

# TheoreticalStudy\_2024

Sakari Lintula, Jaakko Tammilehto, Tom Rosenström

## Notes

- Not meant to be readable necessarily, but to include main analytical results.
- Some proofs may be required if not found from sources.
- Most of mathematics might go to supplement, but this will be easy and fun to write after math part is finished.
- Do we actually need stationarity for VAR(1)?

## Distinguishability of Vector Autoregressive Symptom-Network and Common Factor models

### Previous research

Two central, juxtaposed, models to the comorbidity problem of psychopathology exist in the literature: Symptom-network models as well as common factor models. These are hard to distinguish. Perhaps the first notion of equivalence between mutualistic symptom networks and latent factor models was explicitly stated by (Molenaar 2010). Further studies have shown that cross-sectional networks can be estimated using latent variables (Marsman et al. 2015), that each factor model indeed produces a network model if considering the Gaussian case (Epskamp et al. 2018a) and for categorical variables there exists a map from multidimensional latent factor item response theory models to network models (Epskamp et al. 2018b). Criticism has been put forth for both. Recent Nature review discusses symptom-networks and factor/dimensional models of psychopathology (Eaton et al. 2023), however attempts at identifying distinguishability conditions for the two are scarce if non-existent.

**Common Factor theoretical model** A latent variable  $\eta$  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  $\eta$  is the central proposition of the model. All changes in symptoms are due to changes in  $\eta$  and measurement error, meaning that all covariance is explained by  $\eta$ .  $M_1$  denotes the common factor model.

**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.  $M_2$  denotes the symptom network model.

**Empirical distinguishability** The models are distinguishable up to some moment if  $M_1, M_2$  differ in that moment.

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

For example (Ip 2010) discusses empirical (in)distinguishability of correlated residual vs. multidimensional factor models.

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

Regarding changes in symptoms across time, the simplest models seen applied use that we can use for common factor and symptom-network theory are the strict longitudinal measurement invariance (s-LMI) and the vector autoregressive model of order one (VAR(1)). Notably any discrete time VAR model is a special case of continuous-time VAR model. As such, discrete time VAR works well for a simple theoretical model. 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.

### VAR(1) symptom-network covariance structure

As is typically done we will restrict our analysis to the stationary, non-seasonally trended (hence ergodic (Molenaar 2004)) VAR process. The VAR(1) model is defined in matrix format as  $X_t = C + AX_{t-1} + \Gamma_t$ , where  $\Gamma_t$  is independent ‘error’ or ‘innovation’ 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  $X_{t-1}$ ,  $K$  being the number of observed items (symptoms). In this section the focus is on the  $2K \times 2K$  covariance

matrix where two subsequent measurement time points are observed. All matrices used are real-valued.

First, the covariance matrix (assumed stationary over time) is

$$\begin{aligned}
\text{Cov}(X_t) &= E[X_t X_t^T] \\
&= E[(AX_{t-1} + \Gamma)(AX_{t-1} + \Gamma)^T] \\
&= E[AX_{t-1} X_{t-1}^T A^T] + \underbrace{E[\Gamma \Gamma^T]}_{=:\Psi} \\
&= AE[X_{t-1} X_{t-1}^T]A^T + \Psi \\
&= \Sigma_t = \Sigma
\end{aligned}$$

where stationarity poses that  $\Sigma$  is not dependent of time and so the covariance of VAR(1) is denoted as such from hereon.  $\Gamma$  is a random  $K \times 1$  column vector of (serially) independent innovations at time point  $t$ , with  $E[\Gamma] = 0$ .  $\Psi$  is covariance of the innovations within time point  $t$  - i.e., the contemporaneous covariance. We assume that the innovations are independent and so  $\Psi$  is diagonal. The vectorized covariance matrix can also be solved to equal (Liitkepohl, n.d.)

$$\text{vec}(\Sigma_{VAR(1)}) = (I - A \otimes A)^{-1} \text{vec}(\Psi)$$

(We'll use the above notation for VAR(1) covariance when clarity is needed.) Where  $\text{vec}$  is the vectorization operator and  $\otimes$  is the Kronecker product. In the above the mixed Kronecker matrix vector product is used to obtain the result. This will be handy when doing numerical demonstrations. Let  $(^p)$  denote matrix power of  $p$ .

