

Упорядоченные сети частных корреляций в психологических исследованиях

*Артеменков С.Л.**

Московский государственный психолого-педагогический университет
(ФГБОУ ВО МГППУ), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-1619-2209>
e-mail: slart@inbox.ru

Сетевое моделирование, которое появилось в последние годы, может с успехом применяться к рассмотрению связей между измеряемыми психологическими переменными. В этом контексте психологические переменные понимаются как непосредственно влияющие друг на друга, а не как следствие латентного конструкта. В статье описаны методы регуляризации, которые могут использоваться для эффективной оценки разреженной и интерпретируемой структуры сети на основе частных корреляций психологических измерений. Представлен обзор процедуры регуляризации глассо с использованием выбора модели EBIC для оценки упорядоченной разреженной сети частных корреляций. Рассмотрены вопросы выполнения этого анализа в R при наличии нормального и не нормального распределения данных с учетом влияния гиперпараметра, который устанавливается исследователем вручную. Рассмотренный подход интересен также как способ визуализации возможных каузальных связей между переменными. Данный обзор устраняет пробел, относящийся к отсутствию доступного описания на русском языке этого пока малораспространенного в России и в тоже время многообещающего подхода.

Ключевые слова: корреляционный анализ, сети частных корреляций, регуляризация, моделирование сетей в психологии, язык R.

Для цитаты:

Артеменков С.Л. Упорядоченные сети частных корреляций в психологических исследованиях // Моделирование и анализ данных. 2021. Том 11. № 2. С. 31–50. DOI: <https://doi.org/10.17759/mda.2021110202>

**Артеменков Сергей Львович*, кандидат технических наук, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-1619-2209>, e-mail: slart@inbox.ru



1. ВВЕДЕНИЕ

Корреляционный анализ является одним из самых распространенных методов статистики и представляет собой метод обработки статистических данных, заключающийся в изучении коэффициентов корреляции между переменными. Эти коэффициенты помогают выявить меру линейной связи между соответствующими переменными. «Хотя корреляция прямо не указывает на причинную связь, она может быть ключом к разгадке причин» [7, с. 105].

При существовании ряда связанных между собой переменных, их корреляционный анализ удобно визуализировать различными способами (представления в виде таблиц, коррелограмм, тепловых карт и графов). В частности, весьма удобным и наглядным является представление связей ряда переменных в виде графа, состоящего из сети вершин, соединенных ребрами графа на множестве точек плоскости (или пространства) [1]. При этом вершины графа обозначают исходные переменные, а ребра соответствуют корреляционным связям. В такой общей сетевой модели вес каждого ребра определяется величиной коэффициента корреляции между соответствующими узлами. Величиной веса может определяться и толщина ребра в то время, как длина ребра в сети часто определяется, как обратная величина к весу ребра. Два сильно связанных узла поэтому располагаются ближе друг к другу. Такие сети часто называются взвешенными. Как правило, соединения в сети визуализируются с использованием разных обозначений для отрицательных и положительных корреляций. Например, в цветных изображениях: красные линии указывают отрицательные корреляции, а зеленые (или синие) линии – положительные. В черно-белом варианте положительные корреляции обычно обозначаются непрерывной линией, а отрицательные – пунктирной линией [2].

В последние годы наблюдается все более широкое использование сетевого моделирования для предварительных исследований психологического поведения в качестве альтернативы моделированию конструкторов как латентных переменных [2, 13, 43]. В этих так называемых психологических сетях [20] узлы представляют собой психологические переменные, такие как оценки состояний, настроений, симптомов или отношений, в то время как ребра (связи, соединяющие два узла) представляют неизвестные статистические отношения, которые необходимо оценить, что создает новые проблемы для статистического вывода. В результате этот класс сетевых моделей отличается от социальных сетей, в которых связи известны [49]. Существует большое количество технической литературы по оценке сетевых моделей [24, 25, 27, 34]. Тем не менее, эта литература в основном представлена на английском языке, часто требует довольно глубоких математических знаний и не фокусируется на уникальных проблемах, связанных с анализом психологических данных.

При всей простоте и наглядности корреляционного анализа (в частности, в его сетевом представлении) корреляционное исследование может приводить к ложным выводам о наличии причинно-следственной связи между парами отдельных признаков. Общие коэффициенты корреляции устанавливают лишь статистические взаимосвя-



зи, которые могут иметь место, например, в следствие влияния других переменных. Таким образом, представление сетей общих корреляций мало говорит о реальных каузальных отношениях между переменными даже после проверки условий их значимости, поскольку кроме совместных вариаций переменных они должны иметь правильное хронологическое упорядочение и должно быть исключено возможное влияние других факторов [6].

В связи с этим для установления реальных взаимосвязей между переменными часто важно рассматривать коэффициенты частных корреляций (Partial Correlation), которые показывают степень (тесноту) взаимосвязи двух переменных относительно друг друга, без учета влияния всех других переменных. В настоящее время сеть частных корреляций является распространенной моделью, используемой для оценки психологических сетей на основе непрерывных данных. Частные корреляционные сети обычно оцениваются с использованием методов регуляризации, происходящих из области машинного обучения. Было показано, что эти методы хорошо работают при восстановлении истинной структуры сети [24, 25, 34]. Регуляризация включает в себя оценку статистической модели с дополнительным штрафом за сложность модели. Это приводит к тому, что оцениваемая модель является разреженной: многие параметры оцениваются как точно нулевые. При оценке сетей это означает, что связи, которые могут быть ложными, удаляются из модели, что приводит к сетям, которые проще интерпретировать. Таким образом, регуляризация, позволяет осуществить упорядочение сети, включающие как выбор конкретной модели, так и оценку параметров сети.

Регуляризованная оценка сети используется в значительном количестве публикаций в различных областях, таких как клиническая психология [18, 23], психиатрия [14, 28], исследования личности [8, 17, 29] и медицинские науки [38]. Общим для этих статей является то, что они предполагают, что наблюдаемые переменные каузально влияют друг на друга, что приводит к сетевым моделям, которые позволяют понять связи между симптомами психопатологии (например, печальное настроение, усталость и бессонница), элементами личностных качеств, таких как сознательность (например, импульсный контроль, упорядоченность и трудолюбие), или поведением, связанным со здоровьем (например, ощущение полноты энергии, легкое заболевание и трудности в выполнении повседневных задач). С этой сетевой психометрической точки зрения, корреляции между элементами проистекают из взаимной ассоциации между переменными, которая отличается от традиционной точки зрения, где корреляция между измеряемыми переменными объясняется с помощью введения латентных конструкторов [2, 43].

