О.Г. Янковий

ЛАТЕНТНІ ОЗНАКИ В ЕКОНОМІЦІ

МОНОГРАФІЯ

Одеса 2015

УДК 33 : 316.4. 057.4
ББК 65

JEL O 10

Рекомендовано Вченою радою Одеського національного економічного університету (протокол № 7 від 15 травня 2015 року)

Янковий О.Г. Латентні ознаки в економіці: монографія. – Одеса: Атлант, 2015. – 168 с.

ISBN 978-617-7253-44-9

В монографії представлені наукові результати та напрацювання, пов'язані з поняттями, методами і напрямами оцінювання прихованих властивостей економічних об'єктів, які були отримані безпосередньо автором, а також членами кафедри економіки підприємства, представниками інших кафедр Одеського національного економічного університету.

Наводяться поняття та теоретичні засади латентних ознак в економіці, визначається їх місце серед якісних і кількісних властивостей об'єктів. Доводиться, що для латентних ознак економічних об'єктів можливий дрейф вздовж шкал вимірювання на основі застосування оцінок, одержаних за допомогою багатовимірних статистичних і математичних методів та моделей. Розглядаються методи оцінки латентних ознак економічних об'єктів і їх класифікація. Надається огляд захищених дисертацій та наукових публікацій, в яких застосовувались математико-статистичні методи оцінювання латентних ознак у різних галузях економіки України за останні роки.

Рекомендується студентам, аспірантам, докторантам, викладачам та науковцям, що досліджують проблеми оцінки латентних ознак економічних об'єктів.

ISBN 978-617-7253-44-9

© Атлант, ВОІ СОІУ, 2015

З М І С Т

ПЕРЕДМОВА 5

1. СУТНІСТЬ ЛАТЕНТНОЇ ОЗНАКИ В ЕКОНОМІЦІ 7

1.1 Класифікація ознак і показників економічних об'єктів 7
1.2 Шкали вимірювання ознак економічних об'єктів 11
1.3 Латентні ознаки економічних об'єктів та їх дрейф уздовж шкал вимірювання 19
1.4 Роль оцінки латентних ознак в економічному аналізі та прогнозуванні 22

2. МЕТОДИ ОЦІНКИ ЛАТЕНТНИХ ОЗНАК ЕКОНОМІЧНИХ ОБ'ЄКТІВ 29

2.1 Класифікація методів оцінки латентних ознак 29
2.2 Методи експертного оцінювання 35
2.3 Виділення репрезентативного чинника-симптому 45
2.4 Таксономічний аналіз 48
2.5 Факторний (компонентний) аналіз 58
2.6 Оцінка латентних ознак на основі сполучення кластерного і дискримінантного аналізу 65
2.7 Методи теорії нечіткої логіки 75
2.8 Багатовимірне шкалювання 79
2.8.1 Сутність і основні поняття багатовимірного шкалювання 79
2.8.2 Класична модель метричного шкалювання 86
2.8.3 Неметричне шкалювання 89

3. ОЦІНКА ЛАТЕНТНИХ ОЗНАК В ЕКОНОМІЦІ УКРАЇНИ 96

3.1 Оцінка трудових латентних ознак 96
3.2 Оцінка фінансового стану підприємств 100
3.3 Оцінка конкурентоспроможності підприємств 103
3.4 Оцінка інвестиційної привабливості підприємств 108
3.5 Оцінка глобального економічного критерію при оптимізації виробничої програми підприємства 111
3.6 Оцінка інноваційно-інвестиційних складових розвитку підприємства 113
3.7 Оцінка пріоритетності внутрішньовиробничих інвестиційних заходів та загроз на підприємстві 115
3.8 Оцінка сталого розвитку підприємств 116
3.9 Оцінка рівня розвитку підприємств 118

ЛІТЕРАТУРА 120

ДОДАТКИ 132

А. Властивості середньої арифметичної (простої) 133
Б. Властивості дисперсії (простої) 134
В. Властивості коваріації 136
Г. Властивості регресії 138
Д. Процедура перевірки статистичних гіпотез 141
Е. Алгоритм кластеризації «ФОРЕЛЬ» 155
Ж. Принцип балансу змінних 163

ПЕРЕДМОВА

Традиційно вважається, що використання математики і статистики в економічній науці проявляється в одержанні тільки кількісних характеристик об'єктів, що вивчаються. Таке розуміння виглядає вкрай спрощеним, оскільки кількісні визначеності завжди пов'язані з якісними властивостями. Конкретні дослідження проводяться на різних рівнях: на рівні загальної економічні теорії, спеціальних теорій, наприклад, економіки підприємства та інших, тому важливим завданням є вибір специфічного кількісного апарату, математико-статистичних методів і моделей для кожного рівня окремо.

В історії економічної науки постійно точиться дискусія щодо співвідношення якісного і кількісного аналізу. При цьому пропорція між ними повільно, але постійно змінюється на користь кількісних методів. Так, ідеї емпіричного природознавства суттєво вплинули на розвиток економічної науки, сприяючи проникненню в неї кількісного аналізу, зокрема, вимірювання, математико-статистичного моделювання та прогнозування.

Саме вимірювання, а точніше, неможливість безпосереднього вимірювання в метричній шкалі деяких економічних ознак наштовхнуло автора даної монографії на необхідність застосовувати в процесі їх аналізу багатовимірних статистичних і математичних методів, зокрема, таксономічного, дискримінантного, факторного аналізу, неметричного шкалювання.

До таких прихованих (латентних) ознак відносяться рівні: ефективності праці службовців, фінансового стану, конкурентоспроможності, інвестиційної привабливості, сталого та інвестиційно-інноваційного розвитку підприємств, а також конкурентоздатності продукції, глобального економічного критерію оптимальності виробничої програми, пріоритетності внутрішньовиробничих інвестиційних заходів і загроз на підприємстві та ін.

Перші публікації автора за вказаною тематикою з'явились усередині 80-х років XX ст. [121; 135]. Окрім нього, досить активно застосовували ідеї і методи оцінки латентних ознак в економіці Ю.А. Єгупов [28; 29; 89; 135], А.Ю. Чорний [107-109] та ін. З того часу послідовників даного напряму аналізу прихованих властивостей економічних об'єктів суттєво побільшало. Досить зазначити, що вказані підходи використані в двох захищених докторських дисертаціях (О.Г. Янкового, О.М. Гончаренко) й у більш ніж десятку кандидатських дисертацій (див. автореферати у списку літератури).

Перелік наукових публікацій, дисертаційних робіт, захищених за період 1993-2015 рр. під безпосереднім керівництвом автора, під керівництвом провідних викладачів кафедри економіки підприємства, а також інших кафедр Одеського національного економічного університету, надають підстави вважати, що у даному вищому навчальному закладі існує наукова школа дослідження латентних ознак економічних об'єктів.

Ми сподіваємось, що даний напрямок економічної науки буде й надалі успішно розвиватись, залучаючи в ряди своїх прихильників нових молодих дослідників прихованих властивостей економічних об'єктів. А перелік останніх буде ще більше розширюватись за рахунок проникнення ідей і методів

оцінювання латентних ознак у всі сфери і галузі економіки. Мета цієї монографії полягає в систематизації наукових напрацювань, пов'язаних з поняттями, методами і напрямами оцінювання прихованих властивостей економічних об'єктів, які були здійснені представниками зазначеної наукової школи. Структура книги наступна.

У першому розділі наводяться поняття й теоретичні засади латентних ознак в економіці, визначається їх місце серед якісних і кількісних властивостей об'єктів. Доводиться, що для латентних ознак економічних об'єктів можливий дрейф уздовж шкал вимірювання на основі застосування оцінок, отриманих за допомогою багатовимірних статистичних і математичних методів. Іншими словами, латентні ознаки, які традиційно вимірювались у порядковій шкалі, можуть бути оцінені в метричній шкалі інтервалів, набуваючи тим самим умовно кількісний характер. Це відкриває додаткові широкі можливості для досліджень, адже в процесі аналізу прихованих властивостей економічних об'єктів можна замість градацій «більше – менше», «краще – гірше» оперувати умовно кількісними висновками «на скільки більше», розраховувати величину потенційних резервів і т.п.

Другий розділ монографії присвячений методам оцінки латентних ознак економічних об'єктів та їх класифікації. В ньому розглядаються три основні групи методів: 1. Експертні. 2. Математико-статистичні. 3. Змішані (комбіновані). По кожній з виділених груп надається історична характеристика з указівкою найбільш видатних учених, що зробили помітний внесок у їх становлення й розвиток. Наводиться також короткий формальний апарат запропонованих методів без строгих математичних доведень відомих положень з посиланням на їх програмне забезпечення.

Третій розділ книги представляє собою огляд захищених дисертацій та наукових публікацій, в яких застосовувались методи оцінювання латентних ознак в економіці України за останні роки. Це оцінка трудових латентних показників, фінансового стану, конкурентоспроможності, інвестиційної привабливості рівня і сталого розвитку підприємств. А також результати оцінки глобального економічного критерію в процесі оптимізації виробничої програми, пріоритетності інвестиційних заходів та загроз, інноваційно-інвестиційних складових розвитку окремого підприємства, конкурентоздатності його продукції. У додатках міститься корисний матеріал, який за різними причинами не увійшов до основного тексту роботи: властивості описових статистик, процедура перевірки гіпотез, алгоритм кластеризації «ФОРЕЛЬ», принцип балансу змінних.

На завершення хочеться висловити щиру подяку кандидатам економічних наук Шикіній Ользі Володимирівні та Малютенку Олександру Юр'євичу, які активно посприяли швидкій публікації даної книги.

Зав. кафедри економіки підприємства ОНЕУ,
д.е.н., професор, заслужений діяч науки і
техніки України О.Г. Янковий
травень 2015 р.

1. СУТНІСТЬ ЛАТЕНТНОЇ ОЗНАКИ В ЕКОНОМІЦІ

1.1 Класифікація ознак і показників економічних об'єктів

Сучасна економічна наука вивчає різноманітні відносини між виробничо-фінансовими системами (об'єктами) та між об'єктами і зовнішнім середовищем, що виникають в процесі здійснення ними виробничих, комерційних, інвестиційно-інноваційних, соціальних, екологічних та інших завдань на мікро-, мезо- і макрорівнях управління. Прикладом економічних об'єктів мікрорівня є окремі робітники, працівники підприємств, їх підрозділи та служби, а також самі підприємства – суб'єкти господарювання. На мезорівні у якості економічних об'єктів виступають регіони, кластери, галузі і т.п. Макрорівень представлений такими об'єктами управління, як народне господарство країни, економічний комплекс міждержавних об'єднань типу ЄС, Митний Союз, світове господарство.

Кожен економічний об'єкт на будь-якому рівні управління характеризується певною множиною ознак – властивостей, особливих рис, якостей, котрі відображають природну сутність даного об'єкта. Так, робітники підприємств мають прізвище, ім'я, по батькові, стать, сімейний стан, професію, рівень кваліфікації, величину оплати праці, стаж роботи тощо. Самі підприємства характеризуються назвою, галузевою приналежністю, економічним призначенням продукції, що виробляється, місцем розташування, розміром, обсягом виконаних робіт (послуг), об'ємом прибутку (збитків), рентабельністю, кредиторською і дебіторською заборгованістю і т.п. Аналогічно можна навести безліч ознак (властивостей) економічних об'єктів мезорівня та макрорівня.

