

БАГАТОМІРНЕ ШКАЛЮВАННЯ: ВИКОРИСТАННЯ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ

Анотація.

У статті вивчаються можливості використання багатомірного шкалювання у психології. Робота проводиться на матеріалі реальних психологічних досліджень. Наводяться приклади процесу роботи та інтерпретації отриманих результатів.

Постановка проблеми.

Ми продовжуємо роботу над методами багатомірного математичного аналізу результатів психологічних досліджень. Темою цієї статті є багатомірне шкалювання – метод, який багато дослідників вважають значно простішим за факторний аналіз (як у використанні, так і в інтерпретації) і пишуть, що він є альтернативою факторному аналізу [6, 7]. Що ж робить багатомірне шкалювання таким привабливим? Пригадаємо, що факторний аналіз проводиться на основі кореляційної матриці, для побудови якої ще необхідно зібрати великий масив сирих даних. Багатомірне шкалювання, як виявляється, може працювати з простою матрицею схожості (або відмінності) об'єктів між собою. Як же працює багатомірне шкалювання, яка його внутрішня логіка та як інтерпретувати його результати? Зараз ми ці питання детально і розглянемо. Знову ж таки, будемо спиратися на роботу з програмою Statistica 5.5 (див [9-13]).

Мета багатомірного шкалювання. Багатомірне шкалювання (БШ) дійсно є до певної міри альтернативою факторному аналізу. Перевага БШ – можливість у якості вихідних даних використовувати довільні матриці схожості (це можуть бути і кореляційні матриці, і прості матриці схожості (відмінності) об'єктів у шкалі відношень). Багатомірне шкалювання – це не стільки математична процедура, скільки *спосіб ефективного розміщення об'єктів у просторі певної кількості ознак*, який максимально зберігає відстані між цими об'єктами.

Підготовка даних до БШ. Для проведення багатомірного шкалювання перш за все потрібно мати матрицю схожості (або відмінності) об'єктів. Її можна отримати різними шляхами. Оптимальними є дані, отримані шляхом експертної оцінки близькості об'єктів. Якщо відомості про близькість об'єктів отримані не прямо, а

Дані, підготовлені для БШ у програмі Statistica 5.5 матимуть такий вигляд (табл. 2):

Таблиця 2.

*Матриця близькості психологічних теорій
(підготовлена для обробки у програмі Statistica)*

	ФРОЙД	ЮНГ	АДЛЕР	УОТСОН	СКІННЕР	ХАЛЛ	МАСЛОУ	РОДЖЕРС	МЕЙ
Фройд	10	7	8	1	3	3	2	1	2
Юнг	7	10	6	3	1	4	1	1	2
Адлер	8	6	10	3	3	3	4	4	2
Уотсон	1	3	3	10	8	7	3	4	3
Скіннер	3	1	3	8	10	7	3	4	3
Халл	3	4	3	7	7	10	3	3	3
Маслоу	2	1	4	3	3	3	10	8	9
Роджерс	1	1	4	4	4	3	8	10	7
Мей	2	2	2	3	3	3	9	7	10
matrix	2								

С. 17

Логіка багатомірного шкалювання. Ми з'ясували, що метою БШ є розміщення об'єктів у просторі певної розмірності, максимально зберігши відстані між ними. Логіку цього процесу можна проілюструвати на деяких прикладах.

Нехай маємо матрицю попарних відстаней міст України. Обравши розмірність нового простору рівну 2, метод БШ розташує ці міста у 2-х мірному просторі, максимально зберігши відстані між ними. Отримане розташування точок можна буде використовувати як приблизну карту України. Самі ці 2 виміри можна тепер проінтерпретувати як координати "Пн\Пд" та "Зх\Сх" [8].

Для деталізації прикладу уявимо, що маємо 3 міста, а також відстані між цими містами (табл. 3):

Таблиця 3.