В результате такие психометрические сети могут предложить другое представление о конструкторах, как кластерах определенных взаимосвязанных характеристик: синдромы, такие как депрессия или тревожное расстройство, в области психических расстройств [12, 14], аспекты или области личности, такие как экстраверсия или невротизм в исследованиях личности [44], области исследований здоровья [31] и динамических факторов риска [10]. В отечественных работах сетевые методы были использованы для исследования одаренности [4, 5, 11].



При этом важно отметить, что при использовании методов сетевого корреляционного анализа необязательно придерживаться только представленной выше сетевой перспективы. Психологические сети в целом могут быть мощными инструментами для изучения мультиколлинеарности и прогнозов медиации, а также могут использоваться для выделения латентных переменных и каузальных связей [2, 3].

Цель данной статьи состоит в том, чтобы представить на русском языке введение в регуляризацию сетей частных корреляций, обрисовать возможности для построения и оценки моделей этих сетей, имеющиеся, в частности, в языке программирования R, и рассмотреть некоторые вопросы, возникающие при сетевом корреляционном анализе в области психологии. Представленные положения и результаты основаны на ряде зарубежных работ [16, 20, 21].

2. СЕТИ ЧАСТНЫХ КОРРЕЛЯЦИЙ

Наиболее часто используемая основа для построения психологической сети на основе данных, которые можно считать многомерными нормальными, – это оценка сетевых связей в виде коэффициентов частных корреляций [13]. Эти коэффициенты, также известные как условные независимые ассоциации, варьируются от -1 до 1 и кодируют оставшуюся связь между двумя узлами после контроля всей другой возможной информации. Частные корреляционные сети также называют графами концентрации или Гауссовскими графическими моделями, и они являются частью более общего класса статистических моделей, называемых попарно марковскими случайными полями [30].

После того, как частные корреляции оценены, они могут быть визуализированы во взвешенной сетевой структуре (в виде графа, описанного во введении). Каждый узел представляет переменную, а каждое ребро представляет, что две переменные не являются независимыми после согласования всех переменных в наборе данных. Каждое из ребер сети имеет вес, равный коэффициенту частной корреляции между соответствующими двумя узлами. Всякий раз, когда частная корреляция точно равна нулю, между двумя узлами не проводится ребра, что указывает на то, что две переменные являются независимыми после контроля всех других переменных в сети.

Несколько различных пакетов программного обеспечения могут быть использованы для визуализации сети. Например, можно использовать свободно доступные пакеты программного обеспечения `cytoscape`, `gephi` или R-пакеты `qgraph`, `igraph`, `Rgraphviz` или `ggraph` [40]. Пакет `qgraph` обычно используется в психологической литературе, так как он автоматизирует много шагов для графического построения взвешенных сетей и включает различные методы их оценки. При изображении сетевой модели более широкие и насыщенные линии указывают на более сильные частные корреляции.

2.1. Вычисление частных корреляций

В теории существует несколько способов вычисления коэффициентов частных корреляций. Для быстрого получения частных корреляций обычно используются

два метода. Во-первых, частные корреляции могут быть получены непосредственно из обратной матрицы дисперсии-ковариации. Если y представляет набор переменных измерений, представленных многомерным нормальным распределением, которое, без потери общности, можно считать центрированным, а K обозначает инвертированную матрицу дисперсии-ковариации S , то элементы k_{ij} (строка i , столбец j в K) могут быть стандартизированы для получения коэффициентов частных корреляций между переменной y_i и переменной y_j после согласования со всеми другими переменными [32]:

$$\text{Cor}(y_i, y_j | y_{-(i,j)}) = \frac{k_{ij}}{\sqrt{k_{ii}}\sqrt{k_{jj}}}.$$

Альтернативный способ получения коэффициентов частных корреляций заключается в использовании уравнений линейных регрессий по всем узлам [34]. После проведения множественной регрессии коэффициент частной корреляции между y_i и y_j определяется наклоном линии регрессии, предсказывающей одну переменную из другой или наоборот [41]. Если частная корреляция, полученная из регрессионной модели одного узла, немного отличается от регрессионной модели для другого узла, то получить единственную оценку можно путем усреднения двух оцененных частных корреляций.

2.2. Интерпретация сетей частных корреляций

Сети частных корреляций допускают несколько мощных выводов. Во-первых, они позволяют моделировать уникальные взаимодействия между переменными. Если переменная A коррелирует с B , а B коррелирует с C , то мы естественно ожидаем, что A коррелирует с C . Безусловная нулевая корреляция между A и C будет неожиданной, поскольку к такой корреляционной ситуации приводят немного причинных структур. Два возможных здесь варианта предполагают, что B является общим эффектом от A и C или две ортогональные латентные переменные вызывают ковариацию между A и B и между B и C . Если данные являются нормальными, частные корреляции могут быть интерпретированы как попарные взаимодействия (что не следует путать с эффектами взаимодействия двух переменных на выходной переменной). Для моделирования корреляций достаточно использовать только две пары: взаимодействие между A и B и взаимодействие между B и C . Эта модель будет содержать одну степень свободы и таким образом приводит к проверяемой гипотезе. Такая точка зрения сходна логлинейному моделированию категориальных данных [9], которое структурно сопоставимо с сетью частных корреляций.

Сеть частных корреляций отображает мультиколлинеарность и прогнозирует медиацию, поскольку частные корреляции тесно связаны с коэффициентами, полученными в моделях множественной регрессии: когда независимая переменная не предсказывает зависимую переменную, мы не ожидаем, что в сети будет иметься ребро. Сила частной корреляции, кроме того, напрямую связана с силой коэффициента регрессии. Таким образом, ребра, связанные с одним узлом, показывают исследователю ожидаемый результат анализа множественной регрессии. В отличие от того,



что можно увидеть из анализа множественной регрессии одной зависимой переменной, сеть также показывает, какие переменные будут предсказывать независимые переменные. Связывая отдельные модели множественной регрессии, сети частных корреляций позволяют отображать линейное прогнозирование и мультиколлинеарность среди всех переменных. Это позволяет получить представление о предсказательном посредничестве: сеть, в которой две переменные не связаны напрямую, а косвенно связаны (например, $A - B - C$), указывает на то, что A и C могут быть коррелированы, но любой прогнозирующий эффект от A до C (или наоборот) опосредуется B .

Частные корреляции могут указывать на возможные причинные пути. Отношения условной независимости, например, кодируемые коэффициентами частных корреляций, играют решающую роль в причинно-следственной связи [39]. Когда предполагается, что все соответствующие переменные наблюдаются (то есть, нет латентных переменных), частная корреляция между переменными A и B может быть не нулевой, только если A вызывает B , B вызывает A , существует взаимная связь между A и B или A и B вызывают третью переменную в сети [30, 39].

