VAR modeling, regularization and the SparseTSCGM package

Stijn de Vos

03 - 11 - 2017

Introduction

We start with a small review of vector-autoregressive models. Afterwards, regularization is discussed and the SparseTSCGM package is demonstrated.

Suppose we have four items called 'Positive Affect' (x_1), 'Anhedonia' (x_2), 'Irritability' (x_3) and 'Weightloss' (x_4). Measuring these items multiple times gives a time-series dataset where each observation consists of four elements. Written mathematically, the observation at time t consists of four numbers $(x_{1t}, x_{2t}, x_{3t}, x_{4t})$. As a shorthand notation, people often write $X_t = (x_{1t}, x_{2t}, x_{3t}, x_{4t})$. X_t is called a (four-dimensional) *vector*.

This is why it's called **vector**-autoregressive (VAR) modelling; we regress a vector X_t on past 'versions' of itself, instead of a single variable. The simplest VAR model only regresses X_t on its previous measurement X_{t-1}. This VAR model is said to have a **lag** of 1. In formula form, this model looks as follows:

$$x_{1t} = a_{11}x_{1(t-1)} + a_{12}x_{2(t-1)} + a_{13}x_{3(t-1)} + a_{14}x_{4(t-1)} + \epsilon_{1t} \tag{1}$$
$$x_{2t} = a_{21}x_{1(t-1)} + a_{22}x_{2(t-1)} + a_{23}x_{3(t-1)} + a_{24}x_{4(t-1)} + \epsilon_{2t} \tag{2}$$
$$x_{3t} = a_{31}x_{1(t-1)} + a_{32}x_{2(t-1)} + a_{33}x_{3(t-1)} + a_{34}x_{4(t-1)} + \epsilon_{3t} \tag{3}$$
$$x_{4t} = a_{41}x_{1(t-1)} + a_{42}x_{2(t-1)} + a_{43}x_{3(t-1)} + a_{44}x_{4(t-1)} + \epsilon_{4t} \tag{4}$$

Regularization

Conclusion