УДК 159.9:51

ОСОБЕННОСТИ ФАКТОРНОГО АНАЛИЗА В ПСИХОЛОГИИ

Р.Х. Тугушев

Саратовский государственный университет,
кафедра психологии
E-mail: tugushevrh@info.sgu.ru

Факторный анализ, рожденный в рамках решения психологических проблем, стал популярным во многих науках, имеющих дело со статистическим представлением экспериментальных данных. Сложность и неоднозначность концептуальных подходов методики затрудняют его применение, особенно в работах молодых ученых. Нами была поставлена задача максимально прояснить особенности трудных мест факторного анализа без потери математической точности.
Рекомендации по использованию факторного анализа, с учетом необходимых ограничений и разъяснением причин возможной неоднозначности решений, будут полезны всем, кто заинтересован в усвоении адекватного математического инструментария при решении задач гуманитарного плана, в частности, практической психологии.

Peculiarities of Factor Analysis in Psychology

R.H. Tugushev

Factor analysis appeared in the framework of resolving psychological problems has become popular in many sciences with statistical representation of experimental data. The complexity and ambiguity of its conceptual approaches make its use difficult, especially for young scientists. We have posed a task to clarify special features of difficult points in factor analysis as much as possible without any loss of mathematical accuracy.
Our recommendations for the use of factor analysis, with necessary restrictions and clear reasons of a possible ambiguity of solutions, will be useful to everyone interested in mastering an adequate mathematical tool for solving problems pertaining to the humanities, in particular, those of practical psychology.

Наука является процессом, состоящим из непрерывного цикла построения моделей наблюдаемой действительности, которая вначале кажется необозримой в своем многообразии и поэтому не может быть охвачена подобной моделью. В естественных науках наряду с накапливанием наблюдений происходит формирование теоретических положений, которые проверяются затем на новых данных и постепенно совершенствуются. В последнем столетии наблюдается значительное дифференцирование методов мышления и подходов к решению естественно-научных задач, что обусловлено большим числом проведенных исследований и появлением общих и специальных научных методик.

В разделах знаний, занимающихся изучением поведения человека, к которым относятся психология, социология, частично – медицина и многие смежные области других отраслей знаний, где отсутствует четкая функциональная связь между причиной и следствием, один опыт не играет решающей роли. Присущие гуманитарным явлениям закономерности можно обнаружить лишь путем статистической проверки результатов многих опытов, принимая или отклоняя определенные гипотезы. Важную роль при этом играет так называемый фактор.

Это понятие очень многогранно. Часто говорят о факторе риска, экономическом, политическом, психологическом факторах и др. Факторные теории возникли и в психологии. Их истоки в работах Терстона, Спирмена, Кеттелла и ряда других зарубежных ученых[1]. Впоследствии в разработку теории факторного анализа в его современной форме внесли свой вклад Г. Харман, К. Иберла, А.Г. Шмелев и др.[2]

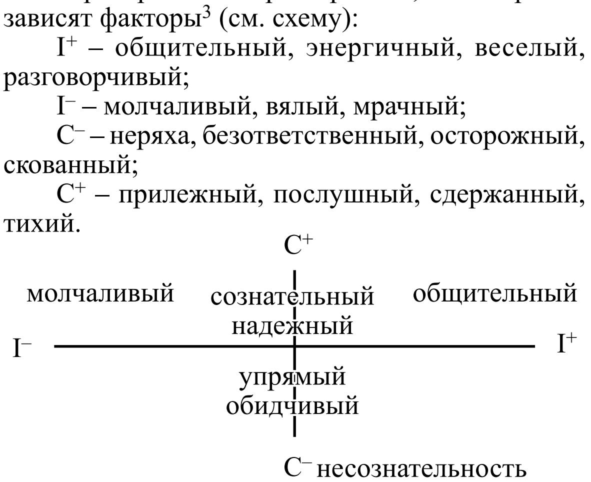
Качественный смысл фактора заключается в том, что определяется некоторая ненаблюдаемая и прямо не измеряемая величина, категория, связанная с множеством близких к ней характеристик, которые могут быть измерены. Например, ищутся факторы сознательности и экстраверсии, с помощью измеренных характеристик искомых факторов C и I. Каждый из них имеет два полюса: I^+ – максимальная выраженность интроверсии, I^- – минимальная интраверсия (и, как следствие, максимальная экстраверсия). Фактор сознательности, обозначенный через C, так же имеет два полюса. Представляя эти факторы в плоскости независимых (ортогональных) координат, получаем пространство характеристик, от которых и зависят факторы[3] (см. схему):

I^+ – общительный, энергичный, веселый, разговорчивый;

I^- – молчаливый, вялый, мрачный;

C^- – неряха, безответственный, осторожный, скованный;

C^+ – прилежный, послушный, сдержанный, тихий.



Представленный пример достаточно прост, но, в целом, проблема взаимосвязи искомых факторов весьма сложна. Как правило, факторы вбирают в себя множество характеристик, при этом неизвестно, какие это факторы, сколько их, каков их удельный вес в общих оценках, ортогональны они или косоугольны и т.д. При этом, забегая вперед, можно сказать, что проблема фактора сходна с поставленной в сказке задачей: «пойди туда, не знаю куда, принеси то, не знаю что». Что касается выяснения направления, то тут кое-что сделано, однако в распространенных методиках недостает единственности и однозначности решений. Это приводит к ряду противоречий в анализе экспериментальных данных.

