

横断データに対する心理ネットワーク分析のチュートリアル

国里愛彦¹・藁科佳奈²・隅田昌孝²・村口ゆら²・大水拓海²

Tutorial of Psychological Network Analysis in Cross-Sectional Data

Yoshihiko Kunisato¹, Kana Warashina², Masataka Sumida², Yura Muraguchi² and Takumi Omizu²

Abstract : 心理学において、心理ネットワークモデルの活用が進んできている。心理ネットワークモデルはさまざまな研究テーマに適用でき、実施のための統計パッケージも充実してきている。今後も、心理ネットワークモデルの活用は広がっていくと予想される。しかし、適切な心理ネットワーク分析の適用については、それがどのような統計モデルに基づくのかを理解して利用することが求められる。本論文では、心理ネットワークモデルの中でもガウシアン・グラフィカルモデルに焦点を当てて、その基礎的な内容の解説とともに統計パッケージの使用方法についてのチュートリアルを提供する。

Keywords : 横断データ, 心理ネットワーク分析, ペアワイズ・マルコフ確率場, ガウシアン・グラフィカルモデル, エッジの推定方法と正確度, 中心性指標と安定性

心理ネットワークモデルは、心理変数などの観察可能な変数を表すノードと統計的関係を表すエッジから構成されるモデルである (Epskamp, Borsboom, and Fried, 2018)。心理ネットワークモデルでは、心理変数間の関係性をネットワークとして表現できる。これは、因子分析やSEM (structural equation modeling) などのように潜在変数を仮定して、それから各項目・症状を説明するモデルとは異なる。例えば、因子分析を用いた抑うつを測定する尺度の開発では、暗黙的に抑うつといった潜在変数を仮定し、その潜在変数の強さによって個々の項目への反応が変わると想定する。一方で、心理ネットワークモデルは、潜在変数を仮定せずに、抑うつに関する個々の症状のネットワーク構造を検討する。例えば、抑うつでは、不眠から頭痛が生じ、それによって集中力が低下し、そのことから自責的になるなどの、症状が連鎖的に影響していくという状況がある。その場合、潜在変数からそれぞれの症状が生じていると考えるよりは、症状間のネットワークを仮定する方が自然であるといえる。

心理ネットワークモデルについての背景、研究上の意義、臨床応用に関しては、すでに椋原 (2019) や椋原・

伊藤 (2022) による優れた解説がある。本論文では、連続変数の横断データにおける心理ネットワーク分析に焦点を当て、そこで頻繁に用いられるガウシアン・グラフィカルモデル (Gaussian graphical model: GGM) についての解説と具体的な解析方法について主に論じる。

心理ネットワークモデルの基礎

心理ネットワークモデルは、心理変数を表すノード (ネットワーク図では円として表現される) と変数間の関係を表すエッジ (ネットワーク図では線として表現される) からなる。ネットワークモデルとあるので、既存のネットワーク科学と関連づけることができるが、心理ネットワークモデルには特有の特徴がある (Epskamp et al., 2018)。一般的なネットワーク科学で扱うネットワークの場合、ネットワークのエッジは観察可能である。例えば、電力網のエッジには電線などが対応し、直接的に観察や測定が可能である。しかし、心理ネットワークのエッジは、データから推定されたものであり、直接的に観察や測定ができない。また、電力網のエッジは存在するかしないかの2値をとるが、心理ネットワークは重み付きネットワークであり、エッジには強弱が存在する。

心理ネットワークモデルでは、エッジは推定されるものである。エッジの推定値は、サンプルサイズの影響を受けるので、ネットワークのエッジの精度も検討する必要がある (Epskamp et al., 2018)。また、一般的なネットワーク科学では、small worldness, density, global clustering などのグローバルなネットワーク特性の指標が使われる。しかし、心理ネットワークのような重み付

受稿日2022年11月24日 受理日2022年12月8日

1 専修大学人間科学部心理学科 (Department of Psychology, Senshu University)

2 専修大学人間科学部 (School of Human Science, Senshu University)

3 本論文は、日本心理学会第86回大会のチュートリアルワークショップ『心理ネットワークアプローチ入門』内の「横断的ネットワーク解析:基礎編」に加筆修正したものである。

4 本研究はJSPS 科研費 (20K20870, 21H05068) の助成を受けたものである。

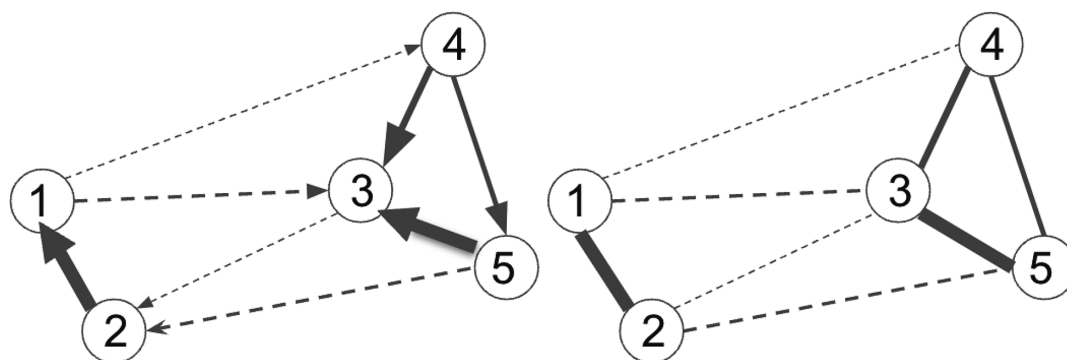


Figure 1：有向ネットワークと無向ネットワーク

きネットワークではこれらは使えない。そのため、1つのノードがどのような影響力をもつのかといったローカルなネットワーク特性を用いる。

ネットワークはエッジの方向性の有無から有向ネットワークと無向ネットワークに分けることができる。有向ネットワークは、エッジに方向性があるネットワークである (Figure 1 左)。有向ネットワークをプロットした際には、特定のノードから特定のノードへのエッジは矢印で表現される。そして、矢印の方向には因果関係が想定される。例えば、「活動性の低下 → 抑うつ気分の増加」なら、活動性の低下が原因となって抑うつ気分が増加するという意味になる。ただし、逆向きの「抑うつ気分の増加 → 活動性の低下」という関係性もありえる。因果関係を論じるには、まず原因は結果に時間的に先行する必要がある、因果に関連する変数がすべて測定された上で、原因と結果が巡回しない非巡回である必要がある (Epskamp et al., 2018)。観察研究によって収集された心理学のデータの場合、時間的な前後関係は縦断データによって示せても、非巡回となるとかなり難しい。

一方、無向ネットワークは、エッジに方向性がないネットワークである (Figure 1 右)。横断データに対する心理ネットワークでは、エッジに方向性がない無向ネットワークが用いられる。無向ネットワークのエッジでは、因果の方向性を示すことはできない。無向ネットワークで、エッジがあるということは、方向性は不明だが因果的な関係がある可能性を示唆しているだけである (Epskamp et al., 2018)。

心理ネットワークモデルでのエッジの色は符号を意味している。エッジの色が青もしくは緑の場合は正、赤の場合は負を表す。なお、論文がカラー印刷ではない場合には、Figure 1のように、エッジが実線の場合は正、点線の場合は負を表す。エッジが正の場合は、2つのノード間には正の関係性が存在する。つまり、片方が増える

と、もう一方も増える。一方、エッジが負の場合は、2つのノード間には負の関係性が存在する。つまり、片方が増えると、もう一方は減る。また、エッジの太さは、関係の強さを表す。この場合の関係の強さとは、全体のノードからの影響を考慮した時の2つのノード間の関係の強さになる。そして、エッジの長さは、ノード間の関係の強さの逆数になる。つまり、ノード間の関係が強ければ強いほど、エッジは短くなる。

心理ネットワーク分析の実施や解釈においては、ネットワークのエッジが推定されたものであることに留意する必要がある。その上で、ネットワークが有向なのか無向なのか検討し、エッジの強さや符号を解釈する。心理ネットワークモデルは、要点を押さえれば、直感的に理解ができるものであり、優れた可視化手法の1つであるといえる。しかし、実際に研究に用いる上では、心理ネットワーク分析で用いられる統計モデル、実施の手順と検討すべきポイントなどを理解する必要がある。以降では、これらについて解説する。

ペアワイズ・マルコフ確率場

変数間の関係がある・ないというグラフを作る場合に、確率分布と結びつけるようなモデルのことをペアワイズ・マルコフグラフ (pairwise Markov graph) と呼び、心理ネットワークで用いられる無向ネットワークモデルの推定では、ペアワイズ・マルコフ確率場 (pairwise Markov random field: PMRF) がよく使われる (Epskamp, Haslbeck, Isvoranu, and van Borkulo, 2022)。心理ネットワーク分析で用いられるペアワイズ・マルコフ確率場は、基本的に重み付きネットワークである。ペアワイズ・マルコフ確率場では、ネットワークのエッジは、ほかのすべてのノードの影響を除外したあとでの2つのノード間の条件付き関連になる。つまり、2つのノード間にエッジがあるということは、条件付き依存