В связи с этим, сети частных корреляций можно рассматривать как исследовательские структуры, которые позволяют генерировать гипотезы, указывающие на возможные причинные эффекты. Хотя существуют исследовательские алгоритмы, целью которых является обнаружение направленных (причинных) сетей, они опираются на сильные предположения, такие как ацикличность (переменная не может в конечном итоге вызывать себя, например, $A \rightarrow B \rightarrow C \rightarrow A$), и на нее сильнее влияют латентные переменные, вызывая ковариацию (латентные переменные будут вызывать направленные ребра между наблюдаемыми переменными, что подразумевает сильную причинную гипотезу). Кроме того, эти модели нелегко идентифицировать или параметризовать: многие эквивалентно ориентированные модели могут в равной степени соответствовать данным, при разной параметризации. Сети с частными корреляциями, с другой стороны, могут быть хорошо идентифицированы (без эквивалентных моделей) и легко параметризуются с использованием коэффициентов частных корреляций.

Таким образом, оперативная оценка неориентированных сетей предлагает привлекательную альтернативу исследовательской оценке направленных сетей, которые характеризуются проблемными и плохо идентифицированными направлениями и величинами воздействий. Следует отметить, что сеть частных корреляций не должна интерпретироваться как сеть, определяющая скелет причинно-следственной модели (направленная сеть со снятыми наконечниками стрелок). В частности, преимущество сети частных корреляций заключается в возможности показа взаимодействий, обусловливающих такой общий эффект, который обычно связывают с существованием латентных переменных. Кроме того, латентные переменные могут приводить к ребрам как в направленных, так и в ненаправленных сетях.

Кластеры в сети могут выделять латентные переменные. Хотя частные корреляции направлены на выделение уникальной дисперсии между двумя переменными, они сохраняют общую дисперсию из-за внешних источников, которые не могут быть



полностью разделены путем контроля других переменных в сети. В результате, если латентная переменная вызывает ковариацию между двумя или более переменными в сети, ожидается, что все эти переменные будут связаны в сети, образуя кластер [2, 26]. Такие кластеры могут указывать на латентные переменные.

3. РЕГУЛЯРИЗАЦИЯ ЛАССО

Как показано выше, частные корреляции можно легко оценить путем инвертирования выборочной дисперсионно-ковариационной матрицы или путем выполнения последовательных множественных регрессий и стандартизации полученных коэффициентов. Однако оценка параметров по данным всегда сопровождается вариациями выборки, что приводит к тому, что оценки никогда не бывают точно равны нулю. Даже когда две переменные условно независимы, в сети могут появляться не нулевые (хотя обычно небольшие) частные корреляции, которые будут представлены как очень слабые ребра в сети. Эти связи условно называют «ложными» или «ложными срабатываниями» [16]. Чтобы предотвратить чрезмерную интерпретацию и сбой в репликации оценочных сетевых структур, важной целью при оценке сети является ограничение количества ложных соединений. Методы, которые позволяют уменьшить количество ложных соединений, называют методами регуляризации сети. При этом можно видеть, что не регуляризованные сети частных корреляций уже позволяют уменьшить количество ложных срабатываний в сети с предельными коэффициентами корреляции [16].

Обычный способ осуществить регуляризацию сети состоит в том, чтобы проверить все частные корреляции на статистическую значимость и удалить все ребра, которые не достигают определенной степени значимости [19]. Однако это создает проблему многократного тестирования, и исправление этой проблемы (например, с помощью коррекции Бонферрони) приводит к потере информативности [16].

Практический способ решения проблемы произвольного выбора сети состоит в построении сетей на основе их различных вариантов и рассмотрении того, насколько стабильными являются основные результаты. В связи с этим все более популярными становятся методы статистической регуляризации, способствующие получению более понятных сетей, которые лучше экстраполируются на новые выборки. В частности, известным методом регуляризации является оператор наименьшей абсолютной усадки и отбора: лассо (lasso; [46]), который, в отличие от других методов регуляризации, может привести к точным оценкам параметров, равным нулю. По сути, лассо ограничивает сумму абсолютных коэффициентов частных корреляций; в результате все оценки уменьшаются, а некоторые становятся точно равными нулю. С технической точки зрения, лассо стремится оценить инверсную матрицу дисперсии-ковариации \mathbf{K} путем максимизации штрафной функции правдоподобия [25]:

$$\log \det (\mathbf{K}) - \text{trace}(\mathbf{SK}) - \lambda \sum_{\langle i, j \rangle} |k_{ij}|.$$



В качестве альтернативы, лассо-регуляризация может применяться к отдельным моделям регрессии, если сеть оценивается с использованием регрессий по узлам [34].

Использование лассо приводит к разреженной сети, в которой удаляются вероятные паразитные ложные ребра. Лассо использует параметр настройки λ (лямбда), который определяет уровень разреженности и напрямую контролирует, насколько частные корреляции штрафуются за сумму абсолютных значений всех коэффициентов корреляции. Когда параметр настройки λ маленький, то удаляются только несколько ребер, что может привести к сохранению ложных ребер. Когда параметр настройки велик, то удаляются многие ребра, что может привести к удалению в сети истинных ребер в дополнение к удалению ложных ребер. Поэтому на практике необходимо тщательно выбирать параметр настройки, чтобы создать сетевую структуру, которая минимизирует количество паразитных ребер при максимальном количестве истинных ребер [24].

При выборе параметра настройки лассо, как правило, несколько сетей оцениваются с разными значениями λ [50]. Различные значения λ могут быть выбраны из логарифмически разнесенного диапазона между максимальным значением λ , для которого не сохраняется ребро (λ равно наибольшему абсолютному значению), и некоторыми скалярными множителями, уменьшающими это максимальное значение в определенное число раз. Текущая версия пакета `qgraph` использует, например, коэффициент 0,01 и оценивает 100 сетей по умолчанию.

Таким образом, метод регуляризации лассо обычно используется для оценки совокупности сетей, а не отдельной сети, начиная от полностью подключенной сети (с максимумом ребер) до полностью отключенной сети (без ребер). Затем из этой коллекции сетей выбирается лучшая сеть. Этот выбор можно сделать путем оптимизации соответствия сети данным, сводя к минимуму некоторый информационный критерий. Было показано, что минимизация расширенного байесовского информационного критерия (EBIC; [15]) хорошо работает для определения истинной структуры сети [24], особенно когда генерируемая сеть разрежена (т.е. не содержит много ребер). Регуляризация лассо с выбором модели EBIC характеризуется высокой специфичностью во всем (т.е. не оценивает ребра, которые не находятся в истинной сети), но изменяющейся чувствительностью (т.е. оценивает ребра, которые находятся в истинной сети) на основе истинной структуры сети и размера выборки. Например, чувствительность обычно меньше, когда истинная сеть плотная (содержит много ребер) или имеет несколько узлов (концентраторов) со многими ребрами.

