

# КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

# Основные понятия корреляционного анализа

***Корреляционный анализ*** – двумерная описательная статистика, количественная мера взаимосвязи (совместной изменчивости) двух переменных.

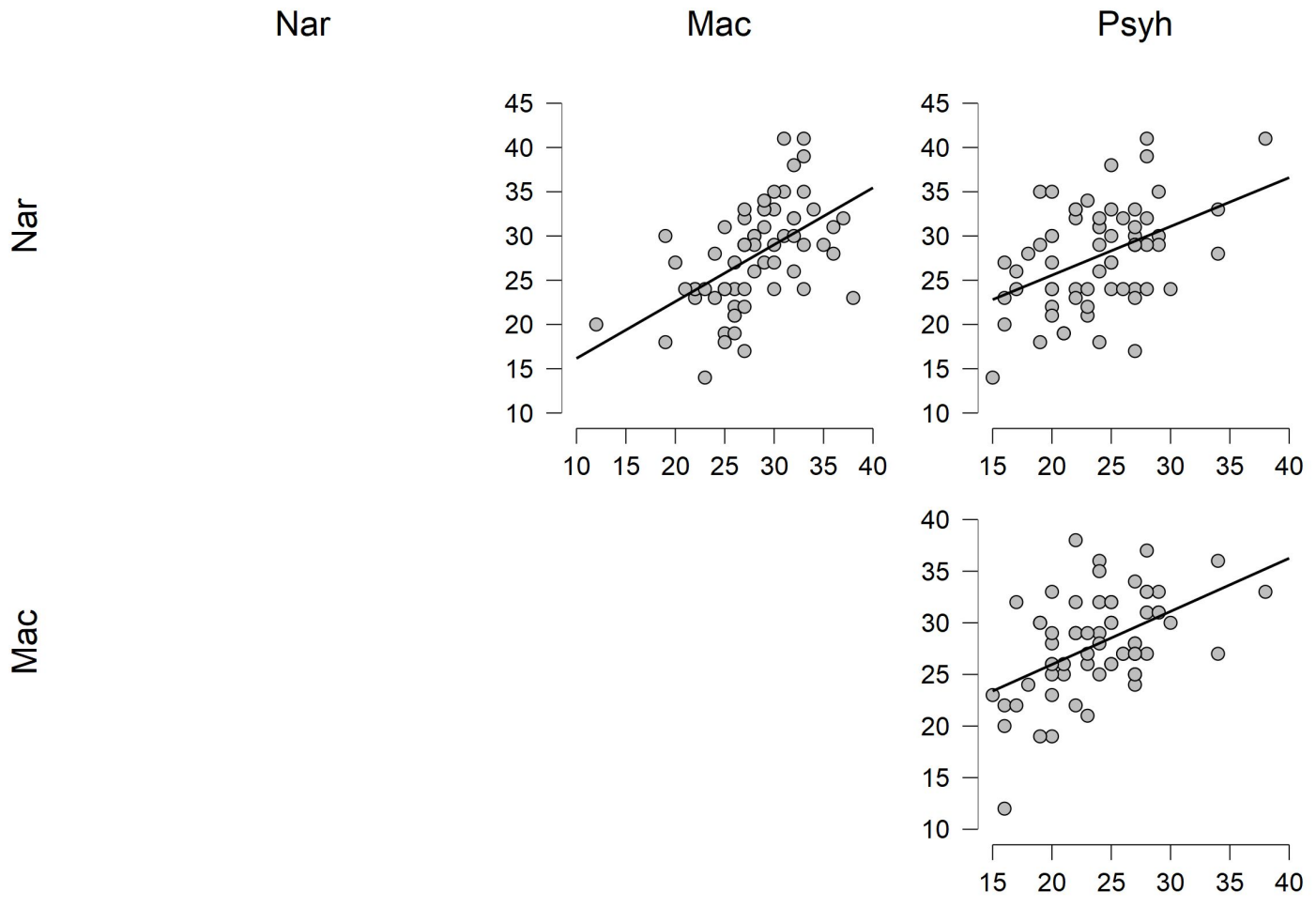
Автор термина и основоположник корреляционного анализа – Френсис Гальтон.

