

Optimal Covariance Matrix Estimation for High-dimensional Noise in High-frequency Data

Jinyuan Chang

Southwestern University of Finance and Economics

(with Qiao Hu, Cheng Liu and Cheng Yong Tang)

We consider high-dimensional measurement errors with high-frequency data. Our objective is on recovering the high-dimensional cross-sectional covariance matrix of the random errors with optimality. In this problem, not all components of the random vector are observed at the same time and the measurement errors are latent variables, leading to major challenges besides high data dimensionality. We propose a new covariance matrix estimator in this context with appropriate localization and thresholding, and then conduct a series of comprehensive theoretical investigations of the proposed estimator. By developing a new technical device integrating the high-frequency data feature with the conventional notion of alpha-mixing, our analysis successfully accommodates the challenging serial dependence in the measurement errors. Our theoretical analysis establishes the minimax optimal convergence rates associated with two commonly used loss functions; and we demonstrate with concrete cases when the proposed localized estimator with thresholding achieves the minimax optimal convergence rates. Considering that the variances and covariances can be small in reality, we conduct a second-order theoretical analysis that further disentangles the dominating bias in the estimator. A bias-corrected estimator is then proposed to ensure its practical finite sample performance. We also extensively analyze our estimator in the setting with jumps, and show that its performance is reasonably robust. We illustrate the promising empirical performance of the proposed estimator with extensive simulation studies and a real data analysis.