Даже в такой, казалось бы, простой задаче по измерению интеллекта, нет однозначности и согласия в его определении, хотя каждый человек на бытовом уровне оценивает интеллект тех, с кем общается. Психологические методы оценки с привлечением тестов могут в некоторой степени усовершенствовать эти оценки. Тесты измеряют в единицах формализованных шкал ряд характеристик: умение считать, логически мыслить, понимать связи и отношения, анализировать, сравнивать, исключать, делать обоснованные выводы и ряд других интеллектуальных операций. Предполагается, что фактор интеллекта находится между свойствами измеренных значений, при этом берутся в расчет и образование, и культура. Но, несмотря на тщательно отобранный перечень характеристик, их измерение и статистическую обработку, нельзя утверждать, что таким образом удалось измерить сам интеллект как фактор. Максимум, чего удается добиться – получить фактор условного интеллекта – IQ – как тестовый показатель.

Тесты способностей, как правило, характеризуются положительными корреляциями. Ч. Спирмен[4] предположил, что все возможные совокупности корреляций обусловливаются одним генеральным фактором – **g**, влияющим на все переменные. Этот генеральный фактор совпадает с тем, что в обиходе называют смышленностью. После его психометрической изоляции можно определить, какие методы измерения более всего пригодны для его оценки, т.е. вычислить нагрузки этого фактора по различным тестам.

С помощью факторного анализа было показано, что интеллектуальные возможности человека можно представить в виде эмпирических показателей, а не только описательно. Эти показатели имеют количественные оценки. Понятие интеллекта заменяется рядом факторов, чьи отображения устанавливаются с помощью методов факторного анализа. На основе психологических тестов интеллектуальных возможностей человека можно выделить набор факторов, которые, большей частью, коррелированы положительно, например, факторы пространственного воображения, вербального восприятия и т.д. С помощью факторного анализа эмпирическим путем проверялись различные психологические теории интеллектуальных возможностей и умственных способностей человека. Такие исследования нашли практическое применение при оказании помощи в выборе профессии, в психотерапии и т.д. Факторный анализ (ФА) играет важнейшую роль во многих психологических исследованиях. Он делает возможным сведение обширного числового материала к нескольким независимым и простым факторам, его методами можно подтвердить существующую гипотезу на основе большого объема наблюдений по выделенным существенным компонентам.

ФА следует рассматривать как статистический метод, независимо от конкретной области его применения. Особую ценность представляет возможность генеза гипотез и их проверки.

Переходя к общей схеме математических методов реализации факторного анализа, сразу заметим, что поскольку нет однозначных решений в приводимых процедурах и примерах, взятых из наиболее известных источников[5], мы встречаемся с проблемой «найти то, не знаю что», поэтому примеры используют прием движения от известного результата к неизвестным характеристикам. Это будет видно из дальнейшего, но сначала нужно определиться с основными формулами ФА.

Введем обозначения, принятые в ФА:

$S_j = \sqrt{\sum_{i=1}^{n} (x_{ij} - \overline{x}_j)^2} = s_j$ – стандартное отклонение,

$$S_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (x_{ij} - \overline{x}_j)(x_{ik} - \overline{x}_k) \text{ – ковариация,}$$

как коэффициент ненормированной корреляции между j и k – столбцами матрицы $x=\{x_{ij}\}$, $j \neq k$. В этих обозначениях, как легко проверить,

$r_{jk} = \dfrac{S_{jk}}{S_j S_k}$, где $S_k = \sigma_k \bullet r_{jk}$ – корреляция стандартного отклонения.

Вводится стандартизация переменных:

$$z_{ij} = \frac{x_{ij} - x_{ij}}{S_j}, \quad z_{ik} = \frac{x_{ik} - \overline{x}_k}{S_k}.$$

Для таких переменных справедливы формулы:

$$\frac{1}{n} \sum_{i=1}^{n} z_{ij} = 0, \quad \frac{1}{n-1} \sum_{i=1}^{n} z_{ij}^2 = 1.$$

Исходя из вышеприведенных соотношений, имеем:

$$r_{jk} = S_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} z_{ij} z_{ik}, \quad z_{ij} = z_{ji}.$$

В матричном виде $\overline{Z}$ – матрица стандартизированных переменных $\{z_{ij}\}$, $\overline{Z}'$ – транспонированная матрица $\overline{Z}$ $\{\overline{z}_{jk}\}$ и $\overline{R}$ – треугольная матрица интеркорреляций $r_{jk}=r_{kj}$. Исходя из этого, получим:

$$\bar{R} = \frac{1}{n-1}\bar{Z}\bar{Z}' = \bar{S}.$$

Для дальнейшего необходимо обратиться к основной формуле ФА, которая выглядит так:

$$z_{ij}=a_{i1}P_{1j}+a_{i2}P_{2j}+\dots a_{ir}P_{rj},$$

где P_{rj} – r факторов по j-й переменной (r-индекс), a_{ir} – величина нагрузок факторов. Становится ясно, что основная формула ФА приводится к задаче поиска неизвестных, число которых больше числа уравнений их определяющих!

Дальше будет видно, как теоретики ФА с этим справляются, а пока перепишем вышеприведенные формулы в новых обозначениях:

$\bar{Z} = \bar{A}\bar{P}$, подставляя $\bar{Z}$ в формулу корреляции, получим:

$$R = \frac{1}{n-1}ZZ' = \frac{1}{n-1}AP(AP)' = \frac{1}{n-1}APP'A' =$$

$$= A(\frac{1}{n-1}PP')A'.$$

Из формулы для корреляции имеем

$$C = \frac{1}{n-1}PP',$$

тогда $R = ACA'$, C полагается равным единичной матрице I, поскольку рассматриваются некоррелирующие факторы. Тогда $R = AA'$.

