust, 2001, 88, 603-621