Відстані між трьома містами

	Місто А	Місто Б	Місто В
Місто А	0		
Місто Б	90 км	0	
Місто В	90 км	90 км	0

Тепер ці міста можна розташувати у певному n -мірному просторі. Нехай це буде двохмірний простір (площина). Тоді всі три міста дійсно можна розташувати на цій площині так, щоб зберегти відстані між ними (рис. 1, зліва):

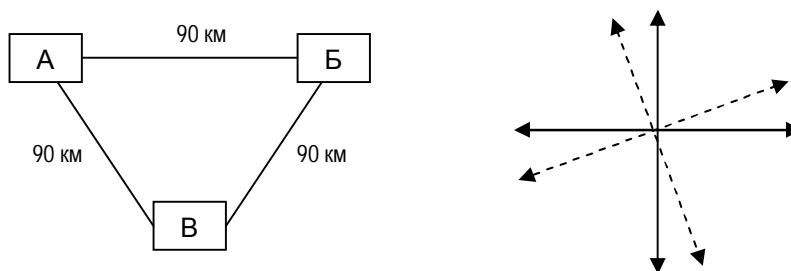


Рис. 1. Міста, розміщені у 2-мірному просторі

При цьому, розмістити ці міста можна так, щоб вони відповідали реальному розміщенню на карті, повертаючи осі координат (рис 1, справа).

Ситуація ускладнюється, якщо маємо не 3, а 4 міста, причому 4-те місто Г знаходиться на відстані 89 км від усіх інших міст. Це може бути тоді, коли відстань між містами виміряна не прямо, а по довжині доріг (табл. 4, рис. 2):

Таблиця 4.

Відстані між чотирма містами

	Місто А	Місто Б	Місто В	Місто Г
Місто А	0			
Місто Б	90 км	0		
Місто В	90 км	90 км	0	
Місто Г	89 км	89 км	89 км	0

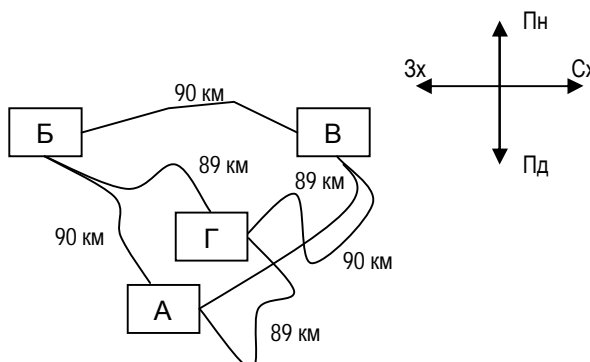


Рис. 2. Реальне розташування 4-х міст на карті

Тоді процедура БШ буде не такою прозорою. Зберегти відстані у просторі осей при БШ вдасться лише у випадку 3-мірного простору (побудується 3-гранна піраміда, у вершинах якої будуть розташовані міста) – рис. 3.

