Jian Cao

Postdoctoral Researcher
Statistics Department & Institute of Data Science
Texas A&M University, 155 Ireland St, College Station, TX 77840, USA
Email: jian.cao@tamu.edu
Website: https://jcatwood.github.io/

Date: June 8, 2023

Education

2020 *Ph.D.* in Statistics, King Abdullah University of Science and Technology
2016 *M.Sc.* in Finance, Shanghai Jiaotong University
2014 *B.Sc.* in Applied Mathematics, University of Science and Technology of China

Areas of Specialization

Gaussian Processes, Variable Selection, Spatial Statistics, Computational Statistics, Low-rank Methods, High-performance Computing

Journal Articles

2023 Cao*, J., Kang, M.*, Jimenez, F., Sang, H., Schäfer, F., & Katzfuss, M. (2023). "Variational Sparse Inverse Cholesky Approximation for Latent Gaussian Processes via Double Kullback-Leibler Minimization," accepted by the 40th International Conference on Machine Learning

2022 Cao, J., Guinness, J., Genton, M. G., & Katzfuss, M. (2022). "Scalable Gaussian-process Regression and Variable Selection using Vecchia Approximations," *Journal of Machine Learning Research*, 2022, **23**(348), pp.1-30

2022 Cao, J., Durante, D., Genton, M. G. (2022). "Scalable Computation of Predictive Probabilities in Probit Models with Gaussian Process Priors," accepted by *Journal of Computational and Graphical Statistics* 2022, **31**(3), pp.709-720

2022 Cao, J., Genton, M. G., Keyes, D. E., & Turkiyyah, G. M. (2022). "tlrmvnmvt: Computing High-Dimensional Multivariate Normal and Student-t Probabilities with Low-rank Methods in R," *Journal of Statistical Software*, **101**, pp.1-25

2022 Abdulah, S., Li, Y., Cao, J., Ltaief, H., Keyes, D. E., Genton, M. G., & Sun, Y. (2022). "Large-scale Environmental Data Science with ExaGeoStatR," accepted by *Environmetrics*

2021 Cao, J., Genton, M. G., Keyes, D. E., & Turkiyyah, G. M. (2021). "Exploiting Low Rank Covariance Structures for Computing High-Dimensional Normal and Student-t Probabilities," *Statistics and Computing*, **31**(1), pp.1-16

2021 Cao, J., Genton, M. G., Keyes, D. E., & Turkiyyah, G. M. (2021). "Sum of Kronecker Products Representation and Its Cholesky Factorization for Spatial Covariance Matrices from Large Grids," *Computational Statistics & Data Analysis*, **157**, pp.107165

2021 Huang, J., Fang, F., Turkiyyah, G., Cao, J., Genton, M. G., & Keyes, D. E. (2021). "An $O(N)$ Algorithm for Computing Expectation of N-dimensional Truncated Multi-variate Normal Distribution I: Fundamentals," *Advances in Computational Mathematics*, **47**(5), pp.1-34

2019 Cao, J., Genton, M. G., Keyes, D. E., & Turkiyyah, G. M. (2019). "Hierarchical-block Conditioning Approximations for High-dimensional Multivariate Normal Probabilities," *Statistics and Computing*, **29**, pp.585-598

2023 Cao, J., Zhang, J., Sun, Z., & Katzfuss, M. (2022). "Locally Anisotropic Covariance Functions on the Sphere," in revision for *Journal of Agricultural, Biological and Environmental Statistics*

Talks & Posters

2023 ***2023 International Indian Statistical Association Conference*** Golden, CO, USA
Invited Session: Variational sparse inverse Cholesky approximation for latent Gaussian processes via double Kullback-Leibler minimization

2023 ***ASA/IMS SPRING RESEARCH CONFERENCE 2023*** Banff, Canada
Contributed Session: Variational sparse inverse Cholesky approximation for latent Gaussian processes via double Kullback-Leibler minimization

2022 ***ENVR 2022 Workshop*** Provo, UT, USA
Poster: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2022 ***IMSI Gaussian Processes Workshop*** Chicago, IL, USA
Poster: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2022 ***Joint Statistical Meetings*** Washington D.C., USA
Contributed Session: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2022 ***ISBA World Meeting*** Montreal, Quebec, Canada
Contributed Talk: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2022

SETCASA Poster Competition College Station, TX, USA
Poster: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2022 ***Texas A&M Statistics Cafe*** College Station, TX, USA
Presentation: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2021 ***TAMIDS Research Conference*** College Station, TX, USA
Presentation: Scalable Gaussian Process Regression and Variable Selection under Automatic Relevance Determination Kernels

2020 ***Joint Statistical Meetings*** Virtual Conference
Contributed Session: Sum of Kronecker Products Representation for Spatial Covariance Matrices and Its Factorization

2019 ***Joint Statistical Meetings*** Denver, CO, USA
Topic-Contributed Session: Exploiting Low Rank Covariance Structures for Computing High-Dimensional Normal and Student-t Probabilities

2018 ***Big Data Meets Large-Scale Computing*** IPAM, Los Angeles, CA, USA
Poster: Exploiting Low Rank Covariance Structures for Computing High-Dimensional Normal and Student-t Probabilities

2018 ***Joint Statistical Meetings*** Vancouver, BC, Canada
Poster: Hierarchical-block Conditioning Approximations for High-dimensional Multivariate Normal Probabilities

2017 ***Joint Statistical Meetings*** Baltimore, MD, USA
Contributed Session: Hierarchical-block Conditioning Approximations for High-dimensional Multivariate Normal Probabilities

Awards

2020 ***Al-Kindi Statistics Student Research Award***
King Abdullah University of Science and Technology

2019 ***Winner of the Student Paper Competition***, Section on Statistical Computing and the Section on Statistical Graphics of ASA
Title: "Exploiting Low Rank Covariance Structures for Computing High-Dimensional Normal and Student-t Probabilities"

Short Courses

2019

A Short Course on Deep Learning, KAUST Saudi Arabia
2017 ***Winter School on Hierarchical Matrices***, Kiel Germany

Teaching

2022 April TAMIDS Webinar "Scalable Gaussian Process Approximation and Optimization
2018 Fall Teaching Assistant for MS level *Probability and Statistics*
2017 Fall Teaching Assistant for MS level *Probability and Statistics*

Programming Languages

R, C++, and Python

R Package

tlrmvnmvt, published on CRAN
Compute high-dimensional multivariate normal (MVN) and multivariate Student-t (MVT) probabilities with tile-low-rank and block reordering (LINK)