Theoretical study Quarto -document

2026-01-28

To do 4.2.

- Main consideration: Should measurement error be dropped from the s-LMI model and just call it autoregressive common factor model. This largely simplifies, and is justifiable in the sense that VAR(1) does not contain across time point measurement errors (innovations).
- Add/integrate previous Rmd file of theoretical result as extension to this one.
- Recheck delta t.
- Check if subindex VAR(1) is necessary for Sigma.
- Integrating Tom's notes (26.1 onwards)
- Serially independent innovations force the
- Many proofs are currently in scratch file, accessible from there.

Empirical distinguishability

Common Factor theoretical model A latent variable η is a causal entity which creates observable symptoms. Conditional independence of variables $P(X_i|\eta, X_j) = P(X_i|\eta) : i \neq j$ given the common factor η is the central proposition of the model. All changes in symptoms are due to changes in η. M_1 will represent an empirical (data-generating) model of the common factor theory.

Mutualistic symptom-network theoretical model Symptoms have causal effects on each others. Symptoms can change by themselves or change due to effects not within the symptom-network itself. The symptom-network is all the symptoms. M_2 will represent an empirical (data-generating) model of the symptom-network theory.

The models need to be kept vague for now, as the key parts of both differ in each scenario to be examined. Empirical distinguishability is here defined up to the fourth order moment:

Empirical distinguishability

$$
\begin{aligned}
E_{M_1}(X) - E_{M_2}(X) &= D_E \\
\mathrm{Cov}_{M_1}(X) - \mathrm{Cov}_{M_2}(X) &= D_C \\
\mathrm{Skew}_{M_1}(X) - \mathrm{Skew}_{M_2}(X) &= D_S \\
\mathrm{Kurt}_{M_1}(X) - \mathrm{Kurt}_{M_2}(X) &= D_S
\end{aligned}
$$

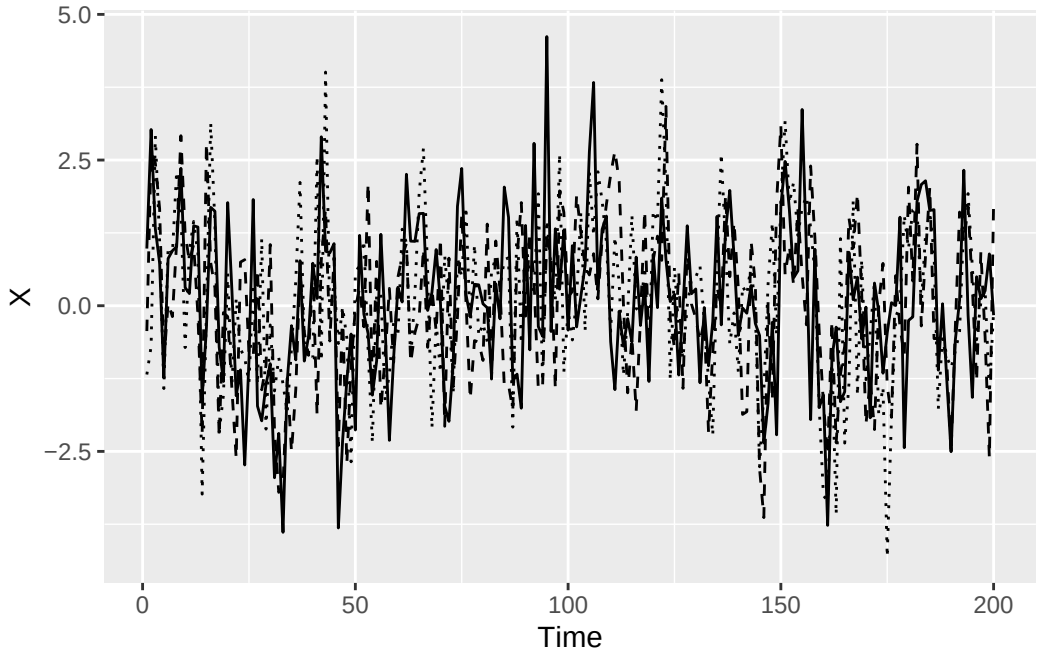
That is, after defining the models, data simulated from both of these models should always be the same w.r.t. their expectation, covariance, skewness and kurtosis. Our aim is to identify practical conditions, where the distinguishability is highest - i.e., when the differences in either expectation or covariance become largest. We do this in X different scenarios. In all scenarios we define a respective common factor theory model as well as a symptom-network mutualism model. We then aim to analytically show the differences between the models, and then through simulation quantify what types of practical difference we might see. We'll be denoting symptoms as random matrix X, of which columns are the symptoms. We use two theoretical models:

The several scenarios relate to study designs, in which both models have been used using real world empirical data. First, we examine changes over time. Second, ..., X:th.