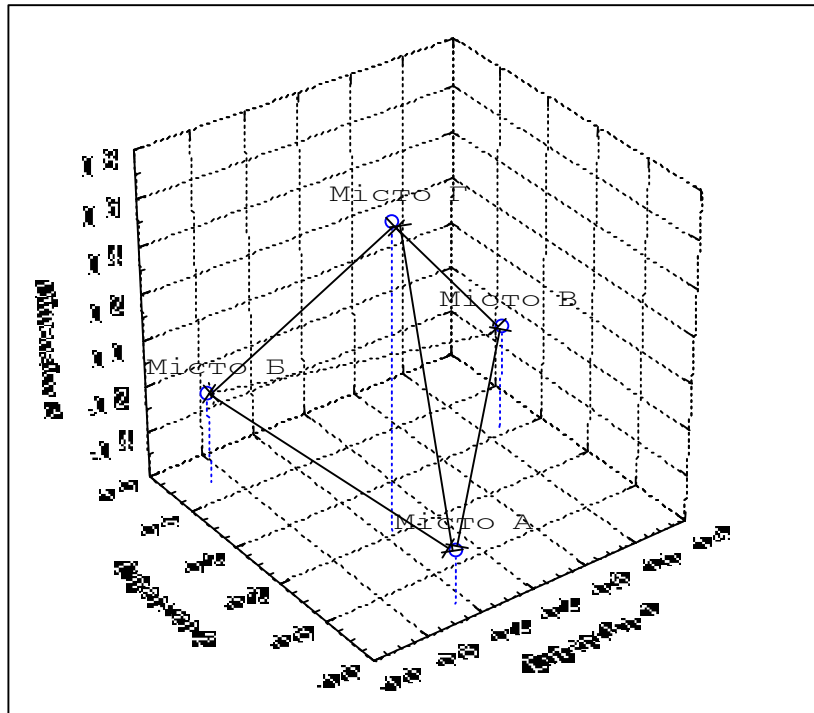


Рис. 3. Розташування 4-х міст у трьохмірному просторі

Але така картина навряд чи буде придатна до інтерпретації – вона не буде мати ніякого смислу. Тоді ми змушені спробувати розмістити ці міста у 2-мірній площині. Наслідком цього буде те, що процедура БШ на основі вихідної матриці відстаней обчислить **оновлену матрицю** відстаней так, щоб оптимізувати розміщення міст – малі відстані можуть бути ще зменшені, і більші – ще збільшені. В результаті **оновлена матриця** буде мати такий вигляд (табл. 5):

Таблиця 5.

Оновлені відстані між містами

	Місто А	Місто Б	Місто В	Місто Г
Місто А	0			
Місто Б	1,86	0		
Місто В	1,98	2,10	0	
Місто Г	0,79	1,22	1,45	0

А розміщення міст на площині буде таке (рис. 4):

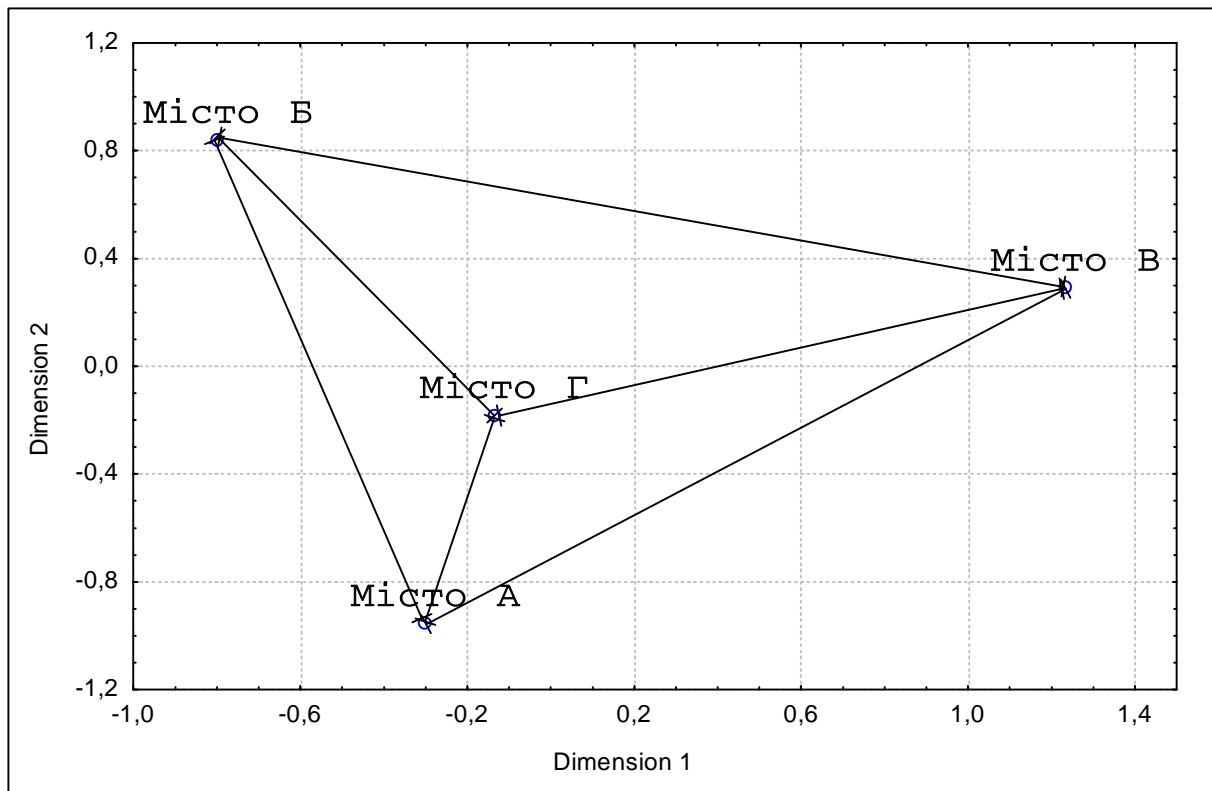


Рис. 4. Розташування 4-х міст у двохмірному просторі на основі оновленої матриці відстаней

Порівняйте результат, отриманий з допомогою БШ та вихідний рисунок – вони дуже схожі. Дійсно, вийшло, що БШ перетворило відстані, виміряні по довжині доріг, у реальні геометричні відстані між містами. Таку картину тепер можна проінтерпретувати як реальне розміщення міст на карті, а осі утвореного простору можна вважати осями сторін світу.