(conditionally dependent) 関係にあることを表し、一方でエッジがない場合は、条件付き独立 (conditionally independent) 関係にあることを表す。

ペアワイズ・マルコフ確率場にはさまざまなモデルがあるが、横断データの心理ネットワーク分析においては、(1) ガウシアン・グラフィカルモデル (Gaussian graphical model: GGM), (2) イジングモデル (Ising model), (3) 混合グラフィカルモデル (mixed graphical models: MGMs) の3つを用いることが多い。ガウシアン・グラフィカルモデルは、多変量正規分布を用いて、連続変数のデータのネットワークを作るのに使える (Epskamp, Maris, Waldorp, and Borsboom, 2018)。イジングモデルは、磁石などの磁性体に関する統計力学的モデルになる (Epskamp et al., 2018)。データが2値の場合のネットワークの記述に使えるのと、相転移などの現象も扱える。混合グラフィカルモデルは、連続、カウント、カテゴリカルデータのような異なるデータを用いてネットワークを作ることができる (Epskamp et al., 2022)。3つのモデルとも、エッジは条件付き関連の強さを表現しており、0では条件付き独立、0以外だと条件付き依存関係になる。ガウシアン・グラフィカルモデルのエッジは-1から1の範囲、イジングモデルのエッジは $-\infty$ から ∞ の範囲、混合グラフィカルモデルのエッジは-1から1の範囲で値をとる。心理学研究においては、連続変数のデータを扱うことが多いので、以下ではガウシアン・グラフィカルモデルについて解説する。

ガウシアン・グラフィカルモデル

ガウシアン・グラフィカルモデルでは、変数間の関係を偏相関係数によってモデル化する。なお、以下で説明する数式と表記は、基本的に Epskamp et al. (2018) の記載に従う。 \mathbf{Y} は、ランダムに選んだ参加者から収集される n 個の反応のベクトルである (確率変数)。1つの反応ではなく、複数の反応からなるベクトルである点に注意する必要がある。 \mathbf{Y} は、平均 $\boldsymbol{\mu}$ と分散共分散行列 $\boldsymbol{\Sigma}$ の多変量正規分布に従う。

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (1)$$

多変量正規分布のため、尤度関数は以下のようになる。なお、 \mathbf{y} は確率変数 \mathbf{Y} の実現値である。

$$f(\mathbf{Y}=\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right) \quad (2)$$

データの中心化 (値から平均値を引く操作) を行うと、平均は0になる。そのため、中心化したデータ \mathbf{y}_c

は、以下のように、平均0、分散共分散行列 $\boldsymbol{\Sigma}$ の多変量正規分布に従う。

$$\mathbf{y}_c \sim N(0, \boldsymbol{\Sigma}) \quad (3)$$

ここで、 \mathbf{y}_c の c は case, つまり人であり、 R などのデータフレームにおける行に対応している。例えば、1列目が抑うつ尺度の項目1、2列目が抑うつ尺度の項目2の場合、 $y [c=1, 1]$ だと1人目の抑うつ尺度の項目1の得点、 $y [c=2, 1]$ だと2人目の抑うつ尺度の項目1の得点、 $y [c=1, 2]$ だと1人目の抑うつ尺度の項目2の得点になる。そして、分散共分散行列 $\boldsymbol{\Sigma}$ は、ベクトル \mathbf{y} に含まれる変数間の分散共分散行列になる。

ガウシアン・グラフィカルモデルでは、実現値 \mathbf{y} は多変量正規分布に従う。平均 $\boldsymbol{\mu}$ と分散共分散行列 $\boldsymbol{\Sigma}$ が定めればデータが生成される。では、どこが心理ネットワークのエッジになるかということ、分散共分散行列 $\boldsymbol{\Sigma}$ から偏相関係行列を計算したものが心理ネットワークのエッジとなる。

分散共分散行列から偏相関係行列へ

ガウシアン・グラフィカルモデルにおいて偏相関係行列を算出する際には、分散共分散行列 $\boldsymbol{\Sigma}$ の逆行列である \mathbf{K} を用いる。具体的な手続きとしては、分散共分散行列からその逆行列である精度行列を計算し、精度行列から偏相関係行列を計算する。以下では、順を追って計算手続きを説明する。まず精度行列は \mathbf{K} と記載する。この精度行列 \mathbf{K} は、分散共分散行列の逆行列になる。

$$\mathbf{K} = \boldsymbol{\Sigma}^{-1} \quad (4)$$

逆行列は、 n 次正方行列 \mathbf{A} に掛けた場合に以下のような単位行列 \mathbf{E} となるような行列である。つまり、分散共分散行列に何か行列を掛け合わせた時に、以下のような単位行列 \mathbf{E} になるようなものがここで知りたい逆行列になる。なお、逆行列が存在する行列は正則行列と呼ばれる。

$$\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (5)$$

以降は R で計算させつつ分散共分散行列から精度行列を経由して、偏相関係行列を計算する。計算にあたり、Epskamp et al. (2018) に記載されている例を参考にし、数値などは一部改変して説明する。これから、以下のような分散共分散行列 \mathbf{A} を偏相関係行列にする。まず、 R で以下のように分散共分散行列 \mathbf{A} を用意する。

```
A <- matrix(c(1, -0.25, 0.34, -0.25, 1, -0.07, 0.34, -0.07,
1), 3, 3)
A
```

```
##      [,1] [,2] [,3]
## [1,] 1.00 -0.25 0.34
## [2,] -0.25 1.00 -0.07
## [3,] 0.34 -0.07 1.00
```

共分散をみると、変数1と変数2は負の関連があり、変数1と変数3は正の関連があり、変数2と変数3は関連が弱い。

この分散共分散行列 \mathbf{A} から逆行列 \mathbf{A}^{-1} を計算する。まず、行列 \mathbf{A} の要素がわかりやすいように、以下のようにする。

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (6)$$

そして、 3×3 の行列の逆行列は以下のように計算ができる。それぞれの要素を計算すると、逆行列になる。

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \quad (7)$$

なお、 \mathbf{A} の行列式 ($\det \mathbf{A}$) は、以下のように計算される。

$$\det \mathbf{A} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \quad (8)$$

手計算でも良いが、上記の計算方法を用いて、R で分散共分散行列 \mathbf{A} の逆行列を計算してみる。 \mathbf{A} の各要素を a_{11} から a_{33} に入れた上で、 \mathbf{A} の行列式 (A_{det}) を計算し、あとは、 \mathbf{A} の逆行列の要素を計算して、 a_{11_i} から a_{33_i} にいれていく。最後に行列として整理する。

```
a_11 <- 1.00
a_12 <- -0.25
a_13 <- 0.34
a_21 <- -0.25
a_22 <- 1.00
a_23 <- -0.07
a_31 <- 0.34
a_32 <- -0.07
a_33 <- 1.00
```

```
A_det <- a_11 * a_22 * a_33 + a_12 * a_23 * a_31
+ a_13 * a_21 * a_32 - a_13 * a_22 * a_31 - a_11 *
a_23 * a_32 - a_12 * a_21 * a_33
a_11_i <- 1/A_det * (a_22 * a_33 - a_23 * a_32)
a_12_i <- 1/A_det * (a_13 * a_32 - a_12 * a_33)
a_13_i <- 1/A_det * (a_12 * a_23 - a_13 * a_22)
a_21_i <- 1/A_det * (a_23 * a_31 - a_21 * a_33)
a_22_i <- 1/A_det * (a_11 * a_33 - a_13 * a_31)
a_23_i <- 1/A_det * (a_13 * a_21 - a_11 * a_23)
a_31_i <- 1/A_det * (a_21 * a_32 - a_22 * a_31)
a_32_i <- 1/A_det * (a_12 * a_31 - a_11 * a_32)
a_33_i <- 1/A_det * (a_11 * a_22 - a_12 * a_21)
matrix (c (a_11_i, a_12_i, a_13_i, a_21_i, a_22_i, a_23_i,
a_31_i, a_32_i, a_33_i) , 3, 3)
```

```
##      [,1]      [,2]      [,3]
## [1,] 1.2005067 0.27289178 -0.38906985
## [2,] 0.2728918 1.06695621 -0.01809627
## [3,] -0.3890699 -0.01809627 1.13101701
```