В модели EBIC используется гиперпараметр γ (гамма), который управляет другими параметрами, и обычно его необходимо устанавливать вручную. Гиперпараметр γ определяет, насколько EBIC предпочитает более простые модели (т.е. имеющие меньше ребер) [15, 24]:

$$EBIC = -2L + E \log(N) + 4\gamma E \log(P).$$

Здесь L обозначает логарифмическую функцию правдоподобия, N – размер выборки, E – число не нулевых ребер и P – количество узлов. Этот гиперпараметр γ



необходимо устанавливать вручную, и обычно он устанавливается в диапазоне от 0 до 0,5 [24], причем более высокие значения указывают на то, что более простые модели (более экономные модели с меньшим числом ребер) являются предпочтительными. При установке гиперпараметра в 0 ошибок обнаружения оценивается большее количество ребер, включая возможные побочные (сеть имеет более высокую чувствительность). Установка гиперпараметра равным 0,5 приводит к ошибкам предосторожности или экономии: получается меньше ребер, что исключает большинство паразитных ребер, но, возможно, при этом пропускаются некоторые ребра (то есть сеть имеет более высокую специфичность). Даже при установке гиперпараметра γ в 0 сеть все равно будет более разреженной по сравнению с сетью частных корреляций, которая не использует какую-либо форму регуляризации; установка γ в 0 указывает на то, что EBIC сводится к стандартному BIC, при котором по-прежнему предпочтение отдается простым моделям.

Какое из этих значений γ лучше всего работает, является сложной функцией (обычно неизвестной) истинной сетевой структуры. На практике рекомендуется первоначально устанавливать среднее значение $\gamma = 0,25$.

В настоящее время многие варианты алгоритмов лассо реализованы в программном обеспечении с открытым исходным кодом [51]. Предлагаемый далее вариант, называемый графическим лассо (glasso; [25], специально предназначен для оценки сетей частных корреляций путем инверсии выборочной матрицы дисперсии-ковариации. Алгоритм графического лассо (glasso) был реализован в пакете glasso для языка статистического программирования R. Функция, которая использует этот пакет в сочетании с выбором модели EBIC, была реализована в пакете qgraph и может также вызываться через пакет начальной загрузки [20].

Алгоритм glasso непосредственно штрафует элементы матрицы дисперсии и ковариации. Это отличает его от других методов оценки сети лассо, которые обычно стремятся оценить структуру сети путем штрафования коэффициентов регрессии в серии моделей множественной регрессии [34]. Процедура glasso может быть задействована с помощью простых команд ввода (потому что она требует только оценки ковариационной матрицы, а не необработанных данных), что позволяет использовать полихорические корреляционные матрицы, когда данные являются порядковыми.

Следует отметить, что хотя лассо-регуляризация приводит к удалению ребер из сети, это не доказывает, что эти ребра фактически равны нулю. Это потому, что лассо стремится максимизировать специфичность; то есть стремится включать как можно меньше ложных срабатываний (ребер, которых нет в истинной модели). В результате рассмотрение предполагаемой сети, которая является разреженной (с отсутствующими ребрами), или даже пустой сетью, никоим образом не является доказательством того, что пропущенные ребра фактически отсутствуют. Оценка лассо может привести к множеству ложных негативов, ребер, которые не присутствуют в предполагаемой сети, но присутствуют в истинной сети. Это связано с хорошо известной проблемой проверки нулевых гипотез: не отказ от нулевой гипотезы не является доказательством того, что нулевая гипотеза верна [48]. Мы не можем включить



ребро либо потому, что данные слишком шумные, либо потому, что нулевая гипотеза верна. Лассо-регуляризация, как и классическая проверка значимости, не может различить эти две причины.

4. НЕ РАСПРЕДЕЛЕННЫЕ НОРМАЛЬНО ДАННЫЕ

Как было отмечено ранее, оценка сетей частных корреляций связана с предположением о многомерной нормальности. Оценка этих сетей тесно связана с моделированием структурными уравнениями [22], где также обычно требуется многомерное нормальное распределение. Это не только означает, что предельные распределения должны быть нормальными, но и все отношения между переменными также должны быть линейными. При этом возникает вопрос, что делать с ненормальными (например, упорядоченными категориальными) данными, которые часто встречаются в психологических данных? Несколько решений, предложенных в литературе по моделированию структурными уравнениями, могут также быть использованы для моделирования сетей.

Предположение о нормальности может быть ослаблено, если положить, что наблюдаемые данные являются производными скрытой многомерной нормально распределенной системы [21, 33]. Рассмотрим эту ситуацию, следуя работе [21], из которой в основном взяты дальнейшие рисунки 1–4.

Пример такой системы показан на рис. 1, где квадратные элементы обозначают наблюдаемые переменные, круглые элементы указывают нормально распределенные латентные переменные, а направленные стрелки указывают на монотонные функции преобразования (каждое значение латентной переменной преобразуется в одно уникальное значение измеряемой переменной с сохранением порядкового номера неизменным; более высокие значения в исходной шкале также выше в преобразованной шкале). Обратите внимание, что здесь не идет речь об ошибке измерения, которая может быть учтена при наличии нескольких индикаторов, приходящихся на одну латентную переменную [22]. Предполагается, что каждая наблюдаемая переменная указывает только на одну латентную переменную [35].

В этом случае два сценария являются наиболее распространенными: наблюдаемые переменные являются непрерывными или они состоят из упорядоченных категорий. Когда наблюдаемые переменные непрерывны, но не распределены нормально, переменные могут быть преобразованы так, чтобы иметь предельное нормальное распределение. Мощный метод, который был использован в оценке сети, заключается в применении непаранормального преобразования [33]. Это преобразование использует кумулятивные распределения (кодирующие вероятность того, что переменная ниже некоторого уровня), чтобы преобразовать распределение наблюдаемой переменной в распределение латентной нормально распределенной переменной.

Рисунок 2 показывает упрощенный пример того, как два распределения X и Y могут быть связаны их совокупными (кумулятивными) распределениями. Предположим, что X нормально распределено, а Y – гамма-распределено (потенциально ис-

кажено). Затем значения X могут быть сопоставлены с кумулятивным распределением с помощью функции вероятности (в R: `pnorm`). Затем эти кумулятивные вероятности могут быть сопоставлены со значениями гамма-распределения с помощью функции квантиля (в R: `qgamma`). На практике распределение Y неизвестно. Плотность и совокупная плотность X , с другой стороны, известны, и совокупное распределение Y может быть оценено путем вычисления эмпирической интегральной функции распределения (в R: `ecdf`).