С. 18

А.Ю. Терьохіна наводить приклад дослідженні візуального сприймання букв російського алфавіту [6]. Учасникам дослідження пропонувалося оцінити пари букв за їх схожістю. Після обробки групової матриці схожості було проведено БШ та отримано таку картину (рис. 5):

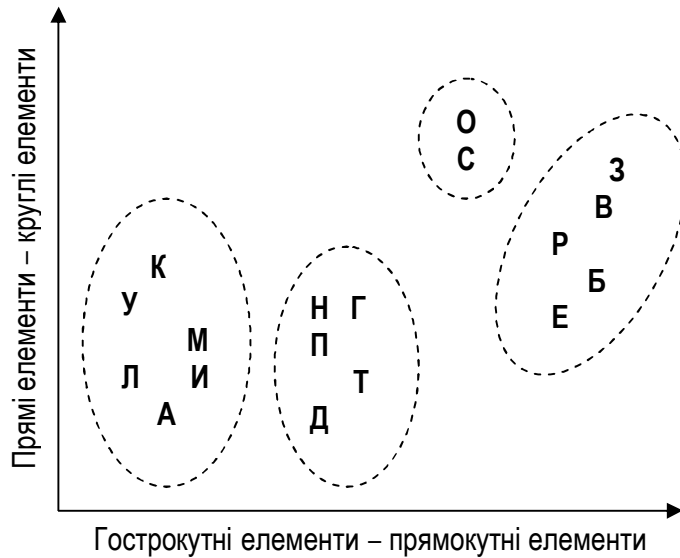


Рис. 5. Результати використання багатомірного шкалювання у дослідженні візуального сприймання букв

Букви згрупувалися за схожістю в просторі двох осей, які на основі змістового аналізу вдалося проінтерпретувати як “прямі-круглі” та “гострокутні-прямокутні”.

Якщо продовжити наш приклад про сприймання психологічних теорій, то за логікою БШ повинні утворитися три групи теорій, найбільш близькі між собою (психоаналітичні – Фройд, Юнг, Адлер; поведінкові – Скіннер, Толмен, Халл; гуманістичні – Маслоу, Роджерс, Мей).

Проблема розмірності при багатомірному шкалюванні. На попередньому прикладі з містами ми побачили, що чим більша розмірність простору, у якому ми проводимо БШ, тим більше буде відповідності між вихідною матрицею даних та оновленою матрицею. Якщо кількість вимірів простору буде рівна кількості змінних, ми тримаємо повну відповідність між цими матрицями. Однак, не це є нашою задачею. Навпаки, нам важливо, щоб кількість вимірів простору БШ була якомога меншою. Але тоді буде більше розбіжностей між вихідною та оновленою матрицею відстаней. З’являється суперечність, яку необхідно вирішити. Виявляється, існує два суттєвих фактори, які впливають на вибір мірності простору БШ.

Перший фактор – інтерпретованість. Необхідно підбирати таку мірність простору, щоб отримана конфігурація точок мала смисл і була якомога простішою та зрозумілою. Якщо конфігурація має вигляд хаотичної хмари, то навіть якщо усі

математичні показники ідеальні – вони мало чим допоможуть. Ми не зможемо зробити жодних висновків. У такому випадку слід спробувати змінити мірність простору (збільшити чи зменшити) – можливо, тоді з’явиться очікувана структура.

Другий фактор – критерій кам’янистого насипу. Цей критерій був запропонований Р. Кеттелом для оптимального вибору кількості факторів у факторному аналізі (там йшлося про графік залежності власних значень факторів від їх кількості). У випадку ж багатомірного шкалювання можна говорити про графік залежності величини стресу (*про стрес більш детально поговоримо далі*) від мірності простору. Необхідно на цьому графіку знайти точку, у якій спадання стресу максимально сповільнюється. Ця кількість вимірів і буде оптимальною.