Исходя из этой формулы строим схему ФА (рис. 1).

Анализ начинается со сбора экспертных данных в виде матрицы $X(m, n)$ и затем производится корреляц ионный анализ R полученной матрицы X.

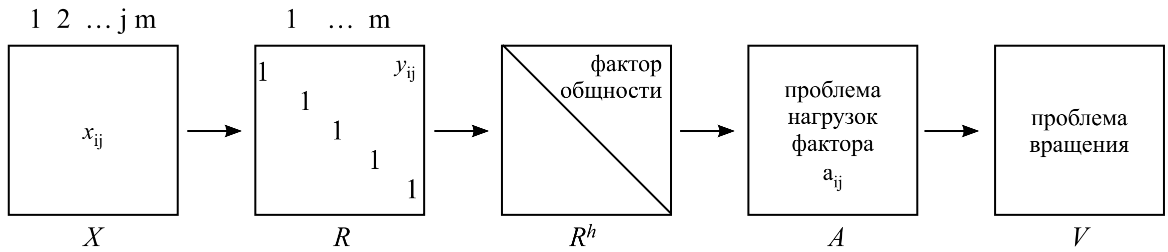


Рис.1. Общая схема факторного анализа

Особая проблема – интерпретация фактора и его название.

В определенной степени, процедура факторного анализа обратима. Так, если в правой части равенства (простого числового примера) находится A, то в левой части обязательно получим R: $R = A \cdot A'$, где A – матрица факторных нагрузок, A' – транспонированная матрица A. Иллюстрация этого уравнения в матричной формуле дана в работе К. Иберла:

$$\underset{A}{\begin{pmatrix} 0,90 & 0,10 \\ 0,80 & 0,05 \\ 0,70 & 0,10 \\ 0,05 & 0,80 \\ 0,10 & 0,70 \\ 0,05 & 0,50 \end{pmatrix}} \cdot \underset{A'}{\begin{pmatrix} 0,90 & 0,80 & 0,70 & 0,05 & 0,10 & 0,05 \\ 0,10 & 0,05 & 0,10 & 0,80 & 0,70 & 0,50 \end{pmatrix}} =$$

$$R^h$$

$$= \begin{pmatrix} (0,8200) & - & - & - & - & - \\ 0,7250 & (0,6425) & - & - & - & - \\ 0,6400 & 0,5650 & (0,5000) & - & - & - \\ 0,1250 & 0,0800 & 0,1150 & (0,6425) & - & - \\ 0,1600 & 0,1150 & 0,1400 & 0,5650 & (0,5000) & - \\ 0,0950 & 0,0650 & 0,0850 & 0,4025 & 0,3550 & (0,2525) \end{pmatrix}.$$

Числа в примере К. Иберла[5] были подобраны так, чтобы можно было легко произвести вычисления. В этом и заключается главная трудность, так как непонятно, как подобрать элементы матриц. Но если известны A и A', то первый элемент матрицы R получается следующим образом:

(0,90·0,90) + 0,01 = (0,82). Это обычное умножение матриц.

Значение первого диагонального элемента заключается в скобки и называется *общностью*. Это понятие мы будем еще обсуждать более детально. Второй элемент первого столбца равен: (-0,80·0,90) – (0,10·0,10) = –0,73. Совершенно аналогично получаются другие элементы матрицы R путем перемножения матриц A и A' по соответствующему правилу. В факторном анализе эта процедура обратима. Может быть задана корреляционная матрица R и по ней находится A. Но прежде чем это удастся сделать, надо выделить редуцированную матрицу $\{r^h\}$, которая позволит найти матрицу A, т.е. решить проблему общности. Реально, в начале схемы мы имеем корреляционную матрицу R, в диагоналях которой записаны единицы, в то время как в матрице R^h – диагональные элементы, квадраты неизвестных нагрузок $\{a_{il}\}$. Задачей факторного анализа и является определение матрицы A.

Матрица A называется *факторным отображением*, а ее элементы a_{il} – *факторными нагрузками*. При ортогональных факторах элементы принимают значения между –1 и +1. Если факторы не ортогональные, то элементы могут принимать большие значения. Каждый фактор характеризуется столбцом, каждая переменная – строкой матрицы A. Если факторная нагрузка значительно больше или меньше нуля, то принята упрощенная форма записи в виде крестика (х) в соответствующем месте факторного отображения.

Таблица 1

Схематическое изображение факторного отображения

Характеристика	A	B	C		U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8
1	X	X			X							
2	X	X				X						
3	X	X					X					
4	X	X	X					X				
5	X	X	X						X			
6	X		X							X		
7	X		X								X	
8	X		X									X
↑	Общие факторы				Характерные факторы							
	Генеральный фактор											

Фактор называется *генеральным*, если все его нагрузки значительно отличаются от нуля. Следовательно, он имеет нагрузки от всех переменных и схематически такой фактор изображается столбцом А в табл. 1. Фактор называется *общим*, если хотя бы две его нагрузки значительно отличаются от нуля. Столбцы А, В и С (см. табл. 1) представляют такие общие факторы. Они имеют нагрузки от двух и более переменных могут взаимно перекрываться, т.е. одни и те же переменные могут давать нагрузки на несколько факторов. Генеральный фактор является частным случаем общих факторов, так как он тоже имеет более двух значимых нагрузок. В противоположность этому, факторы являются индивидуальными, если у них только одна нагрузка значительно отличается от нуля (столбцы U_1–U_8 в табл. 1). В этом случае говорят о *характерных факторах*, которые представляют только одну переменную. По аналогии с факторами можно провести классификацию переменных по числу достаточно высоких нагрузок. Число высоких нагрузок переменной на общие факторы называется ее *сложностью*. Например, переменная 1 (рис. 2) имеет сложность два, а переменная 4 – три.