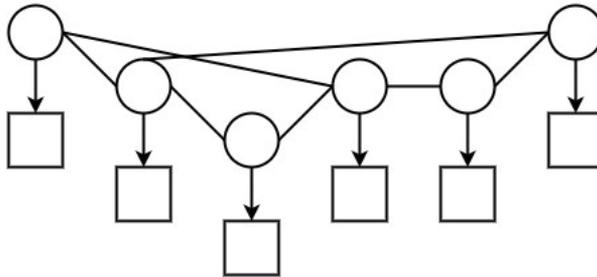


Рис. 1. Метод ослабления предположения о многомерной нормальности. Наблюдаемые переменные (квадраты), предположительно являются преобразованиями многомерных латентных нормальных переменных (круги); связи между кругами заданы произвольно и могут быть другими

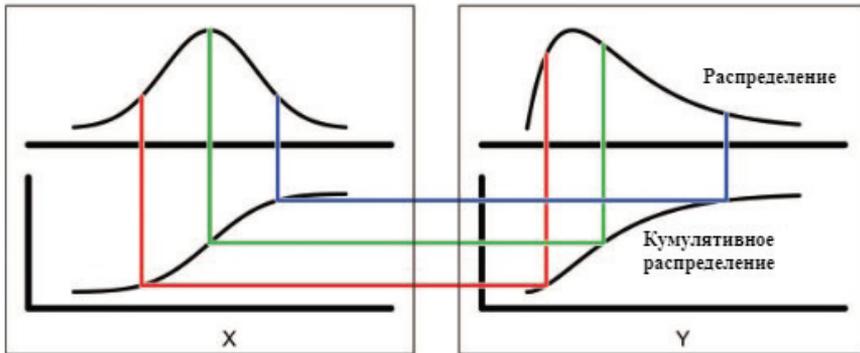


Рис. 2. Схема преобразования кумулятивных распределений для перевода распределения переменной Y в нормальное распределение X

Таким образом, чтобы сопоставить значения Y со значениями нормально распределенной переменной X , необходимо оценить функцию плавного преобразования между двумя кумулятивными распределениями. Это ядро непаранормального преобразования, целью которого является сопоставление каждого уникального результата переменной (например, обозначения 1, 2 или 3 на рисунке 3) с одним уникальным результатом стандартной нормальной переменной (например, 1,96, 0, 1,65). Для этой цели можно использовать функцию `huge.pnr` из пакета `huge` в R [51]. Важно отметить, что это преобразование предполагает плавное увеличение кумулятивных

распределений и поэтому не будет работать, когда присутствуют только несколько возможных вариантов ответа (например, в шкалах Ликерта). Когда данные являются двоичными, преобразованные данные все еще будут двоичными, просто используя метки, отличные от 0 и 1.

Когда доступно только несколько категорий элементов и варианты ответов можно считать порядковыми [45], можно использовать пороговые функции [35] в качестве функций преобразования данных. При этом предполагается, что наблюдаемая оценка снова отражает скрытую нормально распределенную оценку, но корреляции между элементами могут быть непосредственно оценены без необходимости преобразования данных.

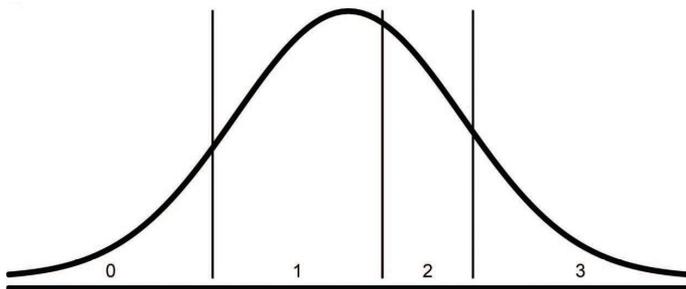


Рис. 3. Визуализация пороговой модели, используемой в полихорических и полисерийных корреляциях

Пример такой пороговой функции показан на рисунке 3, где три порога соответствуют четырем категориям ответа (0, 1, 2 или 3). Нормальное распределение соответствует скрытой оценке предмета, а вертикальные столбцы соответствуют порогам; человек с латентной оценкой ниже первого набрал бы 0, человек с латентной оценкой между первым и вторым порогом набрал бы 1, и так далее. После того, как пороговые значения оценены, корреляции между скрытыми переменными могут быть оценены попарно.

Эти корреляции называются полихорическими корреляциями, когда обе переменные являются порядковыми [36], или полисерийными корреляциями, когда только одна из двух переменных является порядковой [37]. Для вычисления полихорических и полисерийных корреляций в R может использоваться функция `lavCor` из пакета `lavaan` [42]. Эти корреляции затем могут быть использованы в качестве входных данных для алгоритма глассо.

Регуляризация частных корреляций с использованием алгоритма глассо с выбором модели EBIC на основе полихорических корреляций является стандартным выбором при оценке психологических сетей из-за высокой распространенности в этом случае упорядоченных категориальных данных. Важным ограничением является то, что эти методы основаны на предположении, что скрытые переменные, лежащие в основе наблюдаемых порядковых переменных, распределены нормально, что может быть неправдоподобным. Например, некоторые психопатологические симпто-



мы, такие как суицидальные мысли, могут иметь реальную «нулевую точку» – отсутствие симптома. Правильное обращение с такими переменными остается темой будущих исследований.

Когда данные являются двоичными, можно также использовать тетраэрические и бисериальные корреляции (соответственно частные случаи полихорических и полисериальных корреляций). Однако, из-за базового предположения о нормальности, эти данные не всегда стоит обрабатывать с использованием сетей частных корреляций. Когда все переменные являются двоичными, можно оценить модель Изинга с помощью пакета *IsingFit* в R [47]. Результирующая сеть имеет аналогичную интерпретацию, как сети частных корреляций, и также оценивается с использованием лассо с выбором модели EBIC [47]. Когда данные состоят как из категориальных, так и из непрерывных переменных, современную сетевую модель называют смешанной графической моделью, которая реализована в пакете *mgm* также с использованием оценки лассо с выбором модели EBIC.

5. ЗАКЛЮЧЕНИЕ

Настоящая статья посвящена оценке психологических сетей с использованием популярного метода упорядочения оценок связей между узлами сети: регуляризация лассо с выбором модели EBIC. В этом методе могут быть использованы непрерывные или упорядоченные категориальные данные, а связи в сети определяются коэффициентами частных корреляций, что позволяет ограничить в сети число возможных паразитных ребер. Методы, описанные в статье, подходят тогда, когда отдельные значения в данных можно считать независимыми. Представленная сетевая методология приобрела известность в последние годы в зарубежной психологии и используется во все большем количестве публикаций в различных областях психологических исследований в качестве возможной альтернативы структурному моделированию с использованием латентных переменных [2].