精度行列をみると、一部は符号が逆転している。なお、R には逆行列を計算する `solve` 関数があるので、以下のようにすると、簡単に計算できる。

```
inv_A <- solve(A)
inv_A
```

```
##      [,1]      [,2]      [,3]
## [1,] 1.2005067 0.27289178 -0.38906985
## [2,] 0.2728918 1.06695621 -0.01809627
## [3,] -0.3890699 -0.01809627 1.13101701
```

逆行列の計算では、特定の変数間の関連についてほかの変数の影響も考慮することになる。そのため、この精度行列 (分散共分散行列の逆行列) を使って、偏相関行列を求めることができる。以下の式 (9) で、分散共分散行列の逆行列 (精度行列) の要素を用いて、偏相関係数を求めることができる。

$$\text{Cor}(Y_i, Y_j | \mathbf{Y}_{-(i,j)}) = \frac{\kappa_{ij}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}} \quad (9)$$

例えば、変数1と変数2の間の偏相関は、以下のよう
に計算できる。

```
-1 * a_12_i / (sqrt(a_11_i) * sqrt(a_22_i))
## [1] -0.2411209
```

このように、1つずつの要素を計算することもできる

が、以下のように、R の stats パッケージの cov2cor 関数を使えば偏相関行列を計算することができる。なお、符号を逆転させているため、分散が -1 となっているのに注意する。

```
library(stats)
- 1 * cov2cor(inv_A)

##           [,1]      [,2]      [,3]
## [1,] -1.0000000 -0.24112091  0.33389562
## [2,] -0.2411209 -1.00000000  0.01647332
## [3,]  0.3338956  0.01647332 -1.00000000
```

あまり大きな変化はないようにも見えるが、変数 2 と 3 の関係が元々の分散共分散行列よりもさらに小さくなっており、符号も逆になっている (-0.070→0.016)。ガウシアン・グラフィカルモデルでは、この偏相関行列の要素をエッジにしてネットワークを描く。このように考えると、ガウシアン・グラフィカルモデルは特殊なモデルというよりは、心理統計学で学ぶ内容で十分に理解することができる。

心理ネットワーク分析の手順

基本的な心理ネットワーク分析は、(1) 統計的手法によってノード間のエッジを推定する、(2) 推定された統計的關係をもとにネットワークを図示する、(3) ネットワーク特性を表す中心性指標を検討する、(4) エッジの正確度と中心性指標の安定性を検討するの 4 つに分けることができる。

以降では、この 4 つに分けた上で、それぞれについて解説するとともに、R での解析について解説する。それに当たり、Open-Source Psychometrics Project (https://openpsychometrics.org/_rawdata/) において、47,974 名に対して Rosenberg (1965) のローゼンバーグ自尊感情尺度 (Rosenberg Self-Esteem Scale: RSE Scale) への回答を求めたデータを用いる。このデータには、年齢、性別、国籍、ローゼンバーグ自尊感情尺度 10 項目への反応が含まれている。なお、ローゼンバーグ自尊感情尺度には、4 件法 (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree) で回答を求めている。

心理ネットワーク分析に使える R パッケージは複数あるが、心理ネットワーク分析の中心地でもあるアムステルダム大学の研究者が開発した一連の R パッケージが有用である。具体的には、ネットワークの推定 (ガウ

シアン・グラフィカルモデル) やプロットをする qgraph (Epskamp et al., 2012)、ネットワーク推定 (ガウシアン・グラフィカルモデルとイジングモデル) やネットワークの正確度と安定性を検討する bootnet (Epskamp et al., 2018)、より洗練化した書き方でネットワークを推定 (ガウシアン・グラフィカルモデルとイジングモデル) する psychometrics (<http://psychometrics.org/>)、2 値データと連続データが混合したネットワークを推定 (混合グラフィカルモデル) する mgm (Haslbeck and Waldorp, 2020) がある。なお、データハンドリングと psychometrics でパイプ演算子を用いるため、tidyverse も使用する。以下では、bootnet パッケージと qgraph パッケージ、psychometrics パッケージ、tidyverse パッケージを使用するので、以下のように library () で読み込む。もしパッケージが入っていない場合は、install.packages () でインストールする。また、Open-Source Psychometrics Project から R にデータをダウンロードする際に、第 1 著者が作成した openPsychData パッケージ (<https://github.com/ykunisato/openPsychData>) を用いる。openPsychData パッケージがインストールされていない場合は、devtools::install_github ("ykunisato/openPsychData") でインストールする。

```
library(bootnet)
library(qgraph)
library(psychometrics)
library(tidyverse)
library(openPsychData)
```

openPsychData パッケージの load_openPsyData 関数を用いて、Open-Source Psychometrics Project からローゼンバーグ自尊感情尺度のデータをダウンロードする。codebook を TRUE にするとデータについての詳細の情報がダウンロードと同時に開く。read_tsv 関数でデータを読み込んで、ローゼンバーグ自尊感情尺度への回答のみにしぼる (Q1 から Q10)。

```
load_openPsyData(dataset_name = "RSE", codebook = TRUE)
data <- read_tsv("RSE/data.csv") %>%
  select (Q1,Q2,Q3,Q4,Q5,Q6,Q7,Q8,Q9,Q10)
```

エッジの推定法・モデル選択とプロット

ガウシアン・グラフィカルモデルの推定は難しいもの

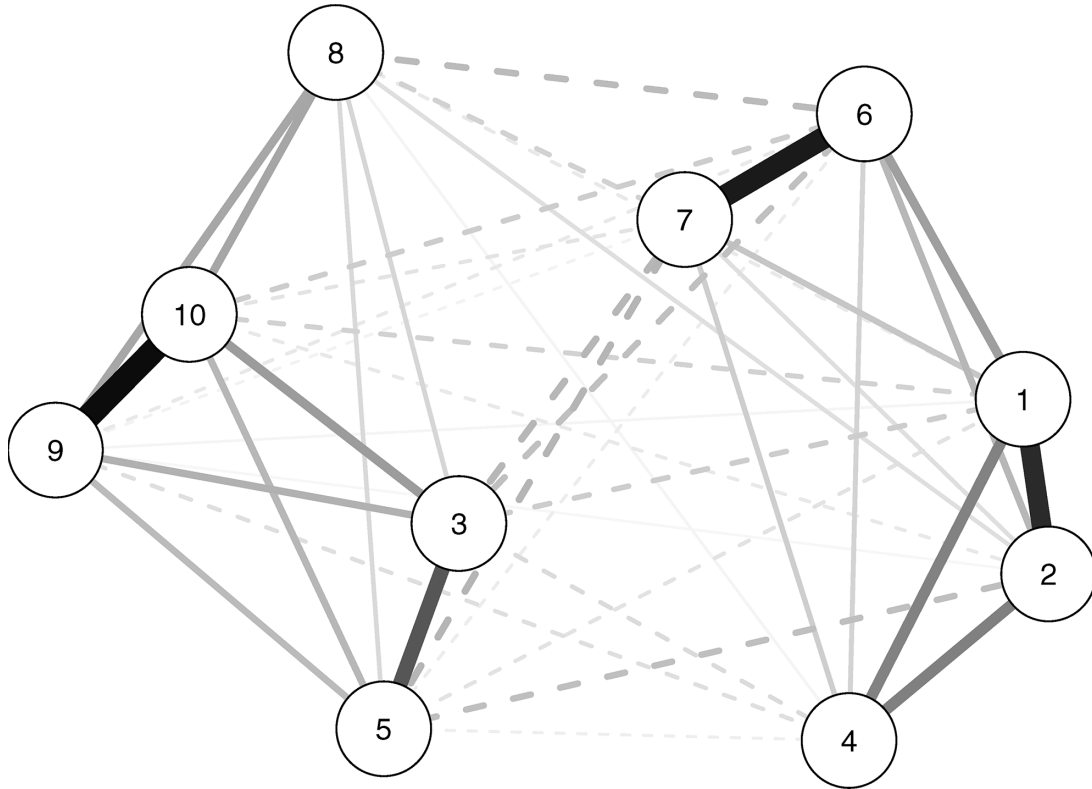


Figure 3 : 刈り込みを用いて推定されたRSEのネットワーク

で、ポリコリック相関を計算する。corMethodで“cor_auto”を指定すれば、自動的に順序変数にはポリコリック相関を計算してエッジを推定する。なお、現状では“cor_auto”を指定すると、7件法以下は順序変数として扱う。また、正規分布していないデータの場合は、スピアマンの順位相関係数を用いて推定を行う(Blanken et al., 2022)。その場合は、corMethod=“spearman”とするかtransform=“rank”とする。

```
result1 <- estimateNetwork(data,default="pcor",
  threshold="sig", alpha=0.05, corMethod="cor_auto")
```

plotでは、theme=“colorblind”, cut=0, layout=“spring”, labels=TRUEなどの指定ができる。themeにはかなり種類があり、雑誌がカラー印刷可能であれば、colorblindなどが良いかもしれない。本論文はカラー印刷ではないので、以下では、grayにしている(Figure 2)。grayの場合、実線は正の関係、点線は負の関係を表し、線の濃さと太さが関係の強さを表す。

```
plot(result1, theme="gray", layout="spring", labels=TRUE)
```