Досліджуючи існуючі в сучасній науковій літературі класифікації ознак економічних об'єктів, слід відмітити, що в більшості з них виділяються два їх класи: атрибутивні й кількісні ознаки. Так, Є.В. Мізіна вказує: «Ознаки поділяють на кількісні і атрибутивні (словесні). Прикладами кількісної ознаки є – стаж роботи, заробітна плата, атрибутивної – наявність вищої (середньої) освіти» [66, с.5].

У роботі [70] можна зустріти наступне ствердження: «Статистика розрізняє два види ознак угрупування: кількісні, що мають цифрове вираження (наприклад, вік, розмір заробітної плати тощо) і якісні (атрибутивні), які не можуть бути виражені цифрами (наприклад, стать, національність і т.д.)».

Аналогічно виглядає така класифікація: «Ознаки поділяються на атрибутивні та кількісні. Атрибутивними називаються ознаки, що мають смислове значення (стать, посада). Кількісними називаються ознаки, які мають числове значення (середній бал, вік)» [49].

При цьому атрибутивні ознаки в економіці фактично ототожнюються з якісними ознаками, а кількісні ознаки зазвичай розглядаються лише з позицій звичайних не масштабованих ознак. Так, у статистичному словнику даються наступні визначення: «Ознака якісна (атрибутивна) – ознака, окремі значення якої виражаються у вигляді понять, найменувань Наприклад, професія робітника (тесляр, слюсар), рівень освіти (начальне, середнє, вище)». І далі:

«Ознака кількісна – ознака, окремі значення якої мають кількісне вираження, наприклад, поголів'я корів: 300 голів, 450 голів; вік людини: 5 років, 10, 50 і т.п.» [94, с.382].

Викладений вище традиційний підхід до класифікації ознак економічних об'єктів можна графічно представити в такий спосіб (рис. 1.1).

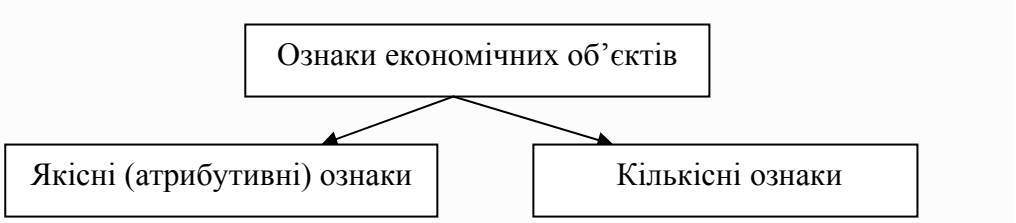


Рис. 1.1. Домінуюча в сучасній науковій літературі концепція щодо класифікації ознак економічних об'єктів

При дослідженні економічних об'єктів поряд з поняттям «ознака» часто використовується поняття «показник», під яким розуміють міру, тобто єдність якісного і кількісного відображення певної властивості соціально-економічного явища чи процесу [93; 97]. Якісна сторона показника відображає сутність властивості (ознаки) в конкретних умовах місця і часу, а кількісна – розмір, абсолютну чи відносну (середню) його величину. Таким, наприклад, є показник об'єму товарної продукції даного підприємства за конкретний період; величина ВВП країни в 2015 р.; розмір заробітної платні конкретного працівника підприємства за певний місяць і т.п. Отже, кожен конкретний показник економічного об'єкта має: 1) кількісну; 2) якісну; 3) просторову; 4) часову визначеність. Число, позбавлене хоча б одної з указаних характеристик, не є показником.

Показники економічних об'єктів можуть бути класифіковані за різними ознаками. За способом обчислення показники поділяються на:

- первинні – визначені шляхом спостереження, зведення та угрупування даних і надані у формі абсолютних величин;
- похідні – обчислені на базі первинних показників у формі середніх чи відносних величин.

За ознакою часу показники поділяються на:

- інтервальні, які характеризують властивості об'єкта за певний період часу;
- моментні, що відображають ознаки об'єкта на певний момент часу.

Показники бувають взаємно-обернені. Вони характеризують одну й ту ж властивість об'єкта і їх добуток дорівнює одиниці. Взаємно-обернені поділяються на:

- прямі показники – зростають із збільшенням властивості;
- обернені показники – зменшуються із збільшенням ознаки.

В залежності від точки зору дослідника вирізняють показники:

- індивідуальні, які характеризують властивості окремого економічного об'єкта – робітника, підприємства, регіону, країни;

- групові (загальні), що відображають ознаки групи (сукупності) однорідних об’єктів (працівників підприємства, господарств краю, підприємств галузі, регіонів країни, країн світу);
- узагальнюючі, які характеризують типовий рівень властивості всієї групи (сукупності) досліджуваних об’єктів, наприклад, середні величини.

Групові (загальні) показники називають також зведеними.

Існує класифікація показників на:

- об’ємні, що відображають обсяг даної кількісної ознаки в групі (сукупності) об’єктів, наприклад, чисельність промислово-виробничого персоналу даного підприємства в конкретному році;
- якісні (не плутати з атрибутивними ознаками), котрі показують об’єм досліджуваної кількісної властивості в розрахунку на один досліджуваний об’єкт (продуктивність праці одного робітника, ціна одиниці продукції даного виду тощо).

Викладена вище загальноприйнята класифікація показників економічних об’єктів представлена на рис. 1.2.

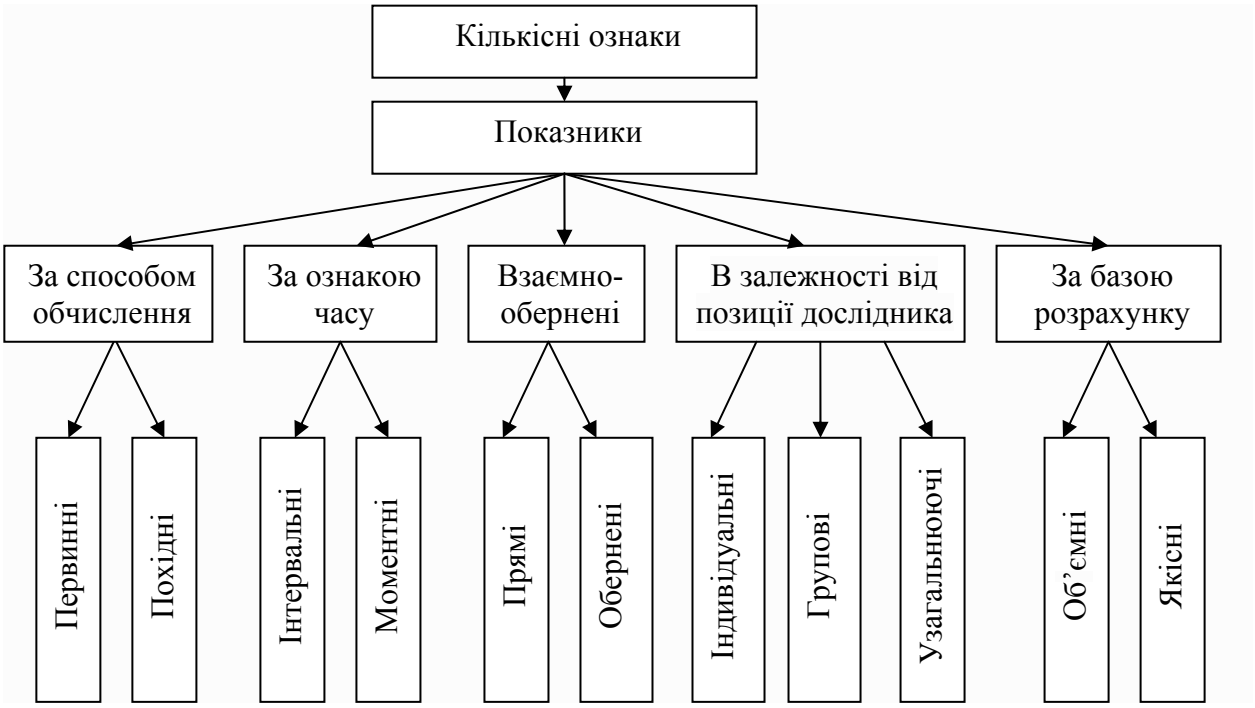


Рис. 1.2. Загальноприйнята класифікація показників економічних об’єктів

Таким чином, зіставляючи характеристики обговорюваних понять, ми приходимо до наступного висновку: ознаки представляють собою більш широкі й абстрактні властивості економічного об’єкта в порівнянні з його показниками. Дійсно, як було показано вище, ознаки бувають якісні (атрибутивні) та кількісні, в той час як показники завжди мають кількісні властивості об’єктів. Ознака, зазвичай, розглядається без конкретної прив’язки до певного місця і часу, на відміну від показника, для якого ці характеристики є іманентними. Хоча, слід відмітити, що в повсякденній економічній практиці, а

також у багатьох наукових дослідженнях поняття «ознака» і «показник» часто ототожнюються й вживаються як синоніми.

Окреме значення ознаки (показника) називається варіантом, а кількість об'єктів, які мають однакове значення варіанту називається частотою. Наприклад, за гендерною ознакою можна визначити статеву структуру працівників даного підприємства, розділив чисельність жінок і чоловіків (частоти обох дихотомічних варіантів) на загальну чисельність персоналу.

Варіанти якісних (атрибутивних) ознак варіюють лише у просторі (від одного економічного об'єкта до іншого). Такими є списки робітників окремих цехів підприємства, галузева приналежність підприємств міста, регіону. А варіанти показників, в основі яких лежать кількісні ознаки, варіюють як у просторі, так і в часі, тобто для одного й того ж об'єкта в різні інтервали чи моменти часу. Так, звітні показники прибутку (збитків), чисельності персоналу, розміру капіталу та інші техніко-економічні ознаки, взяті, наприклад, за рік, змінюються при переході від одного підприємства до іншого. Також спостерігається певна варіація вказаних показників і для даного підприємства по місяцям, по кварталам, по рокам.

Спостерігаючи і вивчаючи два види варіації досліджуваних ознак, у першому випадку створюють ряди розподілу, а в другому – часові ряди (ряди динаміки). У свою чергу, ряди розподілу бувають дискретні, коли окремі варіанти вирізняться один від одного на постійну величину (зазвичай на одиницю), й безперервні, коли окремі варіанти вирізняться один від одного на будь-яку величину. А часові ряди діляться на інтервальні, коли ряд динаміки характеризує результат деякого процесу за певні проміжки часу, і моментні, коли ряд динаміки відображає стан сукупності об'єктів на конкретні моменти часу (рис. 1.3).

