PhD course: “Statistical inference for partially observed stochastic processes”, February–March 2016

http://www.maths.lu.se/index.php?id=110381

This PhD-level course will present an overview of modern inferential methods for partially observed stochastic processes, with emphasis on state-space models (also known as Hidden Markov Models).

Necessary prerequisites: basics of stochastic processes, Bayesian methods and Monte Carlo methods (e.g. Markov Chain Monte Carlo, Metropolis-Hastings method). For example having taken the courses Time series Analysis (FMS051/MASM17) and Monte Carlo and Empirical Methods for Stochastic Inference (FMS091/MASM11).

Students assessment: To pass the course students must solve at least 3 out of 6 home assignments and a final approved written project report/presentation. More in detail, students must solve at least one assignment for each of the following three areas (a)-(b)-(c) where (a) = particle methods and approximate Bayesian computation; (b) = data imputation for diffusions and iterated filtering; (c) = inference for Gaussian Markov random fields. The final project will be about freely choosing a method or article to study more in detail (e.g. studying an extension to some method considered during the course) and implement corresponding simulations.

This document illustrate the course content. For practical info please refer to the course webpage http://www.maths.lu.se/index.php?id=110381.

Inference and data imputation for diffusion and other continuous time processes

Lecturer: Erik Lindström.

The inference problem for diffusion processes is generally difficult due to the lack of closed form expressions for the likelihood function. However, the problems becomes manageable if data is imputed between the observations. This lecture will cover the basic ideas, some important variance reduction techniques (with a focus towards bridge samplers) as well as pointing towards future extensions.

Assignment: Implement the Durham-Gallant sampler for a Cox-Ingersoll Ross diffusion.

Key references

[1] Pedersen, A. R. (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. Scandinavian Journal of Statistics, 55-71.

[2] Durham, G. B., and Gallant, A. R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. Journal of Business & Economic Statistics, 20(3), 297-338.

[3] Lindström, E. (2012). A regularized bridge sampler for sparsely sampled diffusions. Statistics and Computing, 22(2), 615-623.

Iterated filtering

Lecturer: Erik Lindström.

General state space models are defined in terms of a latent Markov process, from which partial observations can be obtained. This typically means that the latent process must be recovered in order to estimate parameters. An old idea, going back at least half a century, is to treat the model parameters as latent processes themselves. This idea has been tested repeatedly with varying success, but no proof was presented before the introduction of iterated filtering.

The lecture will present an historical overview, while trying to explain why older methods have failed. These experiences are then used to introduce the "iterated filtering"' method for which strong consistency can be proved. We also look at some extensions that are far more efficient from a computational point of view.

Assignment: Implement the basic algorithm for a univariate linear, Gaussian model and compare with simpler alternatives.

Key references

[1] Ionides, E. L., Bhadra, A., Atchadé, Y., and King, A. (2011). Iterated filtering. The Annals of Statistics, 39(3), 1776-1802.

[2] Lindström, E., Ionides, E., Frydendall, J., and Madsen, H. (2012, July). Efficient iterated filtering. In System Identification (Vol. 16, No. 1, pp. 1785-1790).

[3] Lindström, E. (2013). Tuned iterated filtering. Statistics & Probability Letters, 83(9), 2077-2080.

[4] Doucet, A., Jacob, P. E., and Rubenthaler, S. (2013). Derivative-free estimation of the score vector and observed information matrix with application to state-space models. arXiv preprint arXiv:1304.5768.

[5] Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S., and King, A. (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. Proceedings of the National Academy of Sciences, 112(3), 719-724.

Particle Marginal Methods for parameter inference

Lecturer: Umberto Picchini.

Sequential Monte Carlo methods (SMC, a.k.a. particle filters) have revolutionised and simplified the problem of filtering for nonlinear, non-Gaussian models. For example SMC can be used to filter the "signal" in state-space models out of noisy measurements by using a set of N computer-generated "trajectories" ("particles") of the system's state. SMC can also be used to construct an approximation to the likelihood function for the parameters θ of the state-space model of interest. A striking result due to Andrieu and Roberts (2009) is when SMC is used in the context of Bayesian inference for θ: if an unbiased SMC approximation to the likelihood function is plugged into the posterior distribution for θ, it is possible to construct a standard MCMC algorithm for sampling *exactly* from such posterior distribution, regardless of the specific (finite) value of N. This is the so called "pseudo-marginal approach" which is a special case of the class of algorithms known as Particle MCMC (PMCMC).

Assignment: TBA

References denoted with an asterisk (*) are recommended and "friendly" (review articles, monographies or blog entries). The remaining ones are the original (notoriously impenetrable) technical references which are not really necessary for a first understanding.

Key references

[1] Andrieu, C., and Roberts, G. O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. The Annals of Statistics, 697-725.

