



Parallel Gaussian Process Regression with Low-Rank Covariance Matrix Approximations

Jie Chen, Nannan Cao, Kian Hsiang Low, Ruofei Ouyang, Colin Keng-Yan Tan, Patrick Jaillet

(Submitted on 24 May 2013)

Gaussian processes (GP) are Bayesian non-parametric models that are widely used for probabilistic regression. Unfortunately, it cannot scale well with large data nor perform real-time predictions due to its cubic time cost in the data size. This paper presents two parallel GP regression methods that exploit low-rank covariance matrix approximations for distributing the computational load among parallel machines to achieve time efficiency and scalability. We theoretically guarantee the predictive performances of our proposed parallel GPs to be equivalent to that of some centralized approximate GP regression methods: The computation of their centralized counterparts can be distributed among parallel machines, hence achieving greater time efficiency and scalability. We analytically compare the properties of our parallel GPs such as time, space, and communication complexity. Empirical evaluation on two real-world datasets in a cluster of 20 computing nodes shows that our parallel GPs are significantly more time-efficient and scalable than their centralized counterparts and exact/full GP while achieving predictive performances comparable to full GP.

Comments: 29th Conference on Uncertainty in Artificial Intelligence (UAI 2013), Extended version with proofs, 13 pages

Subjects: **Machine Learning (stat.ML)**; Distributed, Parallel, and Cluster Computing (cs.DC); Learning (cs.LG)

Download:

- [PDF](#)
- [Other formats](#)

Current browse context:

stat.ML

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1305](#)

Change to browse by:

cs

[cs.DC](#)

[cs.LG](#)

[stat](#)

References & Citations

- [NASA ADS](#)

Bookmark (what is this?)