刈り込みは、psychonetricsパッケージのprune関数を用いる。prune関数では、決めた基準以下のエッジを

0に固定して再推定する。なお、psychonetricsはパイプ演算子を使った書き方ができるので、tidyverseパッケージに慣れているRユーザーには便利かもしれない。また、bootnetパッケージの場合は、欠測値があると推定ができないが、psychonetricsパッケージの場合は完全情報最尤推定法が使える(Blanken et al., 2022)。具体的には、estimator=“FIML”と設定する。以下のように、最初に使用するモデルにデータをいれて、パイプ演算子でrunmodelに送って推定し、prune()で刈り込みをして、getmatrixでエッジの情報を得て、プロットするというように、処理過程をわかりやすく書ける(Figure 3)。

```
ggm(data) %>%
  runmodel %>%
  prune(alpha=0.05) %>%
  getmatrix("omega") %>%
  qqgraph(theme="gray", layout="spring", labels=TRUE)
```

推定したパラメータを確認したい場合は、以下のように、getmatrix()の代わりにparametersを使うと確認できる。なお、parametersは関数ではないので、()は不要である。また、prune(alpha=0.05, adjust=“bon-

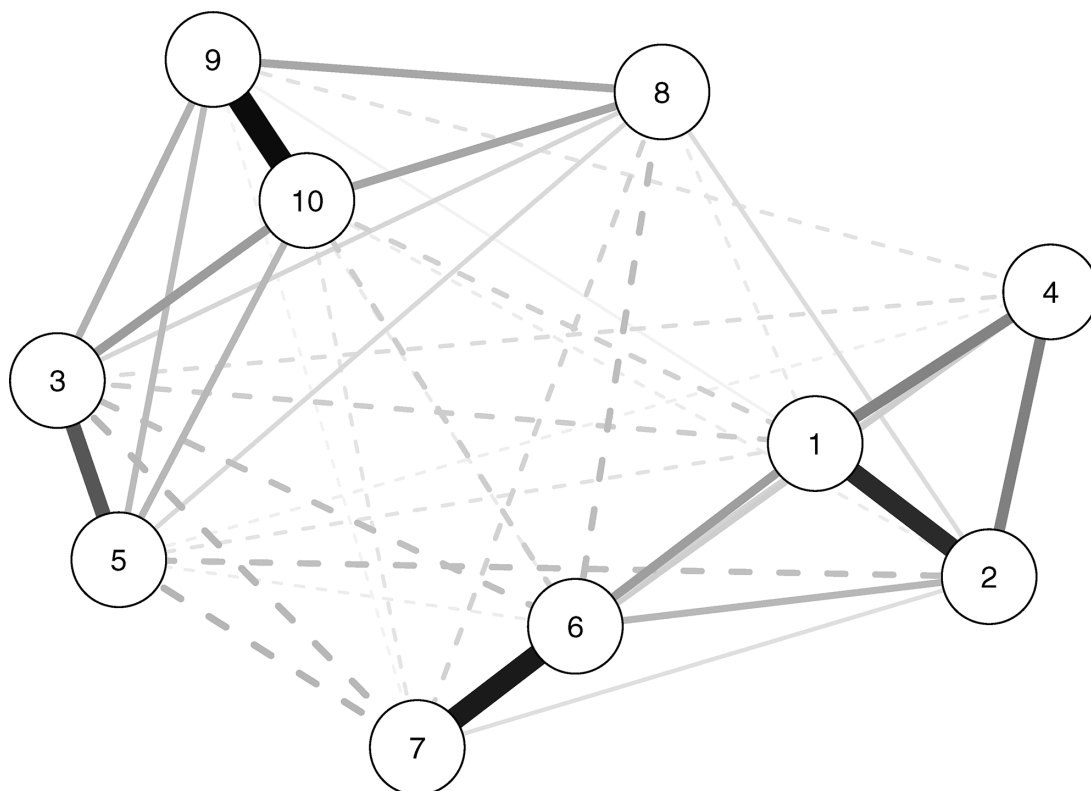


Figure 4：ステップアップ法を用いて推定されたRSEのネットワーク

ferroni”) のようにすると、ボンフェローニ補正もできる。

```
ggm(data) %>%
  runmodel %>%
  prune(alpha = 0.05) %>%
  parameters
```

モデル探索

モデル探索は、可能なネットワークを反復的に探索して最もデータフィットしたモデルを選ぶ方法である (Blanken et al., 2022)。選択の基準として、AIC, BIC, EBIC (extended BIC) などの情報量規準を用いる。ステップアップ (step-up) 法では、反復において、エッジが少ない状態から徐々にエッジを足していった、情報量規準がそれ以上小さくなくなるまで最適なネットワークを探索する (Blanken et al., 2022)。なお、ステップダウン (step-down) 法では、すべてのエッジを含める飽和モデルから徐々にエッジを減らす (Blanken et al., 2022)。ステップアップ法は、psychonetrics パッケージで以下のように実行する (Figure 4)。

```
ggm(data, omega = "empty") %>%
```

```
runmodel %>%
stepup(criterion = "bic") %>%
getmatrix("omega") %>%
qgraph(theme = "gray", layout = "spring", labels =
TRUE)
```

モデル探索における強力なアルゴリズムとして、ggmModSelect アルゴリズムがある (Blanken et al., 2022)。このアルゴリズムでは、正則化手法を使ってモデルを探索してから、ステップワイズなモデル探索法を行う (Blanken et al., 2022)。ggmModSelect アルゴリズムは正確なネットワーク構造を検討する上で有用だが、推定には時間がかかる。psychonetrics パッケージでは、以下のように modelsearch (criterion = "bic") で ggmModSelect アルゴリズムを用いた推定ができる (Figure 5)。

```
ggm(data) %>%
  runmodel %>%
  modelsearch(criterion = "bic") %>%
  getmatrix("omega") %>%
  qgraph(theme = "gray", layout = "spring", labels =
TRUE)
```

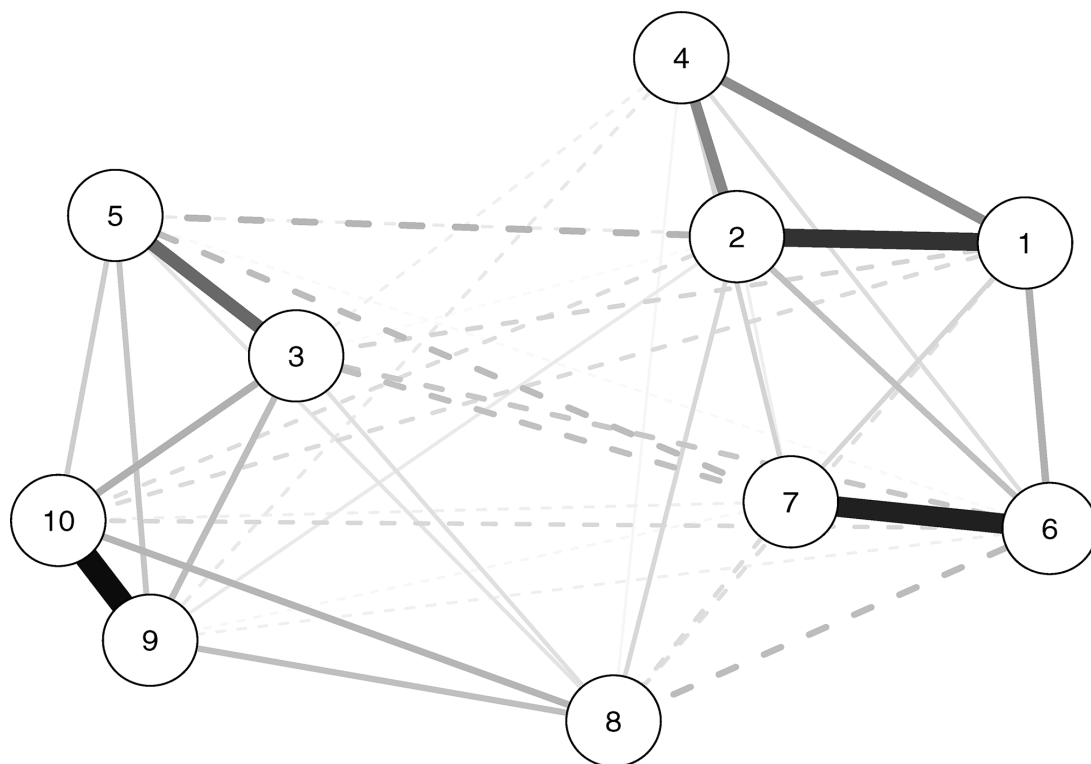



Figure 6 : EBIC を規準とした GLASSO を用いて推定された RSE のネットワーク

らほとんどのモデル選択法でうまくいくが、真のネットワークに似たネットワークの探索には、モデル探索の `ggmModSelect` アルゴリズムが好ましいとされる。次に、サンプルサイズが1000くらいなら、ネットワークの全体構造や強いエッジに興味がある場合は正則化する方法を選び、特定の（ブリッジする）エッジに関心がある場合は `ggmModSelect` アルゴリズムが好ましいとされる。そして、サンプルサイズが300くらいなら、EBICg-lasso などの正則化する方法が一般化可能なネットワーク構造の発見においては好ましいとされる。このように、サンプルサイズと目的に応じて、適切なモデル選択手法は異なる。