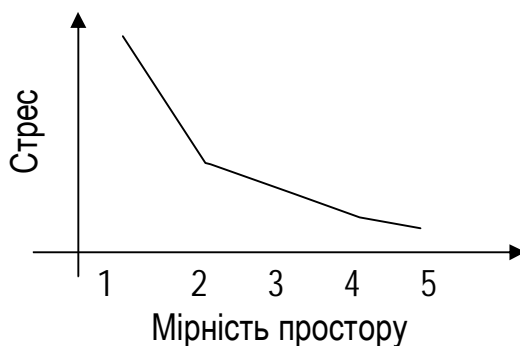


Рис. 6. Залежність стресу від мірності простору багатомірного шкалювання

Однак, Statistica автоматично не будує кам’янистий насип при БШ. Його необхідно побудувати вручну, провівши БШ з різними показниками мірності простору і відмітивши відповідно стрес для кожного БШ.

Інтерпретація результатів багатомірного шкалювання. Для інтерпретації результатів БШ використовують певний набір показників.

Вихідна матриця відстаней – матриця, яку ми створили для проведення БШ. У нашому випадку – це матриця оцінок студентами близькості різних психологічних теорій (табл. 2).

Оновлена матриця відстаней (distance matrix) – матриця відстаней, створена процедурою БШ для розміщення об’єктів у n-мірному просторі. Оновлена матриця відстаней може не співпадати з вихідною матрицею відстаней, оскільки є її певним спрощенням. Для наших даних оновлена матриця відстаней має такий вигляд:

Оновлена матриця відстаней

	ФРОЙД	ЮНГ	АДЛЕР	УОТСОН	СКІННЕР	ХАЛЛ	МАСЛОУ	РОДЖЕРС	МЕЙ
Фройд	0,00	0,44	0,09	1,74	1,75	1,68	1,73	1,74	1,75
Юнг	0,44	0,00	0,53	1,70	1,73	1,54	1,74	1,83	1,74
Адлер	0,09	0,53	0,00	1,73	1,73	1,70	1,70	1,69	1,73
Уотсон	1,74	1,70	1,73	0,00	0,09	0,44	1,73	1,69	1,74
Скіннер	1,75	1,73	1,73	0,09	0,00	0,53	1,74	1,68	1,75
Халл	1,68	1,54	1,70	0,44	0,53	0,00	1,70	1,74	1,68
Маслоу	1,73	1,74	1,70	1,73	1,74	1,70	0,00	0,35	0,09
Роджерс	1,74	1,83	1,69	1,69	1,68	1,74	0,35	0,00	0,43
Мей	1,75	1,74	1,73	1,74	1,75	1,68	0,09	0,43	0,00

Первинний аналіз можна провести уже за цією матрицею. Так, бачимо, що у нас виділяться 3 групи теорій, найбільш близькі між собою.

N-мірна кінцева конфігурація (*Graph final configuration*) – графічне розміщення об’єктів в *n*-мірному просторі. Ця конфігурація є основним результатом БШ, і саме по ній як правило здійснюють інтерпретацію результатів БШ. Наша 3-мірна кінцева конфігурація матиме такий вигляд (рис. 7.):

