Colin Cui

456 Snowden Ln
Princeton, New Jersey 08540

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

Research Interests

Statistics, machine learning, and convex optimization. I am interested in understanding the theoretical results and its applications.

Education

Rutgers University
MS., Statistics

University of California at Davis
B.S., Statistics

Coursework
Probability Theory (Prof. William Strawderman, Ph.D. course),
Statistical Inference (Prof. William Strawderman, Ph.D. course),
Decision Theory (Prof. Harold Sackrowtiz, Ph.D. course),
Data Mining (Prof. Javier Cabrera, Master's course),
Bayesian Data Analysis (Prof. Tong Zhang, Ph.D. course),
Interpretation of Data (Prof. Minge Xie, Ph.D. course),
Statistical Learning and Nonparametric Estimation (Prof. Philippe Rigollet, at Princeton).

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.

Software

Languages: R, Python, Julia, Matlab, Stata

Experience

P1 Consulting Princeton, New Jersey
Statistician/Consultant
- building analytical model for high dimensional statistics using statistical software R
- Running Python code and plotting in ggplot2 in R

Rutgers University *Piscataway, New Jersey*
Research Scholar at Prof. Pelegri's group
- Solving inverse problems using Gaussian Processes as metamodel for Bayesian calibration
- Quantifying parameter uncertainty with simulation on the posterior

- Presented at SIAM Conference regional in Philadelphia, PA

New Jersey Institute of Technology *Newwark, New Jersey*

Adjunct Faculty

- Duties include: teaching statistics course, review student progress, performance, registration

Rutgers University, Department of Statistics *Piscataway, New Jersey*

RA for Professor Zhiqiang Tan

- Simulated Monte Carlo methods for numerical approximation using statistical software R
- Performed stochastic approximation to advanced MCMC algorithm
- Worked on preparing manuscript for publication

Robert Wood Johnson Medical School *New Brunswick, NJ*

Graduate Assistant (Supervisor: Dr. Qingyu Meng)

- Running Lasso regression and penalized regression for high-dimensional data
- implementing feature selection, shrinkage, and sparsity recovery
- code, build, and debug in R software

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