Assignment 3 - Network Analysis 2022

Kyuri Park, 12183881

December 06, 2022

Contents

Conceptual Questions

Question 1 (1 point)

Are the following statements true or false? Explain why (0.5 point per statement).

1. If ergodicity holds, results from between-person analysis are expected to equal results from within-person analysis.

> Yes, if ergodicity holds, there should be no difference between individuals (i.e., no difference in between-person), and correspondingly it aligns with the within-person analysis of every individual (Molenaar, 2004).

2. It is generally recommended to always remove trends (such as linear trends) prior to analyzing your N = 1 time series.

> No, even though some simulation stuides show detrending helps increasing the performance of estimated networks (Epskamp et al., 2018), it is often the case that these changing trends are of main interest. So it is not correct to say it is *always* recommended to detrend (Isvoranu et al., 2022).

Question 2 (1.5 points)

Suppose a therapist measures a patient about 75 times on a set of depression symptoms, including a question on "suicidal thought". Suppose that you estimate a graphical VAR model from this data and find that the node "suicidal thought" is not connected to any of the other nodes in your network, neither in the temporal nor in the contemporaneous network.

List three potential reasons why the node "suicidal thought" may be disconnected in the resulting network.

1. Probably the patient answered "*no*" every time on "suicidal thought, since it is quite an extreme symptom to develop. It is likely that the patient never had suicidal thought and that leads to no variance on that item, which will result in no connections in either of temporal or contemporaneous network (Hoekstra, 2022).
2. It could be the case that the estimated network model was quite large (> 10 nodes) while not having enough observations ($n = 75$ in this case , which is not large). This will result in lack of power, and hence failed to retrieve the edges connecting to "suicidal thought". It is in fact not established whether GVAR model can be reliably estimated from the N = 1 data sets (Isvoranu et al., 2022), especially when the model is large and sample size is relatively small, like in this case ($n = 75$).
3. Or it might be the case that the data is missing a lot on "suicidal thought". For example, the patient felt pressure or some sorts of discomfort to answer the question on "suicidal thought" and left the question unanswered most of the time. Then, the item would have a lot of missing values, which would likely to result in empty temporal/contemporaneous network.
4. Lastly, another potential scenario is that the measurement period was too short to capture the occurrence of "suicidal thought". For example, the 75 observations were measured by 5 times per day over 2 weeks, which is relatively a short period of time. It could be that "suicidal thought" only started occurring after the measurement period had ended, and hence left the item score on "suicidal thought" constant.
   Or there was a mismatch between the chosen time scale and "suicidal thought" processes, such that suicidal thought process is much slower than the chosen measurement frequency. It could be that the therapist measured the symptoms every 1-2 hrs, when the suicidal thought process operates over several days or so, which require more distance between assessments (Isvoranu et al., 2022). This could result in barely any variations on "suicidal thought" item and correspondingly leads to empty temporal and contemporaneous network.

Question 3 (1 point)

Give an example of a relationship that can only be studied at the *between-person* level.

Things that can only be studied at the *between-person* level have to be something that stay stable and do not fluctuate over time. A couple of example relationships that can be studied only in *between-person* level are:

- *genetic structure* and *cancer* : do people with a certain genetic structure have higher risk for cancer?
- *personality trait* and *zodiac sign*: do people with different zodiac signs have different personality traits?
- *average education level* and *race/nationality*: does the average education level differ across different races/nationalities?

Question 4 (2 points)

During the lecture we discussed multiple challenges regarding time-series modeling in the network approach. Pick your favorite challenge and explain this challenge in your own words. Make sure to not only explain what the challenge is, but also why this is a challenge: in what way may this challenge impact your results (i.e., the network model you estimate) and how may this jeopardize your conclusions?

One of the challenges in time-series network modelling is the difficulty of incorporating variables that operate on different time-scales. For example, the symptom network (e.g., network consist of symptoms) evolves relatively fast, as symptoms generally fluctuate over days. Whereas, psychological resilience, which is one of the important protective factors against developing mental disorder, evolves rather slowly and gradually (Lunansky et al., 2020). Obviously, resilience is part of the dynamic of psychopathology, but there exists no such model that can incorporate variables operating on the different time-scales. Absence of such a model prevents us from analyzing the whole network that integrates the slow-changing variables (e.g., resilience) with the fast-changing symptoms. Note that Lunansky et al. (2020) looked into the interaction between slow and fast network processes but it was studied in such a way that the slow changing variables influence the nodes in the symptom network instead of incorporating them all in one network model.

According to Borsboom & Van Borkulo (2021), omitting this relevant factor that operates on slower time-scale from the symptom network could lead to a much denser network. The reasoning is as follows. If the slow-changing variable (SV) is omitted in the symptom network model, then the edges will be estimated without controlling for the SV. Assuming that the SV has consistent influences on the symptoms, not controlling for the SV is likely to exaggerate the edges (i.e., partial correlations) between the symptoms. Therefore, it is expected that the symptom network model excluding the SV would appear denser (i.e., many thicker edges) than the integrated network model including the SV.

Hence, basing the network model only on the symptoms without incorporating the relevant SV can result in misleading model that overestimates the strength of relationships between symptoms. This could be especially detrimental to studies that look into a particular connection between two symptoms. If a SV that has a considerable influence on the symptom dynamics is not included in the model, the researcher may draw a conclusion based on one of the spurious connections in the network, which are resulted from overlooking the effects of the SV.[1]

Question 5 (2 points)

In a recent study by Haslbeck et al., (preprint) it was shown that use of a VAS or Likert scale has an affect on the observed distributions. Take a look at their article. Give (a) a short summary of the problem described in their paper, (b) reflect on this problem (e.g., why is this an issue and how

[1] ***Note***. Reference list can be found at the end of the document.

does it affect the interpretation of our results?) and (c) think of a study to specifically test whether this phenomenon (different scales lead to different distributions) is a methodological artifact or a "true" phenomenon (max 250 words).

(a) Haselbeck et al. assessed the distributional modality and skewness in emotional measurements using seven different emotion ESM datasets and investigated whether there are any patterns of the distributional forms on the level of items, individuals, and measurement designs. They found that skewed unimodal distributions as well as multimodality are highly prevalent across the datasets and the presence of multimodality was found to be strongly associated with the measurement designs such that multimodality was more prevalent in VAS (visual analog scale) than in the Likert scales (e.g., $1-5$, or $1-7$). In addition, they found some other associations on the item level such as negative emotions exhibit higher skewness than position emotions.