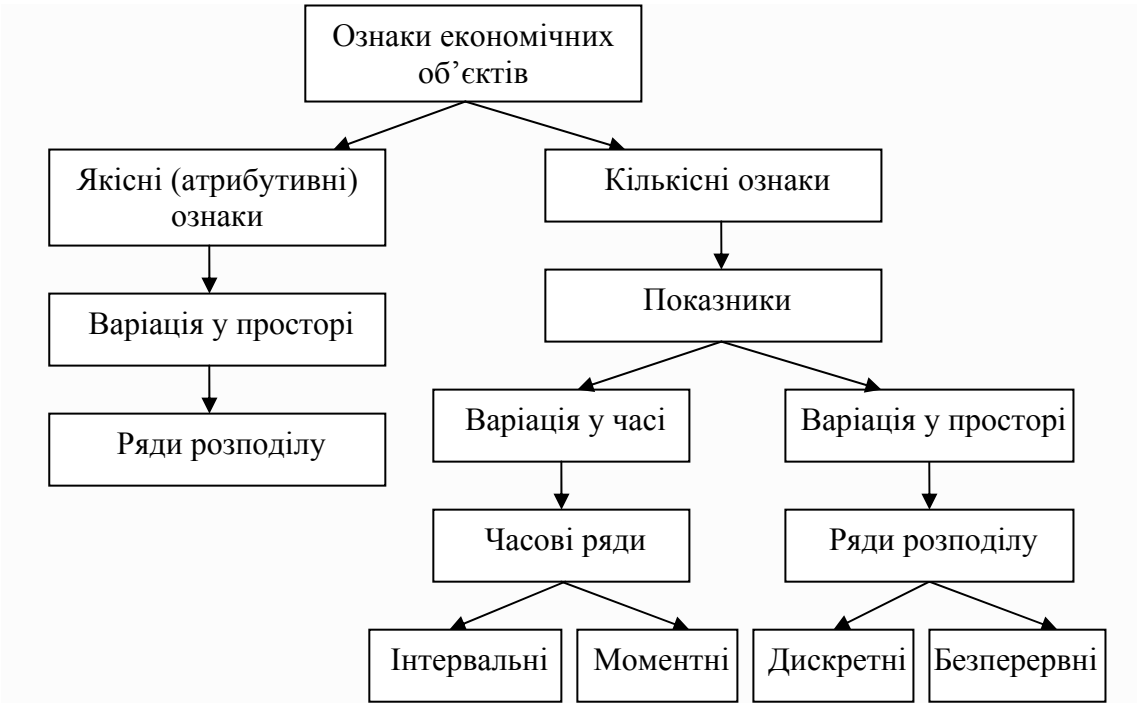


Рис. 1.3. Види варіації ознак і показників економічних об'єктів

Наведені класифікації ознак і показників економічних об’єктів, а також різновиди їх варіації, представлені на рис. 1.1–1.3, детально розглядаються в теорії статистики, використовуються в прикладній і економічній статистиці та в інших областях даного напряму суспільних наук [128, с.27-28]. Вони достатньо детально вивчені, теоретично обґрунтовані і мають широке практичне впровадження у різних галузях наукових досліджень.

У той же час, ми вважаємо, що традиційний, визнаний тепер у науковій літературі підхід до класифікації ознак економічних об’єктів (рис. 1.1) є занадто абстрактним, загальним, який суттєво збіднює уявлення про фактичний спектр властивостей об’єктів в економіці і, відповідно, звужує реальний обрій наукових досліджень у даній області. Він не достатньо деталізований, не враховує підгрупи самих якісних (атрибутивних) і кількісних ознак. Зокрема, повністю випадає з поля зору дослідників такі важливі види економічних ознак, як *латентні* (від лат. *latent* – прихований) та *масштабовані* кількісні ознаки й показники.

Задля усунення вказаних вище недоліків домінуючої в сучасній науковій літературі концепції щодо класифікації ознак розглянемо їх угрупування, поява, формування та історичний розвиток якого для об’єктів будь-якої природи, в тому числі і економічних, базується на теорії вимірюванні. Цей напрямок активно розробляється, перш за все, в таких науках, як психологія, метрологія, соціологія, політологія [19; 40; 52; 71; 72; 77; 116; 117].

1.2 Шкали вимірювання ознак економічних об’єктів

Під вимірюванням взагалі розуміють приписування чисел ознакам об’єктів у відповідності з певними правилами. Вимірювання є експериментальною процедурою, результатом активної взаємодії дослідника з об’єктом пізнання. Перехід від опису об’єкта до його виміру завжди означав перехід до точного знання. Можна констатувати, що вимірювання зробило природничі науки такими, якими вони існують сьогодні, а проникнення вимірювальних процедур у гуманітарні галузі знань наближають їх до точних наук. З цього приводу відомий радянський фахівець в області математичного програмування І.Я. Бірман писав: «Найбільш загальні закономірності й закони економічних явищ і процесів з’ясовуються шляхом якісного аналізу. Але конкретне вираження їх можливо лише за допомогою міри та числа…» [5, с.9].

Мета вимірювання – отримати числову модель, використання якої могло б замінити дослідження самого об’єкта. Це можливо лише тоді, коли властивості моделі відповідають властивостям об’єкта, тобто відносини між числами, що утворюють числову модель, адекватні відносинам між досліджуваними ознаками об’єкта.

Вимірювання дозволяє перевести відмінності між ознаками у відомі, зрозумілі кожному категорії, які називаються числами, і будь-яка вимірювальна процедура, в кінцевому рахунку, обов’язково повинна закінчитися числом. Проте, число, приписане ознаці, ще ні про що не свідчить, якщо не відомі

правила, за якими відбувалося це приписування. Число набуває сенс тільки в тому випадку, якщо відома система, в якій відбувалося вимірювання. Тобто вимірювання – це процедура, за допомогою якої вимірюваний об'єкт порівнюється з деяким еталоном і отримує числове вираження у певній системі – шкалі [119].

Отже, під вимірюванням розуміють особливу процедуру, в результаті якої виникає числова модель досліджуваних ознак об'єкта. При вимірюванні встановлюється відповідність між властивостями об'єкта і властивостями зіставлених їм чисел. Набір властивостей об'єкта і зіставлених їм чисел називають шкалою (властивості об'єкта трактуються тут дуже широко, зокрема, під набором властивостей розуміються також і різні ступені інтенсивності однієї властивості).

Під шкалою розуміють знакову систему, для якої задано деяке однозначне відображення, що ставить у відповідність реальним об'єктам той чи інший елемент шкали. Процес шкалювання полягає в конструюванні шкали за певними правилами і включає в себе два основні етапи. На першому етапі здійснюється збір даних, на другому – їх обробка та аналіз, за результатами якого вибудовується оціночна шкала.

Детальний огляд історичного, філософського, гносеологічного, методологічного і практичного аспектів вимірюванні наведено в роботах [68; 69; 78; 116]. Відмітимо лише той факт, що теорія вимірювання на основі шкалювання була започаткована американським психологом С. Стівенсом у 30-40-і роки XX ст. [158].

За Стівенсом, вимірювання можливо, перш за все, тому, «що існує ізоморфізм між властивостями числових рядів і емпіричними операціями, які ми можемо здійснювати з об'єктами» [95, с.51]. Вимірювання він розумів як процес співвіднесення емпіричної системи з деякою числовий системою. Стівенс висунув чотири типи таких числових систем, які зумовили чотири відповідні шкали вимірювання:

1) шкала найменувань (номінальна шкала);
2) шкала порядку (порядкова або ординальна шкала);
3) шкала інтервалів (абсолютна шкала);
4) шкала відносин (пропорційна шкала).

Ідеї С. Стівенса найшли подальший розвиток у роботах його колег і послідовників: К. Кумбса, Е. Найгела, Н. Кемпбелла, П. Суппеса, Дж. Зіннеса [96; 156], які показали, що кожна шкала характеризується відповідними числовими властивостями. Якщо проводиться вимірювання за визначеною шкалою, то це означає, що здійснюється ізоморфізм між числовою системою шкали і досліджуваними величинами. Числа в цих шкалах можуть свідчити про ступінь вираження вимірюваної ознаки, про кількісні відмінності між об'єктами і т.п.

В залежності від типу шкали до чисел можуть бути застосовані ті чи інші математичні операції. Наведена вище класифікація Стівенса є угрупуванням за ростом допустимих математичних операцій, тобто для даної шкали можуть

бути застосовані всі математичні операції попередніх шкал, але не навпаки. Розглянемо їх дещо докладніше.

1. Шкала найменувань (номінальна шкала) використовується для ідентифікації об'єктів. На ній визначено дві операції – «рівно» і «нерівно». Значення номінальної шкалі надають можливість вирізняти один об'єкт від інших. Номінальна шкала служить передумовою всіх процедур шкалювання і дозволяє ідентифікувати об'єкти, включені в один клас. Градації шкали найменувань – множина еталонів якісної класифікації властивостей об'єктів. Наприклад, це можуть бути власні прізвища, імена, стать, сімейний стан, професія робітників; назва, галузева приналежність, економічне призначення продукції, місце розташування підприємства і т.п. Вони, зазвичай, упорядковані за алфавітом і розглядаються ізольовано одне від одного.

Особливим підвидом шкали найменувань є дихотомічна (альтернативна) шкала, яка складається всього з двох взаємовиключних значень (1; 0). Стать людини є типовою дихотомічною змінною.

Ознаки економічних об'єктів, які можуть бути виміряні лише в номінальній шкалі, називаються якісними, атрибутивними (від лат. *attributio* – приписувати, наділяти). Якісна, атрибутивна ознака – це необхідна, суттєва, невід'ємна властивість об'єкта (табл. 1.1).

Таблиця 1.1

Характеристика ознак економічних об'єктів, виміряних у номінальній шкалі

Класифікація ознак	Приклади ознак	Математичні операції з градаціями шкали
ознаки з двома альтернативними градаціями	стать, сімейний стан, економічне призначення продукції	порівняння та ідентифікація
ознаки з декількома градаціями	список працівників, перелік робочих професій, галузева структура промисловості	

За шкалою найменувань не можна сказати, що один об'єкт більше або менше іншого, на скільки одиниць вони вирізняються, або у скільки разів. Можливі лише операції порівняння та ідентифікації: «вирізняється – не вирізняється». Так, робітники механічного цеху промислового підприємства можуть бути порівняні за ознакою «професія» й ідентифіковані як токарі, слюсарі, фрезерувальники, наладчики тощо.

2. Шкала порядку (порядкова або ординальна шкала) дозволяє не тільки ідентифікувати, але й порівнювати та впорядковувати об'єкти за певною ознакою. Порядкова шкала допускає лише монотонні перетворення, тобто такі, що не порушують порядок визначення вимірюваних величин. Це означає, що їх можна замінити іншими числами зі збереженням колишнього порядку. Саме тому шкали даного типу називають порядковими. Відмінна особливість порядкової шкали – усі її значення впорядковані за допомогою понять «більше – менше», «краще – гірше», «вище – нижче» і т.п.

Зазвичай, порядкова шкала виражає думки експертів щодо ступеня інтенсивності проявлення досліджуваної ознаки економічного об'єкта, яку з певних причин неможливо виміряти в більш потужних (за потенційно можливими математичними операціями) метричних шкалах інтервалів і відносин. Тобто експерт може сказати (і обґрунтувати), що один показник якості продукції більш важливий, ніж інший, перший технологічний об'єкт більш небезпечний, ніж другий, і т.д. Але він не в змозі сказати, на скільки більш важливий або, відповідно, у скільки разів більш небезпечний.

