

Sparsity Learning Formulations for Mining Time-Varying Data

Traditional clustering and feature selection methods consider the data matrix as static. However, the data matrices evolve smoothly over time in many applications. A simple approach to learn from these time-evolving data matrices is to analyze them separately. Such strategy ignores the time-dependent nature of the underlying data. In this paper, we propose two formulations for evolutionary co-clustering and feature selection based on the fused Lasso regularization. The evolutionary co-clustering formulation is able to identify smoothly varying hidden block structures embedded into the matrices along the temporal dimension. Our formulation is very flexible and allows for imposing smoothness constraints over only one dimension of the data matrices. The evolutionary feature selection formulation can uncover shared features in clustering from time-evolving data matrices. We show that the optimization problems involved are non-convex, non-smooth and non-separable. To compute the solutions efficiently, we develop a two-step procedure that optimizes the objective function iteratively. We evaluate the proposed formulations using the Allen Developing MouseBrain Atlas data. Results show that our formulations consistently outperform prior methods.