В статье подробно обсуждаются частные корреляции, сети частных корреляций и как их следует интерпретировать, а также рассматриваются вопросы упорядочения сетей частных корреляций с использованием регуляризации лассо и расширенного байесовского информационного критерия EBIC. Рассмотрена реализация представленных методов в графическом алгоритме лассо (глассо) в пакетах языка статистического программирования R. Показано как методы регуляризации могут работать в случае, когда данные не имеют нормального распределения.

Использование сетевого моделирования в психологии все еще является молодой областью, которая быстро развивается. Настоящая статья устраняет пробел, относящийся к отсутствию доступного описания на русском языке процедуры регуляризации лассо с использованием выбора модели EBIC, и предоставляет обзор пока мало-распространенного и в тоже время многообещающего метода оценки упорядоченной разреженной сети частных корреляций.



Литература

1. Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А. Бинарные отношения, графы и коллективные решения. М.: ФИЗМАТЛИТ, 2017. 344 с.
2. Артеменков С.Л. Сетевое моделирование психологических конструктов // Моделирование и анализ данных. 2017. № 1. С. 9–28.
3. Артеменков С.Л. Инициционно-семантическая модель дивергентной креативности [Электронный ресурс] // Психологическая наука и образование psyedu.ru. 2012. № 3. С. 1–15. URL: http://psyjournals.ru/psyedu_ru/2012/n3/55540.shtml
4. Жукова Е.С., Артеменков С.Л., Богоявленская Д.Б. К вопросу о соотношении одаренности и осознанной саморегуляции. Личностные и регуляторные ресурсы достижения образовательных и профессиональных целей в эпоху цифровизации. Москва: Знание-М, 2020. С. 104–115. DOI: 10.38006/907345–50–8.2020.104.115.
5. Жукова Е.С., Артеменков С.Л., Богоявленская Д.Б. Исследование интеллектуальной активности в младшем школьном и подростковом возрасте / Моделирование и анализ данных. 2019. № 1. С. 11–29.
6. Мутина О.В. Методы исследования каузальных связей // Экспериментальная психология в России: традиции и перспективы. М.: «Институт психологии РАН», 2010. С. 139–143.
7. Шишляникова Л.М. Применение корреляционного анализа в психологии // Психологическая наука и образование. 2009. Том 14. № 1. С. 98–107.
8. Abacioglu C.S., Isvoranu A.M., Verkuyten M., Thijs J. & Epskamp S. Exploring teachers' influence on student motivation in multicultural classrooms: A comparative network analysis. *Journal of School Psychology*, 2019. 74, 90–105. <https://doi.org/10.1016/j.jsp.2019.02.003>.
9. Agresti A. Categorical data analysis. New York, NY: Wiley, Inc. 1990.
10. Berg J.W., Smid W., Kossakowski J.J., Beek D.V., Borsboom D., Janssen E., & Gijs L. The Application of Network Analysis to Dynamic Risk Factors in Adult Male Sex Offenders. *Clinical Psychological Science*, 2020. 8, 539–554.
11. Bogoyavlenskaya D., Joukova E., Artemenkov S. Longitudinal Study Of The Creative Abilities // The European Proceedings of Social & Behavioural Sciences (EpSBS), 2018. 14: 125–131. doi: <https://dx.doi.org/10.15405/epsbs.2018.11.02.14>
12. Borsboom D. A network theory of mental disorders. *World Psychiatry*, 2017. 16(1), 5–13.
13. Borsboom D., & Cramer A.O.J. Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 2013. 9, 91–121.
14. Borsboom D., Cramer A., & Kalis A. Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, 2019. 42, 1–54. doi:10.1017/S0140525X17002266
15. Chen J., & Chen Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 2008. 95, 759–771.
16. Costantini G., Epskamp S., Borsboom D., Perugini M., Mõttus R., Waldorp L.J., & Cramer A.O.J. State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 2015. 54, 13–29.
17. Costantini G., Richetin J., Emanuele P., Casini E., Epskamp S. & Perugini M. Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, 2019. 136, 68–78.
18. de Ron J, Fried E.I., Epskamp S. Psychological networks in clinical populations: investigating the consequences of Berkson's bias. *Psychological Medicine*, 2019. 1–9. <https://doi.org/10.1017/S0033291719003209>
19. Drton M., & Perlman M.D. Model selection for gaussian concentration graphs. *Biometrika*, 2004. 91, 591–602.



20. *Epskamp S., Borsboom D., Fried E.I.* Estimating psychological networks and their accuracy: a tutorial paper. *Behavior Research Methods*. 2018. 50, 195–212.
21. *Epskamp S., Fried E.I.* A tutorial on regularized partial correlation networks. *Psychological Methods*, 2018. 23(4), 617–634. <https://doi.org/10.1037/met0000167>
22. *Epskamp S., Rhemtulla M., & Borsboom D.* Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 2017. 82, 904–927. <http://dx.doi.org/10.1007/s11336-017-9557-x>
23. *Epskamp S., van Borkulo C.D., van der Veen D.C., Servaas M.N., Isvoranu A.M., Riese H. & Cramer A.O.J.* Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections. *Clinical Psychological Science*, 2018. 6(3), 416–427.
24. *Foygel R., & Drton M.* Extended Bayesian information criteria for Gaussian graphical models. *Advances in Neural Information Processing Systems*, 2010. 23, 2020–2028.
25. *Friedman J.H., Hastie T., & Tibshirani R.* Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 2008. 9, 432–441.
26. *Golino H.F., & Epskamp S.* Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One*, 2017. 12, e0174035.
27. *Hastie T., Tibshirani R., & Wainwright M.* Statistical learning with sparsity: The lasso and generalizations. Boca Raton, FL: CRC Press. 2015.
28. *Isvoranu A-M., Guloksuz S., Epskamp S., van Os J., Borsboom D., GROUP Investigators.* Toward incorporating genetic risk scores into symptom networks of psychosis. *Psychological Medicine*, 2020. 50, 636–643. <https://doi.org/10.1017/S003329171900045X>
29. *Kan K-J., de Jonge H., van der Maas H.L.J., Levine S.Z., & Epskamp S.* How to Compare Psychometric Factor and Network Models. *Journal of Intelligence*, 2020. 8(4), 35.
30. *Koller D., & Friedman N.* Probabilistic graphical models: Principles and techniques. Cambridge, MA: MIT Press. 2009.
31. *Kossakowski J.J., Epskamp S., Kieffer J.M., van Borkulo C.D., Rhemtulla M., & Borsboom D.* The application of a network approach to health-related quality of life (HRQoL): Introducing a new method for assessing HRQoL in healthy adults and cancer patient. *Quality of Life Research*, 2015. 25, 781–792.
32. *Lauritzen S.L.* Graphical models. Oxford, UK: Clarendon Press. 1996.
33. *Liu H., Lafferty J.D., & Wasserman L.* The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *The Journal of Machine Learning Research*, 2009. 10, 2295–2328.
34. *Meinshausen N., & Bühlmann P.* High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, 2006. 34, 1436–1462.
35. *Muthén B.* A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 1984. 49, 115–132.
36. *Olsson U.* Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 1979. 44, 443–460.
37. *Olsson U., Drasgow F., & Dorans N.J.* The polyserial correlation coefficient. *Psychometrika*, 1982. 47, 337–347.
38. *Oreel T.H., Borsboom D., Epskamp S., Hartog I.D., Netjes J.E., Nieuwekerk P.T., Henriques J.P.S., Scherer-Rath M., Van Laarhoven H.W.M. & Sprangers M.A.G.* The dynamics in health-related quality of life of patients with stable coronary artery disease were revealed: a network analysis. *Journal of clinical epidemiology*. 2019. 107, 116–123.
39. *Pearl J.* Causality: Models, reasoning, and inference. Cambridge, UK: Cambridge University Press. 2000.
40. *Pedersen T.L.* ggraph: An Implementation of grammar of graphics for graphs and networks (R package version 1.0.0). 2017. Retrieved from <https://CRAN.R-project.org/package=ggraph>