Досить часто на практиці вживається різновид порядкових шкал – рангові шкали. Справа в тім, що експертів часто просять здійснити ранжирування (впорядкування) об'єктів, тобто розташувати їх у порядку зростання (убування) інтенсивності досліджуваної властивості, яка цікавить організаторів експертизи. Ранг – це номер об'єкта в упорядкованому ряду значень ознаки у різних об'єктів. Рангова шкала задається позитивними числами 1, 2, 3, …, *n*, що відображають зниження або зростання вимірюваної властивості, і яких може бути стільки, скільки досліджуваних об'єктів. Але з цими числами не можна робити звичні арифметичні операції. Вони лише визначають відносну інтенсивність властивості, але не абсолютну її величину.

Цінність шкал даного типу полягає в тім, що вони встановлюють порядок економічних об'єктів за певною ознакою, а головний недолік – цей порядок не є метричним. Числа, присвоєні об'єктам у цій шкалі, свідчать про ступінь вираження вимірюваної ознаки у них, але при цьому рівні різниці чисел не будуть означати рівних різниць у кількостях вимірюваної властивості. Наприклад, можна ранжирувати рівень кваліфікації робітників механічного цеху підприємства за допомогою 6-и розрядної тарифної сітки, але зрозуміло, що кваліфікація окремих індивідуумів усередині певного розряду може бути різною.

Окрім того, відомо, що експерти при впорядкуванні економічних об'єктів за певною невимірною ознакою враховують одну або декілька характеристик – так званих чинників-симптомів, які зазвичай вимірюються в метричних шкалах інтервалів і відносин (табл. 1.2).

Таблиця 1.2

Характеристика ознак економічних об'єктів, виміряних у порядковій шкалі

Класифікація ознак	Приклади ознак	Математичні операції з градаціями шкали
упорядковані за одним чинником-симптомом	ранги підприємств за величиною їх: 1) прибутку; 2) обсягу продукції; 3) капіталу	порівняння та упорядкування
упорядковані за декількома чинниками-симптомами	ранги підприємств за рівнем їх: 1) фінансового стану; 3) інвестиційної привабливості; 4) конкурентоздатності	

3. Шкала інтервалів (абсолютна шкала) відноситься до класу метричних шкал. На відміну від шкали порядку, вона дозволяє не тільки ранжирувати

об'єкти за досліджуваними ознаками, але і задавати відомі інтервали між пунктами цієї шкали. Інтервальна шкала допускає лінійні перетворення виду

$$Z = aX + в, \tag{1.1}$$

де X – вихідне значення ознаки;
Z – перетворене значення ознаки;
a – параметр масштабу (додатне число);
$в$ – довільна точка відліку (додатне або від'ємне число).

Зміна величини a призводить до зміни масштабу шкали, а зміна значення $в$ призводить до зсуву за шкалою праворуч або ліворуч, тобто положення нуля на інтервальній метричній шкалі не зафіксовано. Метрична шкала інтервалів служить основою масштабованих показників економічних об'єктів.

На відміну від двох попередніх шкал, у цій шкалі існує одиниця виміру – реальна, або умовна, за допомогою якої можна встановити кількісні відмінності між об'єктами щодо вимірюваної ознаки. Однакові різниці чисел у цій шкалі будуть означати рівні відмінності в кількостях вимірюваної властивості у різних об'єктів, або у одного і того ж об'єкта в різні інтервали (моменти) часу. Однак, те, що одне число виявляється в кілька разів більше іншого не обов'язково говорить про такі ж відносини в кількостях вимірюваних ознак.

У шкалі інтервалів може бути задіяна вся числова вісь, але при цьому нуль не вказує на відсутність вимірюваної властивості, тому нульова точка часто є умовною, наприклад, як у шкалі стандартизованих чисел. У цьому випадку формула (1.1) набуває вигляду:

$$Z = \frac{X - \overline{X}}{\sigma} = \frac{1}{\sigma} X - \frac{\overline{X}}{\sigma}, \tag{1.2}$$

де $\overline{X}$ – середнє значення ознаки;
σ – стандартне відхилення ознаки.

Порівнюючи вираження (1.1) і (1.2), легко бачити, що у разі стандартизації ознак $a = 1/\sigma$, $в = -\overline{X}/\sigma$. При цьому $Z = 0$ при $X = \overline{X}$. Стандартизована шкала застосовується в багатьох математико-статистичних процедурах для уникнення масштабу вимірюваних змінних і приведення їх до безрозмірного виду.

Відомо, що окрім стандартизації в таких процедурах використовують також й інші методи масштабування, зокрема такі, що враховують екстремальні значення показників. Наприклад,

$$Z = \frac{\left| X - X_{\max(\min)} \right|}{X_{\max} - X_{\min}}, \tag{1.3}$$

де $X_{\max(\min)}$ – найбільше (найменше) значення показника X.

Окремим випадком масштабування ознак економічних об'єктів є їхнє нормування за формулою

$$Z = \frac{X}{X_{\text{НОРМ}}}, \tag{1.4}$$

де $X_{\text{НОРМ}}$ – нормативне значення змінної X.

Ознаки економічних об'єктів, які можуть бути виміряні в метричній інтервальній шкалі, називаються кількісними масштабованими ознаками (табл. 1.3).

Таблиця 1.3

Характеристика ознак економічних об'єктів, виміряних у метричній інтервальній шкалі (з умовною нульовою точкою)

Класифікація ознак	Приклади ознак	Математичні операції з градаціями шкали
обмежені	стандартизовані	порівняння та впорядкування, лінійні перетворення виду $Z = aX + в$
необмежені	нормовані	

Вони застосовуються в основному для елімінування масштабу одиниць вимірюваних показників, приведення всіх їх значень у заданий інтервал, наприклад, від -3 до 3 за формулою (1.2), від 0 до 1 за формулою (1.3) і т.п.

4. Шкала відносин (пропорційна шкала) теж відноситься до класу метричних шкал і допускає лінійні перетворення виду

$$Y = aX. \tag{1.5}$$

Шкала відносин, на відміну від інтервальної шкали, має точку нульового відліку. В ній також існує одиниця виміру, за допомогою якої об'єкти можна впорядкувати щодо вимірюваної властивості і встановити кількісні відмінності між ними. Особливістю шкали відносин є те, що у ній обов'язково (принаймні, теоретично) є нуль, який свідчить про абсолютну відсутність вимірюваної ознаки.

Більшість нині існуючих фізичних шкал (довжини, маси, часу, температури тощо) є яскравими прикладами шкал відносин. Наприклад, довжина може вимірюватися в метрах, футах, дюймах – це визначається параметром масштабу a. Більшість фізичних ознак в економіці вимірюють саме за шкалами відносин.

Якщо $a = 1$, то шкалу відносин називають абсолютною. В якості прикладу такої шкали приводять, зазвичай, шкалу рахунку (якщо рахувати одиницями, а не десятками, сотнями і т.д.). При цьому дослідник має справу з ідеальною метричною шкалою, що нагадує шкалу інтервалів, але з однією перевагою: відлік у цій шкалі починається не з довільної точки, а з експериментально встановленого нульового пункту. Для таких шкал застосовні всі, без винятку, математичні операції з числами, так як можна визначити, на скільки або у скільки разів даний пункт на шкалі перевищує (нижче) іншого. Подібні

метричні шкали прийняті в точних науках (фізиці, хімії), де нульовий пункт експериментально зафіксований.

Усі ознаки метричної шкали відносин за характером варіації її числових значень діляться на два класи: дискретні та неперервні. При дискретній варіації ознак економічних об'єктів окремі варіанти вирізняються один від одного на заздалегідь відому постійну величину, зазвичай, на одиницю. Такою, наприклад, є варіація чисельності працівників підприємства на кінець року, кількості самих підприємств на території району тощо. При неперервній варіації ознак окремі варіанти вирізняються один від одного на будь-яку величину: цілу або дрібну. Класичною неперервною ознакою є час. Тому такі ознаки, як вік, стаж роботи працівника відносяться до неперервних.

Ознаки економічних об'єктів, які можуть бути виміряні в метричній шкалі відносин, називаються кількісними (табл. 1.4).

Таблиця 1.4

Характеристика ознак економічних об'єктів, виміряних у метричній шкалі відносин (з фіксованою нульовою точкою)

Класифікація ознак	Приклади ознак	Математичні операції з градаціями шкали
дискретні	чисельність працівників, кількість одиниць виробленої продукції	порівняння та впорядкування, лінійні перетворення виду $Z = aX + в$, $Y = aX$
неперервні	вік, стаж роботи	

Як відмічалось вище, між самими шкалами теж існують відносини порядку. Кожна з перерахованих шкал є шкалою більш високого порядку по відношенню до попередньої шкали. Так, вимірювання, здійснені в шкалі відносин, можна перевести в шкалу інтервалів, з шкали інтервалів – у шкалу порядку, але зворотна процедура неможлива, тому що при переході до шкал більш низького порядку частина інформації (про одиниці виміру, кількості властивостей) втрачається.

Тим не менш, це не завжди означає, що шкали більш високого порядку кращі по відношенню до шкал більш низького порядку, а в ряді випадків – навіть, навпаки. Нарешті, існують ознаки об'єктів, які можна вимірювати будь-якою шкалою, наприклад, вік працівників. І такі, для вимірювання яких підходить тільки одна шкала, як, наприклад, їх стать. На вибір вимірювальної шкали, таким чином, можуть впливати багато факторів – як переваги самої шкали, так і специфіка об'єкта вимірювання.

У теперішній час отримала поширення точка зору відомого фахівця з теорії вимірювання У. Торгерсона, який не відносить номінальну шкалу до систем вимірювання, а розглядає її як просто реалізацію класифікації. З вимірюванням він пов'язує процес, для якого важливі три елементи: порядок, початок відліку і одиниця виміру. В залежності від наявності або відсутності цих елементів виникають різні шкали вимірювання. Важливо відзначити, що за Торгерсоном, порядок притаманний будь-якому вимірюванню [160].

Деякі дослідники, навпаки, виділяють більшу кількість шкал вимірювання у порівнянні с класичними чотирма шкалами С. Стівенса. Так,

А.І. Орлов пропонує користуватися шістьома типами шкал: 1) номінальною; 2) порядковою; 3) інтервалів; 4) відносин; 5) різностей; 6) абсолютною [67, с.43-49]. Тут звичайна метрична шкала відносин розбита на дві частинні шкали в залежності від параметра *a* у вираженні (1.5), а в шкалі інтервалів виділена спеціальна шкала різниць для вимірювання часу.

Окрім того, він пише: «Порядкова шкала і шкала найменувань – основні шкали якісних ознак. Тому в багатьох конкретних областях результати якісного аналізу можна розглядати як вимірювання за цими шкалами» [67, с.43]. Ми не розділяємо вказану позицію, виходячи з наступних міркувань.

На нашу думку, тільки ті ознаки економічних об'єктів, які можуть бути виміряні в номінальній шкалі, відносяться до якісних (атрибутивних). А ознаки, що вимірюються в метричних шкалах інтервалів і відносин, слід вважати кількісними ознаками. Що стосується ознак порядкової шкали, то, на наш погляд, їх треба відносити до умовних кількісних ознак. Їх умовність проявляється в даному випадку в неможливості застосування арифметичних дій до градацій порядкової шкали. Дійсно, не має сенсу підсумовувати чи перемножувати ранги (або оцінки експертів) властивостей досліджуваних економічних об'єктів.

