LDATS: Latent Dirichlet Allocation Coupled with Time Series Analyses

Juniper L. Simonis, Erica M. Christensen, David J. Harris, Renata M. Diaz, Hao Ye, Ethan P. White, S.K. Morgan Ernest

We analyze high-dimensional temporal data to determine if and when systems undergo abrupt regime shifts in the context of continuous change, autocorrelation, and stochasticity. Our LDATS package combines two established but disparate methods to tackle this problem: Latent Dirichlet Allocation (LDA) and Bayesian multivariate Time Series (TS) models. LDA first decomposes the high-dimension data into a lower-dimension latent representation while retaining key information. The proportions of the latent groups are then analyzed using generalized Bayesian TS models that include changepoints and smooth dynamics. Recent expansions to the LDATS package include cross-validation-based evaluation, alternative-to-LDA decomposition models, and simplex-based TS methods. We illustrate LDATS using a compendium of long-term ecological data sets, focusing on desert rodents and plants. LDATS is a well-designed, user-friendly package with a clearly-defined top-level API, application vignettes, and functions to produce attractive and accessible publication-ready figures. It represents a significant contribution to multivariate time-series analyses and has broad applicability across research modalities.