# 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 both the theoretical results of statistical learning 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 in R: Basis Pursuit/Lasso**
Features dimensionality reduction by solving the quadratic minimization problem using $\ell_0$-norm to recover its sparsity. Since solving $\ell_0$-norm is NP-hard, we give convex relaxation using $\ell_1$-norm as a surrogate. PIMA Indians was the dataset.

**Classification in Python: 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. Python was the software used.

**Mathematical Analysis (current)**
Proving differentiable functions, lipchitz continuity, uniform convergence, and integrability of series of functions.

**Image Processing in Matlab: Eigenvalue Decomposition**
Solving singular value problems with top k singular values and singular vectors to minimize Frobenius norm objective for image compression. Matlab was the software used.

**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 Gaussian Processes as a latent model for Bayesian inference to quantify parameter uncertainty
- SIAM Conference presentation

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

- Tempering, resampling, and markov moving using proposed general framework of

resampling MCMC
- I worked on the manuscript for journal publication

**Rutgers Medical School (former UMDNJ)** *Piscataway, NJ*

*Data Analyst* (Supervisor: Dr. Hsu)
- Provide statistical analysis on prospect sclerosis patients at multiple clinical centers
- Prepare manuscript results on scheroderma hospital patients

**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