Наиболее распространенные коэффициенты корреляции разработал Карл Пирсон.

# Виды корреляционных связей: линейные и нелинейные

***Линейная корреляция*** – если с увеличением или уменьшением одной переменной  $X$ , вторая переменная  $Y$  в среднем либо также растет, либо убывает.

***Нелинейная корреляция*** – при увеличении одной величины, характер изменения второй не линеен, а описывается другими законами.



# Основные характеристики

- **Коэффициент корреляции** – это количественная мера силы и направления вероятностной взаимосвязи переменных; принимает значения в диапазоне от  $-1$  до  $+1$ .
- **Сила связи** достигает максимума при условии взаимно однозначного соответствия. Показателем силы связи является абсолютная (без учета знака) величина коэффициента корреляции.
- **Направление связи** определяется прямым или обратным соотношением значений двух переменных. Показателем направления связи является знак коэффициента корреляции.

# Виды корреляционных связей: положительные и отрицательные

***Положительная корреляция*** – если с увеличением переменной  $X$  переменная  $Y$  в среднем также увеличивается.

***Отрицательная корреляция*** – если с увеличением переменной  $X$  переменная  $Y$  в среднем имеет тенденцию к уменьшению.

- **Выбросы** – экстремально большие или малые значения признака. В большей степени влияют на корреляцию Пирсона, т.к. величина этого коэффициента прямо пропорциональна отклонению значения переменной от среднего.
- **Влияние «третьей» переменной** – корреляция между двумя переменными обусловлена не связью между соответствующими свойствами, а влиянием некоторой общей причины совместной изменчивости этих переменных (например: обратная связь роста и длины волос, третья переменная – пол).
- **Нелинейные связи** – при уменьшении или увеличении значений одной из переменной связь меняет свое направление (например: связь активации и продуктивности деятельности). В подобных случаях выборку разделяют на подгруппы

# Выбор коэффициента корреляции в зависимости от типов шкал

| Типы шкал                            |                                      | Коэффициент                              |
|--------------------------------------|--------------------------------------|--|
| Переменная X                         | Переменная Y                         |  |
| Интервальная или абсолютная          | Интервальная или абсолютная          | Коэффициент r-Пирсона                    |
| Ранговая, интервальная или отношений | Ранговая, интервальная или отношений | Коэффициент ρ-Спирмена                   |
| Ранговая                             | Ранговая                             | Коэффициент τ-Кендалла                   |
| Номинативная                         | Номинативная                         | Коэффициент φ и V Крамера                |
| Дихотомическая                       | Ранговая                             | Рангово-бисериальный $\rho_{\text{бис}}$ |
| Дихотомическая                       | Интервальная или отношений           | Бисериальный $R_{\text{бис}}$            |
| Интервальная                         | Ранговая                             | Не разработан                            |



# Коэффициент корреляции r-Пирсона

- это мера прямолинейной взаимосвязи; он не чувствителен к криволинейным связям.

## ***Условия применения:***

1. Сравнимые переменные должны быть получены в интервальной шкале или шкале отношений.
2. Распределения переменных должны быть близки к нормальному.
3. Число варьирующих признаков в сравниваемых переменных должно быть одинаковым.
4. Таблицы уровней значимости рассчитаны от 5 до 1000 человек. Оценка уровня значимости осуществляется при числе степеней свободы  $k = n - 2$
5. Выборка должна составлять от 30 респондентов

# Коэффициент корреляции ρ-Спирмена

- это коэффициент корреляции рангов, предложенный К. Спирменом, относится к непараметрическим показателям связи между переменными, измеренными в ранговой шкале.