41. Pourahmadi M. Covariance estimation: The glm and regularization perspectives. *Statistical Science*, 2011. 26, 369–387.
42. Rosseel Y. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 2012. 48(2), 1–36.
43. Schmittmann V.D., Cramer A.O.J., Waldorp L.J., Epskamp S., Kievit R.A., & Borsboom D. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 2013. 31, 43–53.
44. Soutter A.R.B., Bates T.C., Möttus R. Big Five and HEXACO Personality Traits, Proenvironmental Attitudes, and Behaviors: A Meta-Analysis. *Perspect Psychol Sci*. 2020. 15(4), 913–941. doi: 10.1177/1745691620903019.
45. Stevens S.S. On the theory of scales of measurement. *Science, New Series*, 1946. 103, 677–680.
46. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 1996. 58, 267–288.
47. van Borkulo C.D., & Epskamp S. IsingFit: Fitting Ising models using the elasso method. R package version 0.2.0. 2014.
48. Wagenmakers E.-J. A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 2007. 14, 779–804.
49. Wasserman S., & Faust K. *Social network analysis: Methods and applications*. New York, NY: Cambridge University Press. 1994.
50. Zhao P., & Yu B. On model selection consistency of lasso. *The Journal of Machine Learning Research*, 2006. 7, 2541–2563.
51. Zhao T., Li X., Liu H., Roeder K., Lafferty J., & Wasserman L. huge: High-dimensional undirected graph estimation (R package version 1.2.7). 2015. Retrieved from <https://CRAN.R-project.org/package=huge>



Ordered Partial Correlation Networks in Psychological Research

Sergei L. Artemenkov*

Moscow State University of Psychology and Education, Moscow, Russia

ORCID: <https://orcid.org/0000-0002-1619-2209>

e-mail: slart@inbox.ru

Network modeling, which has emerged in recent years, can be successfully applied to the consideration of relationships between measurable psychological variables. In this context, psychological variables are understood as directly affecting each other, and not as a consequence of a latent construct. The article describes regularization methods that can be used to effectively assess the sparse and interpretable network structure based on partial correlations of psychological indicators. An overview of the glasso regularization procedure using EBIC model selection for evaluating an ordered sparse network of partial correlations is presented. The issues of performing this analysis in R in the presence of normal and non-normal data distribution are considered, taking into account the influence of the hyperparameter, which is manually set by the researcher. The considered approach is also interesting as a way to visualize possible causal connections between variables. This review bridges the gap related to the lack of an accessible description in Russian of this approach, which is still uncommon in Russia and at the same time promising.

Keywords: correlation analysis, partial correlation networks, regularization, network modeling in psychology, language R.

For citation:

Artemenkov S.L. Ordered Partial Correlation Networks in Psychological Research. *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2021. Vol. 11, no. 2, pp. 31–50. DOI: <https://doi.org/10.17759/mda.2021110202> (In Russ., abstr. in Engl.).

References

1. Aleskerov F.T., Khabina E.L., Shvarts D.A. Binarnyye otnosheniya, grafy i kollektivnyye resheniya. M.: FIZMATLIT, 2017. 344 s.
2. Artemenkov S.L. Setevoye modelirovaniye psikhologicheskikh konstruktov // Modelirovaniye i analiz dannykh. 2017. № 1. S. 9–28.
3. Artemenkov S.L. Initsionno-semanticheskaya model' divergentnoy kreativnosti [Elektronnyy resurs] // Psikhologicheskaya nauka i obrazovaniye psyedu.ru. 2012. № 3. S. 1–15. URL: http://psyjournals.ru/psyedu_ru/2012/n3/55540.shtml.
4. Zhukova E.S., Artemenkov S.L., Bogoyavlenskaya D.B. K voprosu o sootnoshenii odarennosti i osoznannoy samoregulyatsii. Lichnostnyye i regulyatornyye resursy dostizheniya obrazovatel'nykh i professional'nykh tseyev v epokhu tsifrovizatsii. Moskva: Znaniye-M, 2020. S. 104–115. DOI: 10.38006/907345–50–8.2020.104.115.

***Sergei L. Artemenkov**, PhD in Engineering, Moscow State University of Psychology and Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-1619-2209>, e-mail: slart@inbox.ru