На користь даної точки зору свідчить і той факт, що якісні (атрибутивні) ознаки економічних об'єктів є відносно постійними у часі і варіюють лише у просторі, тобто при переході від одного об'єкта до іншого. Кількісні ж ознаки, в тому числі й умовно кількісні, змінюються як у просторі, так і в часі – для одного і того ж об'єкта в різні інтервали часу, або на різні моменти часу (див. рис. 1.3).

Оскільки ознаки, оцінені експертами, вимірюються саме в порядковій шкалі, але при цьому вони варіюють як в просторі, так і в часі, то їх слід віднести до класу умовних кількісних ознак. Дійсно, сортність певного виробу конкретного підприємства може змінюватися по рокам під впливом внутрішніх та зовнішніх чинників. З часом ранги підприємств, визначені експертами за рівнем їх фінансового стану, інвестиційної привабливості, конкурентоздатності тощо, варіюють у більшу або в меншу сторону. Тобто їх відносна величина (кількісна характеристика), на відміну від якісних (атрибутивних) ознак, не залишається постійною, змінюється, що притаманно саме кількісним ознакам.

Методи шкалювання, які використовуються в окремих економічних дисциплінах, мають певну специфіку, пов'язану з особливими властивостями досліджуваних об'єктів. Наприклад, у маркетингу вони тяжіють до методів вимірювання, характерних для психології і соціології, оскільки часто базуються на результатах опитувань, анкетувань тощо. Ці методи зводяться, головним чином, до різновидів порядкової шкали, які пропонуються експерту (у даному випадку – респонденту), й умовно діляться на порівняльні й не порівняльні [120].

Порівняльні шкали припускають пряме порівняння розглянутих об'єктів. Дані порівняльних шкал мають властивості тільки порядкових і рангових величин. Тому порівняльне шкалювання також називають неметричним.

При застосуванні непорівняльних шкал кожен об’єкт оцінюється незалежно від інших. Такі шкали вважаються умовно метричними, а отримані дані – інтервально або рейтингово шкальованими.

Методи непорівняльного шкалювання використовують: 1) неперервні рейтингові шкали; 2) деталізовані рейтингові шкали (Лайкерта, семантичний диференціал, Степела та ін.)

Шкала Лайкерта застосовується тоді, коли від експертів потребують визначити ступінь згоди або незгоди для кожного набору тверджень про розглянуті об’єкти. Зазвичай, кожен пункт шкали у вигляді певного твердження має п’ять категорій для відповіді – від «абсолютної незгоди» до «повної згоди».

Семантичний диференціал – семибальна шкала з крайніми точками, які є протилежними оцінками. Експерти оцінюють об’єкти за декількома пунктами семибальних шкал, обмежених по краях протилежними прикметниками.

Шкала Степела будується на використанні значень від -5 до +5 без нейтральної (нульової) точки і, зазвичай, зображується вертикально. Експертів просять визначити, наскільки вірно або невірно кожен термін описує об’єкт шляхом вибору відповідного числа на шкалі відповідей. Чим вище число, тим ближче термін до опису об’єкта

1.3 Латентні ознаки економічних об’єктів та їх дрейф уздовж шкал вимірювання

Під латентною (прихованою) ознакою в економіці ми будемо розуміти таку властивість економічних об’єктів, яка може бути виміряною лише в порядковій шкалі, тобто на основі суджень експертів. Зазначимо, що для латентних ознак економічних об’єктів, які вимірюються в порядковій шкалі, зазвичай, виконуються наступні умови:

1) приховані властивості, за визначенням, не можуть бути виміряні безпосередньо, тобто за допомогою шкал більш високого порядку – шкал інтервалів і відносин;

2) ранжирування величини латентних ознак, зазвичай, проводиться за результатами експертної оцінки відповідних фахівців, які про силу проявлення прихованих властивостей судять на основі значень одного чи множини чинників-симптомів, що характеризують різні проявлення досліджуваних властивостей економічних об’єктів;

3) у процесі оцінювання сили проявлення латентної ознаки експертами застосовуються градації порядкової шкали типу «більше – менше», «ближче – дальше», «краще – гірше» тощо.

Так, при оцінці якості продукції (робіт, послуг) в кваліметрії (від лат. *gualitas* – якість, грецьк. *metro* – міра) часто застосовують порядкові шкали, що надає підставу розглядати вказану властивість товару як латентну. А саме, одиниця продукції оцінюється як придатна або не годна за наявністю в ній дефектів. При більш детальному аналізі якості продукції використовується шкала з трьома градаціями: 1) є значні дефекти; 2) мають місце тільки незначні дефекти; 3) немає дефектів. Іноді застосовують чотири градації: «є критичні

дефекти, що роблять неможливим використання продукції»; «є значні дефекти»; «присутні тільки незначні дефекти»; «немає дефектів».

Це приклади застосування експертами (працівниками відділу контролінгу підприємства) порядкових шкал якості продукції, побудованих за одним чинником-симптомом – за наявністю дефектів у виробах. У більш складних ситуаціях, наприклад, при визначенні сортності продукції в машинобудуванні, її якість оцінюється експертами за множиною показників (чинників-симптомів) у відповідності до діючих стандартів і технічних вимог. Серед них можна назвати: 1) гарантійний термін роботи; 2) кількість і вартість гарантійних безкоштовних для споживачів ремонтів у розрахунку на один виріб; 3) наявність рекламацій, їхня кількість і вартість; 4) кількість і сума штрафів за поставку неякісної продукції; 5) відсоток браку; 6) відсоток повернення продукції для виправлення дефектів; 7) відповідність сучасній моді; 8) наявність і рівень попиту на даний виріб.

У групі однорідних підприємств про рівень конкурентоздатності кожного судять на основі аналізу множини таких показників, як об’єм прибутку (збитку), конкурентна позиція на відповідному ринку, продуктивність праці, капіталовіддача, рентабельність активів тощо. Аналогічно, інвестиційну привабливість економічних об’єктів характеризують наступні чинники-симптоми: інвестиційний клімат країни, галузі, регіону; місце розташування; розмір і стан основних виробничих фондів; показники ділової активності і т. ін.

Таким чином, ознаки порядкової шкали, ранжирування величини яких проводиться за результатами експертної оцінки відповідних фахівців на базі значень множини чинників-симптомів, що характеризують різні проявлення досліджуваних властивостей економічних об’єктів, називаються латентними, прихованими

З урахуванням викладеного вище матеріалу (див. табл. 1.1–1.4) ми пропонуємо наступну класифікацію ознак об’єктів в економіці [118], в основу якої покладено різноманітні шкали вимірювання їх властивостей (рис. 1.4).

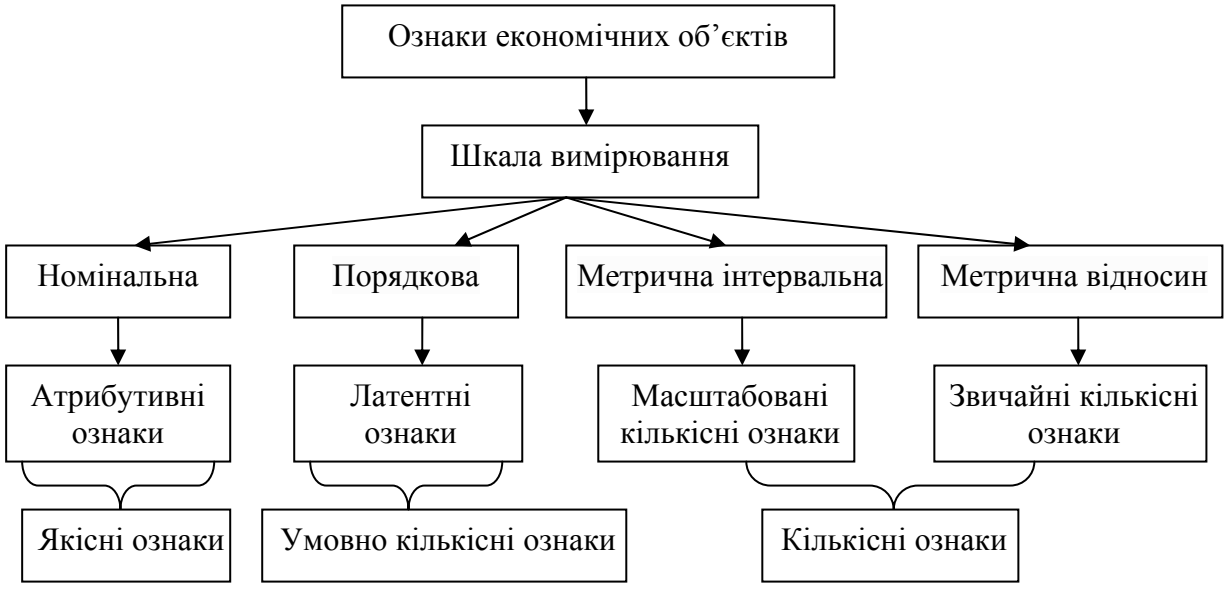


Рис. 1.4. Запропонована класифікація ознак економічних об’єктів

В процесі розвитку відповідної області знань тип шкали для одної і той же ознаки може змінюватись. Тому слід зазначити, що наведені приклади економічних ознак у табл. 1.1–1.4 та їх класифікація на рис. 1.4 не є чимось незмінним і зафіксованим у часі.

На наш погляд, з розвитком сучасної економічної науки, застосуванням у ній математико-статистичних моделей деякі ознаки можуть переходити з одної шкали до іншої (зазвичай, з ліва на право на рис. 1.4) у відповідності до процесу наближення даної галузі знань до точних наук [120].

Класичним історичним прикладом такого зрушення ознак від найбільш «слабких» до самих «сильних» шкал є вимірювання температури фізичного об'єкта. На зорі історії людської цивілізації в якості єдиного джерела вимірювання температури виступали особисті почуття. Приймаючи до уваги суб'єктивне почуття більшості населення можна було дійти одного з двох висновків: тепло чи холодно (градації дихотомічної номінальної шкали).

З розвитком і удосконаленням спостережливості люди навчились проводити вимірювання температури, зіставляючи об'єкт з незмінними або строго циклічними явищами. Для визначення температури таким явищем став агрегатний стан води в природі – крига, рідина, пар. Тут вже використовуються ранжирувані значення порядкової шкали – холодно, тепло, жарко.

Наступним кроком у розвитку процедури вимірювання температури був винахід температурних шкал, у яких здійснювався довільний вибір точки відліку – нуля. В шкалі Цельсія, абсолютно умовно, це температура замерзання води, довільний і масштаб (ціна поділки різна у шкалах Цельсія, Фаренгейта, Реомюра). Так, температура за Фаренгейтом пов'язана з температурою за Цельсієм співвідношенням $F = 32 + 1{,}8°C$. Ці шкали являють собою вже метричні інтервальні шкали вимірювання температури. Інтервальними є всі температурні шкали, окрім абсолютної шкали Кельвіна*, яка є метричною шкалою відносин.

