Bayesian inference for stationary autoregressions

Autoregressive models are among the most widely used classes of model for time-series data that exhibit dependence over time. A very common assumption is that of stationarity which, in the context of Gaussian processes, asserts that the mean, variance and covariances governing the process remain constant over time. Since this assumption can be violated in real time-series, for example due to systematic changes in the mean over time or periodic seasonal variations, stationary autoregressive processes are also used as building blocks in the construction of more sophisticated non-stationary models. As such, their importance to the field cannot be overstated.

An autoregressive model of order p, often abbreviated to an AR(p) model, expresses the random variable y_t at time t as a linear combination of its previous p values and an error term ϵ_t, that is, $y_t = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \epsilon_t$. Conditional on $y_{t-1}, \ldots, y_{t-p}$, the random variable y_t is independent of any observations further back in the history of the time series. In Gaussian models, the error terms ϵ_t are modelled as normal random variables, independent of past values $\epsilon_{t-1}, \epsilon_{t-2}, \ldots$. A sufficient condition for an autoregressive process to be stationary is that the roots of the so-called characteristic polynomial lie outside the unit circle. This restricts the autoregressive coefficients $\phi_1, \ldots, \phi_p$ to lie in a constrained space called the stationary region, for example $\phi_1 \in (-1, 1)$ if $p = 1$. When the order of the process is greater than $p = 2$, the stationary region has a complicated shape which presents challenges for Bayesian inference. Fortunately, a simple reparameterisation of the model in terms of the partial autocorrelations at lags 1 through p eliminates this problem, requiring only that each partial autocorrelation lies in the open interval $(-1, 1)$.

In this project, students will take a Bayesian approach to inference, fitting stationary models using Markov chain Monte Carlo methods in R. This could involve writing bespoke code or it could involve a probabilistic programming language like JAGS or Stan, interfaced through R. Beyond this, there is flexibility over the precise topic of research. Possible areas for exploration include

- Development and comparison of time-series models built from a stationary autoregression for a particular real-world data set,
- Determination of the order of the process,
- Identification and modelling of change points in a piecewise stationary process,
- Evaluating evidence for unit roots (non-stationarity),
- Extension to binary outcomes using a probit model,
- Sparsity in order 1 vector autoregressions where the time-series is multivariate.

Essential prior modules

- MATH3421: Bayesian Computation and Modelling.

Essential companion modules

- MATH4341: Spatio-Temporal Statistics.

Suggested resources

The following resources provide references, examples and a comprehensive treatment of inference for autoregressions:

- Barnett, G., R. Kohn, and S. Sheather (1996). Bayesian estimation of an autoregressive model using Markov chain Monte Carlo. *Journal of Econometrics.* **74**, 237–254.
- Chib, S. (1993). Bayes estimation of regressions with autoregressive errors: A Gibbs sampling approach. *Journal of Econometrics.* **58**, 275–294
- Congdon, P. (2003). *Applied Bayesian Modelling.* John Wiley & Sons.
- Marriott, J., N. Ravishanker, A. Gelfand, and J. Pai (1996). Bayesian analysis of ARMA processes: complete sampling-based inference under exact likelihoods. In D. A. Berry, K. M. Chaloner, and J. K. Geweke (Eds.), *Bayesian Analysis in Statistics and Econometrics*, pp. 243–256. John Wiley & Sons.
- Prado, R. and West, M. (2010). *Time Series: Modelling, Computation, and Inference.* CRC Press.

Supervisor

Sarah Heaps (`sarah.e.heaps@durham.ac.uk`)