(b) This has a great implication on multiple aspects. First, with regard to theorizing emotion dynamics, if the found multimodality and high skewness is indeed true dynamics of emotion, then it implies that a person experiences emotions by going through multiple states (intensity) instead of experiencing it in one typical intensity all the time (i.e., unimodal). Additionally, if we presume the multimodality, then the measurement design should preferably include a broader scale (e.g., VAS with $0-100$ scale) rather than a limited scale such as Likert scale of $1-5$, so that it can capture the variabilities across multiple peaks. Lastly, the VAR model might not be deemed appropriate if the true underlying dynamic follows a skewed/multimodal distribution as the VAR model assumes a single equilibrium state that is drawn from a normal distribution. Accordingly, a VAR model will fit poorly on the skewed or multimodally-distributed time-series emotion data, and correspondingly interpreting its parameters would be misleading.

(c) However, it is difficult to reason whether the found result by Haselbeck et al. actually reflects the true emotion dynamics or it might be just some sorts of methodological artifacts (e.g., specific measurement design induce multimodality). In order to test this, we could design a study such that participants are asked to report on each item with different measurement scales. For example, a participant reports each item twice, once with VAS (1 - 100 scale) and once with Likert scale (e.g., 1 - 7) at every occasion and in the end we can compare the distribution of items to see if they approximate to each other or not. If the distributions of item based on VAS and Likert-scale are turned out to be similar (i.e., multimodal), then we can be more confident in presuming that the found multimodality and skewness do reflect the true phenomenon.

Question 6 (1 point)

What do the *day* and *beep* arguments in the packages `graphicalVAR`, `psychonetrics`, and `mlVAR` do? Suppose you have a dataset containing only one observation per day for every weekday (but not weekends) in a year long study. Which argument would you use (and why) to make sure that measures from Mondays are not regressed on measures from Fridays?

- `dayvar` argument indicates assessment day and when you add this, it ensures that the first measurement of a day is not regressed on the last measurement of the previous day (it removes pairs of observations that cross a night).
- `beepvar` argument indicates the assessment beep per day, and when you add this, it will treat the non-consecutive beeps as missing (removes pairs of observations that are not consecutive).
- In this case, we want to regress measures of each weekdays on measures of the previous days but there are breaks during the weekends. We can then use `dayvar` argument and specify the same number to each week (e.g., Monday - Friday of week 1: `dayvar = 1` and Monday - Friday of week 2: `dayvar = 2`, …) then the measures from Mondays are not going to be regressed on the measures from Fridays of the previous week.

Practical Questions

```
# load data
load("clean_network.RData")

# Variables to investigate:
vars <- paste0("Q",1:18)

# Labels:
varLabs <- c("Relax","Irritable","Worry","Nervous","Future","Anhedonia",
"Tired","Hungry","Alone","Angry","Social_offline","Social_online",
"Music","Procrastinate","Outdoors","C19_occupied","C19_worry",
"Home")

# Rename columns in data:
names(Data2)[names(Data2) %in% vars] <- varLabs

# Remove items:
Data2 <- Data2 %>% select(-Hungry,-Angry,-Music,-Procrastinate)
varLabs <- varLabs[!varLabs %in% c("Hungry","Angry","Music","Procrastinate")]
```

N =1 time series

```
## select a random participant
student_number <- 12183881
set.seed(student_number)
subject <- sample(Data2$id, 1)
my_data_copy <- my_data <- Data2[which(Data2$id == subject),]
```

I chose 4 variables: *"Irritable", "Worry", "Nervous", "Alone".*

```
chosen_var <- c("Irritable", "Worry", "Nervous", "Alone")
```

Question 7 (1.5 points)

Estimate a *saturated* (no model selection) GVAR model on your sampled subject using the 3 or 4 variables you chose above using `psychonetrics`. Plot the estimated temporal (partial directed correlations) and contemporaneous networks with the same layout (1 point). Which edges are significant (0.5 point)?

```
## estimate GVAR model
res1 <- gvar(my_data, vars = chosen_var, dayvar = "day", beepvar = "beep", estimator="FIML") %>% runmodel
## temporal network
temporal1 <- getmatrix(res1, "PDC") # PDC = Partial Directed Correlations
## contemporaneous network
contemporaneous1 <- getmatrix(res1, "omega_zeta")
```

```
## get average layout
L1 <- averageLayout(temporal1, contemporaneous1)
## get labels
labs <- chosen_var
## plot networks
layout(t(1:2))
qgraph(temporal1, layout = L1, theme = "colorblind", directed=TRUE, diag=TRUE,
       title = "Temporal", vsize = 12, mar = rep(6,4), asize = 5, labels = labs)
qgraph(contemporaneous1, layout = L1, theme = "colorblind",
       title = "Contemporaneous", vsize = 12, mar = rep(6,4), asize = 5, labels = labs)
```

