Robust estimation of high-dimensional vector autoregressive models.

Abstract

High-dimensional time series data appear in many scientific areas in the current data-rich environment. Analysis of such data poses new challenges to data analysts because of not only the complicated dynamic dependence between the series, but also the existence of aberrant observations, such as missing values, contaminated observations, and heavy-tailed distributions. For high-dimensional vector autoregressive (VAR) models, we introduce a unified estimation procedure that is robust to model misspecification, heavy-tailed noise contamination, and conditional heteroscedasticity. The proposed methodology enjoys both statistical optimality and computational efficiency, and can handle many popular high-dimensional models, such as sparse, reduced-rank, banded, and network-structured VAR models. With proper regularization and data truncation, the estimation convergence rates are shown to be nearly optimal under a bounded fourth moment condition. Consistency of the proposed estimators is also established under a relaxed bounded $(2 + 2\varepsilon)$-th moment condition, for some $\varepsilon \in (0, 1)$, with slower convergence rates associated with $\varepsilon$. The efficacy of the proposed estimation methods is demonstrated by simulation and a real example. This talk is based on the joint work with Ruey S. Tsay.