Таким чином, упродовж історичного прогресу практики і науки вимірювання спостерігається перехід ознаки «температура» буд-якого фізичного об'єкта від простішої номінальної шкали найменувань до найбільш складної й точної метричної шкали відносин з фіксованою нульовою точкою.

Ми вважаємо, що подібна ситуація спостерігається зараз і в економіці. Справа в тім, що деякі латентні ознаки економічних об'єктів можуть бути оцінені за допомогою багатовимірних статистичних методів і представлені у вигляді масштабованих метричних показників інтервальної шкали. Тобто цілком можливим і реальним виглядає дрейф ряду ознак економічних об'єктів у бік точних наук.

---

*Часто вживане найменування термодинамічної температурної шкали. Названа на честь лорда Кельвіна (У. Томсона), який запропонував у 1848 р. принцип побудови температурної шкали, за початок відліку якої прийнятий абсолютний нуль (-273,15°*C*), одиницею відліку є 1 Кельвін (1*K* = 1°*C*).

Якщо масштабовані кількісні ознаки економічних об'єктів в основному є похідними від звичайних кількісних ознак на базі перетворень (1.2) – (1.4) і використовуються здебільшого в математико-статистичних процедурах для елімінування впливу одиниць їх виміру на результати таксономічного, кластерного, дисримінантного аналізу та інших багатовимірних методів, то латентні ознаки в буквальному сенсі цього слова оточують дослідника, особливо на рівні підприємства.

Прикладами порядкових шкал в економіці, що відображають співвідношення між прихованими властивостями окремих підприємств, є рейтинги рівня якості однойменної продукції різних товаровиробників, ранги їх сталого розвитку, фінансового стану, інвестиційної привабливості, конкурентоздатності, пріоритетності інвестиційних проектів і т.п.

Науковий термін «латентні ознаки» використовуються дослідниками для позначення складних багатогранних понять, які неможливо кількісно виміряти у метричній шкалі інтервалів і відносин. В силу певних обставин не вдається дати однозначну кількісну характеристику зазначених властивостей економічних об'єктів і про їх рівні зазвичай судять експерти за допомогою градацій порядкової шкали типу «краще – гірше», «більше – менше», «легше – тяжче» і т. ін.

Базою такої оцінки служить зовнішнє проявлення прихованої ознаки у вигляді множини значень чинників-симптомів, що характеризують її різні сторони. Наприклад, рівень науково-технічного прогресу на підприємстві відображають такі техніко-економічні показники, як фондоозброєність робітників, частка сучасних технологій у загальному технологічному циклі, ступінь комп'ютеризації виробництва і управління, матеріалоємність та енергоємність продукції та ін. [125, с.51, 52].

1.4 Роль оцінки латентних ознак в економічному аналізі та прогнозуванні

Проблеми кількісного визначення рівня шуканих прихованих властивостей досліджуваних об'єктів, особливо, промислових підприємств, завжди цікавили науковців та практиків. І цей інтерес різко підвищився в умовах ринкової економіки. Справа в тім, що своєчасні сигнали про негативні тенденції в розвитку таких ознак, як конкурентоспроможність, інвестиційна привабливість, фінансовий стан, сталий розвиток дозволяють прийняти відповідні заходи в тактичному антикризовому управлінні, а постійний їх моніторинг служить основою для розробки довгострокової стратегії забезпечення сталого розвитку підприємств у перспективі.

Головною метою оцінки латентної ознаки економічних об'єктів є одержання достовірної інформації для прийняття управлінських рішень, з одного боку, керівництвом підприємства про корегування концепції розвитку даної прихованої властивості і зміну стратегії в цьому напрямку діяльності, а, з іншого – зовнішніми користувачами (інвесторами, кредиторами, органами державної влади) при реалізації конкретних планів по відношенню до даного підприємства – злиття, поглинання, інвестування, надання позик тощо.

На рис. 1.5 представлені основні користувачі інформації, яка відображає рівень певних прихованих властивості підприємства. З нього видно, що в сучасному соціально-економічному середовищі існує достатньо велика кількість фізичних на юридичних осіб, зацікавлених в одержанні об'єктивної і точної інформації про рівень латентних ознак конкретного промислового підприємства.

Рис. 1.5. Класифікація користувачів інформації щодо прихованих властивостей підприємства

У разі успішного вирішення проблеми визначенні рівня ознаки, що вивчається, подальший аналіз включає наступні важливіші завдання, що потрібно розв'язати при дослідженні, моделюванні та прогнозуванні шуканої властивості підприємства:

- виявлення тенденцій і закономірностей у підвищенні (зниженні) латентної ознаки за певний період;

- аналіз чинників та «вузьких» місць, що негативно впливають на рівень прихованої властивості підприємства;
- визначення резервів, які підприємство може використати для підвищення (зниження) латентної ознаки, що досліджується;
- розробка стратегії підприємства, спрямованої на позитивну зміну даної прихованої властивості підприємства.

У залежності від вирішення конкретної задачі вказані завдання можуть мати різну міру деталізації і глибину аналізу, але вони повинні проводитися в усіх напрямках – як внутрішньовиробничих, так і зовнішніх, що пояснюється високим ступенем взаємозв'язку показників господарської діяльності підприємства й навколишнього середовища, які визначають рівень латентних ознак підприємства.

Як було показано вище, латентні ознаки відносяться до умовно кількісних ознак і складають зміст так званої порядкової шкали. Вона, відрізняється від звичайної метричної шкали тим, що допускає тільки логічні операції порівняння типу «більше-менше», «краще-гірше» і т.п. Відомо, що латентні показники кількісному вимірюванню не підлягають: їх можна тільки оцінити на основі значень звичайних чинників-симптомів, які розглядаються як кількісні прояви прихованого економічного явища. Причому, таке оцінювання здійснюється за допомогою як експертних, так і математико-статистичних методів.

В деяких ситуаціях виникає проблема кількісного порівняння (ранжирування, класифікації) різних об'єктів за величиною латентної ознаки. Наприклад, потенціальні інвестори, зазвичай, намагаються виявити найбільш привабливе підприємство, споживачі завжди цікавляться співвідношенням «ціна/якість», певна інтенсивність праці робітника потребує відповідної матеріальної винагороди. Іншими словами, необхідно кількісно оцінити величину латентної ознаки, що вивчається, у кожного з об'єктів дослідження на основі чинників-симптомів, інформація про яких представлена у вигляді відповідних матриць вихідних даних.

Сьогодні в економічній науці гостро постало питання створення універсальної процедури оцінки прихованих властивостей підприємства, яка б ураховувала кількісні та якісні характеристики об'єкта дослідження, давала об'єктивну, прозору та більш-менш однозначну оцінку даної латентної ознаки, нівелювала розбіжності в результатах опитування та аналізу думок різних експертів.

Актуальність поставленої проблеми тим більше зростає у зв'язку з тим, що в сучасних наукових дискусіях з економіки досить часто можна чути риторичне запитання: а навіщо взагалі оцінювати якісні, тобто латентні (за загальноприйнятою класифікацією) ознаки економічних об'єктів? Якщо вони якісні, то нехай собі залишаються не цифрованими!

На наш погляд, з такою думкою можна погодитись лише частково. Дійсно, по відношенню до атрибутивних ознак немає ніякого сенсу казати про якусь кількісну оцінку статі, сімейного стану, професії робітників; назви,

галузевої приналежності, економічного призначення продукції підприємства тощо [132].

Однак, по відношенню до прихованих економічних властивостей підприємства – таких, як ефективність праці службовців, фінансовий стан, конкурентоспроможність, інвестиційна привабливість, глобальний економічний критерій оптимізації виробничої програми, інноваційно-інвестиційний рівень розвитку та його складових, пріоритетність проекту інвестиційного портфеля, рівень сталого розвитку, конкурентоздатність продукції підприємства та інші цілком слушно застосовувати різноманітні методи оцінювання. Це, перш за все, надасть змогу перейти від упорядковування зазначених ознак окремих об'єктів за допомогою понять «більше – менше», «краще – гірше», «вище – нижче» і т.п. до порівняння кількісних відмінностей між об'єктами щодо вимірюваних латентних економічних ознак на основі інтервальної масштабованої шкали.

Нехай при стандартизації даних за формулою (1.2) шкала впорядкованої латентної економічної ознаки, що оцінена за допомогою багатовимірних статистичних методів для об'єктів *A*, *B*, *C*, має вигляд, показаний на рис. 1.6.

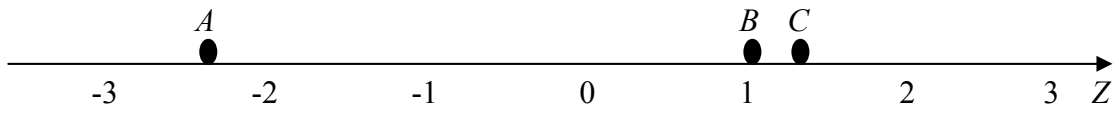


Рис. 1.6. Числова вісь стандартизованих і впорядкованих значень латентної економічної ознаки для об'єктів *A*, *B*, *C*

Якщо прихована економічна ознака стимулятор, підвищення якої є бажаним, наприклад, рівень фінансового стану, конкурентоспроможності, інвестиційної привабливості підприємства тощо, то позитивна частина осі на рис. 1.6 містить об'єкти-лідери (*B*, *C*), в околиці нуля – об'єкти-середняки, негативна частина вісі – об'єкти-аутсайдери (*A*). При цьому однакові відстані об'єктів у цій шкалі будуть означати рівні відмінності в кількостях оціненої латентної ознаки і, навпаки. Іноді досліджується так звана часова варіація прихованої властивості одного й того ж економічного об'єкта в різні періоди або моменти часу.

При вимірюванні латентної ознаки, що показана на рис. 1.6, у порядковій шкалі, наприклад, на базі експертного опитування, два об'єкти-лідери *B*, *C*, які знаходяться в правій частині осі, можуть виявитись нерозрізненими. Це означає, що експерти з імовірністю, близькою до 1, встановлять нерівності $A < B$, $A < C$, а нерівність $B < C$ викличе в них певні труднощі й помилки типу $B \geq C$.

Отже, отримання в результаті оцінювання латентних економічних ознак масштабованої шкали, наприклад, тієї, що показана на рис. 1.6, є надійною запорукою недопущення хибних висновків при використанні порядкових шкал, які формуються на основі різноманітних експертних методів і прийомів.

У разі масштабування даних за формулою (1.3) шкала впорядкованої прихованої економічної властивості, що оцінена за допомогою багатовимірних

статистичних методів для об'єктів *A*, *B*, *C*, має вигляд, показаний на рис. 1.7, з аналогічними висновками.

Друга важлива причина, за якою впорядковані значення латентної економічної ознаки на основі масштабованих даних, отриманих у результаті оцінювання за допомогою багатовимірних статистичних методів, мають суттєву перевагу над судженнями «більше – менше», «краще – гірше», «вище – нижче», характерними для порядкових шкал, полягає в наступному.