Решающее значение в факторном отображении (см. рис. 2) имеют общие факторы А, В, С, так как характерные факторы получаются автоматически, если общие факторы установлены. Гипотеза, которая содержит факторное отображение в табл. 1, может быть представлена в другой форме (см. рис. 2). Становится ясно, что одинаковые характеристики могут быть связаны с различными факторами.

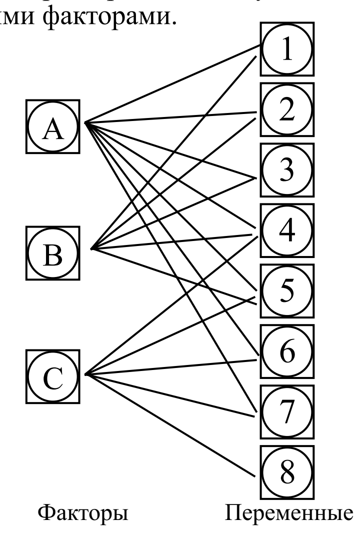


Рис. 2. Пример связей характеристик

Любая матрица, которая содержит общие факторы типа A, B, C, отражает дифференцированные гипотезы о структуре величин, которые, отчасти, стоят за наблюдаемыми переменными (т.е. являются факторами), а отчасти – сами и переменными.

При любом факторном отображении легко перейти от одного способа изображения к другому. Общие факторы для отличия их от переменных обозначаются здесь буквами A, B и C, однако чаще всего они обозначаются римскими цифрами. Поскольку статистических данных в формулах ФА недостаточно, ищутся внестатистические критерии.

Нагрузки общих и характерных факторов связаны определенным соотношением через единичную дисперсию переменных. Действительно, эта единичная дисперсия равна:

$$S_i^2 = 1 = \frac{1}{n-1}\sum_{j=1}^{n} {}_i z_j^2 \text{ для всех } i.$$

Факторы должны быть стандартизованы и некоррелированы, тогда

$$s_i^2 = a_{i1}^2 + a_{i2}^2 + \ldots + a_{ig}^2 = 1. \quad (1)$$

Это равенство выполняется при условии, что переменные стандартизованы. Соблюдение условий стандартизации и некоррелированности является необходимым требованием при использовании наиболее употребляемых сегодня методов. Приведенное выше равенство можно записать в модифицированном виде, где в скобках сначала приведены нагрузки a_r общих факторов, а затем – доля дисперсии характерного фактора. Равенство (1) отличается от последующего лишь последними двумя членами:

$$S_i^2 = 1 = (a_{i1}^2 + a_{i2}^2 + \ldots + a_{ir}^2) + b_i^2 + e_i^2. \quad (2)$$

Полная дисперсия переменной по равенству, только что приведенному, раскладывается на отдельные компоненты, которые представляют собой квадраты факторных нагрузок. Суммы квадратов нагрузок общих факторов называются *общностью* h_i^2. Это та доля дисперсии, которая указана в последнем равенстве:

$$h_i^2 = a_{i1}^2 + \ldots + a_{ir}^2. \quad (3)$$

Общность представляет собой часть единичной дисперсии переменной, которую можно приписать общим факторам. Она равна квадрату коэффициента множественной корреляции между переменной и общими факторами. Если из 1 вычесть h_i^2, то останется доля дисперсии, обозначаемая u_i^2, которая соответствует квадрату нагрузки определенного характерного фактора. Часть дисперсии u_i^2 называется *характерностью*. Она представляет собой часть единичной дисперсии переменной, которая не связана с общими факторами:

$$U_i^2 = 1 - h_i^2 = b_i^2 + e_i^2. \quad (4)$$

Как показывает вторая половина равенства (4), характерность U_i^2 можно разбить на две составляющие, одна из которых, b_i^2, называется *специфичностью*, а другая, e_i^2, является дисперсией, *обусловленной ошибкой*. Такое разделение на практике проводится редко. Специфичность bi2 является той долей единичной дисперсии переменной, которая не связана с общими факторами, не может быть также сведена к ошибке и присуща лишь одной определенной переменной. Специфичность и общность образуют надежность r_{ii}^2 ($\geq h_i^2$).

Надежность, являясь долей единичной дисперсии, дополняет дисперсию ошибки до единицы. Она может быть измерена различными способами:

$$r_{ii}^2 = h_i^2 + b_i^2 = 1 - e_i^2, \text{ т.е. } r_{ii}^2 \geq h_i^2. \quad (5)$$

Отсюда следует, что общность не превышает надежности и равна ей только в случае нулевой специфичности.

Теперь поставим задачу – получить общность и соответствующий характерный фактор из корреляционной матрицы. Основная модель факторного анализа записывается следующим образом: $Z = AP$. В этой модели не проводится разницы между общими и характерными факторами. В нижеследующем равенстве постулируется для каждой переменной характерный фактор. Эта специальная модель многофакторного метода имеет следующий вид:

$$Z_{ij} = a_{i1}p_{1j} + \ldots + a_{ir}p_{rj} + u_i p(r+i)j, \quad (6)$$

или в матричной форме: $Z = FP^+$.