## ***Условия применения:***

1. Сравнимые переменные должны быть получены в порядковой (ранговой) шкале, но могут быть также измерены в шкале интервалов и отношений.
2. Характер распределения не имеет значения.
3. Число варьирующих признаков в сравниваемых переменных должно быть одинаковым.
4. Таблицы для определения критических значений расчитаны от 5 до 40 человек, при большем числе используются таблицы Пирсона. Нахождение критических значений осуществляется при  $k=n$ .
5. Выборка от 5 человек и более.

# Корреляционная матрица

- это результат вычисления корреляций одного типа для каждой пары из множества  $P$  переменных, измеренных в количественной шкале на одной выборке.


## ***Основные характеристики корреляционной матрицы:***

- 1) она является квадратной;
- 2) она симметрична относительно главной диагонали;
- 3) на главной диагонали располагаются единицы, т. к. корреляция признака с самим собой равна единице.

### Корреляции

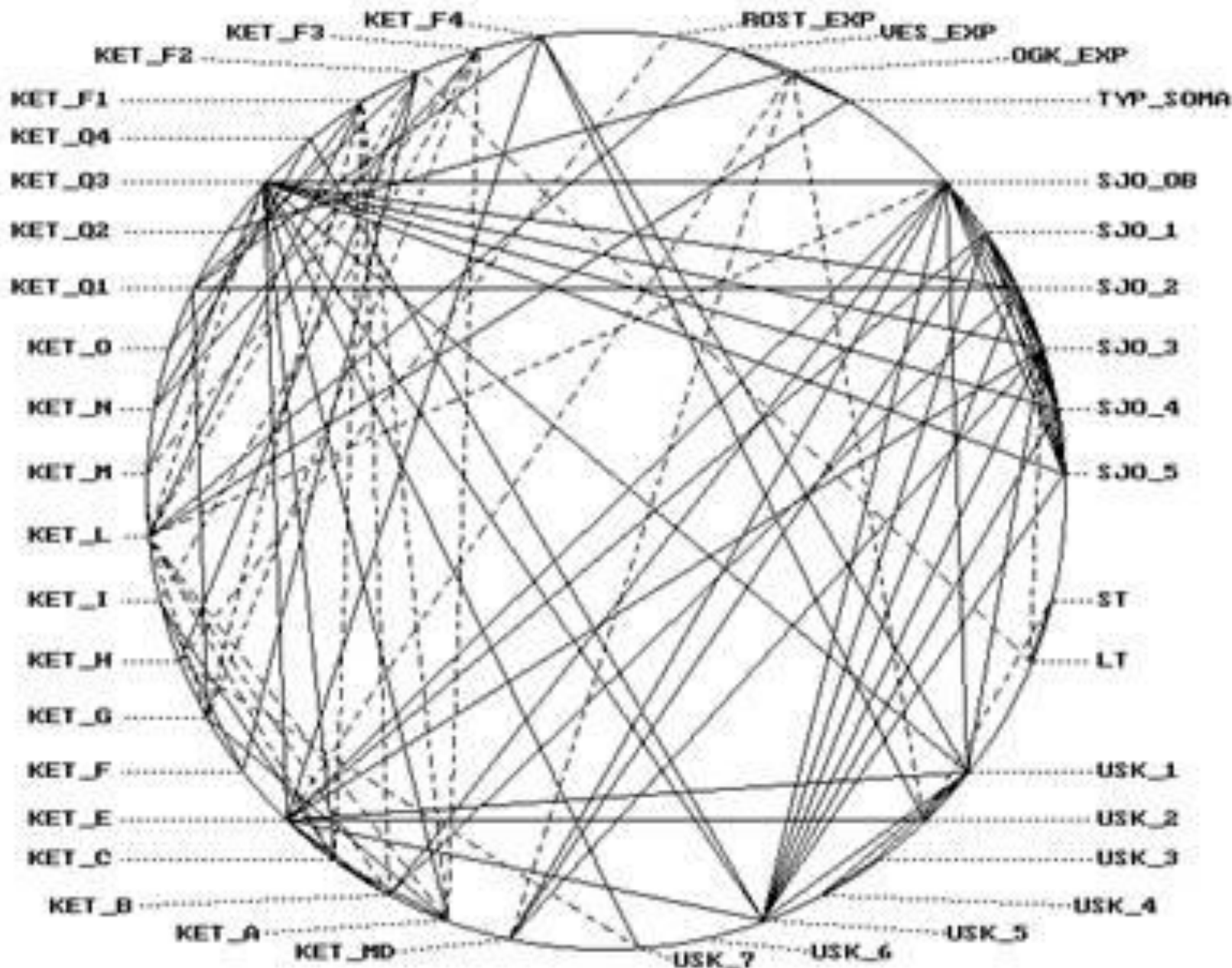
|      |                    | Nar  | Mac  | Psyh |
|------|--------------------|------|------|------|
| Nar  | Корреляция Пирсона | 1    | ,529 | ,435 |
|      | Знч.(2-сторон)     |      | ,000 | ,000 |
|      | N                  | 64   | 64   | 64   |
| Mac  | Корреляция Пирсона | ,529 | 1    | ,493 |
|      | Знч.(2-сторон)     | ,000 |      | ,000 |
|      | N                  | 64   | 64   | 64   |
| Psyh | Корреляция Пирсона | ,435 | ,493 | 1    |
|      | Знч.(2-сторон)     | ,000 | ,000 |      |
|      | N                  | 64   | 64   | 64   |