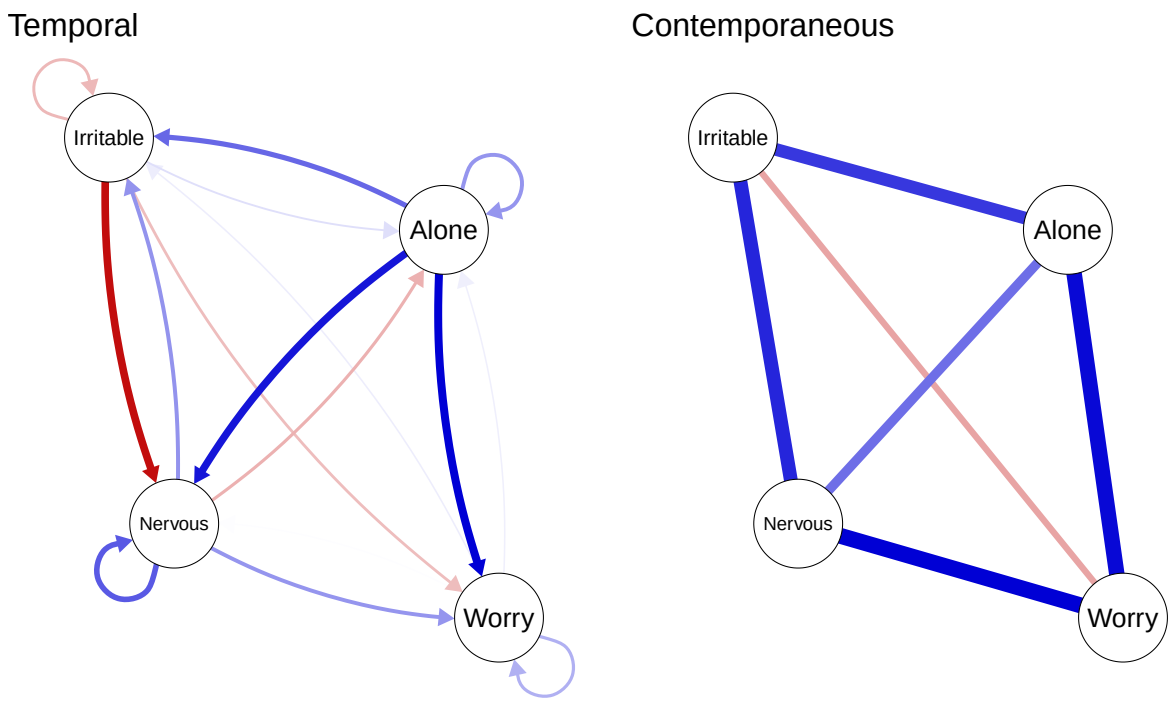

Figure 1: Estimated graphical vector auto-regression (GVAR) model

Assuming that $\alpha = 0.05$,

- there is no significant edge found in the temporal network as all $p > 0.05$ in the `beta` matrix, as shown below.
- the significant edges in the contemporaneous network are: `Nervous -- Irritable` and `Alone--Worry` as $p < 0.05$ in the `omega-zeta` matrix.

```
Parameters for group fullsample
  -  mu
          var1 op var2  est    se        p row col par
Irritable_lag1 ~1      1.62  0.13 < 0.0001   1   1   1
    Worry_lag1 ~1      2.02  0.10 < 0.0001   2   1   2
  Nervous_lag1 ~1      1.95  0.13 < 0.0001   3   1   3
   Alone_lag1 ~1       1.39 0.089 < 0.0001   4   1   4
     Irritable ~1      1.49  0.10 < 0.0001   5   1   5
         Worry ~1      1.86 0.096 < 0.0001   6   1   6
```

```
      Nervous ~1      1.82  0.11 < 0.0001   7   1   7
        Alone ~1      1.29 0.070 < 0.0001   8   1   8
```

- exo_cholesky

```
         var1     op           var2   est    se        p row col par
Irritable_lag1 ~chol~ Irritable_lag1  0.80 0.096 < 0.0001   1   1   9
    Worry_lag1 ~chol~ Irritable_lag1  0.17  0.11     0.11   2   1  10
  Nervous_lag1 ~chol~ Irritable_lag1  0.41  0.12  0.00097   3   1  11
    Alone_lag1 ~chol~ Irritable_lag1  0.22 0.090    0.015   4   1  12
    Worry_lag1 ~chol~     Worry_lag1  0.62 0.075 < 0.0001   2   2  13
  Nervous_lag1 ~chol~     Worry_lag1  0.34  0.11   0.0014   3   2  14
    Alone_lag1 ~chol~     Worry_lag1  0.24 0.080   0.0030   4   2  15
  Nervous_lag1 ~chol~   Nervous_lag1  0.58 0.070 < 0.0001   3   3  16
    Alone_lag1 ~chol~   Nervous_lag1 0.078 0.075     0.30   4   3  17
    Alone_lag1 ~chol~     Alone_lag1  0.44 0.053 < 0.0001   4   4  18
```

- beta

```
     var1 op      var2    est   se     p row col par
Irritable <- Irritable  -0.13 0.25  0.61   1   1  19
    Worry <- Irritable -0.099 0.24  0.68   2   1  20
  Nervous <- Irritable  -0.46 0.28  0.10   3   1  21
    Alone <- Irritable  0.043 0.16  0.79   4   1  22
Irritable <-     Worry  0.044 0.25  0.86   1   2  23
    Worry <-     Worry   0.16 0.25  0.53   2   2  24
  Nervous <-     Worry 0.0054 0.27  0.98   3   2  25
    Alone <-     Worry  0.029 0.17  0.86   4   2  26
Irritable <-   Nervous   0.24 0.26  0.36   1   3  27
    Worry <-   Nervous   0.19 0.25  0.45   2   3  28
  Nervous <-   Nervous   0.35 0.28  0.21   3   3  29
    Alone <-   Nervous  -0.12 0.17  0.46   4   3  30
Irritable <-     Alone   0.43 0.34  0.20   1   4  31
    Worry <-     Alone   0.62 0.35 0.075   2   4  32
  Nervous <-     Alone   0.66 0.38 0.079   3   4  33
    Alone <-     Alone   0.22 0.22  0.32   4   4  34
```

- omega_zeta (symmetric)

```
   var1 op      var2   est   se     p row col par
  Worry -- Irritable -0.13 0.18  0.46   2   1  35
Nervous -- Irritable  0.32 0.15 0.033   3   1  36
  Alone -- Irritable  0.29 0.16 0.061   4   1  37
Nervous --     Worry  0.37 0.20 0.063   3   2  38
  Alone --     Worry  0.36 0.17 0.032   4   2  39
  Alone --   Nervous  0.21 0.16  0.19   4   3  40
```

- delta_zeta (diagonal)

```
     var1  op      var2  est    se        p row col par
Irritable ~/~ Irritable 0.56 0.062 < 0.0001   1   1  41
    Worry ~/~     Worry 0.42 0.067 < 0.0001   2   2  42
  Nervous ~/~   Nervous 0.48 0.063 < 0.0001   3   3  43
    Alone ~/~     Alone 0.37 0.051 < 0.0001   4   4  44
```

Question 8 (1 point)

Estimate a GVAR model on your sampled subject using the 3 or 4 variables you chose above using the graphicalVAR package (use $\gamma = 0$), and compare your results to the results of the previous question.

The resulting GVAR model estimated by `graphicalVAR` (Figure 2) is much more sparse compared to the GVAR model estimated by `psychonetrics::gvar` (Figure 1). That is because regularization is applied in `graphicalVAR` by default: the temporal network model is estimated with LASSO regularization and the contemporaneous network model is estimated with GLASSO using BIC (as $\gamma = 0$). On the contrary, the networks in Figure 1 are estimated without any model selection/regularization.

```
## estimate GVAR model using graphicalVAR
res2 <- graphicalVAR(my_data, vars = chosen_var, gamma = 0, dayvar = "day", beepvar = "beep")
## plot networks
layout(t(1:2))
plot(res2, "PDC", layout = L1, main = "Temporal", vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme="colorblind")
plot(res2, "PCC", layout = L1, main = "Contemporaneous", vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme = "colorblind")
```