Change across time: Strict Longitudinal Measurement Invariance vs VAR(1) models

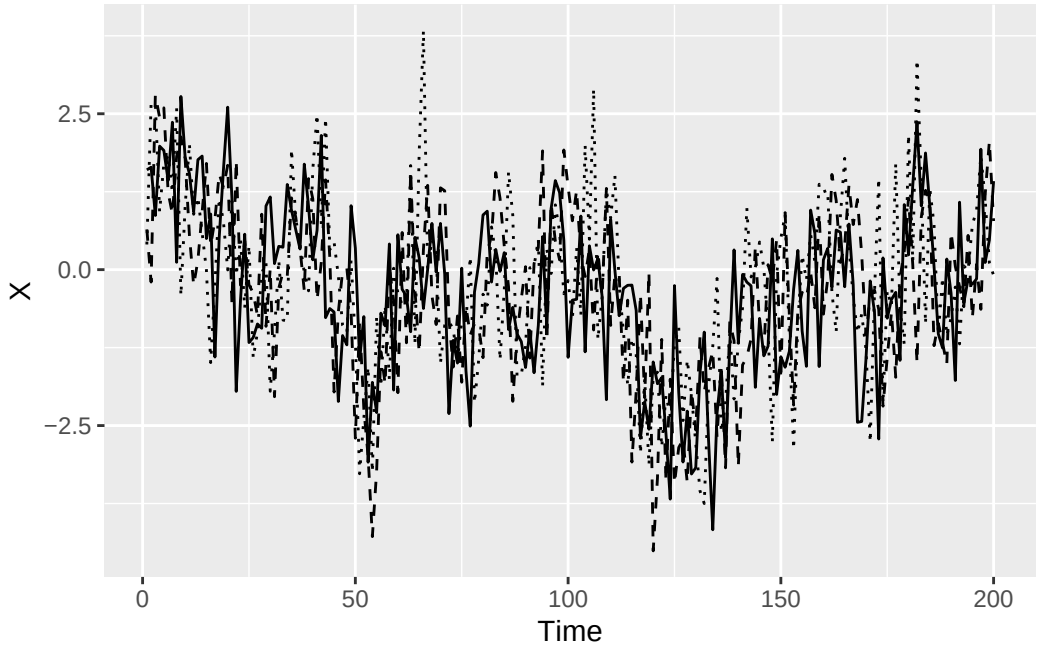
Regarding changes in symptoms across time, the most basic models that we can use for common factor and symptom-network theory are the strict longitudinal measurement invariance (s-LMI) referred to as and the vector autoregressive model of order one (VAR(1)). Here we will analytically approach differences between these models and then perform simulations to quantify how well empirically the models differ when generating from a mixture distribution. We will first inspect what type of covariance structure VAR(1) imposes, and then similarly s-LMI.

Their difference is not always evident. See for example the figure below, where data was generated from both models.



The model was set up so that there is a true common factor η causing three variables X. The common factor had an autoregressive coefficient (of order 1) of $\sqrt{0.5} \approx 0.7071$ to itself and $\psi_t \sim N(0, \sqrt{0.5})$ innovations (0.5 variance). Each of the $X_{i,t}$ also had their own measurement errors $\omega_{i,t} \sim N(0, \sqrt{0.5})$. X_t has a covariance matrix approximately such that the diagonal elements are 2, and off-diagonal elements are 1.

If simulating from another numerical example, this time a VAR(1) process, we can observe a fairly similar looking pattern.



The difference between the two example time series in D_C is

```
D_C = (cov(X_VAR) - cov(X_CF))
dimnames(D_C) = list(paste("X",1:3,sep = ""),paste("X",1:3,sep = ""))
round(D_C, 2)
```

```
      X1    X2    X3
X1 -0.23  0.06 -0.14
X2  0.06 -0.09 -0.07
X3 -0.14 -0.07 -0.12
```

VAR(1) covariance structure

Before the analysis, a brief theoretical consideration. We will assume that the symptom-network represented as the VAR(1) model is a state. This means that VAR(1) is stationary, and that there are no external causes meaning that the innovations are uncorrelated - only symptoms can correlate to each other. The VAR(1) model is defined in matrix format for K symptoms as $X_{t+1} = C + AX_t + \Gamma_t$, where Γ_t is independent error column vector with $E[\Gamma_t] = 0$, C is a constant assumed zero. Also assume centered X, $E[X_t] = 0$, in our case. Centering makes covariance calculations easier as the products of expected values can be mostly ignored (they become 0). A is $K \times K$ (borrowing from CT-VAR terminology) 'drift' matrix that includes all lagged effects of $K \times 1$ column vectors X_t to X_{t-1}. A is independent of time. All matrices used are real-valued.