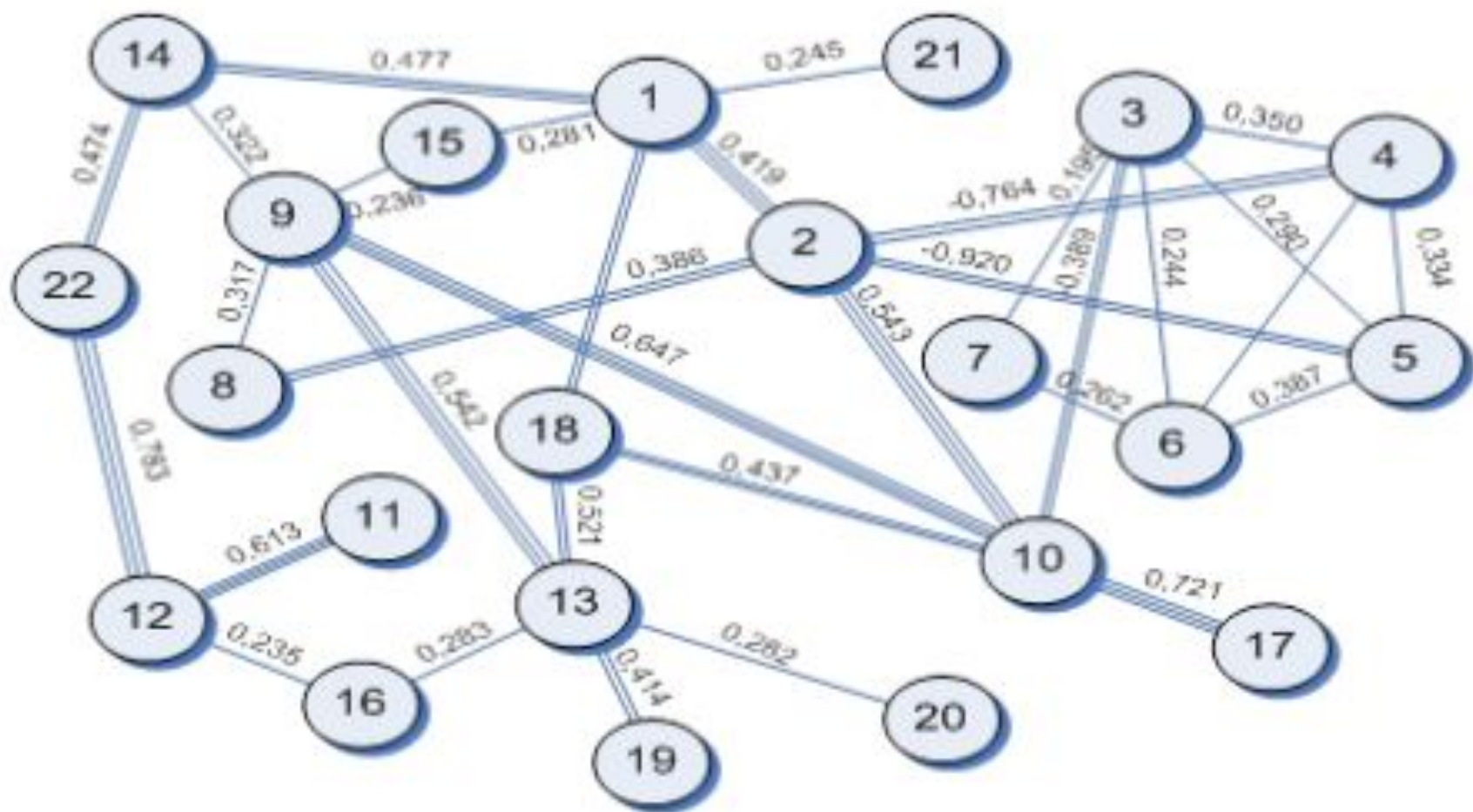


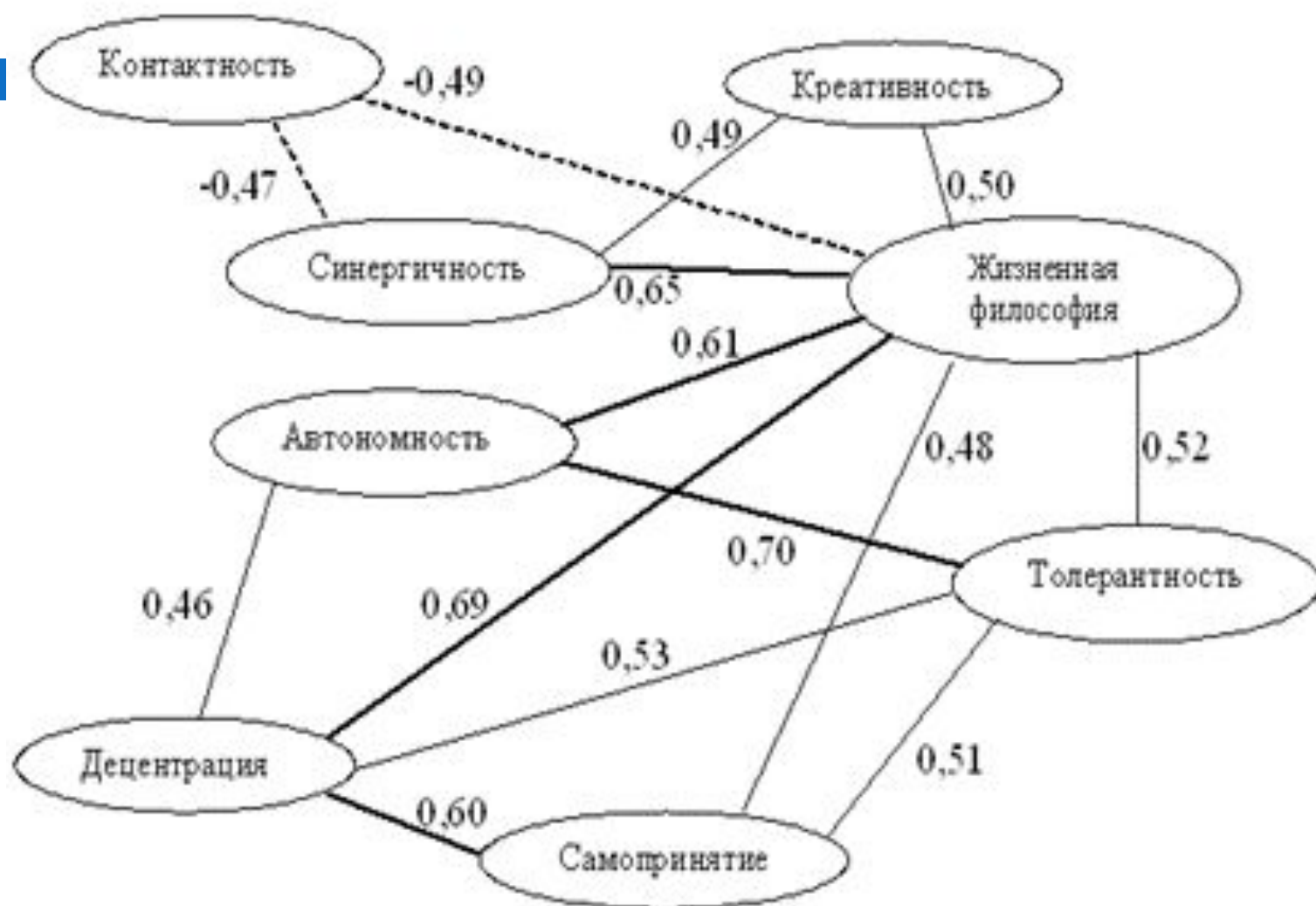


***Основная задача*** корреляционной