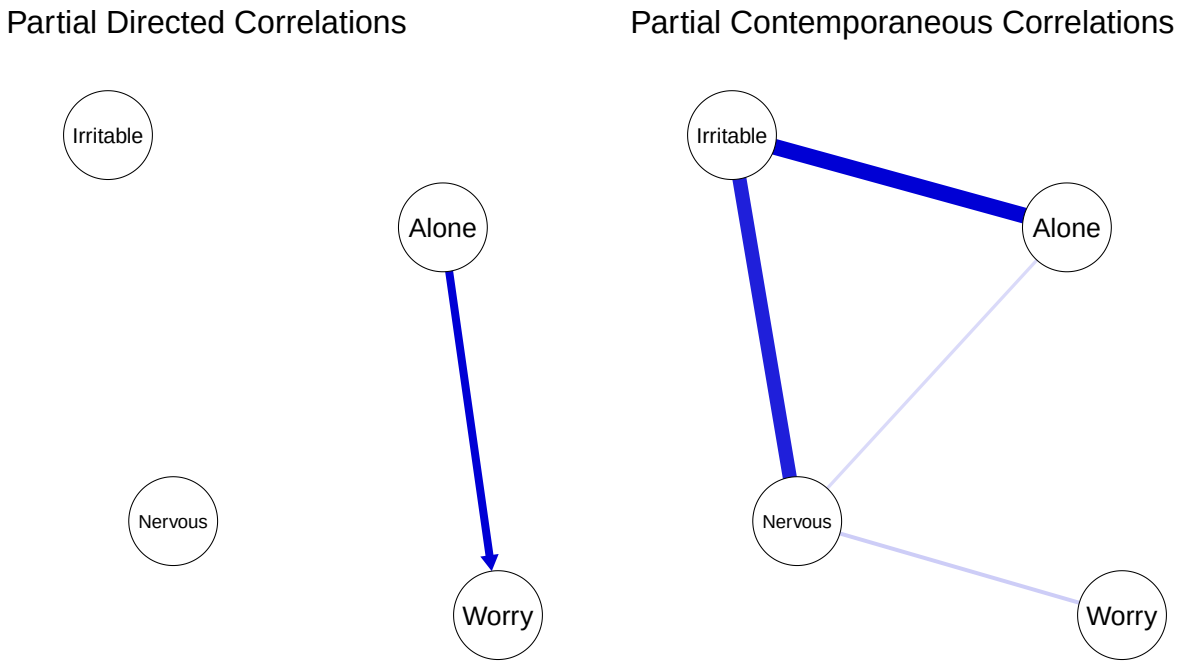

Figure 2: Estimated graphical vector auto-regression (GVAR) model using graphicavlVAR function

Question 9 (1.5 points)

Test for significant trends ($\alpha = 0.05$) for your selected variables and detrend these variables if the trends are significant. Then, re-estimate the networks using either `graphicalVAR` or `psychonetrics`. Did your estimated networks change?

```
## Detrending significant linear trends
for (v in seq_along(chosen_var)){
```

```
  ff <- as.formula(paste0(chosen_var[[v]]," ~ conc"))
  fit <- lm(ff, data = my_data)
  # find significant coefficient of "conc"
  if (anova(fit)$P[1] < 0.05){
    message(paste("Detrending variable:", chosen_var[v], "as the p-value is", round(anova(fit)$P[1],3)))
    # detrend the found variables with significant linear trends
    my_data[[chosen_var[v]]][!is.na(my_data[[chosen_var[[v]]]])] <- residuals(fit)
  }
}
```

```
Detrending variable: Worry as the p-value is 0.049

Detrending variable: Nervous as the p-value is 0.008
```

As shown above, upon having found the significant effects in `Worry` ($p < 0.05$) and `Nervous` ($p < 0.01$)[2], we detrended the corresponding variables. Re-estimated networks after detrending the variable `Worry` and `Nervous` are shown below in Figure 3.
Overall, the networks estimated after detrending barely show any changes. The ones estimated with `psychonetrics` seems more or less the same as the networks estimated on non-detrended data (Figure 1), except for a few changes in the weak edges (faded ones) of the temporal network. The networks estimated with `grahicalVAR()` are also shown to be almost the same as the ones estimated on non-detrended data (Figure 2). The only visible change is that the edge between `Nervous -- Worry` is gone in the contemporaneous network after detrending.

```
## after detrending, estimate GVAR model using psychonetrics:
res3 <- gvar(my_data, vars = chosen_var, dayvar = "day", beepvar = "beep", estimator="FIML") %>% runmodel
# temporal network
temporal3 <- getmatrix(res3, "PDC")
# contemporaneous network
contemporaneous3 <- getmatrix(res3, "omega_zeta")

## after detrending, estimate GVAR model using graphicalVAR:
res4 <- graphicalVAR(my_data, vars = chosen_var, gamma = 0, dayvar = "day", beepvar = "beep")

## plot networks after detrending
layout(t(matrix(1:4, 2, 2)))
# psychonetrics GVAR networks
qgraph(temporal3, layout = L1, theme = "colorblind", directed=TRUE, diag=TRUE,
       vsize = 12, mar = rep(6,4), asize = 5, labels = labs)
title(main = "Detrended Temporal (psychonetrics)", font.main = 1, cex.main=1)
qgraph(contemporaneous3, layout = L1, theme = "colorblind", vsize = 12,
       mar = rep(6,4), asize = 5, labels = labs)
title(main = "Detrended Contemporaneous (psychonetrics)", font.main = 1, cex.main=1)
# graphicalVAR networks
plot(res4, "PDC", layout = L1, main = "Temporal", vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme="colorblind", titles = FALSE)
title(main = "Detrended Temporal (graphicalVAR)", font.main = 1, cex.main=1)
plot(res4, "PCC", layout = L1, main = "Contemporaneous", vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme = "colorblind", titles=FALSE)
title(main = "Detrended Contemporaneous (graphicalVAR)", font.main = 1, cex.main=1)
```

[2] ***Note.*** See the *Appendix* for the full summary results of linear models.