心理ネットワーク分析でサンプルサイズを決める場合、`bootnet` パッケージの `netSimulator` 関数でシミュレーションもできるが、ネットワークの構造が不明なこともあるので実際は使いにくい (Blanken et al., 2022)。ここでも、Isvoranu and Epskamp (in press) の大規模シミュレーション研究が参考になり、ノードは30より少ないのが望ましく、できるだけサンプルサイズが大きいほうが望ましいとされる。実際にいくつかのデータで心理ネットワーク分析をしてみると、ノードが数十ある場合に、サンプルサイズが100名以下だと推定が不安定になる。また、研究目的がネットワークの構造

を調べるものより、特定のエッジに焦点を絞ったものの方が大きなサンプルサイズが必要になる (Blanken et al., 2022)。

エッジの解釈法

ガウシアン・グラフィカルモデルのエッジは、偏相関になるので、通常の2変数間の相関係数ではなく、2変数以外の影響も考慮した2変数の関係になる。ネットワークに含まれる変数の違いによっては、値が異なる可能性もあり、2変数だけでなくネットワーク全体を考慮した解釈が必要になる。その際の解釈としては、つい「変数 A によって変数 B が引き起こされた」という因果に踏み込んだ解釈をしてしまうが、横断データの心理ネットワーク分析で得られたエッジからそのような因果の方向性をもった解釈は難しい。

ペアワイズ・マルコフ確率場のエッジの解釈としては、(1) 予測可能性の示唆と (2) 因果関係の示唆がある (Epskamp et al., 2018)。因果関係は不明だが、ほかの変数の影響を考慮しても、ある変数から別の変数が予測できているとはいえる。そのため、A-B-C という形でエッジがある場合に、B は A か C を予測する可能性がある。もちろん、最終的には縦断調査によって予測ができるかどうか検証する必要はある。

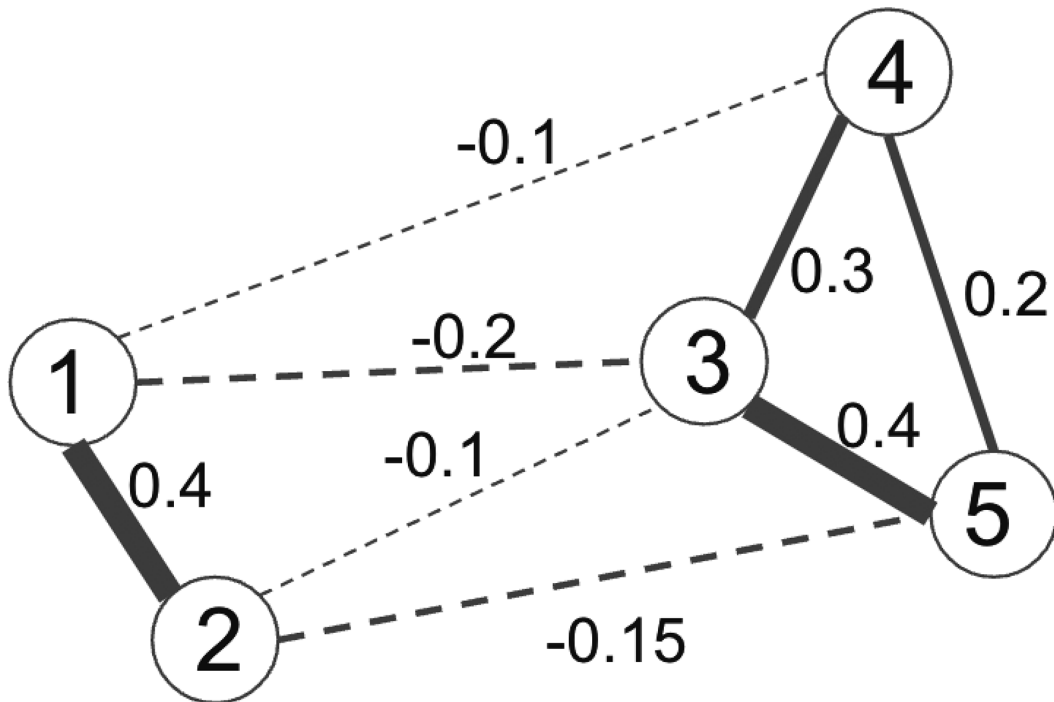


Figure 7 : 中心性指標のためのネットワークの例

次に、因果関係の仮定を満たすことはできないし、因果の方向性は不明だが、そのエッジには何らかの因果的な効果がある可能性はある。そこで、横断データのネットワーク分析の結果を踏まえて、より因果関係に踏み込めるような研究を計画することができる。若干、奥歯に物が挟まったような表現だが、多くの場合心理ネットワーク分析は探索的に実施されることから、強い理論的予測による仮説検証の結果のような解釈をしないように十分に気をつける必要がある。

ネットワーク特性の指標（中心性指標）

心理ネットワークは、重み付きネットワークのため、その他のネットワークと違って、グローバルなネットワーク特性の指標（small worldness, density, global clustering など）が使えない。そこで、各ノードのほかのノードへの影響性のようなローカルなネットワーク特性（中心性指標）を用いる。心理ネットワーク分析で用いられる主な中心性指標として、Strength, Closeness, Betweenness の3つがある。

Strength

Strength は、あるノードがつながっているすべてのエッジの強さを合計したものである。あるノードがほかのノードとどのくらい強くつながっているのかを表す指標であり、そのノードの全体に対する影響力を表してい

る。Figure 7 のネットワークを例にして、ノード 5 の Strength を計算してみる。ノード 5 とノード 3 の間には 0.4 のエッジがあり、ノード 5 とノード 4 の間には 0.2 のエッジがあり、ノード 5 とノード 2 の間には、 -0.15 のエッジがある。それぞれのエッジの絶対値を合計した 0.75 がノード 5 の Strength になる。この計算では、エッジの正負は問われない。

Strength は接続しているノードのエッジの強さの合計なので、接続しているノードが増えるほど大きくなる。そのため、かつては中心性指標は z 得点に変換されてプロットされていたが、最近では変換されていない値をプロットするのが推奨されており、qgraph パッケージでもデフォルトが z 変換していないものになっている。

Closeness

Closeness は、あるノードとほかのすべてのノード間の最短経路長の合計の逆数になる。まず、ノード間の最短経路長とは、あるノードから別のノードまで最短距離でいける距離になる。この距離というのは、エッジの強さの逆数になる。つまり、エッジの強さが 0.1 ならばその逆数は $1/0.1$ で 10 になり、エッジの強さが 0.5 ならば $1/0.5$ で 2 になる。最短経路長は、この距離を使って、あるノードから別のノードまでの距離を計算して、最も短い距離でいける距離を計算する。そして、あるノードへのほかのすべてのノードとの最短経路長を計算して、

合計したものの逆数が Closeness になる。

Figure 7において、ノード5からノード4への最短経路長を調べる。ノード5からノード4へいくには、(1) ノード5からノード4へ直接移動する、(2) ノード5からノード3を経由してノード4へ移動する、(3) ノード5からノード2を経由して最終的にノード4へ到達する(複数経路がある)などがある。経路自体は複数あるが、Figure 7の場合は、エッジの強さから、(1)か(2)が最短経路と考えられる。まず、(1)については、ノード5とノード4のエッジは0.2なので、その逆数の距離は、5になる。次に、(2)については、ノード5からノード3のエッジが0.4で距離は2.5、ノード3からノード4のエッジは0.3で距離は3.33になり、その合計は5.83となる。(1)と(2)を比較すると、(1)のノード5からノード4への直接経路の方が短く、この長さの5が最短経路長となる。