При этом Z является матрицей исходных данных, записанных в стандартной форме. F – факторное отображение, включающее характерные факторы, т.е. F является матрицей порядка $m \cdot (r + m)$. P^+ имеет размер $(r + m) \cdot n$ и содержит значения факторов у отдельных индивидуумов, включая значения характерных факторов.

F можно представить в виде суммы двух матриц, а именно $A + U$. F представляет собой так называемую *полную факторную матрицу*. Она содержит нагрузки общих и характерных факторов. A является редуцированной факторной матрицей, содержащей нагрузки общих факторов. Матрица U является диагональной и содержит нагрузки характеристик факторов на главной диагонали.

Для коэффициентов корреляции между переменными и факторами имеет место соотношение:

$$r_{i1} = a_{i1}c_{1l} + a_{i2}c_{2l} + \ldots + a_{ir}c_{rl}. \quad (7)$$

В случае ортогональных факторов факторное отображение идентично структуре: $r_{il} = r_{li}$. Тогда в правой строке равенства (6) остается только один член, а именно тот, множителем которого является $c_u = 1$, так как $c_{il} = 0$ при всех $i \neq l$.

Все представленные до сих пор равенства и определения содержат важнейшие концепции факторного анализа, а также основную и специальную модели многофакторного анализа. С их помощью теперь можно описать общую процедуру решения. При проведении факторного анализа все расчеты носят последовательный характер.

Проблема общности состоит в установлении оценок h_i^2. Это самая первая проблема, которая возникает в ходе факторного анализа. Из R_h с помощью определенных способов извлекают факторы, получая в результате матрицу A. Столбцы матрицы A ортогональны и занимают произвольную позицию в отношении переменных, определяемую методом выделения факторов. Возможно большое число матриц A, которые будут одинаково хорошо воспроизводить R_h. Выбор единственной матрицы составляет *проблему вращения*. Решение проблемы вращения одним из нескольких способов приводит к матрице $V(v_{il}$ – факторная матрица после поворота, $l = 1,2, \ldots, n$ – факторы; $i = 1,2, \ldots, m$ – переменные). И, наконец, последняя проблема касается *оценки значений факторов* для каждого индивидуума.

Отметим, что факторный анализ не всегда выполним, так как для его осуществления есть определенные условия. Во-первых, надо определить число факторов. Здесь могут быть варианты: из них нужно выбрать такой, который базируется на теории, здравом смысле и прошлом опыте:

1) нельзя факторизовать качественные переменные;
2) все переменные должны быть независимыми, распределение близкое к нормальному;
3) все связи имеют приблизительно линейный характер;
4) в исходных корреляционных матрицах должно быть несколько корреляций, не меньше 0,3;
5) выборка должна быть достаточно большой.

Затем нужно установить:

1) не способствует ли увеличение числа факторов уменьшению доли нагрузок в диапазоне ±0,4. Если это так, то увеличение, скорее всего, не имеет смысла;
2) нет ли больших корреляций между факторами при осуществлении вращений, т.е. присутствует слишком много факторов и два из них проходят через общий кластер переменных;
3) не разделились ли какие-либо хорошо известные факторы на две или большее количество частей, т.е. возникло слишком много факторов.

Обычно ограничиваются первыми m факторами, дающими 90–95% дисперсии в сумме.

Геометрическая иллюстрация алгебраических зависимостей факторного анализа облегчает усвоение отдельных проблем. Геометрическое представление обладает рядом важных преимуществ при знакомстве с математической стороной факторного анализа, но при этом пространственную модель можно строить, учитывая не более трех измерений. При выделении четырех и более факторов, что чаще всего и встречается на практике, невозможность наглядного пространственного изображения не имеет решающего значения. В этом случае при численном решении используются чисто алгебраические приемы матричного исчисления.

Иногда удобнее пользоваться геометрическими приемами отображения матричных данных. В этом случае каждый индивидуум обозначен точкой в двухмерной системе координат, причем по осям координат откладываются значения переменных. В общем случае наблюдается более двух переменных, тогда исходные данные задаются в виде матрицы. Ниже приведен пример такой матрицы только для трех лиц (табл. 2).

Таблица 2

Матрица, отображающая расположение личностей в координатах характеристик

Личности / Переменные	A	B	C
1	7,3	3,0	0,0
2	3,4	3,1	0,9
3	0,6	5,3	5,4
4	0,8	-1,5	2,7
5	6,5	-2,5	1,0

Данная матрица может быть отражена графически. Каждая диаграмма будет содержать только три точки, так как мы располагаем данными только трех лиц. Точки при этом соответствуют лицам, т.е. столбцам таблицы. Координатные оси каждый раз соответствуют двум строкам. При подобном графическом изображении для матрицы исходных данных потребовалось бы несколько корреляционных диаграмм, а именно столько, сколько имеется коэффициентов корреляции, т.е. $(m/2)\cdot(m-1)$. Очевидно, что такой способ геометрического представления не очень удобен из-за большого числа графиков. Отказываясь от двухмерного изображения, можно пойти по одному из следующих путей.

Во-первых, можно столбцы таблицы представить в виде трех точек в пятимерном пространстве. Обычные корреляционные диаграммы являются тогда проекциями этого пространства на соответствующие плоскости.

Во-вторых, можно пять переменных, или строки таблицы представить в виде пяти точек в трехмерном пространстве. Три лица здесь соответствуют координатным осям. Вся информация этой таблицы переходит в рисунок. Каждая переменная представляется вектором или стрелкой, координаты концов которых берутся из таблицы. Итак, матрицу исходных данных можно рассматривать как *n*-мерное пространство тестов, в котором находится *m* точек – переменных (пример таких отображений будет показан при анализе проблемы вращений на рис. 3–5). Главная задача факторного анализа в психологических исследованиях – извлечение фундаментальных факторов личности.