Second, the covariance between two time points called cross covariance is

$$\Sigma_{t,t+\Delta t} = \text{Cov}(X_t, X_{t+\Delta t}) = A^{(\Delta t)} \Sigma$$

*(Proof can be done by induction if no reference is found. I checked for Delta = 1,2. - Sakari)*

The symmetric  $TK \times TK$  cross covariance matrix is

$$\begin{pmatrix}
\Sigma & A\Sigma & A^2\Sigma & \dots & A^{\Delta t}\Sigma \\
\Sigma_{t,t+1} & \Sigma & A\Sigma & \dots & \vdots \\
\Sigma_{t,t+2} & \Sigma_{t+1,t+2} & \Sigma & & \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\Sigma_{t,t+\Delta t} & \dots & \dots & \dots & \Sigma
\end{pmatrix}$$

which is the most important result as we will compare this to C-F imposed covariance. On a brief note we can further decompose the above equation, using the power method of eigenvalues, into

$$A^{(\Delta t)} \Sigma = P D^{(\Delta t)} P^{-1} \Sigma$$

where  $D$  is a diagonal matrix of eigenvalues of  $A$ ,  $P$  is an orthonormal matrix of eigenvectors of  $A$  as columns. Because eigenvalues of (stationary)  $A$  are less than one, i.e.,  $\text{diag}(D) : |d_{ii}| < 1$  as well, and hence the matrix power converges to  $D^{(\Delta t)} \xrightarrow[n]{\infty} 0$ . (*This restrains the analysis to psychopathological states as desired.*)

### S-LMI common factor covariance structure

Notably, there also exists the dynamic factor model (Molenaar 1985) which is a general case of a dynamic latent process. S-LMI differs in that it does not require stationarity to be estimable, the residuals (or ‘errors’ or ‘innovations’) are assumed to be independent of each other with diagonal covariance function (i.e., cross-sectional covariance and between time points covariance is diagonal) and that the cross-covariance between any two time points is symmetric as we will see below. s-LMI with 1 common factor decomposes the covariance as

$$\Sigma = \Lambda\Lambda^T + \Omega$$

where by definition of s-LMI  $\Omega$  is assumed diagonal and  $\Lambda$  is a  $K \times 1$  column vector of ‘factor loadings’ assumed constant over time. We also need the cross covariance of the common factor. Let  $\delta$  be the latent regression coefficient which links the common factor to itself at a previous time point such that  $\eta_t = \delta\eta_{t-1} + \psi_t$ , where  $\psi_t$  is independent random term (‘innovation’, ‘error’, ‘disturbance’) with  $E[\psi_t] = 0$ . Assuming standardized common factor such that  $E[\eta_{t-1}] = 0$ ,  $\text{Var}(\eta_{t-1}) = 1$  covariance of the common factor at two subsequent time points is

The s-LMI imposed  $TK \times TK$  covariance matrix is

$$\begin{pmatrix} \Lambda\Lambda^T + \Omega & \dots & \Lambda\Lambda^T \prod \delta^{(\Delta)} + \Omega_{\Delta} \\ \vdots & \ddots & \vdots \\ \Lambda\Lambda^T \prod \delta^{(\Delta)} + \Omega_{\Delta}^T & \dots & \Lambda\Lambda^T (\prod \delta + \text{Var}(\psi_t)) + \Omega \end{pmatrix}$$

## References

- Eaton, Nicholas R., Laura F. Bringmann, Timon Elmer, Eiko I. Fried, Miriam K. Forbes, Ashley L. Greene, Robert F. Krueger, et al. 2023. “A Review of Approaches and Models in Psychopathology Conceptualization Research.” *Nature Reviews Psychology* 2 (10): 622–36. <https://doi.org/10.1038/s44159-023-00218-4>.
- Epskamp, Sacha, Gunter Maris, Lourens J. Waldorp, and Denny Borsboom. 2018b. “Network Psychometrics.” In, 953–86. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118489772.ch30>.
- . 2018a. “Network Psychometrics.” In, 953–86. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118489772.ch30>.
- Ip, Edward Haksing. 2010. “Empirically Indistinguishable Multidimensional IRT and Locally Dependent Unidimensional Item Response Models.” *British Journal of Mathematical and Statistical Psychology* 63 (2): 395–416. <https://doi.org/10.1348/000711009X466835>.
- Liitkepohl, Helmut. n.d. “New Introduction to Multiple Time Series Analysis.”
- Marsman, Maarten, Gunter Maris, Timo Bechger, and Cees Glas. 2015. “Bayesian Inference for Low-Rank Ising Networks.” *Scientific Reports* 5 (1): 9050. <https://doi.org/10.1038/srep09050>.
- Molenaar, Peter C. M. 1985. “A Dynamic Factor Model for the Analysis of Multivariate Time Series.” *Psychometrika* 50 (2): 181–202. <https://doi.org/10.1007/BF02294246>.
- . 2004. “A Manifesto on Psychology as Idiographic Science: Bringing the Person Back into Scientific Psychology, This Time Forever.” *Measurement: Interdisciplinary Research and Perspectives* 2 (4): 201–18. [https://doi.org/10.1207/s15366359mea0204\\_1](https://doi.org/10.1207/s15366359mea0204_1).
- . 2010. “Latent Variable Models Are Network Models.” *Behavioral and Brain Sciences* 33 (2-3): 166–66. <https://doi.org/10.1017/S0140525X10000798>.