матрицы – выявление структуры взаимосвязей множества признаков.

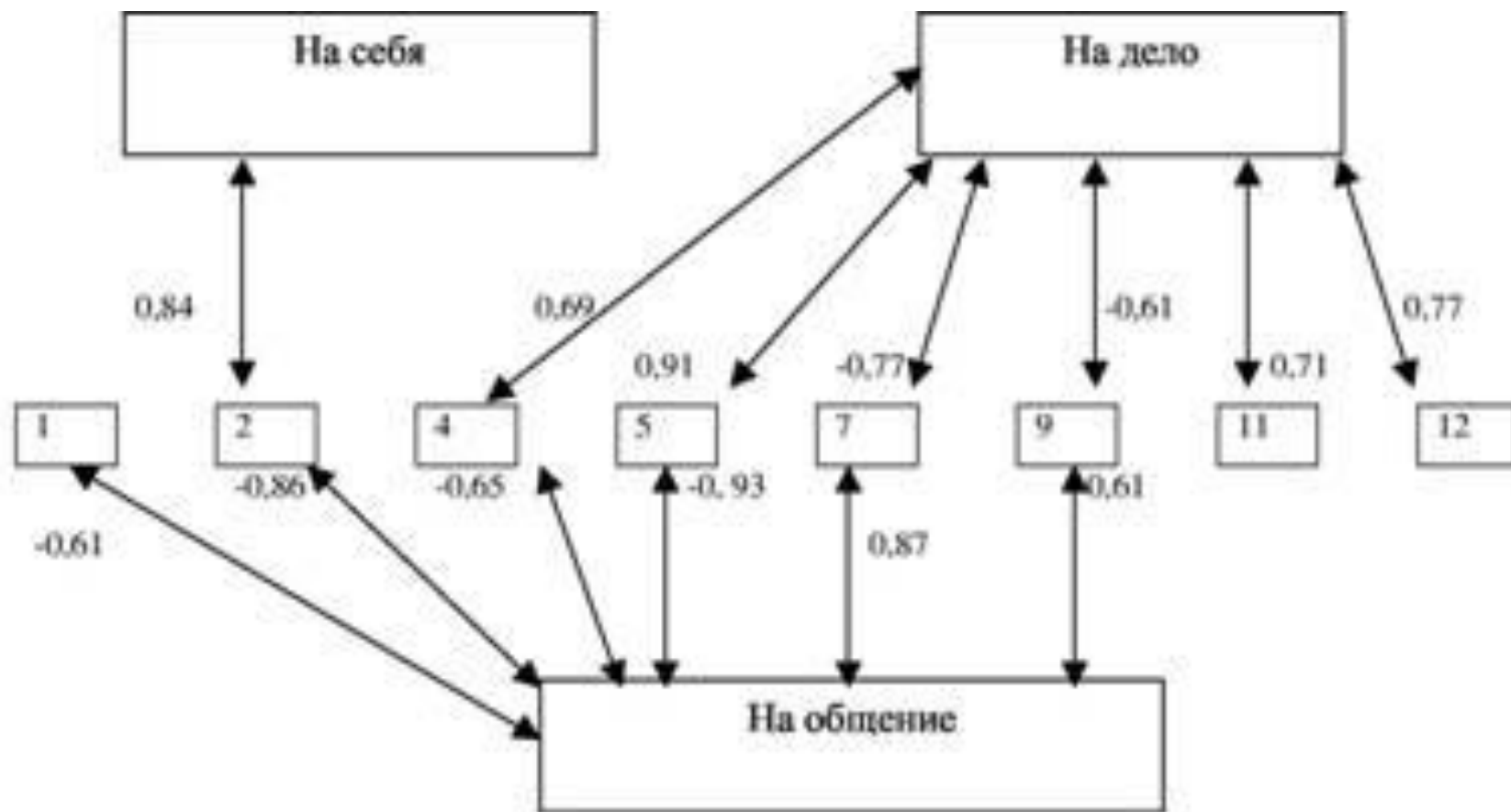
Возможен визуальный анализ ***корреляционных плеяд*** – графического изображения структуры статистически значимых связей.