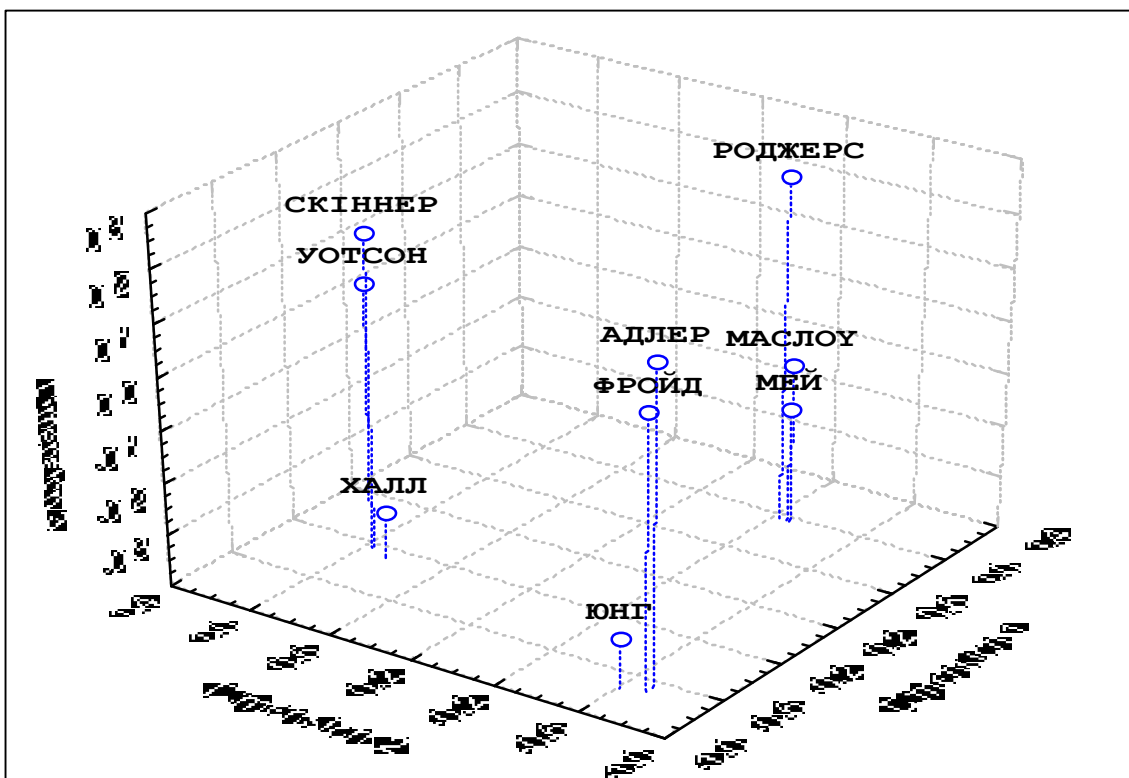


Рис. 7. 3-мірна кінцева конфігурація

При аналізі кінцевої конфігурації необхідно:

1. Проаналізувати виявлені групи точок, а також конфігурацію цих груп, їх кількість тощо.
2. Проінтерпретувати осі координат (виміри).

У нашому випадку виділилося 3 групи теорій (психоаналітична, гуманістична та поведінкова). Осі координат можна описати як: 1 вісь – психоаналіз-гуманістичний підхід, 2 вісь – поведінковий підхід – гуманістичний підхід, 3 вісь – слабо піддається інтерпретації, і очевидно, є лишньою. Можна провести повторне багатомірне шкалювання у 2-мірному просторі.

Таблиця кінцевої конфігурації (final configuration) – координати кожного об'єкта по виділених осях. Це числове представлення кінцевої конфігурації, і схоже на таблицю факторних навантажень. Дійсно, з таблиці кінцевої конфігурації видно, що 3 вимір (3-тя вісь) є зайвою, оскільки жоден з об'єктів сильно його не навантажує.

Таблиця 7.

Таблиця кінцевої конфігурації

D-star: Raw stress = ,0027762; Alienation = ,0058544			
D-hat: Raw stress = ,0012195; Stress = ,0038801			
	DIM. 1	DIM. 2	DIM. 3
ФРОЙД	-0,70577	-0,71113	0,121468
ЮНГ	-0,74058	-0,60869	-0,30383
АДЛЕР	-0,66259	-0,71264	0,204988
УОТСОН	-0,24198	0,963722	0,109729
СКІННЕР	-0,22722	0,96582	0,200653
ХАЛЛ	-0,2872	0,86177	-0,31107
МАСЛОУ	0,952193	-0,27553	-0,09334
РОДЖЕРС	0,956395	-0,22415	0,252479
МЕЙ	0,956752	-0,25918	-0,18108

Стрес – показник, який характеризує ступінь відмінності між вихідною матрицею даних та оновленою матрицею. Чим більше це число, тим менш надійними є результати БШ. У нашому випадку стрес рівний 0,0038801, що є достатньо непоганим показником.

