Varlam Kutateladze

varlam@varlam.net • varlam.net • varlamnet • @varlamnet
Economist • Data Scientist • Mathematician
900 University Avenue, Riverside, CA

Summary

PhD researcher with 4+ years' experience in predictive modeling & statistical inference using data science/machine learning methods and 3+ years of teaching quantitative classes. Check out my website for more info!

Research Area Machine Learning · High-Dimensional Statistics · Nonlinear Methods · Factor Modeling · Forecasting

Education

University of California, Riverside

PhD Econometrics (STEM) *Sep 2016 - Jun 2021*
MS Mathematics (GPA 4.0/4.0) *Sep 2019 - Jun 2021*

Research

"The Kernel Trick for Nonlinear Factor Modeling" *[Job Market Paper]* *Aug 2020*
Under review in the International Journal of Forecasting
"Nonlinear Shrinkage Covariance Matrix Estimation" *Oct 2020*
"Fast and Efficient Data Science Techniques for Covid-19 Group Testing" *(with E. Seregina)* *Sep 2020*

Presentations

- Econometric Society EWMES2020 *Dec 2020*
- 40th International Symposium on Forecasting *Oct 2020*
- OCLB American Statistical Association *Sep 2020*
- 2020 NABE Tech Econ Conference *Nov 2020*
- Department of Economics, UCR *Oct 2020*
- Data Science Conference on Covid-19 *Sep 2020*

Experience

University of California, Riverside

Instructor *Summer 2019, 2020*
Money, Banking & Credit (35 students, Eval: 5.0/5.0) · Introductory Econometrics (35 students, Eval: 4.0/5.0)
Teaching Assistant *Sep 2017 - Jun 2021*
Advanced Econometrics (PhD) · Statistics · Introductory Econometrics · Macroeconomics · Stock Market

Honors & Awards

2018 AI competition by Two Sigma (Halite), Top ~1% *New York, NY*
2016 University of California Dean's Distinguished Fellowship (~$220'000) *Riverside, CA*
2013 Oxford Russia Fellowship (~$10'000) *Moscow, Russia*
2010 Moscow Mathematical Olympiad, Winner *Moscow, Russia*

Skills

Python	SciPy · scikit-learn · statsmodels · Pandas · PyTorch · Keras · Tensorflow · Selenium
Data	MySQL · SQLite · BigQuery · Apache Spark · AWS · Docker
Visualization	Tableau · Plotly · Matplotlib · Seaborn
Other	R · Matlab · SAS · STATA · TEX · Git · Scala · JavaScript

Additional Information

Referee	Journal of Quantitative Economics · Studies in Nonlinear Dynamics & Econometrics
Affiliations	AEA · AMS · ASA · IIF · NABE
Other	CFA Level 1 (2014) · STEM OPT Work Authorization

References

Professor Jang-Ting Guo
Department of Economics, UCR
+1 951 827-1588
guojt@ucr.edu

Professor Tae-Hwy Lee (Chair)
Department of Economics, UCR
+1 951 827-1509
tae.lee@ucr.edu

Distinguished Professor Aman Ullah
Department of Economics, UCR
+1 951 827-1591
aman.ullah@ucr.edu

Paper Abstracts

"The Kernel Trick for Nonlinear Factor Modeling" *[Job Market Paper]* *Aug 2020*

Factor modeling is a powerful statistical technique that permits to capture the common dynamics in a large panel of data with a few latent variables, or factors, thus alleviating the curse of dimensionality. Despite its popularity and widespread use for various applications ranging from genomics to finance, this methodology has predominantly remained linear. This study estimates factors nonlinearly through the kernel method, which allows flexible nonlinearities while still avoiding the curse of dimensionality. We focus on factor-augmented forecasting of a single time series in a high-dimensional setting, known as diffusion index forecasting in macroeconomics literature. Our main contribution is twofold. First, we show that the proposed estimator is consistent and it nests linear PCA estimator as well as some nonlinear estimators introduced in the literature as specific examples. Second, our empirical application to a classical macroeconomic dataset demonstrates that this approach can offer substantial advantages over mainstream methods.

Keywords: Forecasting · Latent factor model · Nonlinear time series · kernel PCA · Neural networks · Econometric models

"Nonlinear Shrinkage Covariance Matrix Estimation" *Oct 2020*

Covariance matrix estimates are required in a wide range of applied problems in multivariate data analysis, including portfolio and risk management in finance, factor models and testing in economics, and graphical models and classification in machine learning. In modern applications, where often the model dimensionality is comparable or even larger than the sample size, the classical sample covariance estimator lacks desirable properties, such as consistency, and suffers from eigenvalue spreading. In recent years, improved estimators have been proposed based on the idea of regularization. Specifically, such estimators, known as rotation-equivariant estimators, shrink the sample eigenvalues, while keeping the eigenvectors of the sample covariance estimator. In high dimensions, however, the sample eigenvectors will generally be strongly inconsistent, rendering eigenvalue shrinkage estimators suboptimal. We consider an estimator that goes beyond mere eigenvalue shrinkage and employs recent advancements in random matrix theory to account for eigenvector inconsistency in a large-dimensional setting. We provide the theoretical guarantees and an empirical evaluation demonstrating the superior performance of the proposed estimator.

Keywords: Shrinkage estimator · Rotation equivariance · Random matrix theory · Large-dimensional asymptotics Bias correction · Principal components

"Fast and Efficient Data Science Techniques for Covid-19 Group Testing" *(with E. Seregina)* *Oct 2020*

Researchers and public officials tend to agree that until a vaccine is developed, stopping SARS-CoV-2 transmission is the name of the game. Testing is the key to preventing the spread, especially by asymptomatic individuals. With testing capacity restricted, group testing is an appealing alternative for comprehensive screening and has recently received FDA emergency authorization. This technique tests pools of individual samples, thereby often requiring fewer testing resources while potentially providing multiple folds of speedup. We approach group testing from a data science perspective and offer two contributions. First, we provide an extensive empirical comparison of modern group testing techniques based on simulated data. Second, we propose a simple one-round method based on ℓ_1-norm sparse recovery, which outperforms current state-of-the-art approaches at certain disease prevalence rates.

Keywords: Pooled Testing · Compressed Sensing · Sparse Recovery · Lasso · Sensing Matrix · SARS-CoV-2