

Network Analysis 2017
Assignment 7

Sacha Epskamp

15-12-2017

$N = 1$ time-series

Download the file "Patient1.csv" from Blackboard and read it into R:

```
Data1 <- read.csv("Patient1.csv")
```

This is the data of the patient described in the paper uploaded the Blackboard. The variables `relaxed`, `sad`, `nervous`, `concentration`, `tired`, `ruminatation` and `bodily.discomfort` represent symptoms measured over time. These can be collected in a vector as follows:

```
Vars <- c("relaxed", "sad", "nervous",
          "concentration", "tired",
          "ruminatation", "bodily.discomfort")
```

The variable `time` represents the moment of measurement. We can use this to compute which measurements come from the same day:

```
# Encode time variable in a way R understands:
Data1$time <- as.POSIXct(Data1$time)

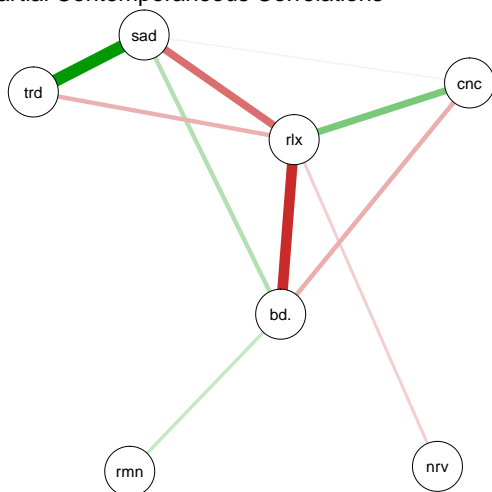
# Extract days:
Data1$Day <- as.Date(Data1$time)
```

Exercise 1 Install and load the `graphicalVAR` package and look at the help file for the function `graphicalVAR`.

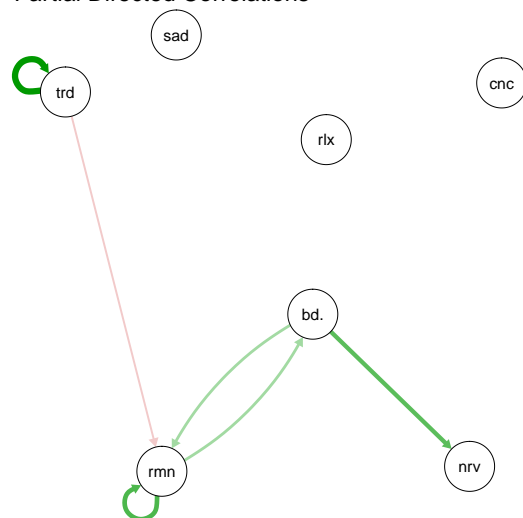
Exercise 2 Estimate a graphical VAR model on the dataset using EBIC tuning parameter $\gamma = 0$. Tip: use the arguments `vars = Vars` to set the variables to be estimated and `dayvar = "Day"` to remove nights.

We can use the `plot` method to plot the temporal and contemporaneous networks using the same layout, assuming the results are stored in the object `Results`:

Partial Contemporaneous Correlations



Partial Directed Correlations



Exercise 3

Interpret the estimated contemporaneous and temporal networks. There is no right solution to this

question, but you can read our thoughts on this in the paper in Blackboard. ■

Advanced: detrending

We can also use the time variable to check for trends in the data. To test if the variable “tired” features a trend, we can run a linear regression on time:

```
lm_tired <- lm(tired ~ time, data = Data1)
summary(lm_tired)

##
## Call:
## lm(formula = tired ~ time, data = Data1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36927 -0.32883 -0.01885  0.32075  2.50944
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.741e+03  4.062e+02  -4.286 6.34e-05 ***
## time         1.248e-06  2.904e-07   4.298 6.10e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7846 on 63 degrees of freedom
## (5 observations deleted due to missingness)
## Multiple R-squared:  0.2267, Adjusted R-squared:  0.2144
## F-statistic: 18.47 on 1 and 63 DF,  p-value: 6.1e-05
```

This effect is significant! We can now “detrend” tired as follows:

```
Data1$tired[!is.na(Data1$tired)] <- residuals(lm_tired)
```