5. Zhukova E.S., Artemenkov S.L., Bogoyavlenskaya D.B. Issledovaniye intellektual'noy aktivnosti v mladshem shkol'nom i podrostkovom vuzraste / Modelirovaniye i analiz dannykh. 2019. № 1. S. 11–29.
6. Mitina O.V. Metody issledovaniya kauzal'nykh svyazey // Eksperimental'naya psikhologiya v Rossii: traditsii i perspektivy. M.: «Institut psikhologii RAN», 2010. S. 139–143.
7. Shishlyannikova L.M. Primeneniye korrelyatsionnogo analiza v psikhologii // Psikhologicheskaya nauka i obrazovaniye. 2009. Tom 14. № 1. S. 98–107.
8. Abacioglu C.S., Isvoranu A.M., Verkuyten M., Thijs J. & Epskamp S. Exploring teachers' influence on student motivation in multicultural classrooms: A comparative network analysis. *Journal of School Psychology*, 2019. 74, 90–105. <https://doi.org/10.1016/j.jsp.2019.02.003>.
9. Agresti A. *Categorical data analysis*. New York, NY: Wiley, Inc. 1990.
10. Berg J.W., Smid W., Kossakowski J.J., Beek D.V., Borsboom D., Janssen E., & Gijs L. The Application of Network Analysis to Dynamic Risk Factors in Adult Male Sex Offenders. *Clinical Psychological Science*, 2020. 8, 539–554.
11. Bogoyavlenskaya D., Joukova E., Artemenkov S. Longitudinal Study Of The Creative Abilities // The European Proceedings of Social & Behavioural Sciences (EpSBS), 2018. 14: 125–131. doi: <https://dx.doi.org/10.15405/epsbs.2018.11.02.14>.
12. Borsboom D. A network theory of mental disorders. *World Psychiatry*, 2017. 16(1), 5–13.
13. Borsboom D., & Cramer A.O.J. Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 2013. 9, 91–121.
14. Borsboom D., Cramer A., & Kalis A. Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, 2019. 42, 1–54. doi:10.1017/S0140525X17002266
15. Chen J., & Chen Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 2008. 95, 759–771.
16. Costantini G., Epskamp S., Borsboom D., Perugini M., Möttus R., Waldorp L.J., & Cramer A.O.J. State of the aRT personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 2015. 54, 13–29.
17. Costantini G., Richetin J., Emanuele P., Casini E., Epskamp S. & Perugini M. Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, 2019. 136, 68–78.
18. de Ron J, Fried E.I., Epskamp S. Psychological networks in clinical populations: investigating the consequences of Berkson's bias. *Psychological Medicine*, 2019. 1–9. <https://doi.org/10.1017/S0033291719003209>
19. Drton M., & Perlman M.D. Model selection for gaussian concentration graphs. *Biometrika*, 2004. 91, 591–602.
20. Epskamp S., Borsboom D., Fried E.I. Estimating psychological networks and their accuracy: a tutorial paper. *Behavior Research Methods*. 2018. 50, 195–212.
21. Epskamp S., Fried E.I. A tutorial on regularized partial correlation networks. *Psychological Methods*, 2018. 23(4), 617–634. <https://doi.org/10.1037/met0000167>.
22. Epskamp S., Rhemtula M., & Borsboom D. Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 2017. 82, 904–927. <http://dx.doi.org/10.1007/s11336-017-9557-x>
23. Epskamp S., van Borkulo C.D., van der Veen D.C., Servaas M.N., Isvoranu A.M., Riese H. & Cramer A.O.J. Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections. *Clinical Psychological Science*, 2018. 6(3), 416–427.
24. Foygel R., & Drton M. Extended Bayesian information criteria for Gaussian graphical models. *Advances in Neural Information Processing Systems*, 2010. 23, 2020–2028.



25. Friedman J.H., Hastie T., & Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 2008. 9, 432–441.
26. Golino H.F., & Epskamp S. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One*, 2017. 12, e0174035.
27. Hastie T., Tibshirani R., & Wainwright M. *Statistical learning with sparsity: The lasso and generalizations*. Boca Raton, FL: CRC Press. 2015.
28. Isvoranu A-M., Guloksuz S., Epskamp S., van Os J., Borsboom D., GROUP Investigators. Toward incorporating genetic risk scores into symptom networks of psychosis. *Psychological Medicine*, 2020. 50, 636–643. <https://doi.org/10.1017/S003329171900045X>
29. Kan K-J., de Jonge H., van der Maas H.L.J., Levine S.Z., & Epskamp S. How to Compare Psychometric Factor and Network Models. *Journal of Intelligence*, 2020. 8(4), 35.
30. Koller D., & Friedman N. *Probabilistic graphical models: Principles and techniques*. Cambridge, MA: MIT Press. 2009.
31. Kossakowski J.J., Epskamp S., Kieffer J.M., van Borkulo C.D., Rhemtulla M., & Borsboom D. The application of a network approach to health-related quality of life (HRQoL): Introducing a new method for assessing HRQoL in healthy adults and cancer patient. *Quality of Life Research*, 2015. 25, 781–792.
32. Lauritzen S.L. *Graphical models*. Oxford, UK: Clarendon Press. 1996.
33. Liu H., Lafferty J.D., & Wasserman L. The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *The Journal of Machine Learning Research*, 2009. 10, 2295–2328.
34. Meinshausen N., & Bühlmann P. High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, 2006. 34, 1436–1462.
35. Muthén B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 1984. 49, 115–132.
36. Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 1979. 44, 443–460.
37. Olsson U., Drasgow F., & Dorans N.J. The polyserial correlation coefficient. *Psychometrika*, 1982. 47, 337–347.
38. Oreel T.H., Borsboom D., Epskamp S., Hartog I.D., Netjes J.E., Niewekerker P.T., Henriques J.P.S., Scherer-Rath M., Van Laarhoven H.W.M. & Sprangers M.A.G. The dynamics in health-related quality of life of patients with stable coronary artery disease were revealed: a network analysis. *Journal of clinical epidemiology*. 2019. 107, 116–123.
39. Pearl J. *Causality: Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press. 2000.
40. Pedersen T.L. *ggraph: An Implementation of grammar of graphics for graphs and networks (R package version 1.0.0)*. 2017. Retrieved from <https://CRAN.R-project.org/package=ggraph>
41. Pourahmadi M. Covariance estimation: The glm and regularization perspectives. *Statistical Science*, 2011. 26, 369–387.
42. Rosseel Y. *lavaan: An R package for structural equation modeling*. *Journal of Statistical Software*, 2012. 48(2), 1–36.
43. Schmittmann V.D., Cramer A.O.J., Waldorp L.J., Epskamp S., Kievit R.A., & Borsboom D. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 2013. 31, 43–53.
44. Soutter A.R.B., Bates T.C., Möttus R. Big Five and HEXACO Personality Traits, Proenvironmental Attitudes, and Behaviors: A Meta-Analysis. *Perspect Psychol Sci*. 2020. 15(4), 913–941. doi: 10.1177/1745691620903019.



45. Stevens S.S. On the theory of scales of measurement. *Science, New Series*, 1946. 103, 677–680.
46. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 1996. 58, 267–288.
47. van Borkulo C.D., & Epskamp S. IsingFit: Fitting Ising models using the elasso method. R package version 0.2.0. 2014.
48. Wagenmakers E.-J. A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 2007. 14, 779–804.
49. Wasserman S., & Faust K. *Social network analysis: Methods and applications*. New York, NY: Cambridge University Press. 1994.
50. Zhao P., & Yu B. On model selection consistency of lasso. *The Journal of Machine Learning Research*, 2006. 7, 2541–2563.
51. Zhao T., Li X., Liu H., Roeder K., Lafferty J., & Wasserman L. huge: High-dimensional undirected graph estimation (R package version 1.2.7). 2015. Retrieved from <https://CRAN.R-project.org/package=huge>