Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > stat > arXiv:1305.4886

Search or Article-id

All papers

(Help | Advanced search)

Go!

Statistics > Computation

Parallelizing Gaussian Process Calculations in R

Christopher J. Paciorek, Benjamin Lipshitz, Wei Zhuo, Prabhat, Cari G. Kaufman, Rollin C. Thomas

(Submitted on 21 May 2013)

We consider parallel computation for Gaussian process calculations to overcome computational and memory constraints on the size of datasets that can be analyzed. Using a hybrid parallelization approach that uses both threading (shared memory) and message-passing (distributed memory), we implement the core linear algebra operations used in spatial statistics and Gaussian process regression in an R package called bigGP that relies on C and MPI. The approach divides the matrix into blocks such that the computational load is balanced across processes while communication between processes is limited. The package provides an API enabling R programmers to implement Gaussian process-based methods by using the distributed linear algebra operations without any C or MPI coding. We illustrate the approach and software by analyzing an astrophysics dataset with n=67,275 observations.

Comments:	21 pages, 8 figures
Subjects:	**Computation (stat.CO)**; Mathematical Software (cs.MS)
Cite as:	**arXiv:1305.4886 [stat.CO]**
	(or **arXiv:1305.4886v1 [stat.CO]** for this version)

Submission history

Which authors of this paper are endorsers?

Download:

- PDF
- Other formats

Current browse context:
stat.CO
< prev | next >
new | recent | 1305

Change to browse by:

cs
 cs.MS
stat

References & Citations

- NASA ADS

Bookmark (what is this?)