[2] (*) D. Wilkinson's blog. The pseudo-marginal approach to "exact approximate" MCMC algorithms, `https://goo.gl/W4bvH2`

[3] (*) D. Wilkinson's blog. Marginalisation, importance sampling and the bootstrap particle filter, `https://goo.gl/R8BT3q`

[4] (*) P. Jacob (2015) Sequential Bayesian inference for implicit hidden Markov models and current limitations. ESAIM: Proceedings and Surveys, Vol. 51, p. 24-48, `http://goo.gl/O3sPRV`

[5] (*) N. Kantas, A. Doucet, S. Singh, J. Maciejowski, N. Chopin (2015) On Particle Methods for Parameter Estimation in State-Space Models. Statistical Science 30(3) 328–351, `http://goo.gl/Hg1aCW`

[6] (*) S. Särkkä (2013). Bayesian Filtering and Smoothing. Cambridge University Press. This monography is freely available at `http://goo.gl/1mt5CM`

Approximate Bayesian Computation

Lecturer: Umberto Picchini.

Approximate Bayesian Computation (ABC) is a class of algorithms allowing for inference in complex models with "intractable likelihoods". Specifically, with "complex" we mean models for which we are unable to make use of the likelihood function (because it is analytically unavailable or computationally too expensive to evaluate). However, it is often the case that it is possible – and computationally "cheap" – to simulate from the data generating model, and this implies sampling "simulated data" from the likelihood function. By repeatedly drawing from the likelihood we can construct "likelihood-free" methods for Bayesian inference even when we cannot evaluate the likelihood pointwise (but we can somehow sample from it!). In the most typical scenarios, such methods only result in approximate Bayesian inference, though they can also produce exact inference under some very stringent conditions.

Assignment: TBA

All the following references are friendly, introductory review papers.

Key references

[1] Marin, J. M., Pudlo, P., Robert, C. P., & Ryder, R. J. (2012). Approximate Bayesian computational methods. Statistics and Computing, 22(6), 1167-1180.

[2] Sisson, S. A., & Fan, Y. (2011). chapter "Likelihood-free MCMC" from Handbook of Markov Chain Monte Carlo, Chapman & Hall. Freely available at `http://arxiv.org/pdf/1001.2058.pdf`

[3] Sunnåker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M., & Dessimoz, C. (2013). Approximate Bayesian computation. PLoS Comput Biol, 9(1), e1002803.

Gaussian Markov Random Fields

Lecturer: Johan Lindström.

A common model for spatial data consists of a latent Gaussian field with (non-)Gaussian observations. The dependence structure in the latent process is often described using a parametric covariance function. For large datasets the computation, storage and inversion of the covariance matrix becomes a major issue. Replacing the covariance with a suitable Markov random field representation leads to latent fields with sparse precision matrices which have computational benefits. This lecture will cover the basic ideas of models with latent Gaussian processes, discussing alternatives to covariance matrices for large data. We will then discuss the formulation of latent Gaussian processes as solutions to stochastic partial differential equations (SPDE). The links between SPDEs and older conditional and simultaneous autoregressive models (CAR & SAR), the spectral interpretation of the SPDE and the construction of solutions using basis functions, and the selection of different basis functions will be discussed.

Assignment: Compute precision matrix elements for an irregularly spaced SPDE approximation on the real line, compare results to an AR(1) process. (i.e. solve the SPDE in 1D)

Key references

[1] Lindgren, Rue & Lindström (2011) "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach" JRSSB

[2] Lindgren & Rue (2008) "On the Second-Order Random Walk Model for Irregular Locations" Scand. J. Stat.

[3] Whittle (1963) "Stochastic processes in several dimensions" Bull. Int. Stat. Ins.

[4] Rue (2001) "Fast sampling from Gaussian Markov random fields" JRSSB

Inference for Gaussian Markov Random Fields

Lecturer: Johan Lindström.

The inference for GMRF-based models (and other latent field models) is often based on MCMC computations. This lecture will cover the general framework of Hierarchical Bayesian modelling, i.e. partially observed latent process with unknown parameters governing process and observations. The inference for these models will be discussed highlighting: 1) The advantage of blocking MCMC updates, 2) Construction of Laplace approximations to the posterior, 3) using Laplace approximations to construct MCMC proposals and 4) replacing the MCMC step with numerical integration, resulting in INLA. If time allows other methods for estimating latent fields, mainly expected maximisation (possibly something about SA-EM) and expected conjugate gradient algorithms van be discussed.

Assignment: Assuming a latent AR(1) process with non-Gaussian observations implement a Laplace-approximation based parameter estimation (i.e. the optimisation part of INLA), possibly consider MCMC algorithms based on the approximation.

Key references

[1] Wikle, Berliner & Cressie (1998) "Hierarchical Bayesian space-time models" Environ. Ecol. Stat,

[2] Knorr-Held & Rue (2002) "On Block Updating in Markov Random Field Models for Disease Mapping" Scand. J. Stat.

[3] Rue, Martino & Chopin (2009) "Approximate Bayesian inference for hierarchical Gaussian Markov random field models" JRSSB

[4] Givens & Hoeting (2013) "Chapter 4: EM Optimization Methods" in Computational Statistics 2nd Ed.

[5] Lange (1995) "A gradient algorithm locally equivalent to the EM algorithm" JRSSB