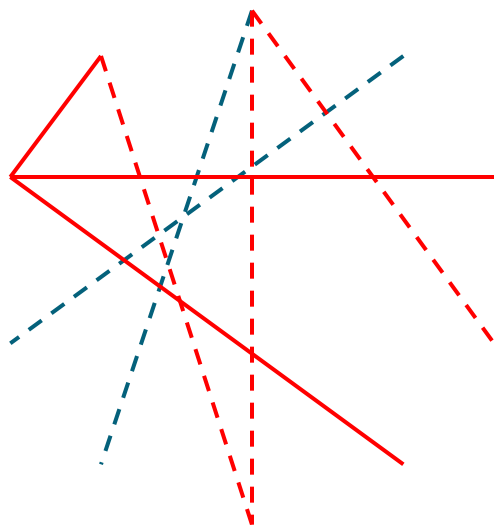












————— высокозначимая  
положительная взаимо-  
связь ( $p \leq 0,01$ )

- - - - - среднезначимая  
положительная взаимо-  
связь ( $p \leq 0,05$ )

————— высокозначимая  
отрицательная взаимо-  
связь ( $p \leq 0,01$ )

- - - - - среднезначимая  
отрицательная взаимо-  
связь ( $p \leq 0,05$ )

| Возрастная группа | Вариант диспозиции              | Взаимосвязи с показателями личностных особенностей: |  |
|-------------------|---------------------------------|---|--|
|                   |                                 | прямая  | обратная                                     |
| Юность            | Блокирование информации о среде | -   | Индекс общей активности (r= -0,299, p=0,005) |
|                   | Изменение дистанции: уход       | Мотивация боязни неудачи (r=0,290, p=0,006)         | -  |
|                   | Самоконтроль                    | Мотивация успеха (r=0,332, p=0,002)                 | -  |
|                   | Преобразование себя             | Индекс общей активности (r=0,334, p=0,001)          | -  |

# Описание корреляционных связей

- **Интеркорреляционные взаимосвязи** – связи между показателями одной методики. Данные связи ожидаемы, так как подтверждают внутреннюю согласованность и валидность теста.
- **Межкорреляционные взаимосвязи** – связи между показателями разных методик. Это основная задача данного вида математической процедуры.