Colin Cui

456 Snowden Ln
Princeton, New Jersey 08540

Phone: 415-518-3959
E-mail: colstat@gmail.com

Summary

I work at the intersection of statistics, machine learning, and convex optimization problems. I have experience in understanding theoretical results as well as its applications.

Education

Rutgers University
MS., Statistics

University of California at Davis
B.S., Statistics

Coursework
Probability Theory (Prof.William Strawderman), Statistical Inference (Prof.William Strawderman), Decision Theory (Prof.Harold Sackrowtiz), Data Mining (Prof.Javier Cabrera), Bayesian Data Analysis (Prof.Tong Zhang), Interpretation of Data (Prof. Minge Xie), Machine Learning (Coursera Andrew Ng), Statistical Learning and Nonparametric Estimation (Princeton, Prof. Philippe Rigollet).

Projects

Sparsity Recovery: Basis Pursuit/Lasso
Features dimensionality reduction by solving the quadratic minimization problem to recover sparsity. Since solving ℓ_0-norm is NP-hard, we gave ℓ_1-norm convex relaxation as surrogate for sparsity recovery.

Classification: Random Forest
Imported data, performed exploratory data analysis, and plotted heatmap using *seaborn* package. Built random forest using *scikit-learn*, and evaluate model accuracy performance.

Image Processing: Eigenvalue Decomposition
Solving singular value problems with top k singular values and singular vectors to minimize Frobenius norm objective for image compression.

Analysis Proofs
Proving differentiability, lipchitz continuity, uniform convergence, and integrability of series of functions.

Software

Languages: R, Python, Julia, Matlab, Stata

Experience

Rutgers University, Department of Engineering *Piscataway, New Jersey*
Research Scholar (Supervisor: Prof. Pelegri)

- Solving inverse problems using GP (Gaussian Processes) as metamodel for Bayesian calibration
- Quantifying parameter uncertainty with simulation on the posterior
- Presented at SIAM Conference regional in Philadelphia, PA

Rutgers University, Department of Statistics *Piscataway, New Jersey*
Research Assistant (Supervisor: Prof. Tan)

- Resampling MCMC results with tempering, resampling, and Markov moving using proposed general framework
- Literature review and simulation
- Worked on preparing manuscript for publication

Rutgers Medical School (former UMDNJ) *New Brunswick, NJ*

Data Analyst (Supervisor: Dr. Vivian Hsu)

- Provided statistical analysis on prospect sclerosis patients at multiple clinical centers
- Prepared results on scheroderma hospital patients for publication

Conference Bayesian Inference Using Gaussian Process Metamodel in Biomedical Imaging (*with* A. Pelegri, and X. Zhao). *Conf. Mathematical Aspect of Material Science,* 2016 link

Papers GoldBerg, et al. Clinical Outcomes of Scleroderma Patients At High Risk for Pulmonary Hypertension. Analysis of the Pulmonary Hypertension Assessment and Recognition of Outcomes in Scleroderma Registry. *ACR/ARHP Annual Meeting*, 2012. (**acknowledged**)

Participation Tan, Zhiqiang. Resampling Markov Chain Monte Carlo Algorithms: Basic Analysis and Empirical Comparisons, *Journal of Computational and Graphical Statistics*, 24, 328-356