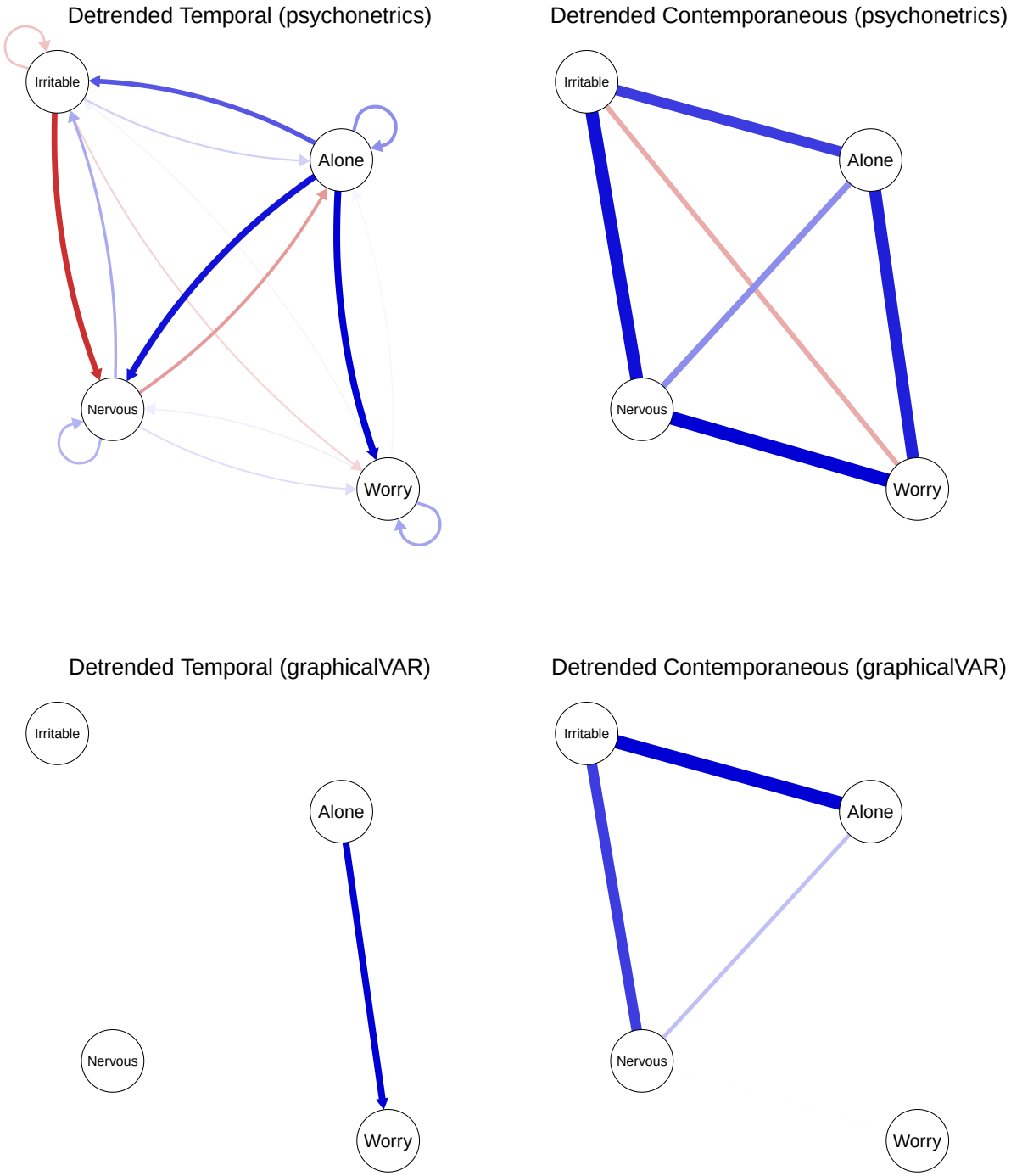

Figure 3: Estimated GVAR model after detrending

Question 10 (1 point)

For question 8 you estimated a graphical VAR network containing 3 or 4 variables for a sampled subject using the `graphicalVAR` package. Repeat this process for another subject. Now that you have estimated two graphical VAR models on non-detrended data, plot the resulting networks and inspect the similarities and differences between these estimated graphical VAR networks. Can you conclude there is heterogeneity between these two people? Why (not)?

Just at a glance, it seems two people are different as the estimated network structures look different (see Figure 4). Two subjects show different connections in the network: subject id of 66 has one edge `Alone -> Worry` in the temporal network, on the other hand subject id of 33 has one additional edge `Alone -> Irritable` in the temporal network. They also show differences in contemporaneous networks (e.g., subject id of 66 has four connections in total, whereas subject id of 33 has two connections with different weights from subject id of 66). Overall, there aren't really much of similarities observed other than the common presence of couple of edges. Note that the `layout` and `maximum` absolute value of edge weights are set to be the same across the estimated network models.

However, despite of these clear differences observed in visual comparison of network structures, we cannot really conclude that there is heterogeneity based on this, because we cannot ensure that these estimated GVAR models are reliable. It could be that they are just random noises, then we would compare a set of random noises to another set of random noises. According to Hoekstra et al. (2022), even when the generating structure is the same, it is likely that we would find different network structures. Hence, it is not advisable to interpret the differences between individual network models as an evidence for heterogeneity, especially when the networks are sparse as in this case.

```
## Another seed: select another random subject
set.seed(13294889)
subject2 <- sample(Data2$id, 1)
my_data2 <- Data2[which(Data2$id == subject2),]

## Estimate GVAR model for another subject
res5 <- graphicalVAR(my_data2, vars = chosen_var, gamma = 0, dayvar = "day", beepvar = "beep")

## plot networks and compare
layout(t(matrix(1:4, 2, 2)))
# extract the max value and specify it the same across the networks
Max <- max(res2$PCC, res2$PDC, res5$PCC, res5$PDC)
# subject from Q8 (id = 68)
plot(res2, "PDC", layout = L1, vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme="colorblind", titles=FALSE, maximum=Max)
title(main = "Temporal (subject id=68)", cex.main = 1, font.main = 1)

plot(res2, "PCC", layout = L1, vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme = "colorblind", titles=FALSE, maximum=Max)
title(main = "Contemporaneous (subject id=68)", cex.main = 1, font.main = 1)

# another subject (id = 33)
plot(res5, "PDC", layout = L1, vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme="colorblind", titles=FALSE, maximum=Max)
title(main = "Temporal (subject id=33)", cex.main = 1, font.main = 1)

plot(res5, "PCC", layout = L1, vsize = 12, mar = rep(6,4), asize = 5,
       labels = labs, theme = "colorblind", titles=FALSE, maximum=Max)
title("Contemporaneous (subject id=33)", cex.main = 1, font.main = 1)
```