あるノードから他のノードまでの最短経路長をそれぞれ計算していった、それらを合計したものの逆数が Closeness になる。Closeness は、あるノードがどのくらい間接的な影響を含めて他のノードとつながっているのかを示している。一方、Strength は、あるノードがどのくらい直接的に他のノードとつながっているのかを示している。つまり、Closeness は、あるノードから他のノードへの拡散の速さを表現していると言える。

Betweenness

Betweenness は、2つのノード間の最短経路上に、あるノードが何回あるのかを示している。Closeness で最短経路を計算したが、その最短経路上に、特定のノードが登場する回数を見ることで、そのノードが影響性の流れにおいてどの程度の影響力を持つのかを検討できる。

Figure 7において、ノード5とノード1との最短経路は、ノード5からノード3を経由してノード1に直接いくというものである(複数ある経路の距離を計算するとよい)。このように、ノード5とノード1の最短経路上には、ノード3がある。このようにして、特定のノードが最短経路上に何回出てくるのかカウントすることで Betweenness が計算できる。

Betweenness は、あるノードが2つのノードの接続においてどのくらい重要かを示すものである。このノードによっては、影響性の流れを変えることができるので、心理的介入などにおいても重視される指標といえるかもしれない。

中心性の指標は、qgraph パッケージの centralityPlot 関数でプロットできる。中心性指標は複数あるので、include で“Strength”, “Closeness”, “Betweenness”の3つを指定する(デフォルトでは Strength のみが出力される)。ローゼンバーク自尊感情尺度においては (Fig-

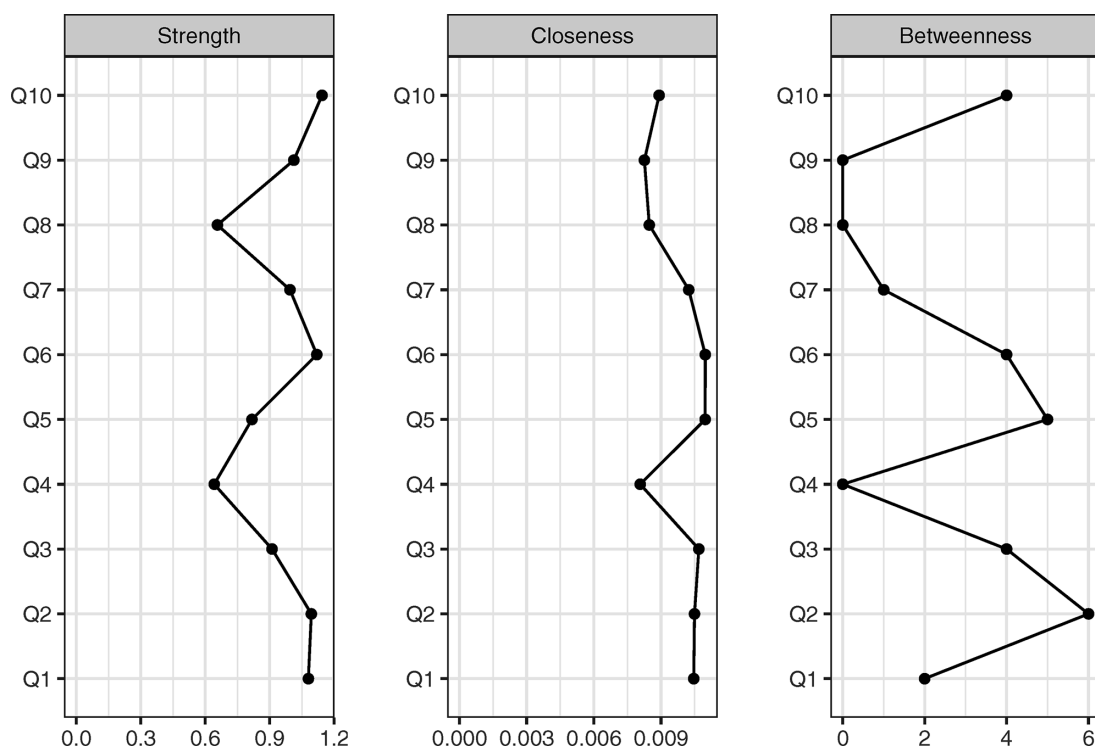


Figure 8 : ローゼンバーク自尊感情尺度の中心性指標

ure 8), Q1, Q2, Q6, Q10が Strength が高く, Q5, Q6が Closeness が高く, Q2が Betweenness が高いことが分かる。

```
centralityPlot(result2, include = c("Strength", "Closeness", "Betweenness"))
```

エッジの正確度と中心性指標の安定性

エッジは統計的に推定されるものなので、サンプルやサンプルサイズの影響をうける。サンプルサイズが足りないと、ネットワークのエッジの正確度 (accuracy) や中心性指標の安定性 (stability) が低くなることがある。そのため、ネットワーク分析では、正確度や安定性についても検討しておく必要がある (ネットワークの正確度については、Epskamp et al. (2018) や Fried, Epskamp, Veenman, and van Borkulo (2022) が詳しい)。

エッジの重みの正確度と中心性指標の安定性の検証方法は、(1) ブートストラップ法による信頼区間をプロットすることでエッジの重みの正確度を検討する、(2) データのサブセットを用いた場合の中心性指標の安定性 (stability) を検討する、(3) エッジの重みと中心性指標間でブートストラップ差異検定を行って、それらが有意に異なるのかを検討するの3ステップからなる (Epskamp et al., 2018)。心理ネットワーク分析を用いた際

には (1) を必ず実施するが、(2) と (3) については中心性に関心があるか差異に関心があるかによって実施を検討する。

エッジの重みの正確度

エッジの重みの正確度の評価には、ブートストラップ法によって95%信頼区間を推定する。ブートストラップ法は、データ元から復元抽出 (無作為に取り出したデータを戻してから取り出すことを繰り返すこと、つまり同じデータを取り出すことがある) を繰り返して、何度も取り出したデータから推定値を計算して、誤差などの性質を明らかにする方法になる。エッジの重みの95%信頼区間もブートストラップ法を用いて計算できる。なお、ブートストラップ法で推定した95%信頼区間は、0をまたぐかどうかで有意性の確認にも使えるが、心理ネットワークモデルの場合は正規化されているので、避けるほうが良い。あくまで、エッジの重みの正確度を調べるために使う。

ブートストラップ法には、ノンパラメトリックとパラメトリックがある。ノンパラメトリック・ブートストラップは常に使えるが、パラメトリック・ブートストラップは、データのパラメトリックモデルがある場合のみ使用できる。基本的にはノンパラメトリック・ブートストラップを使う。パラメトリック・ブートストラップ

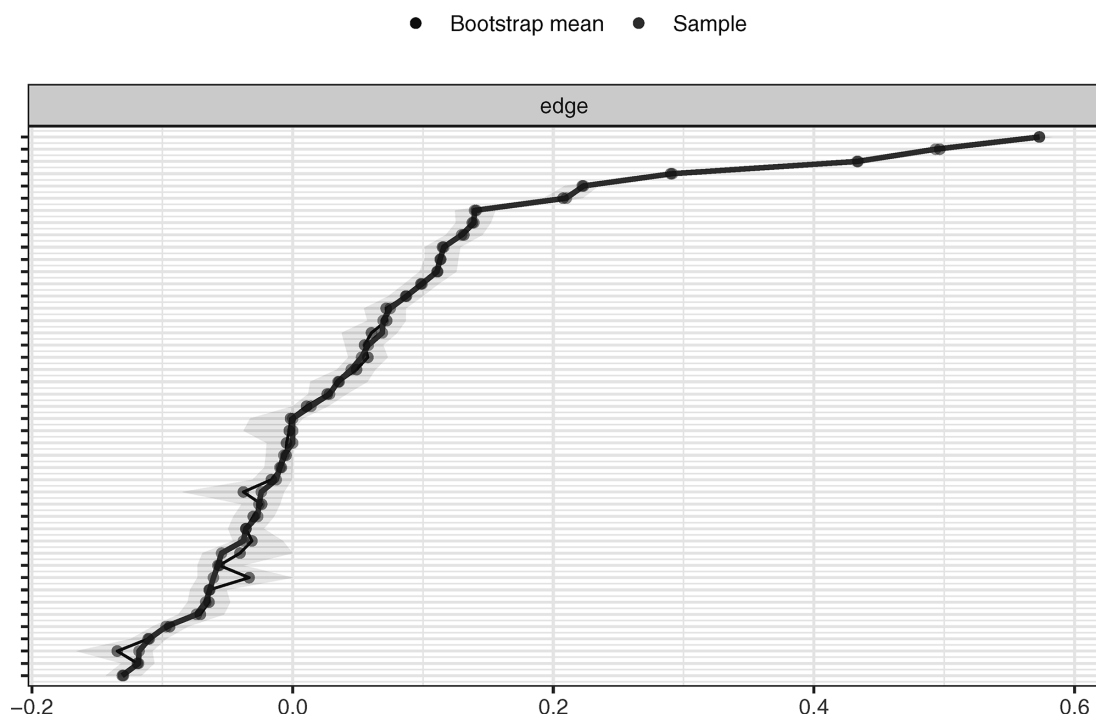


Figure 9 : 各エッジの正確度

を使うのは、正則化されていない、ノンパラメトリックでは結果が不安定である、ノンパラメトリックとパラメトリックで比較したい場合のみになる。

エッジの重みの正確度の推定には、bootnet パッケージの bootnet 関数を使う。nBoots でサンプル数を設定するが、ここでは1000サンプルとする。nCores で並列処理をして計算を高速化する。以下では、4 コアに設定しているが、解析を実施する環境に合わせて変更する必要がある。中心性指標も計算するために、statistics で "strength", "closeness", "betweenness" を指定する。

```
accuracy_edge <- bootnet(result2, nBoots = 1000,
nCores = 4, statistics = c("edge", "strength", "closeness",
"betweenness"))
```