Діаграма Шепарда (Shepard diagram) – діаграма, за якою візуально можна побачити ступінь відмінності між вихідною та оновленою матрицею даних. Діаграма Шепарда для нашого випадку має такий вигляд:

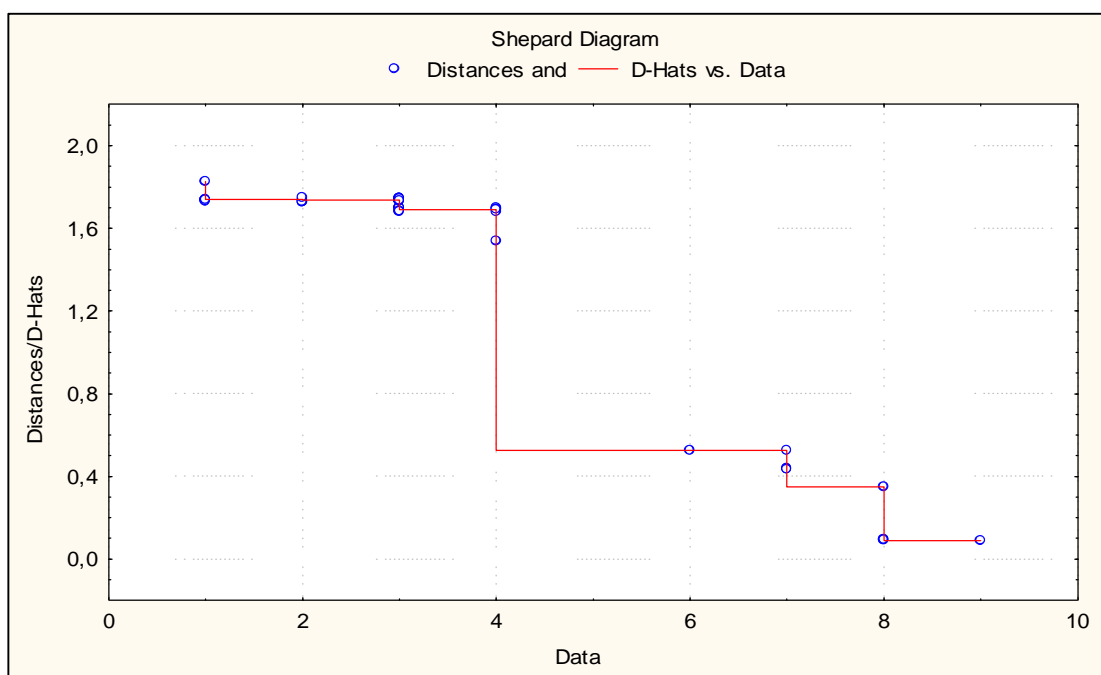


Рис. 8. Діаграма Шепарда

Для аналізу діаграми Шепарда дивляться, наскільки добре точки на ній лягають на ступінчасту лінію. Якщо точки і лінія майже співпадають, то аналіз можна вважати надійним – оновлена матриця не спотворила вихідні дані.

Багатомірне шкалювання та факторний аналіз. На початку нашої роботи ми відмітили, що БШ може слугувати певною альтернативою факторному аналізу. Для кращого орієнтування у проблемі вибору аналізу слід описати деякі відмінності цих методів.

1. Факторний аналіз накладає на вихідні дані вимогу відповідності критеріям нормального розподілу та статистичної значимості коефіцієнтів кореляції. В БШ такого обмеження немає.
2. Факторний аналіз вимагає, щоб змінні були оцінені за певним списком характеристик, і лише тоді обчислює кореляційну матрицю. У БШ можна прямо оцінювати попарну схожість між об'єктами.

3. Для роботи з факторним аналізом необхідно мати коефіцієнти кореляцій між змінними. Для БШ крім коефіцієнтів кореляцій можна використовувати будь-які міри схожості чи відмінності між об'єктами – бали чи ранги.
4. Факторний аналіз ґрунтується на уявленні про лінійний зв'язок між аналізованими змінними. БШ такого обмеження теж не має.
5. Факторний аналіз намагається виділити більшу кількість факторів, ніж БШ. З цієї точки зору БШ дає більш прості для інтерпретації рішення.