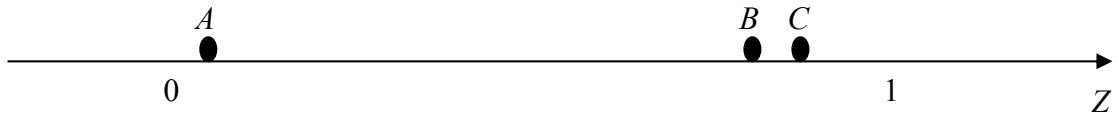


Рис. 1.7. Числова вісь масштабованих за формулою (1.3) і впорядкованих значень латентної економічної ознаки для об'єктів *A*, *B*, *C*

Дуже часто при проведенні факторного економічного аналізу латентних ознак виникає завдання виявлення важливіших чинників, що визначають варіацію прихованої властивості, яка вивчається. Наприклад: кількісно оцінити вплив різноманітних (економічних, технічних, організаційних та ін.) факторів, детермінуючих рівень фінансового стану, конкурентоспроможності, інвестиційної привабливості промислових підприємств конкретної галузі.

Подібне завдання при виконанні умови, що результативна економічна ознака Y і чинники X_1, X_2, ..., X_m, що її визначають, належать до кількісних, виміряних у метричній шкалі відносин або інтервалів, вирішується на базі кореляційно-регресійного аналізу [21; 22; 123; 127-129], шляхом побудови рівняння

$$\hat{Y} = a_0 + a_1X_1 + a_2X_2 + \ldots + a_mX_m, \qquad (1.6)$$

де $\hat{Y}$ – розрахункове значення результативної ознаки;

$a_0, a_1, \ldots, a_m$ – невідомі коефіцієнти регресії, які зазвичай оцінюються за допомогою методу найменших квадратів.

Множинний кореляційно-регресійний аналіз і побудова рівняння (1.6) виконуються виключно на персональному комп'ютері з використанням стандартних програм редактора Excel або ж інших пакетів програм, наприклад, STATISTICA [11-13]. Важливіші властивості середньої величини, дисперсії, коваріації, регресії, які використовуються в процесі даного аналізу, наведено в додатках А – Г.

Застосування методу найменших квадратів у якості оцінювача невідомих коефіцієнтів регресії базується на ряді його особливих властивостей, які випливають із вибіркового підходу до розглядання сутності кореляційно-регресійного аналізу взаємозв'язку економічних ознак. Ці властивості висувають його на перший план серед множини інших методів оцінки параметрів моделі в генеральній сукупності. Наведемо основні з них.

1. Незміщеність, яка означає, що знайдені коефіцієнти регресії a_0, a_1, …, a_m не містять систематичних помилок, тобто математичне очікування коефіцієнтів регресії, знайдених за даними нескінченного числа вибірок з генеральної сукупності, дорівнює їх істинним значенням.

2. Ефективність, яка полягає в тому, що випадкові помилки знайдених коефіцієнтів регресії a_0, a_1, …, a_m мінімальні в порівнянні з випадковими помилками параметрів рівняння (1.7), що забезпечуються всіма іншими методами оцінювання.

3. Спроможність, яка означає, що із зростанням обсягу статистичної сукупності спостережень N коефіцієнти регресії a_0, a_1, …, a_m прагнуть до своїх істинних значень у генеральній сукупності. Іншими словами, зростання інформаційної бази економіко-статистичної моделі забезпечує збільшення точності і достовірності оцінюваних регресійних взаємозв'язків.

Оцінки, що володіють властивостями 1 – 3, називаються найкращими лінійними оцінками без зміщень (best liner unbiased estimates) або BLUE-оцінками. Отже, коефіцієнти регресії a_0, a_1, …, a_m є BLUE-оцінками невідомих параметрів у генеральній сукупності.

4. Нормальність розподілу коефіцієнти регресії a_0, a_1, …, a_m, яка випливає з центральної граничної теореми теорії ймовірностей і математичної статистики. Властивість 4 є особливо важливим для проведення процедури перевірки різних статистичних гіпотез у процесі аналізу отриманого рівняння (1.6) і побудови довірчих інтервалів його параметрів, а також довірчих інтервалів прогнозу.

Знайдені коефіцієнти регресії a_1, a_2, …, a_m показують, на скільки в середньому змінюється досліджувана економічна ознака Y при зміні відповідного фактора на одиницю. Вони представляють собою абсолютну міру впливу чинників X_1, X_2, ..., X_m на Y.

На основі величин a_1, a_2, …, a_m розраховуються коефіцієнти еластичності за формулою

$$E_j = a_j \frac{\overline{X}_j}{\overline{Y}}. \qquad (1.7)$$

Коефіцієнти еластичності E_j (j = 1, 2, …, m) показують, на скільки відсотків у середньому змінюється досліджувана економічна ознака Y зі зміною відповідного фактора X_j на один відсоток (або в 1,01 разу). Вони є відносними показниками ступеню впливу чинників X_1, X_2, ..., X_m на Y.

Далі розраховуються бета-коефіцієнти (β_j) за формулою

$$\beta_j = a_j \frac{\sigma_j}{\sigma_Y}. \qquad (1.8)$$

β_j-коефіцієнт показує, на скільки стандартних відхилень у середньому зміниться досліджувана економічна ознака Y зі зміною фактора X_j на одне своє стандартне відхилення. Абсолютні значення β_j характеризують ступінь «регульованості» відповідних чинників по відношенню до змінної Y. Дійсно, чим вище σ_j, тим більше по абсолютній величині β_j, тим значніші можливості

зміни досліджуваної ознаки Y за рахунок зміни фактора X_j. Якщо змінна X_j є заданою, тобто константою, то $\sigma_j = 0$ и $\beta_j = 0$. Тому β_j-коефіцієнт відображає вплив на Y чинника X_j з урахуванням міри його коливання, яка вимірюється величиною стандартного відхилення σ_j.

Бета-коефіцієнти використовуються в множинному кореляційно-регресійному аналізі в ролі кількісних оцінок величин резервів (ефектів) зміни Y за рахунок фактора X_j. Це означає, що більшому значенню β_j відповідає більша величина резерву (ефекту) зміни Y.

І, нарешті, даний підхід, що базується на застосуванні рівняння регресії (1.6), дозволяє не тільки визначити чинники, що детермінує найбільший абсолютний, відносний вплив і забезпечують максимальний резерв (ефект) зміни досліджуваної економічної ознаки Y, а й приблизно оцінити величину цього резерву (ефекту) в розрізі кожного фактора X_1, X_2, ..., X_m за допомогою порівняльного аналізу об'єктів-лідерів та об'єктів-аутсайдерів [123, с.132-135; 128, с.60-62; 129, с.98-100].

Однак, у випадку, коли досліджувана економічна ознака Y є латентною, тобто вимірюється в порядковій шкалі, то вираження (1.6) – (1.8) залишаються невизначеними, оскільки судженнями експертів типу «більше – менше», «краще – ліпше», «вище – нижче» неможливо підставити у відповідні формули і поставлене завдання стає невирішеним.

Таким чином, перш ніж приступати до виявлення важливіших чинників, що визначають варіацію прихованої ознаки, що вивчається, треба спочатку її оцінити, надати прихованій властивості конкретні числові значення для кожного економічного об'єкта сукупності спостережень. І тут на допомогу приходять багатовимірні статистичні методи оцінки латентних ознак в економіці.

2. МЕТОДИ ОЦІНКИ ЛАТЕНТНИХ ОЗНАК ЕКОНОМІЧНИХ ОБ'ЄКТІВ

2.1 Класифікація методів оцінки латентних ознак

В останні роки бурхливий розвиток досліджень латентних ознак економічних об'єктів [23; 74; 103; 107-109] привів до використання множини прийомів, алгоритмів і процедур, досить нерівнозначних за своїм призначенням. Відсутність їх чіткої систематизації викликає подальше розширення інструментарію вивчення прихованих властивостей часто неефективними і суб'єктивними методами. Тому, важливо, щоб на даному етапі розвитку науки була створена базова класифікація методів оцінки латентних ознак економічних об'єктів.

Вона повинна відповідати ряду вимог, важливішими з яких є:

- достатня повнота охоплення методів оцінювання;
- єдність класифікаційної ознаки на кожному рівні угрупування;
- не пересічність розділів класифікації;
- відкритість класифікаційної схеми (можливість доповнення новими методами).

На рис. 2.1 наведено авторську класифікацію методів оцінки латентних ознак економічних об'єктів. Ми не претендуємо на універсальність і завершеність запропонованого угрупування методів оцінки латентних ознак економічних об'єктів, що складається з трьох основних рівнів.

Кожний рівень деталізації визначається своєю класифікаційною ознакою: ступенем формалізації вихідної інформації, загальним принципом застосування, різновидом частинних процедур і алгоритмів.

За ступенем формалізації вихідної інформації всі методи діляться на експертні, змішані й математико-статистичні (рис. 2.1).

Експертні методи оцінки латентних ознак є історично першими, традиційними, найбільш популярними. Їх використання почалося відтоді, коли з'явився науковий інтерес дослідників до таких ознак економічних об'єктів, як фінансовий стан, конкурентоспроможність, інвестиційна привабливість підприємства; ефективність, продуктивність та інтенсивність праці працівників, конкурентоздатність продукції тощо. При цьому, зазвичай, звертаються до опитування фахівців у відповідній сфері економіки, які на основі власного досвіду, знань, інформації про величину чинників-симптомів визначають пріоритети досліджуваних об'єктів за певними прихованими властивостями. Отримані індивідуальні й колективні експертні оцінки застосовують як кінцеві результати або в якості вихідних даних у змішаних (комбінованих) методах оцінки латентних ознак.

До експертних методів (як індивідуальних, так і колективних) відносяться:

1) метод сум (простих чи зважених);
2) метод добутку (простого чи зваженого);
3) побудований на базі методів сум або добутку метод відповідних середніх – арифметичних або геометричних.