summary(accuracy_edge) で推定結果を確認できる。ただ、以下のプロットのほうが確認しやすい。サンプルサイズが大きいためか、95%信頼区間の幅がかなり狭いことが確認できる (Figure 9)。

```
plot(accuracy_edge, labels = FALSE, order = "sample")
```

中心性指標の安定性

中心性指標の安定性を評価するために、データのサブ

セットを用いて、中心性指標の順序がどのくらい安定しているか検討する。具体的には、参加者を減らしていった場合 (ケースドロップ) もしくはノードを減らしていった場合 (ノードドロップ) にどのくらい中心性指標の順序が安定しているかを検討する。ノードドロップは解釈が難しくなるので、ケースドロップを使用することが多い。

中心性指標の安定性は、CS 係数 (correlation stability coefficient) によって評価する。CS 係数はドロップできるケースの最大比率を表現している。具体的には、参加者数を減らしていった場合に、元の中心性指標とドロップした中心性指標間に0.7以上の相関がブートストラップしたサンプルの95%にあるような最大のドロップケース比率である。多少恣意的な基準にはなるが、中心性の解釈をする上では、CS 係数が0.25未満は不適切であり、できればCS 係数は0.5を超えている必要がある (Epskamp et al., 2018)。

中心性指標の安定性は、エッジの重みの正確度と同様に bootnet 関数を使う。type で "case" を指定するとケースドロップ時の安定性を推定できる。なお、中心性の指標は複数あるので、statistics で "strength", "closeness", "betweenness" を指定する。

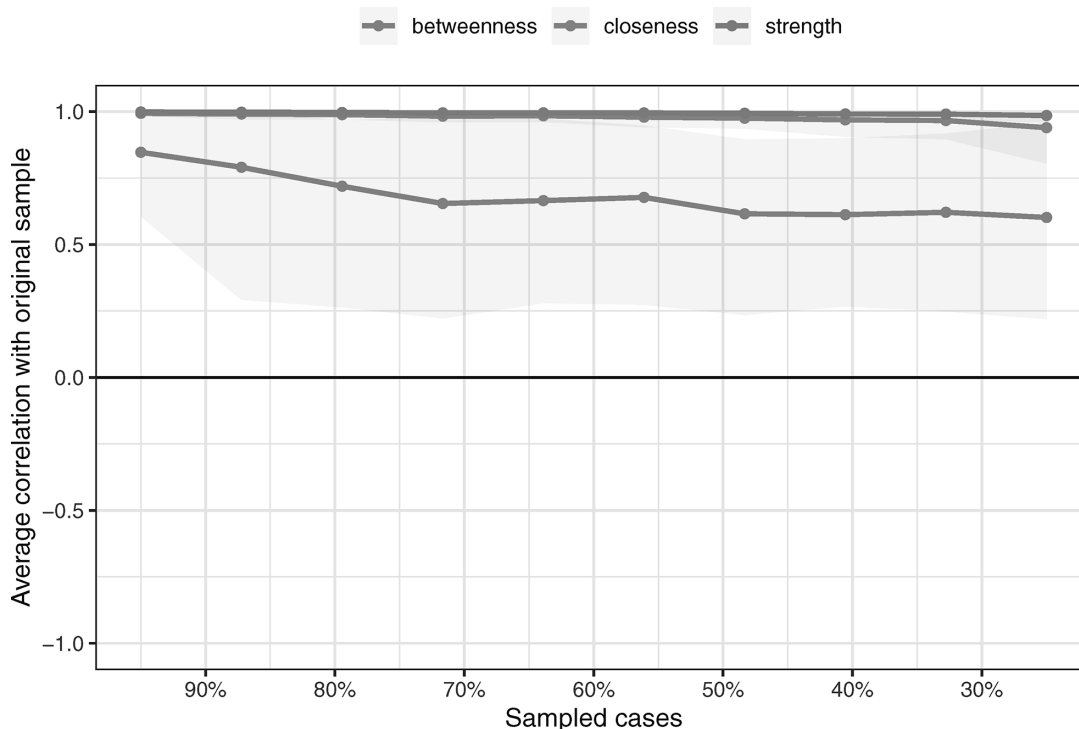


Figure 10 : ケースドロップ時の中心性指標の安定性

```
stability Centrality <- bootnet(result2, nBoots=1000,
type="case", nCores=4, statistics=c("strength",
"closeness", "betweenness"))
```

結果の以下のようにプロットする (Figure 10)。 Closeness や Strength はケースドロップしても安定しているが、 Betweenness は安定しない。今回は、47,974名のデータを用いているので、30%まで減らしても1万名以上のデータがある。それを考慮すると、 Betweenness の不安定さは際立っている。

```
plot(stability_Centrality, c("strength", "closeness",
"betweenness"))
```

以下のように、 corStability 関数で、 CS 係数を確認できる。実行すると、 Strength と Closeness は CS 係数が 0.7を超えているが、 Betweenness はかなり低いことがわかる。 Betweenness については、 解釈を控えた方がよいことがわかる。

```
corStability(stability_Centrality)
```

有意差検定

ブートストラップを活用して、 エッジの重みの比較や中心性指標の比較もできる。項目 1 と項目 2 のエッジ

は、 項目 1 と項目 3 のエッジよりも大きいことを確認したい場合や項目 1 のノードの Strength は項目 2 のノードの Strength よりも大きいことを確認したい場合に使える。これは、 ブートストラップ差異検定 (bootstrapped difference test) と呼ばれる。具体的には、 比較したいエッジ (中心性指標) のブートストラップ値の差を計算して、 その差の95%信頼区間を算出し、 0をまたぐかどうかで差の有無を判断できる。なお、 この際に、 0をまたがないことで、 そこに差があることを示すが、 0をまたぐことでエッジや中心性指標が等しいとはいえない点に注意する必要がある。

Strength の高い Q10と低い Q 4 を比較する。以下のように、 bootnet パッケージの differenceTest 関数で検討できる。実行すると、 95%CI が0をまたいでいないことが示され、 Q 4 と Q10は有意に Strength の高さが異なることが示される。

```
differenceTest(accuracy_edge, 4, 10, "strength")
```

エッジ間の差の検定結果を一気にプロットすることもできる (Figure 11)。それぞれのエッジ間で有意な差がある場合は黒く塗りつぶされ、 有意ではない場合は灰色に塗りつぶされる。