Для иллюстрации изложенных нами выше положений представляется целесообразным воспользоваться некоторыми примерами, полученными в работах названных авторов. Вначале рассмотрим пример факторизации психических типов в их связи с конституциональной теорией В. Шелдона[7]. У Шелдона коэффициенты корреляции между соответствующими физическими и психологическими типами равнялись, соответственно: «эндоморфизм» – «висцеротония» – $r_{эв}= +0,79$, «эктоморфизм» – «церебротония» – $r_{экц}= +0,63$, «мезоморфизм» – «соматотония» – $r_{м.с}= +0,82$.

Таким образом, это факторы типов темперамента с высокой нагрузкой. Интересно, что у других исследователей (например у Дж. Гилфорда) эти нагрузки значительно меньше. Это пример того, что данные факторного анализа у разных авторов могут быть неоднозначны.

Большинство личностных методик определяет черты и свойства личности с помощью проекции ответов испытуемого на специально сконструированные вопросы. В ключи шкал входит много переменных, имеющих незначительную причастность к измеряемому свойству. Например, шкала шизофрении (№181 в MMPI)[8] имеет 32 вопроса, ориентированных на выявление этого отклонения, среди которых есть такие: «моя половая жизнь удовлетворительна», «вся пища мне кажется одинаковой на вкус», «я бы хотел избавиться от мыслей, связанных с половым вопросом, так как они меня беспокоят» и т.п. Заметим сразу, что подобные вопросы, как правило, не выявляют шизоидность, т.е. в оригинале MMPI сами шкалы если и выступают в виде факторов, нагрузка на которые дает смысловые согласия с вопросами, большинство этих нагрузок близко к нулю. Более точный анализ показывает, что для шизофрении как фактора большие значения нагрузок дают положительные ответы на вопросы, представленные в табл. 3.

Таблица 3

Скорректированные вопросы к шкале шизофрении по MMPI

№ п/п	Вместо №	Вопрос
1	179	Анализировать собственные мысли привычнее, чем вести беседу с кем-либо
2	202	Окружающий мир слишком беден и неинтересен, по сравнению с миром моего воображения
3	210	Меня иногда обвиняют в эмоциональной тупости
4	212	Глупо переживать о несчастьях других людей, да и своих тоже
5	238	Недавно начал чувствовать какое-то недомогание
6	259	Ничего не хочется делать
7	266	Часто все окружающее вызывает неприязнь
8	282	Бывает, что у меня быстро меняется симпатия на антипатию и наоборот
9	297	Меня часто интересуют проблемы, которые кажутся окружающим неразрешимыми
10	301	Часто замечаю, что могу мыслить одновременно на разные темы
11	303	Из-за безразличия ко всему хочется подолгу не вставать с постели
12	307	Иногда у меня возникают агрессивные чувства к своим близким
13		Часто хочется сделать все наоборот
14	320	Бывает, что при беседе мои ответы на вопросы оценивают как бы «не по существу»
15	339	Объекты окружающего мира легче классифицировать не по смыслу, а по выбранной символике

Ответы «Да» суммируются в баллах шкалы шизофрении. Сюда можно было бы добавить еще вопросы, относящиеся к катастрофической боязни больного темноты и фантастических сущностей. Известный литературный персонаж М.А. Булгакова Мастер заболел этой «странной» болезнью и главное, что он чувствовал – всеразрушающий страх. Этот симптом и есть сигнал начала распада личности.

Данный пример приведен для того, чтобы показать, что фактор, косвенно определяемый различными характеристиками, будет тем более точен, чем ближе эти характеристики друг к другу и «невидимому» фактору.

Р. Кеттелл с последователями и учениками очень много сделал для факторного анализа в психологии. Результаты его исследований представляют значительный интерес, несмотря на некоторую противоречивость. Каждый из его 16 факторов дихотомичен, у него есть положительный и отрицательный полюсы. Для некоторых факторов Р. Кеттелл сконструировал названия, так как в известных словарях нет подходящих терминов. Приведем здесь в качестве примера конечного этапа факторного анализа в психологии личности результаты, полученные им самим (табл. 4, 5)[9]. Из смыслового анализа соответствия характеристик личности самим факторам видно, что встречаются неоднозначности определений. Это связано с особенностями факторного анализа – математическая статистика линейных связей не во всем совпадает с семантическим пространством описательных прилагательных.

Таблица 4

Фактор Н.
Parmia (Parasympathetic immunity – Threctia (Theat reactivity).
Благотворительная (филантропическая) циклотимия –
обструктивно очерченный шизотимический темперамент

Н+	Н–
Сотрудничающий	Обструктивный
Дружелюбный	Враждебный
Идеалист	Циничный
Благодарный	Неблагодарный
Мягкосердечный	Жесткосердечный
Оживленный	Подозрительный
Импульсивный	Привередливый
Естественный	Осторожный
Эмоциональный	Аффективный
Нетерпеливый	Терпеливый
Авантюрист	Робкий
Наглый	Вялый
С музыкальными и эстетическими интересами	Безжизненный
	Эксгибиционист

Таблица 5

Фактор М.
Autia – Praxernia.
Беспечность – практичность

М+	М–
Эксцентричный	Конвенциальный
Высокомерный	Тенденция быть «как все»
Гордый	Чуждый эстетическим интересам
Независимый	Зависимый
Порывистый	Сварливый
Чувствительный	Неэмоциональный
ЭстетичныйНетрадиционный	Неартистичный

Анализ Р. Кеттелла показывает как положительные стороны факторного анализа, так и его недостатки. Множество работ в этом направлении подтверждает перспективность применения факторного анализа в психологии, например у А.Г. Шмелева[10].