Реплікуюче багатомірне шкалювання. Часто дослідник стикається з проблемою багатомірного результатів багатьох індивідуальних матриць. В такому випадку він змушений або на основі усереднення цих матриць будувати одну групову, і з нею проводити БШ, або проводити БШ окремо для кожної матриці, і потім узагальнювати результати. Однак, не так давно було розроблено метод *реплікуючого багатомірного шкалювання*. Він полягає у тому, що програма одночасно обробляє набір індивідуальних матриць, і будує комплексний груповий простір об'єктів. Поки що у програмі Statistica ця можливість не реалізована, однак, вона є у програмі SPSS.

Висновки.

Багатомірне шкалювання є ефективною альтернативою факторному аналізу у випадку, коли необхідно використати пряму оцінку схожості або відмінності об'єктів між собою. Так можна отримати просту конфігурацію, зручну для інтерпретації. Можливості використання багатомірного шкалювання у психологічних дослідженнях ще далеко не вичерпані. Можна говорити про використання його у психофізіологічних дослідженнях (коли необхідно оцінювати та порівнювати стимули), у соціально-психологічних дослідженнях (коли оцінюють близькість між собою членів групи), у психології реклами (аналіз параметрів сприймання споживачами різних видів продукції), у педагогічній психології (для розробки систем комплексної оцінки знань студентів) тощо. Однак, у якій би сфері не проводилося багатомірне шкалювання, слід пам'ятати, що успіх можливий лише тоді, коли дослідник опирається на теоретичну модель явища, яке вивчає. Без цього багатомірне шкалювання перетворюється на сліпе оперування масою цифр, а сама

наукова робота – на псевдонаукову, коли дослідник перетворюється на фокусника і про себе думає: “Спритність рук і ніякого шахрайства!”.

Звісно, що хотілося б запропонувати читачам простий алгоритм, який би дозволив провести багатомірне шкалювання вручну (як це можна зробити з факторним або кластерним аналізами). І такий алгоритм є – це алгоритм У.С. Торгерсона [1]. Однак, для його використання необхідно перед тим освоїти у повному обсязі курс матричної алгебри і математичного аналізу. Тому, поки що, оптимальним варіантом лишається освоєння комп’ютерних статистичних програм, як то Statistica, SPSS, NCSS тощо.

С. 20

Література.

1. Дронов С.В. Многомерный статистический анализ. – Барнаул: Изд-во Алтайского гос. ун-та, 2003. – 213 с.
2. Крылов В.Ю. О многомерном шкалировании в неметрических пространствах // Психол. журн. – 1987. – Т.8. – №5. – С. 140-142.
3. Лосик Г.В. Исследование восприятия гласных методом многомерного шкалирования // Психол. журн. – 1992. – Т.13. – №2. – С. 89-95.
4. Наследов А. Д. SPSS: Компьютерный анализ данных в психологии и социальных науках. – СПб.: Питер, 2004. – 416 с.
5. Петренко В.Ф. Основы психосемантики: Учеб. пособие. – М.: Изд-во Моск. ун-та, 1997. – 400 с.
6. Терехина А.Ю. Многомерное шкалирование в психологии // Психол. журн. – 1983. – Т.4. – №1. – С. 76-88.
7. Шепард Р.Н. Многомерное шкалирование и безразмерное представление различий // Психол. журн. – 1980. – Т.1. – №4. – С. 72-84.
8. <http://www.statsoft.ru>

Робота з програмою STATISTICA

9. Боровиков В. STATISTICA: искусство анализа данных на компьютере (с CD-ROM). – СПб.: Питер, 2003. – www.statsoft.ru, www.biblion.ru, www.glagol.ru.
10. Боровиков В. Популярное введение в программу Statistica. – КомпьютерПресс,

2000. – www.statsoft.ru

11. Боровиков В. Программа Statistica для студентов и инженеров. – КомпьютерПресс, 2001. – www.statsoft.ru, www.mistral.ru, www.ozon.ru, www.biblio-globus.ru, www.books.ru, www.bolero.ru.
12. Боровиков В., Боровиков И. STATISTICA® – Статистический анализ и обработка данных в среде Windows. – 1997. – www.statsoft.ru
13. Салин В., Чурилова Э. Практикум по курсу "Статистика". В системе Statistica. – М.: Перспектива, 2002 – www.ozon.ru, www.bolero.ru, www.colibri.ru.