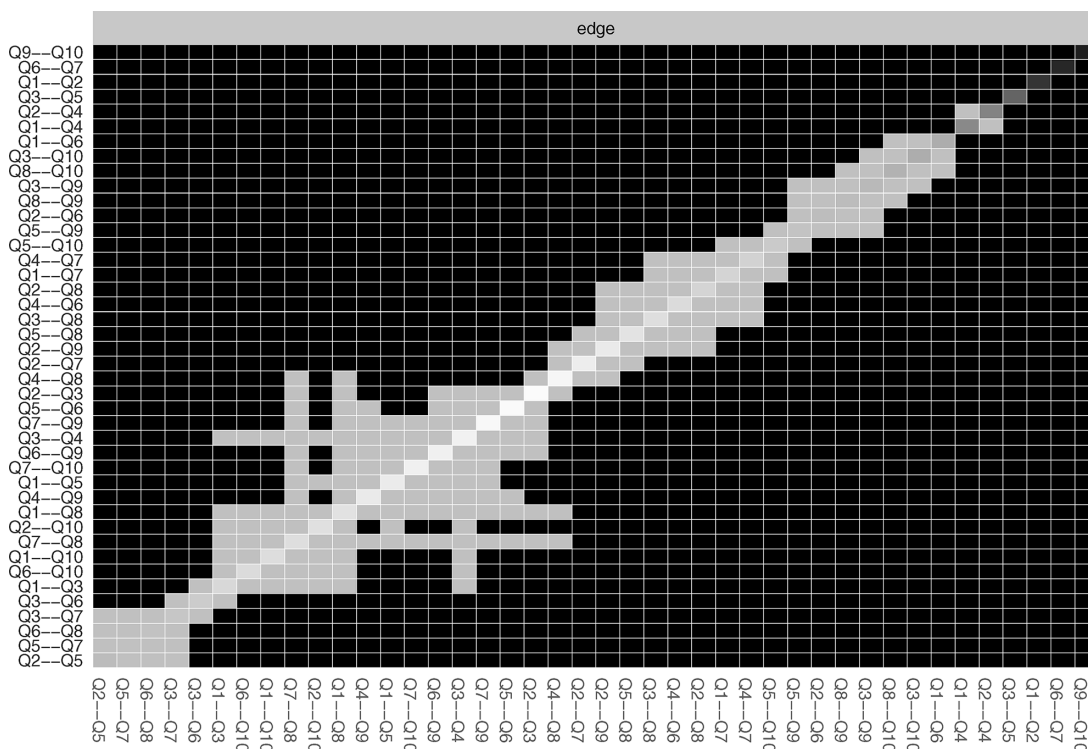


Figure 11 : エッジ間の差の検定結果

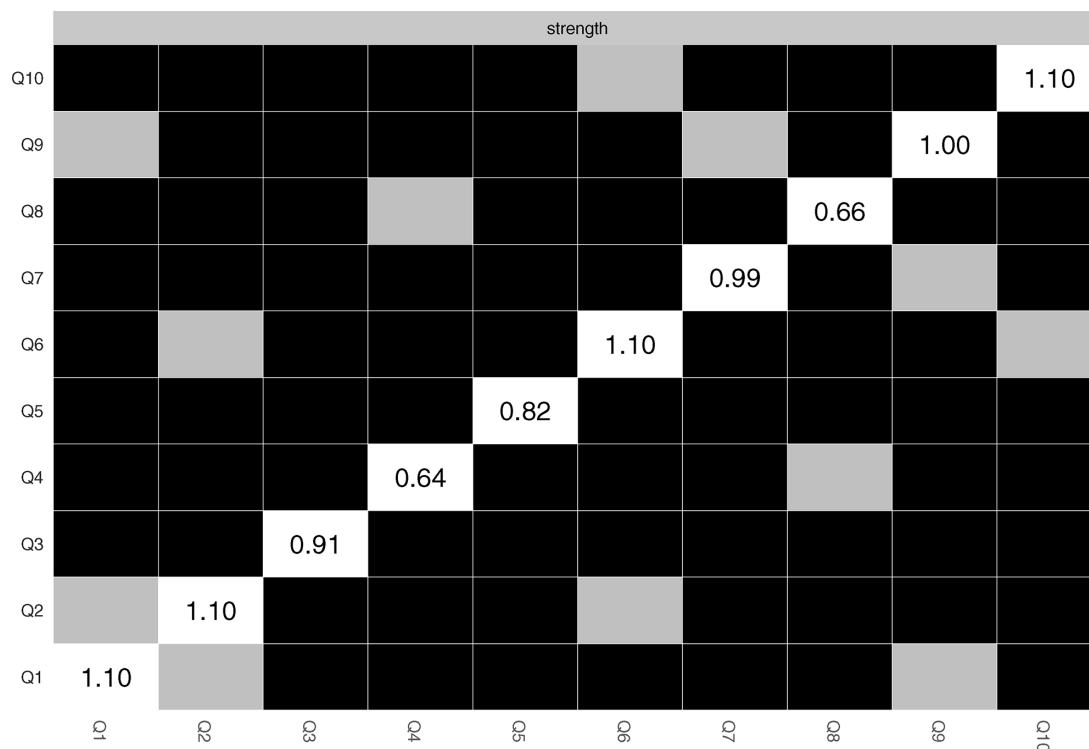


Figure 12 : Strength に関するノード間の差の検定結果

```
plot(accuracy_edge, "edge", plot="difference", only-
NonZero=TRUE, order="sample")
```

各ノード間の Strength の差の検定結果をプロットすることもできる。それぞれのエッジ間で有意な差がある場合は黒く塗りつぶされ、有意ではない場合は灰色に塗りつぶされる (Figure 12)。

```
plot(accuracy_edge, "strength")
```

おわりに

本論文では、横断データの心理ネットワーク分析における基本的な統計モデルであるガウシアン・グラフィカルモデルを中心に解説を行った。心理ネットワーク分析は、便利な R パッケージも充実してきている。その一方で、心理ネットワーク分析がどのように行われているのか中身についての理解がないと、誤用に繋がりやすい手法でもある。また、心理ネットワーク分析の方法は、日々新しい方法の開発がなされており、解析法の推奨も変化していつている。そのような状況があるので、アムステルダム大学の Psych Systems (<http://psychosystems.org/>) のサイトを確認したり、適宜最新の動向を確認しておく必要がある。

本論文では紹介できなかったが、横断データの心理

ネットワーク分析に関して、研究報告ガイドラインも作成されてきている (Burger et al., in press)。心理ネットワーク分析は比較的新しい手法ということもあり、研究報告についての決まりはなく、各研究者ごとに報告方法が異なっていた。Burger et al. (in press) は、横断データの心理ネットワーク分析に関して、方法と結果に分けて、報告すべき事項について整理をしている。Burger et al. (in press) には本論文では扱っていない部分についても言及があるが、本論文で整理した基本事項を押さえれば、求められる報告内容について理解できると思われる。また、このガイドラインをベースにした、心理ネットワーク分析のための事前登録フォーム (<https://osf.io/p9wn2/>) も提案されている。ガイドラインと合わせて、手元において、研究計画の立案、論文執筆することが望まれる。

引用文献

- Blanken, T.F., Isvoranu, A.M. and Epskamp, S. (2022). Estimating network structures using model selection In A.M. Isvoranu, S. Epskamp, L.J. Waldorp & D. Borsboom (Ed.), *Network psychometrics with R: A guide for behavioral and social scientists* (pp.111-

- 132). London : Routledge, Taylor & Francis Group.
- Burger, J., Isvoranu, A., Lunansky, G., Haslbeck, J.M., Epskamp, S., Hoekstra, R.H., Blanken, T.F. (in press). Reporting standards for psychological network analyses in cross-sectional data, *Psychological methods*.
- Epskamp, S., Haslbeck, J.M., Isvoranu, A.M. and Van Borkulo, C.D. (2022). Pairwise Markov random fields In A.M. Isvoranu, S. Epskamp, L.J. Waldorp & D. Borsboom (Ed.), *Network psychometrics with R: A guide for behavioral and social scientists* (pp.93-110). London : Routledge, Taylor & Francis Group.
- Epskamp, S., Borsboom, D. and Fried, E.I. (2018). Estimating psychological networks and their accuracy: A tutorial paper, *Behavior research methods*, 50 (1), 195-212.
- Epskamp, S., Cramer, A.O., Waldorp, L.J., Schmittmann, V.D. and Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data, *Journal of Statistical Software*, 48 (4), 1-18.10.18637/jss.v048.i04
- Epskamp, S., Maris, G., Waldorp, L.J. and Borsboom, D. (2018). Network Psychometrics In P. Irwing, T. Booth & D.J. Hughes (Ed.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp.953-986). Chichester, UK : John Wiley & Sons, Ltd.
- Fried, E.I., Epskamp, S., Veenman, M. and van Borkulo, C.D. (2022). Network stability, comparison, and replicability In A.M. Isvoranu, S. Epskamp, L.J. Waldorp & D. Borsboom (Ed.), *Network psychometrics with R: A guide for behavioral and social scientists* (pp.133-153). London : Routledge, Taylor & Francis Group.
- Isvoranu, A. and Epskamp, S. (in press). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers, *Psychological methods*.
- Haslbeck, J.M. and Waldorp, L.J. (2020). mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data, *Journal of Statistical Software*, 93 (8), 1-46.10.18637/jss.v093.i08
- 檜原潤 (2019). 精神病理ネットワークの応用可能性心理学評論, 62 (2), 143-165. 10.24602/sjpr.62.2_143
- 檜原潤・伊藤正哉 (2022). 心理ネットワークアプローチがもたらす「臨床革命」—認知行動療法の文脈に基づく展望—認知行動療法研究, 48 (1), 35-45. 10.24468/jjbct.20-015
- Rosenberg, M. (1965). *Society and the adolescent self-image*. NA: Princeton, NJ: Princeton University Press.