Частично проблему неоднозначности ФА решает процедура матричного вращения факторов в пространстве их нагрузок. Вращение в ФА реализуется по специальным машинным программам. Получаемые при этом результаты не всегда наглядны и приходится приводить дополнительные, так называемые «ручные повороты». Для этой цели необходимо рассматривать алгебро-геометрические соотношения в простых матрицах.

Рассмотрим матрицу координат стимулов $\overline{X}$ и ее графического представления.

$$\bar{x} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{I} & \mathrm{II} \end{array} \\ \begin{bmatrix} 1{,}00 & 1{,}00 \\ 1{,}00 & -1{,}00 \\ -1{,}00 & -1{,}00 \\ -1{,}00 & 1{,}00 \end{bmatrix} \end{array}$$

На рис. 3 представлены точки *a, b, c, d* в координатах осей I, II.

На рис. 4 показано, что при повороте координатных осей на 45° против часовой стрелки, они проходят через точки *a, b, c, d.*

На рис. 4 новая система координат имеет положения, аналогичные первоначальному, получена она после поворота на 45° против часовой стрелки.

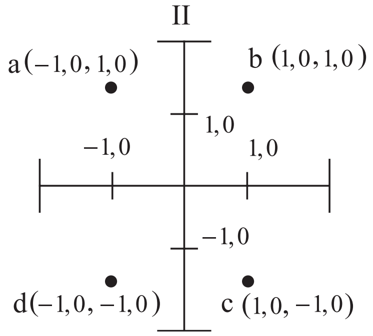


Рис. 3. Графическое представление точек в первоначальной матрице

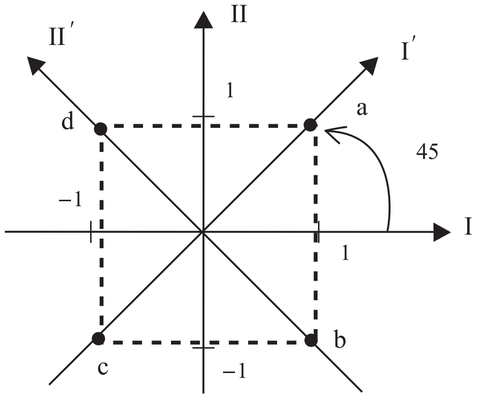


Рис. 4. Графическое представление операции вращения

Матрица координат точек, относящаяся к новой (после вращения) системе координат I′ и II′, примет вид:

$$\begin{array}{c|cc} & \mathrm{II}' & \mathrm{I}' \\ \hline a & 0{,}0 & \sqrt{2} \\ b & \sqrt{2} & 0{,}0 \\ c & 0{,}0 & -\sqrt{2} \\ d & -\sqrt{2} & 0{,}0 \end{array}$$

Оставляя связи точек с новой координатной системой неизменными, получим конфигурацию точек на рис. 5.

На рис. 5 точки a, b, c, d имеют по одной нулевой координате и одной, равной $\sqrt{2}$. После поворота осей на 45° вместе с точками, привязанными к новой системе, мы видим координатную систему на рис. 5. Эту же операцию вращения можно произвести с помощью умножения первичной матрицы на матрицу преобразования T:

$T = \begin{vmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{vmatrix}$, где $t_{rc} = \cos\Theta_{rc}$ – косинусы углов между соответствующими координатными осями.

Матрица $\overline{Y}$ получается из матрицы $\overline{X}$ после умножения на матрицу вращения $\overline{T}$:

$$\begin{vmatrix} 1{,}00 & 1{,}00 \\ 1{,}00 & -1{,}00 \\ -1{,}00 & -1{,}00 \\ -1{,}00 & 1{,}00 \end{vmatrix} \cdot \begin{vmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{vmatrix} = \begin{vmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \\ -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{vmatrix}$$

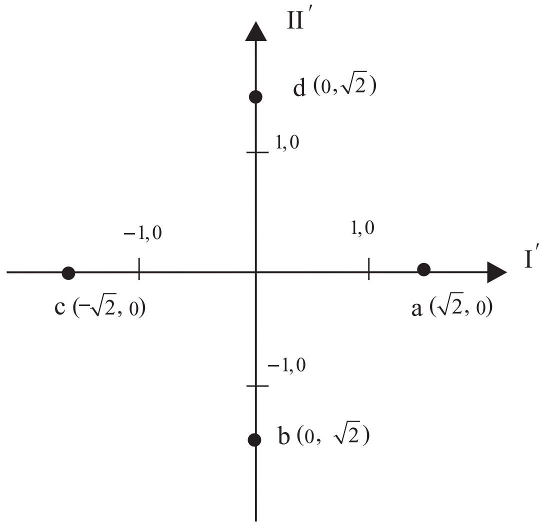


Рис. 5. Графическое представление точек матрицы после поворота

Таким образом, дополнительные ручные повороты позволяют получить такую конфигурацию пространства факторов, когда в преобразованной матрице происходит своеобразное «адресное» получение нагрузок, их общее число уменьшается.

