Estimating Dynamic Graphical Dependencies from Multivariate Time-Series

Alex J. Gibberd & James D. B. Nelson

September 28, 2015

In our data-hungry society we are not only harvesting more data points but also measuring an ever-increasing number of variables. The complex systems represented by such data-sets arise in many socio-scientific domains, such as: cyber-security, neurology, genetics and economics. A dynamic, multivariate approach which models systems jointly can help us focus our analytic and experimental resources on the most significant inter-dependencies as they evolve through time. Unfortunately, such an approach comes with a cost, namely that the number of possible graphs which encode conditional dependencies become exponentially large as the number of variables increase. This problem is only compounded when we permit this dependency structure to vary through time.

We present and discuss a computationally efficient method that: (i) recovers the key dependency structure within a time-series and (ii) detects structural changepoints where the dependencies between variables appear to change at a systematic level (across many data-streams simultaneously). Building on foundations in regularised learning, we formulate a convex relaxed estimator for jointly selecting and estimating the partial correlation structure of a piecewise-stationary multi- variate Gaussian graphical model. This estimator, which we refer to as the Group-Fused Graphical Lasso (GFGL), can be understood as a semi-Bayesian approach where a sparsity inducing prior is applied to a set of precision matrices. The output of this model is a dynamic undirected graphical model encoding the conditional dependency structure of a multivariate time-series.

We solve the GFGL problem in the MAP setting, develop an efficient proximal splitting algorithm and demonstrate estimator performance on synthetic data. We further illustrate the potential utility of our method by applying it to pressing real-world problems, such as identifying dependencies and changepoints in gene activation.