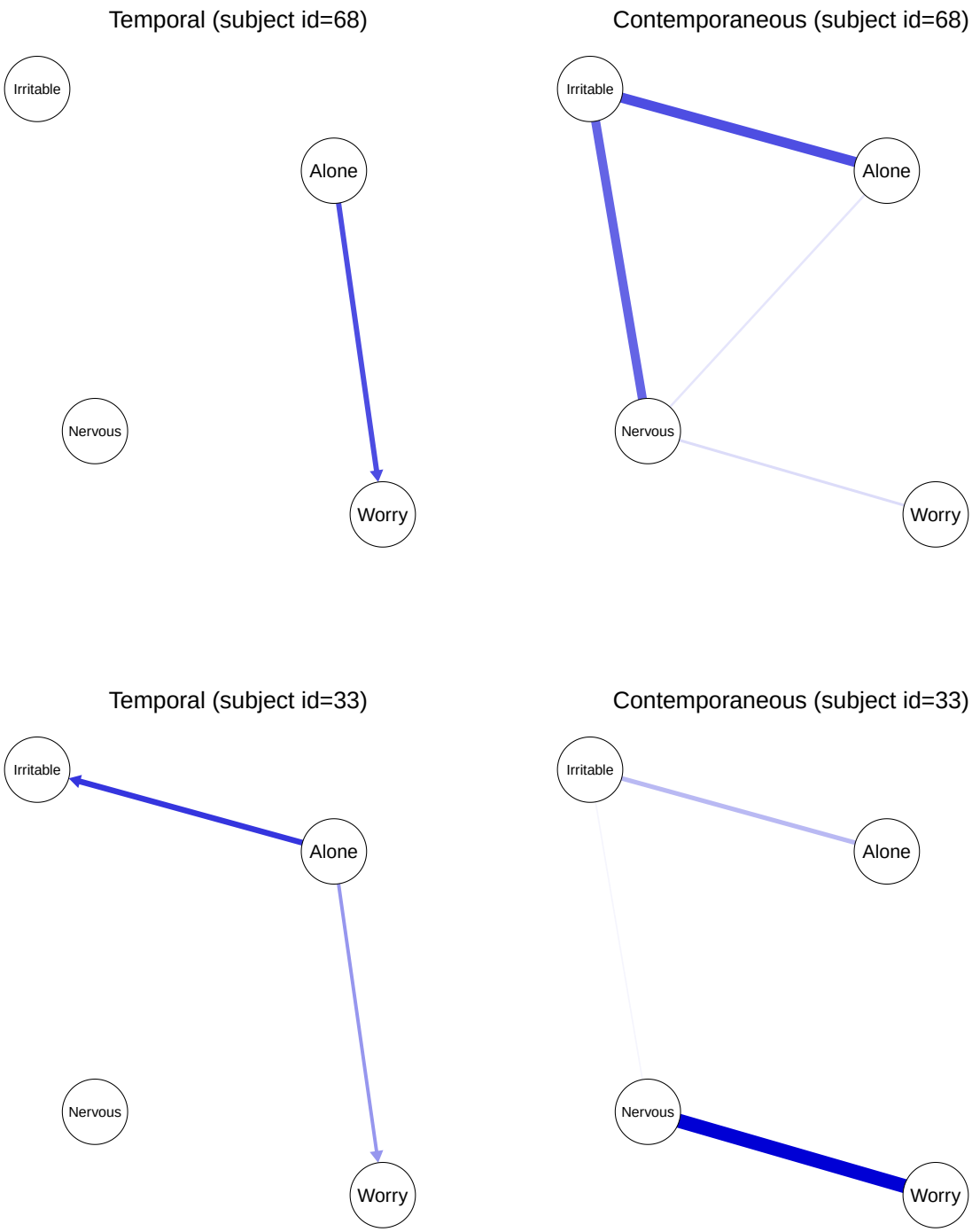

Figure 4: Comparison of GVAR model

N > 1 time-series

Question 11 (1 point)

Look at the help file for the `mlVAR` function and estimate a multilevel GVAR model on the entire dataset on 4 to 6 variables of your choice.

```
## choose 5 variables
chosen_var2 <- c("Irritable", "Worry", "Nervous", "Alone", "C19_worry")
## estimate mlGVAR model
mlVAR_res <- mlVAR(Data2, vars = chosen_var2, idvar = "id" , beepvar = "beep",
                   dayvar = "day", contemporaneous = "correlated", temporal = "correlated")
```

Question 12 (1 point)

Look at the mlVAR plot method help file (`?plot.mlVAR`), and plot the estimated fixed-effect temporal, contemporaneous and between-subjects networks. Plot all networks with the same layout (circle layout or average layout), and hide (threshold) edges that are not significant at $\alpha = 0.05$. For the contemporaneous and between-subjects networks, use an "and"-rule to minimize type-1 error rate in showing edges.

```
# Get networks:
temp <- getNet(mlVAR_res, "temporal", nonsig = "hide")
cont <- getNet(mlVAR_res, "contemporaneous", nonsig = "hide", rule = "and")
bet  <- getNet(mlVAR_res, "between", nonsig = "hide", rule = "and")
# Get average layout
L <- averageLayout(cont,bet,temp)

## Plot networks
par(mfrow=c(1,3))
plot(mlVAR_res, "temporal", title="(a) Temporal (Lag-1)", layout = L, nonsig = "hide",
     alpha = 0.05, theme='colorblind', vsize = 16, asize = 5,
     mar = rep(7,4), title.cex = 1.3)

plot(mlVAR_res, "contemporaneous", title="(b) Contemporaneous", layout = L, nonsig = "hide",
     alpha = 0.05, rule = "and", theme='colorblind', vsize = 16,  asize = 5,
     mar = rep(7,4), title.cex = 1.3)

plot(mlVAR_res, "between", title="(c) Between-subjects", layout = L, nonsig = "hide",
     alpha = 0.05, rule = "and", theme='colorblind', vsize = 16, asize = 5,
     mar = rep(7,4), title.cex = 1.3)
```

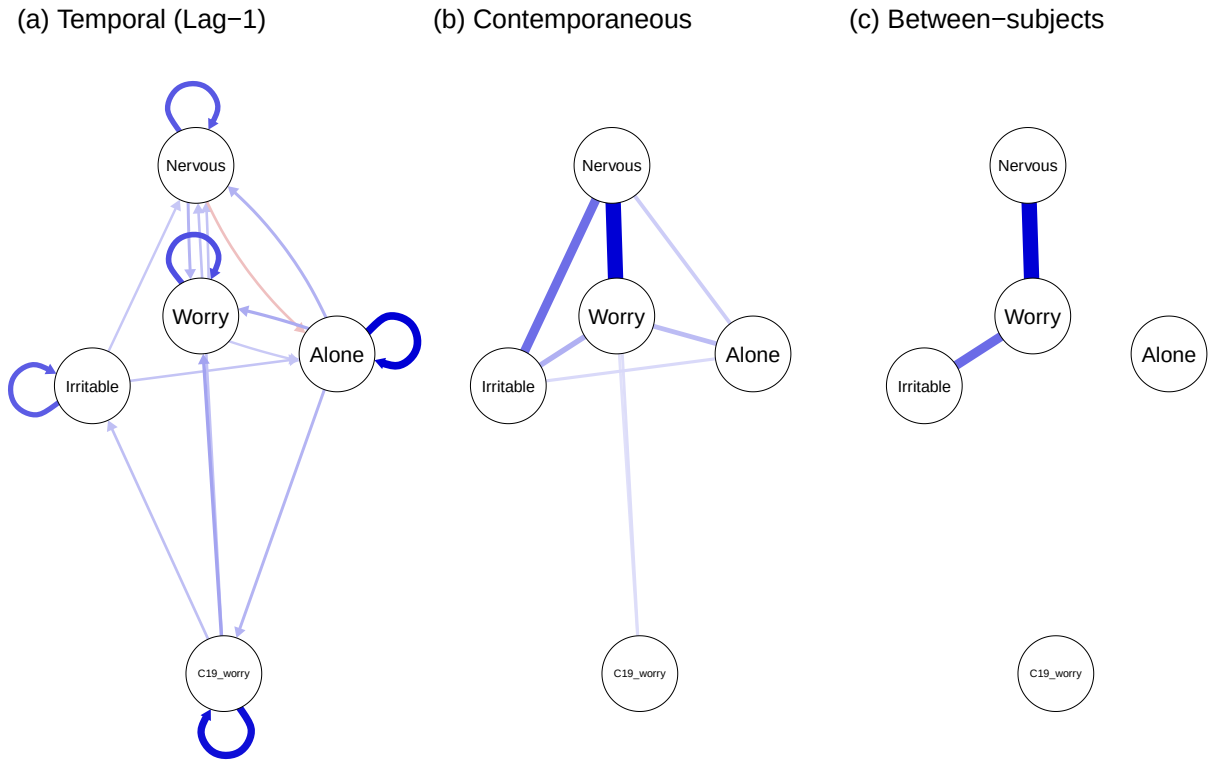

Figure 5: Estimated fixed effects of network structures

Question 13 (2 points)

Plot the estimated individual differences for both the temporal and the contemporaneous network. Explain what the edges in these networks represent. Inspect the plotted networks, what conclusion can you draw about heterogeneity within this sample based on these network structures?

