Conference Report

12th International Conference on Bayesian Nonparametrics

This past 24-28 of June, Oxford hosted the 12th International Conference on Bayesian Nonparametrics. An impressive number of 220 participants from 21 countries (per affiliation) gave life to a diverse and high quality programme which covered all the recent advances in the field. The programme included three plenary talks given by Aad van der Vaart, Tamara Broderick and Long Nguyen, 9 invited sessions, 12 contributed sessions and 2 poster sessions. The scientific committee, led by Antonio Lijoi, put together an exciting programme which covered topics such as Bayesian inverse problems, BNP asymptotics and credible sets, random measures and predictive inference in BNP, BNP computations, foundational aspects of BNP, BNP for discrete structures, BNP for Biostats and BNP for high dimension. The poster sessions displayed 65 posters of high quality. 82 Phd students attended the conference and 43 travel funds were awarded to early career researchers. This included 8 ISBA travel funds awarded to Cecilia Balocchi, Diana Cai, Amine Hadji, Didong Li, Zacharie Naulet, Bo Ning, Miriam Shiffman and Giacomo Zanella. A junior ISBA reception, intended for students or researchers within 5 years of having completed their degree, was held on Wednesday evening. Non ISBA members were very welcome to attend. A childcare service was provided during the conference.

The 13th edition of the conference will be held in Chile in 2021. I am sure we are looking forward to this meeting in 2021!

– François Caron and Fabrizio Leisen

Members' News

Childcare at BNP

This year, the Bayesian Nonparametrics conference has fostered novel collaborations of sorts by attracting a new generation of early Bayesians. They interacted intensely every day, 9am-5pm, under the supervision of two fantastic day care women. Their different mother tongues, Spanish, Italian and French, leaves open the question of the language of their actual communication. Day care activities included daily visits to Oxford's Museum of Natural History, manual activities, games of any sorts, etc. From our parents point of view, we are indebted to the organizers Fabrizio, François, and all the local organizing committee for proposing this service, and to the whole community for supporting it. Thanks to it, we were able to attend the conference sessions seamlessly, with the extra benefit of sharing with our children this key and joyful moment in researchers' life that is a (BNP) conference!

SOFTWARE HIGHLIGHT

SEMIPARAMETRIC REGRESSION AND THE HRW R PACKAGE

Matt P. Wand
University of Technology Sydney
matt.wand@uts.edu.au

HRW is a new R package on the Comprehensive R Archive Network (`cran.r-project.org`) and supports the new book *Semiparametric Regression with R* by Jarek Harezlak, David Ruppert and Matt P. Wand (`http://semiparametric-regression-with-r.net`). The package contains datasets, scripts for illustrative semiparametric regression analyses and some support functions.

A feature of Harezlak, Ruppert and Wand (2019) is that it is one of the first books to make extensive use of the Stan Bayesian inference engine (Carpenter *et al.*, 2017) via the R package rstan. The HRW

package contains 20 scripts that perform Bayesian semiparametric regression analyses via rstan. The first of these is a script named `WarsawAptsBayes.R` that performs Bayesian nonparametric regression for some scatterplot data concerning apartments in the city of Warsaw. The R commands:

```
library(HRW) ; demo(WarsawAptsBayes,package="HRW")
```

run the script. If a user wants to study the code for analysis of a different data set then the R command:

```
system.file("demo","WarsawAptsBayes.R",package = "HRW")
```

returns the full path name of the script file on the user's computer. The user has the option of copying `WarsawAptsBayes.R` to a new file and modifying the code for analyses of other datasets.

Even though the HRW package and corresponding book contain several Bayesian examples, it is for pragmatic rather than philosophical reasons. Indeed, most of the book is on non-Bayesian methodology and R software such as the gam (Hastie, 2018) and mgcv (Wood, 2019) packages which support generalized additive model analyses with barely a prior or credible interval in sight. Bayesian inference engines are included and recommended due to the extra flexibility they provide. Examples (with HRW package illustrative script) are:

- contrast curves with variability bands (pointwise credible intervals) for the *difference* between two regression functions. (`WarsawAptsSimpFacByCurv.R`),
- less common response distributions such the skew extension of the t-distribution developed by Jones and Faddy (2003) (`MichIncMCMCskewt.R`),
- marginal generalized additive mixed for longitudinal data as described in Al Kadiri, Carroll & Wand (2010) (`margAddMod.R`),
- nonparametric regression with missing predictor data (`PIDana.R`),
- nonparametric regression with data subject to measurement error (`BCRana.R`), and
- additive models that cater for heteroscedasticity (`ozoneAna.R`).

The HRW package has only a few functions, but the one named `summMCMC()` should be of interest to anyone carrying out Bayesian analyses using Markov chain Monte Carlo methodology. Given Markov chain Monte Carlo samples, or *chains*, for one or more parameters of interest `summMCMC()` provides a single figure summary. Its use can be explained without even doing a Bayesian analysis. Simply type:

```
library(HRW)
xListSingleChain <- list(cbind(rnorm(100),rnorm(100),
                         rnorm(100),rnorm(100)))
summMCMC(xListSingleChain,parNames = list("par1","par2","par3","par4"))
```

and then sit back and marvel at the resulting visual feast. If your Bayesian analysis involves multiple chains then the commands:

```
xListMultipleChains <- list(chain1 = cbind(rnorm(100),rnorm(100),
                                           rnorm(100),rnorm(100)),
                            chain2 = cbind(rnorm(100),rnorm(100),
                                           rnorm(100),rnorm(100)))
summMCMC(xListMultipleChains,
         parNames = list("par1","par2","par3","par4"))
```

add Brooks-Gelman-Rubin (Brooks & Gelman, 1998; Gelman & Rubin, 1992) diagnostic plots. The help page for `summMCMC()` provides full details on tailored use of this function.

References

Brooks, S.P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, **7**, 434–455.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, **76**, Issue 1, 1–32.

Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–511.

Harezlak, J., Ruppert, D. and Wand, M.P. (2018). *Semiparametric Regression with R*. New York: Springer.

Hastie, T. (2018). gam: Generalized additive models. R package version 1.14. `http://www.r-project.org`.

Jones, M.C. and Faddy, M.J. (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society, Series B*, **65**, 159–174.

Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.18.2. `https://mc-stan.org/`.

Wood, S.N. (2019). mgcv: Mixed GAM computation vehicle with GCV/AIC /REML smoothness estimation. R package version 1.8. `http://cran.r-project.org`.