OBANSoft: integrated software for Bayesian statistics and high performance computing with R

Manuel Quesada[1,⋆], **Domingo Giménez**[1], **Asunción Martínez**[2]

1. Universidad de Murcia
2. Universidad Miguel Hernández
⋆Contact author: manuel.quesada@um.es

Keywords: Bayesian statistics, parallelism, OpenMP, MPI, CUDA

Among all the applications for Bayesian analysis, there is none that integrates the whole process of Objective Bayesian Analysis. The aim of the "Software for Objective Bayesian Analysis" (*OBANSoft*) is to cover this gap. The first version includes the easiest models (Yang and Berger (1996)) but its design will allow for more complex algorithms in time.

The statistics engine has been implemented using *R*. Implementing the algorithms with this language lets us take advantage of the increasing number of new routines that the *R* community is developing, so we can focus the effort on other purposes such as high performance computing. For this reason, *OBANSoft* has been designed taking into account that it is going to incorporate a variety of parallelism levels, and it will be the first *R* parallel in a shared memory architecture. Moreover, the heterogeneity of parallel architectures will obstruct non advance users in parallelism optimizing the performance of their programs without the help of an expert. Concerning parallel performance, *OBANSoft* is ready to include an auto tuning module that configures the best parameters to execute the parallel algorithms (Katagiri et al. (2004)).

Although the statistics engine is developed in *R*, this is envolved in a top layer defined like a *Java* interface (using ***JRI*** to link them). This main layer coordinates the lower layers where the different engines based on several architectures are implemented. Indeed, this top layer will implement the auto-tuning algorithms. The final user will only use a desktop application that will solve a problem in the most efficient way for each platform.

The first version is a complete and integrated *Java* Desktop Application (*OBANSoft*) implementing the first models used to teach Bayesian analysis (Quesada (2010)). All the models are implemented using *R* libraries that have been reorganized to compose our *R* engine. The ***SnowFall*** library (SnowFall (2011)) allows exploitation of parallelism in multicore systems. However, the performance of this library is not satisfactory enough and other parallelism strategies need to be considered.

In conclusion, we have developed the base application where we are going to include more complex models with higher computational needs. Our concern is to implement and link the *R* engine with other parallel languages and high performance libraries (BLAS, OpenMP, MPI, CUDA, etc.). *OBANSoft* will assume all those new routines, including auto-tuning decisions. Our goal is to distance the final user from the problems of parallel computation.

Referencias

Katagiri, T., K. Kise, H. Honda, and T. Yuba (2004). Effect of auto-tuning with user's knowledge for numerical software. In *Proceedings of the 1st conference on Computing frontiers*, pp. 12–25. ACM.

Quesada, M. (2010, Julio). Obansoft: aplicación para el análisis bayesiano objetivo y subjetivo. estudio de su optimización y paralelización. Master's thesis, Universidad de Murcia.

SnowFall (2011). Url. http://cran.r-project.org/web/packages/snowfall/.

Yang, R. and J. O. Berger (1996). A catalog on noninformative priors. *Discussion Paper, 97-42, ISDS, Duke University, Durham, NC*.