Figure 6 shows the estimated networks of the individual differences, where the edges represent the standard deviation of random effects in the temporal and in the contemporaneous network, respectively. Hence, it indicates how much variance there is in the estimated relations in our sample.

In the temporal network, the `minimum` argument is set to 0.1, meaning that only the standard deviation of temporal random effects above 0.1 are shown with a non-transparent arrow, following the advice of Bringmann et al. (2013). From Figure 6 (a), we can see that the largest individual differences are found in the auto-correlations, given that the self-loops of `Alone` and `C19-worry` are the most pronounced. It implies that there is a high individual variability over those items such that for some individuals, once they feel alone they tend to feel lonely for long time, but for the other individuals, it is rather momentary. The same interpretation goes for `C19_worry`. Among the cross-lagged relations, `Nervous -> Irritable` turns out to have relatively larger individual differences, meaning that the extent to which `Nervous` at time t affects `Irritable` at time $t+1$ differs quite some across the individuals in our sample. When looking at the contemporaneous network in Figure 6 (b), the largest individual differences are again found in the relationship between `Nervous` and `Irritable`, followed by the relationship between `Nervous` and `Worry`.

All in all, we can conclude that there is some amount of individual heterogeneity observed in our data. In general, relatively higher individual differences are shown to present in the auto-correlations (e.g., `Alone`, `C19_worry`) compared to cross-lagged relations. Additionally, the relation between `Nervous` and `Irritable` seems to vary relatively a lot across individuals in this sample as compared to the other relations, considering that it is pronounced in both temporal and contemporaneous networks.

```
layout(t(1:2))
## temporal: random effects SD
plot(mlVAR_res , type = "temporal", SD = TRUE, title = "Individual differences - Temporal",
     layout = L, labels = chosen_var2, theme='colorblind', vsize = 12, asize = 4, mar = rep(7,4),
     minimum = 0.1)

## contemporaneous: random effects SD
plot(mlVAR_res , type = "contemporaneous", SD = TRUE, title ="Individual differences - Contemporaneous",
     layout = L, labels = chosen_var2, theme='colorblind', vsize = 12, asize = 4, mar = rep(7,4))
```

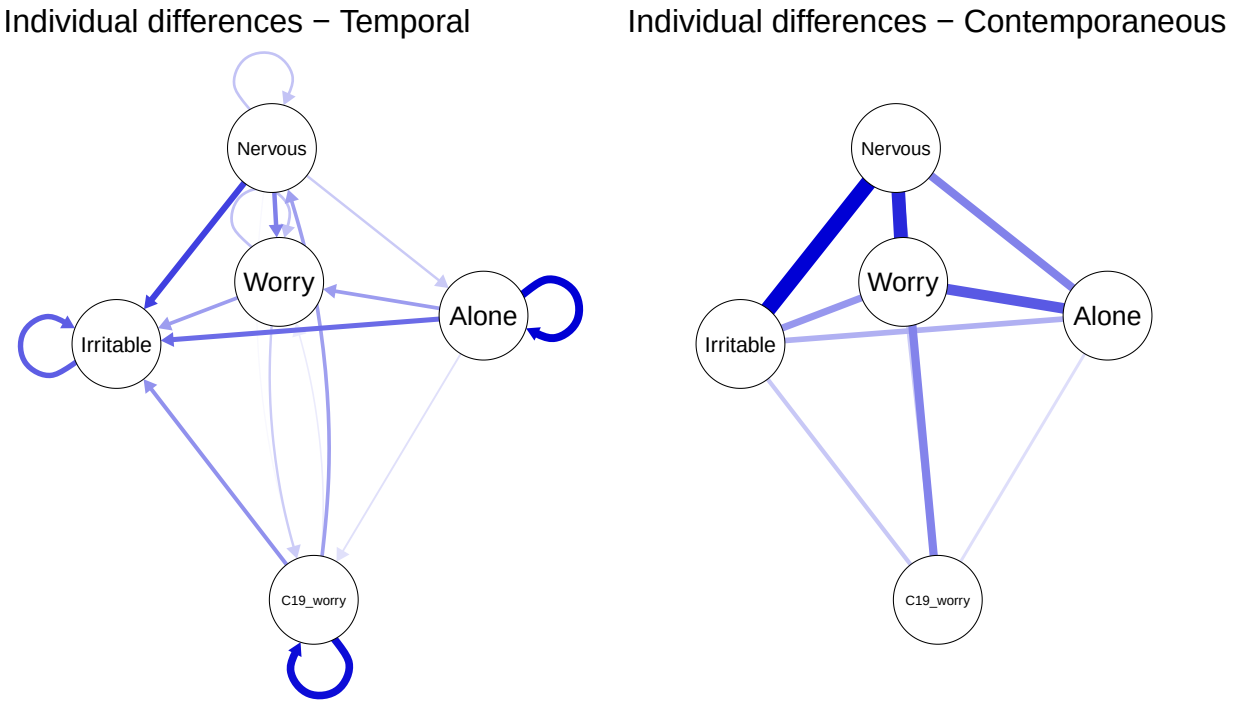

Figure 6: Individual differences networks

Essay Question

Question 14 (2 points)

Inspect these network measures and choose one that would be interesting to compute on one of the estimated networks from any of the previous questions. Report your chosen measure and what it represents for the network you have chosen, explain your choice, and what research question it could answer in the form of a short essay (max 250 words).

After inspecting various measures from Brain Connectivity Toolbox, I decided to look at *in-strength & out-strength* for the fixed effect *temporal network* from Figure 5 (a). The overall strength is the sum of absolute edge weights for a given node (Costantini et al., 2015). For the temporal network, which includes directionality of edges, strength can be sub-divided into *in-strength* and *out-strength*. In-strength is the sum of inward edge weights and out-strength is the sum of outward edge weights. In the temporal network, in-strength represents the *relative susceptibility* (i.e., the extent to which a node is influenced by the previous state of the other nodes) and out-strength represents the *relative predictability* (i.e., the extent to which a node has an influence on the following state of the other nodes) (Black et al., 2022).

The motivation to look at in-/out-strength for the temporal network is that in temporal networks, it is interesting to see whether a node is a strong affector (activate other nodes) or an effector (activated by other nodes). Given that a temporal network contains information not only about the direction of influence but also the strength of influence, in-/out-strength is appropriate measure as it can capture both aspects.