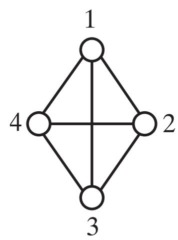


Рис. 6. Корреляционная плеяда «Общая слабость нервной системы»

Располагая характеристики в углах ромба (см. рис. 6), связывая характеристики прямыми линиями, получаем полный граф, показывающий, что все характеристики группируются вокруг какого-то фактора. В данном случае этот фактор – «слабость нервной системы».

Еще один пример связан с корреляционными плеядами, полученными из значимых корреляций ряда шкал MMPI у испытуемых различного пола, различных групп крови в двух выборках испытуемых (здоровых и больных). Такие графы были получены совместно с медицинскими работниками и показали, что существуют яркие различия в конфигурациях связей. Подобные графы можно построить при наличии в матрице интеркорреляций своеобразных «сгущений», образованных наличием значимых корреляций, в то время как незначимые считаются нулями корреляционной матрицы. Подобные графы можно усовершенствовать, вводя размерности расстояний или близостей между свойствами фактора. Например, расстояния можно выбрать как обратные значения корреляций. Тогда наибольшая близость будет равна 1, а максимальная – ∞.

Есть еще один вариант приблизительной факторизации без обращения к сложным процедурам ФА, когда учитывается, что направление искомого фактора находится в пространстве многомерных векторов на этом направлении.

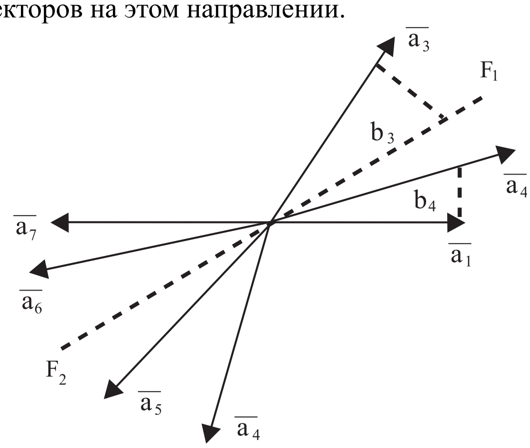


Рис. 7. Пучок многомерных векторов характеристик фактора F

Искомые факторы F_1 и F_2 (см. рис. 7) как бы свободно «плавают» в пространстве векторов $\overline{a}_1, \overline{a}_2, ..., \overline{a}_7$. На направления F_1, F_2 проецируются перпендикуляры из концов векторов, образовывая нагрузки $b_1, b_2, \ldots, b_7$. Отношения этих проекций b к длинам векторов есть не что иное как косинусы углов между векторами характеристик и направлениями искомых факторов, они же численно равны корреляциям между факторами и их свойствами.

Зная корреляции между векторами, можно приблизительно найти и корреляции между ними и факторами. Эти корреляции будут зависеть от направления фактора по отношению ко всем векторам, отсюда и возникает неоднозначность. Полагая, что фактор делит угол, находясь между векторами a_1, a_2, пополам, можно найти его нагрузки. Конечно, это будет лишь приближение, но не лишенное смысла, так как соблюдается главное условие соотношения характеристик и фактора. При этом следует помнить, что вектора располагаются в многомерном пространстве, а на рис. 7 представлена лишь их проекция на плоскость.

Интересные результаты с применением семантического дифференциала и факторного анализа получил В.Ф. Петренко[11]. С применением личностного семантического дифференциала проанализированы особенности вербальной семантики, образной репрезентации, визуальной семантики, структур обыденного сознания и эффективности речевого коммуникативного воздействия. Примеры из практики ФА позволяют проводить далекоидущие междисциплинарные аналогии и наряду с проверкой научно-практических гипотез формулировать новые.

Примечания

1 *Allport G.W.* Personality: a psychological interpretation. N.Y., 1937; *Cattell R.B.* Personality. A systematic theoretical and Factual Study. N.Y.; Toronto; L., 1950; *Eysenck H.J.* The logical basis of factor analysis //Amer. Psychol. J. 1953. V. 8. P. 105–113; *Guilford J.P., Guilford R.B.* An analysis of the Factors in atypical test of intovers-extravers // J. Abnorm. Soc. Psychol. 1934. V. 28. P. 277–393; *Hathaway S.R., Mc`Kinley J. Ch.* Minnesota Multiphasic Personality Inventory. N.Y., 1971; *Spearman Ch.* Theory of general factor // Brit. J. Psychol. 1946. V. 36. P. 117–131; *Thurstone L.L.* Multiple factor analysis. Chicago, 1961.

2 *Иберла К.* Факторный анализ. М., 1998; *Небылицын В.Д.* Современное состояние факторного анализа // Вопр. психологии. 1960. №4; *Петренко В.Ф.* Психосемантика сознания. М., 1988; *Тугушев Р.Х.* Психика как система, ее структура, компоненты и элементы // Материалы юбилейной науч-практ. конф., посвященной 30-летию отделения психологии Саратовского государственного университета. Саратов, 2001.

3 См.: *Шмелев А.Г.* Психодиагностика личностных черт. СПб., 2002.

4 См.: *Spearman Ch.* Op. cit.

5 См.: *Харман Г.* Современный факторный анализ. М., 1972; *Иберла К.* Указ. соч.

6 *Иберла К.* Указ. соч.

7 Цит. по: *Фресс П., Пиаже Ж.* Экспериментальная психология. М., 1975. Вып. 5. С. 120.

8 См.: *Hathaway S.R., Mc`Kinley J.Ch.* Op. cit.

9 См.: *Cattell R.B.* Op. cit.

10 *Шмелев А.Г.* Указ. соч.

11 *Петренко В.Ф.* Основы психодиагностики. М., 1997.