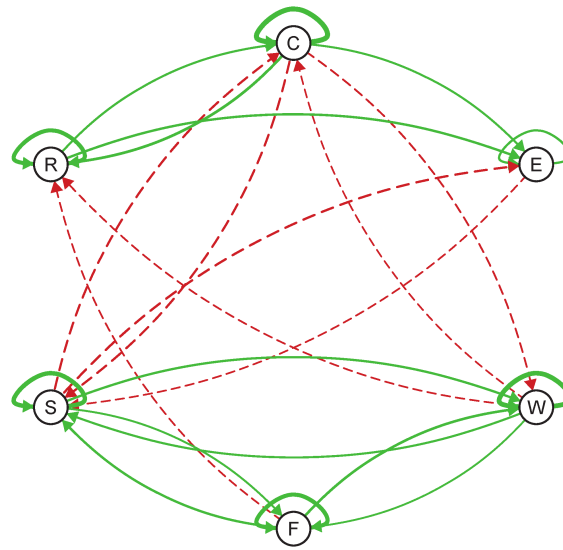
Exercise 4 (bonus)

Check all variables for trends and detrend if necessary. Note that this is a more advanced R question (bonus point) and you might want to skip it. ■

$N > 1$ time-series

In 2013, PlosOne published [a paper by Laura Bringmann and colleagues](#)¹ on the estimation of multi-level vector autoregression (VAR) models and their use in psychopathology. Bringmann and colleagues only investigated temporal networks, and reported the following temporal network at baseline:

¹Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., ... & Tuerlinckx, F. (2013). A network approach to psychopathology: new insights into clinical longitudinal data. *PLoS one*, 8(4), e60188.



C = cheerful, E = pleasant event, W = worry, F = fearful, S = sad and R = relaxed. This network was made with an older version of qgraph (self-loops are now drawn differently and edges curve opposite). The data used by Bringmann and colleagues was also published. On Blackboard you can find the subset of the data of patients at baseline. This can be loaded in R as follows:

```
Data2 <- read.table(file="BringmannData.csv",header=TRUE,sep=",")
head(Data2)

##      id day beep cheerful pleasant worry fearful sad relaxed
## 1 10720  1   1      NA        NA     NA      NA   NA     NA
## 2 10720  1   2      NA        NA     NA      NA   NA     NA
## 3 10720  1   3      NA        NA     NA      NA   NA     NA
## 4 10720  1   4         4         3      6       5   4     2
## 5 10720  1   5         2         1      5       4   4     2
## 6 10720  1   6         3         2      5       3   5     5
```

We can use this data to estimate a multi-level VAR model:

```
vars <- c("cheerful", "pleasant", "worry", "fearful", "sad", "relaxed")
idvar <- "id"
beepvar <- "beep"
dayvar <- "day"

library("mlVAR")

res <- mlVAR(Data2, vars, idvar, dayvar, beepvar, lags = 1,
             temporal = "correlated", contemporaneous = "correlated")
```

Exercise 5

Plot the average (fixed effects) temporal network, using, similarly to Bringmann and colleagues a circular layout and hiding non-significant edges.

The mlVAR package was originally based on the R codes supplied in the supplementary materials of this paper. Since then, centering of VAR models to distangle within- from between-subject processes has been worked out², allowing for a slightly different estimation of the VAR model now implemented in mlVAR.

²Hamaker, E. L., & Grasman, R. P. (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in psychology*, 5, 1492.

Finally, the underlying software used (the R package lme4) has gone through some changes as well. As a result, mlVAR now returns a slightly different network structure.

Exercise 6

Compare your network structure to that of Bringmann and colleagues. Do the networks differ? ■

As described in the lecture, mlVAR now also estimates contemporaneous networks. These networks show the relationships at the same time point after controlling for temporal effects.

Exercise 7

Plot the average (fixed effects) contemporaneous partial correlation network. Use a circle layout and hide non-significant edges. Use the "and" rule to minimize type-1 error rate in showing edges. Compare the network to the temporal network. Are the results what you expect? ■

In addition to the contemporaneous network, mlVAR also estimates a between-subjects network by investigating between-subject prediction of the mean of one variable given the means of other variables.

Exercise 8

Plot the between-subjects partial correlation network. Use a circle layout and hide non-significant edges. Use the "and" rule to minimize type-1 error rate in showing edges. Can you describe in words what a negative relationship between sad and experiencing pleasant events at the between-subjects level means? Compare the network to the temporal and contemporaneous networks, are the results what you expect? ■

The between-subjects network shows less connections, as this is a network between means there generally are less observations making it harder to reach statistical significance.