Using in-/out-strength we can answer research questions such as: *what are the nodes that have highest predictive power* or *what are the nodes with highest susceptibility?* And based on the result, we may be able to draw an inference on the dynamics of the system by figuring out, for example, which node takes a main role when it comes to activating the system.

Figure 7 shows the in-strength and out-strength of the temporal network from Figure 5 (a). We can see that `Alone` and `C19_worry` have relatively high predictive power as they have stronger influence on the following states of other nodes, while `Nervous` and `Worry` have high susceptibility as they are affected strongly by the previous state of the other nodes. Based on that, we can assume that the central dynamic in this network consists of `Alone` and `C19Pworry` mainly activating `Nervous` and `Worry` variables. Lastly, `Irritable` is shown to be low on both out-strength and in-strength, which implies that it probably does not play a crucial role in the corresponding temporal network dynamic. Note that the stability of centrality measures is not tested. Hence, any inferences based on in-/out-strength centrality values should be drawn with caution.

```
## compute in-/out-strength
inoutStrength <- temp %>% centrality_auto() %>%
  .$node.centrality %>%
  select(InStrength, OutStrength)
## plot in-/out-strength
p1 <- inoutStrength %>%
  ggplot(aes(y = OutStrength, x=reorder(rownames(.), OutStrength))) +
  geom_line(aes(group=1)) + geom_point() +  coord_flip() +
  labs(x="", y = "", title = "Out-Strength") + ylim(0.1, 0.26)+ theme_bw()
p2 <- inoutStrength %>%
  ggplot(aes(y = InStrength, x=reorder(rownames(.), OutStrength))) +
  geom_line(aes(group=1)) + geom_point() + coord_flip()+
  labs(x="", y = "", title = "In-Strength") + ylim(0.05, 0.26)+ theme_bw()
ggarrange(p1, p2)
```

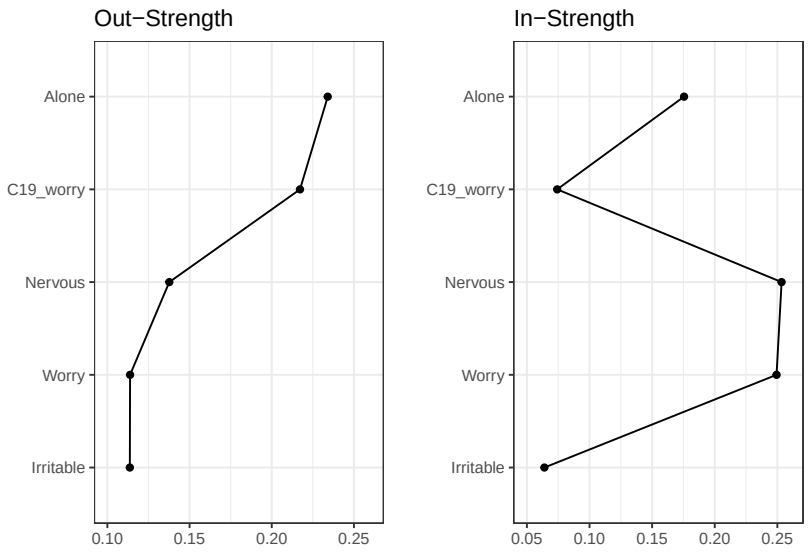

Figure 7: In-/Out-strengh centrality plot

Appendix

Summary results of linear models (Question 9)

```
## run linear regression on chosen variables
lms <- chosen_var %>%
  paste(., '~ conc') %>%
  map(as.formula) %>%
  map(lm, data = my_data_copy)
names(lms) <- chosen_var

## get the summary results
lapply(lms, summary)
```

```
$Irritable

Call:
.f(formula = .x[[i]], data = ..1)

Residuals:
    Min      1Q  Median      3Q     Max
-0.6114 -0.4959 -0.4076  0.4905  2.4611

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.624942   0.223492   7.271 2.53e-09 ***
conc        -0.004527   0.006664  -0.679      0.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7353 on 49 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.00933,   Adjusted R-squared:  -0.01089
F-statistic: 0.4615 on 1 and 49 DF,  p-value: 0.5001


$Worry

Call:
.f(formula = .x[[i]], data = ..1)

Residuals:
     Min       1Q   Median       3Q      Max
-1.19171 -0.57716  0.04182  0.28764  1.28764

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.22859    0.20458  10.894 1.09e-14 ***
conc        -0.01229    0.00610  -2.015   0.0494 *
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6731 on 49 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.07651,   Adjusted R-squared:  0.05767
F-statistic:  4.06 on 1 and 49 DF,  p-value: 0.04942


$Nervous

Call:
.f(formula = .x[[i]], data = ..1)

Residuals:
     Min       1Q   Median       3Q      Max
-1.27329 -0.48327  0.01729  0.33512  2.25339

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.364097   0.218806  10.805 1.45e-14 ***
conc        -0.018161   0.006524  -2.784  0.00762 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7199 on 49 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.1365,    Adjusted R-squared:  0.1189
F-statistic: 7.748 on 1 and 49 DF,  p-value: 0.007617


$Alone

Call:
.f(formula = .x[[i]], data = ..1)

Residuals:
    Min      1Q  Median      3Q     Max
-0.3968 -0.3236 -0.2433  0.6550  1.6675

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.400390   0.153096   9.147 3.56e-12 ***
conc        -0.003570   0.004565  -0.782    0.438
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5037 on 49 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.01233,   Adjusted R-squared:  -0.007827
F-statistic: 0.6117 on 1 and 49 DF,  p-value: 0.4379
```

References

Black, L., Panayiotou, M., & Humphrey, N. (2022). Internalizing symptoms, well-being, and correlates in adolescence: A multiverse exploration via cross-lagged panel network models. *Development and Psychopathology*, *34*(4), 1477–1491. https://doi.org/10.1017/S0954579421000225

Borsboom, D., & Van Borkulo, C. D. (2021). *Bachelor thesis project in psychological methods*. University of Amsterdam.

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. *PLOS ONE*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in r. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Epskamp, S., Borkulo, C. D. van, Veen, D. C. van der, Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Hoekstra, R. H. A. (2022). *Lecture in network analysis*. https://canvas.uva.nl/courses/31498/modules

Hoekstra, R. H. A., Epskamp, S., & Borsboom, D. (2022). Heterogeneity in Individual Network Analysis: Reality or Illusion? *Multivariate Behavioral Research*, 1–25. https://doi.org/10.1080/00273171.2022.2128020

Isvoranu, A.-M., Epskamp, S., Waldorp, L., & Borsboom, D. (2022). *Network psychometrics with r: A guide for behavioral and social scientists*. Routledge. https://doi.org/10.4324/9781003111238

Lunansky, G., Borkulo, C. van, & Borsboom, D. (2020). Personality, Resilience, and Psychopathology: A Model for the Interaction between Slow and Fast Network Processes in the Context of Mental Health. *European Journal of Personality*, *34*(6), 969–987. https://doi.org/10.1002/per.2263

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1