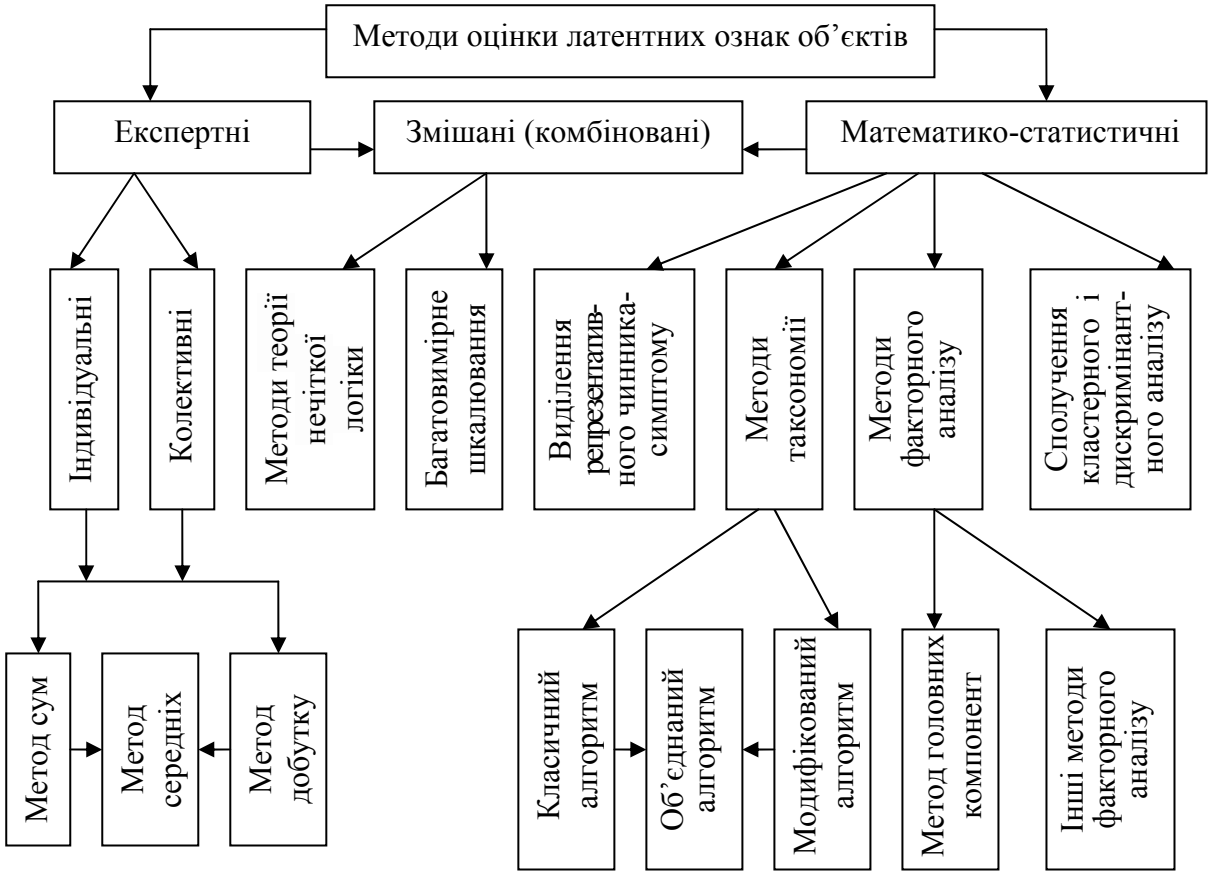


Рис. 2.1 Класифікація методів оцінки латентних ознак економічних об'єктів

Індивідуальні експертні методи базуються на використанні думок експертів – спеціалістів відповідної галузі економіки незалежно один від одного. Найбільш популярними є наступні два методи експертного оцінювання латентних ознак економічних об'єктів: 1) інтерв'ювання; 2) аналітичні експертні оцінки.

Метод інтерв'ю заснований на проведені бесіди дослідника з експертом, у ході якої перший у відповідності до заздалегідь розробленої програми ставить перед другим запитання щодо рівня прихованої властивості економічних об'єктів або її чинників-симптомів. Успіх такої процедури в значній мірі залежить від здатності експерта експромтом надавати висновки і судження за запитаннями дослідника.

Аналітичні експертні оцінки передбачають тривалу і ретельну самостійну роботу експерта над поставленою проблемою оцінки досліджуваної латентної ознаки економічних об'єктів. Цей метод дозволяє експерту застосовувати всю необхідну йому інформацію щодо прихованої властивості об'єкта оцінки.

Головними перевагами обговорюваних методів є можливість максимального використання індивідуальних здібностей експерта й незначущість психологічного тиску, що, зазвичай, вчиняється на нього. Однак, ці методи не завжди вільні від певного суб'єктивізму і волюнтаризму.

В основі методів колективних експертних оцінок лежать принципи виявлення колективної думки експертів щодо рівня досліджуваної латентної ознаки економічного об'єкта або її чинників-симптомів. При цьому висувається гіпотеза про наявність у експертів уміння з високим ступенем достовірності зрозуміти важливість поставленого завдання, доцільність вибору з множини можливих оцінок певного компромісного варіанту.

Сьогодні широке розповсюдження отримали методи колективних експертних оцінок прихованих властивостей економічних об'єктів, що базуються на роботі спеціальних комісій, коли групи фахівців за «круглим столом» обговорюють поставлену проблему з метою узгодження думок і вироблення єдиної шуканої оцінки (або прийнятного діапазону оцінок). Даний метод поряд із суб'єктивізмом, притаманним будь-якому експертному оцінюванню, має також такий недолік, як керівництво групи експертів в своїх судженнях, головним чином, логікою компромісу з досить суттєвим психологічним тиском колективної думки на окремих спеціалістів.

Як правило, важливішими задачами при оцінюванні латентної ознаки за допомогою колективу експертів є формування репрезентативної експертної групи, підготовка і проведення експертизи, обробка отриманих результатів. При формуванні групи експертів головними є питання її якісного й кількісного складу. Для отримання якісного складу колективу експертів до учасників групи висувається ряд вимог, серед яких важливішими є наступні:

- високий рівень загальної економічної ерудиції;
- глибокі спеціальні знання в області даної прихованої ознаки;
- наявність наукового інтересу або практичних навичок щодо оцінюваної властивості об'єктів.

Оптимальний кількісний склад групи експертів визначити достатньо важко. В теперішній час розроблені формалізовані підходи до розрахунку цієї величини. На основі математичних формул, що відкривають можливість оцінити компетентність потенційного експерта, максимальної і мінімальної межі чисельності групи, знаходять границі шуканого кількісного складу групи [82, с.157-162].

Математико-статистичні процедури й алгоритми є порівняно «молодими» методами оцінки латентних ознак економічних об'єктів, виникнення яких пов'язане з розвитком і впровадженням у практику досліджень багатовимірних статистичних і математичних моделей, зокрема, кореляційно-регресійного, факторного, кластерного, дискримінантного, таксономічного аналізу, а також багатовимірного шкалювання. Не зважаючи на те, що математико-статистичний апарат цих методів був розроблений досить давно (наприклад, основи теорії кореляції були закладені К. Пірсоном ще наприкінці XIX ст.), їх впровадження в різних галузях науки і практики було загальмоване відсутністю технічних засобів вирішення складних задач високої розмірності. І лише з другої половини минулого століття з розвитком швидкодіючих ЕОМ та комп'ютерної техніки вказані методи отримали широке розповсюдження в дослідження прихованих властивостей будь-яких об'єктів, в тому числі й економічних.

Таксономічний аналіз або метод таксономії представляє собою одну з простіших багатовимірних процедур, що дозволяє отримати уявлення про величину досліджуваної латентної ознаки. Він базується на розрахунку відстаней (схожості) всіх точок (об'єктів) до еталона (антиеталона) в просторі спостережуваних чинників-симптомів, які є зовнішнім проявленням прихованої властивості об'єктів. Даний метод утілений у дві головні алгоритми таксономічного аналізу – класичний і модифікований. Якщо класична процедура передбачає визначення схожості з еталоном, то модифікована процедура побудована на розрахунку відстаней до антиеталона. При цьому схожість з еталоном або відстань до антиеталона розглядаються в якості шуканих оцінок латентних ознак, що вивчаються.

Слід відмітити, що метод таксономії не є абсолютно вільним від можливих похибок: класичний алгоритм іноді приводить до помилок при визначенні схожості з об'єктами-аутсайдерами (за дослідженою прихованою властивістю об'єктів), а модифікований алгоритм, навпаки, страждає неточністю при ідентифікації об'єктів-лідерів. В теперішній час задля усунення або хоча б зменшення вірогідності вказаних похибок спостерігаються спроби розробити змішаний алгоритм таксономічного аналізу, котрий об'єднує позитивні сторони двох основних процедур.

В основі факторного аналізу (в математичному розумінні даного терміну) як методу оцінки латентних ознак лежить припущення про те, що спостережувані значення чинників-симптомів – показників метричної шкали x_1, x_2,…, x_m є зовнішнім проявом деякої прихованої властивості, що характеризує економічні об'єкти (рис. 2.2). Завдання факторного аналізу полягає в тім, щоб, змоделювавши значення x_1, x_2, …, x_m, кількісно оцінити цю латентну характеристику за допомогою штучних змінних F_1, F_2, …, F_m, які представляють собою лінійні комбінації чинників-симптомів x_1, x_2, …, x_m.

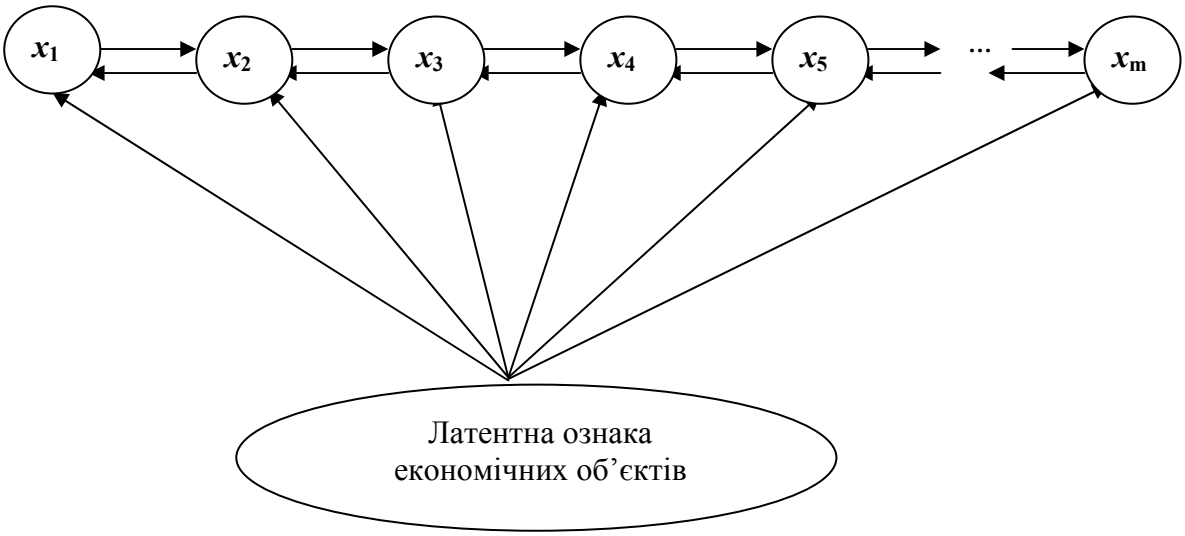


Рис. 2.2. Взаємозв'язок чинників-симптомів між собою і з латентною ознакою економічних об'єктів в моделі факторного аналізу

Обговорювану ситуацію можна прокоментувати в такий спосіб: показники метричної шкали x_1, x_2,…, x_m економічних об'єктів корелюють між собою тому, що вони перебувають під впливом прихованої властивості. Але якщо чинники-симптоми об'єктів x_1, x_2,…, x_m безпосередньо спостерігаються й вимірюються, то про величину латентної ознаки можна судити лише опосередковано – за розміром й тіснотою взаємозв'язку змінних x_1, x_2,…, x_m.

Серед алгоритмів факторного аналізу, які досить детально і всебічно описані в математико-статистичній літературі, можна вказати наступні найбільш відомі методи, що вирізняються за способом визначення факторного рішення:

1) головних компонент;
2) максимальної правдоподібності;
3) центроїдів;
4) головних осей;
5) наближених загальностей.

Доведено і це підтверджується практикою, що всі вони надають добре узгоджені результати з досить високим коефіцієнтом парної кореляції. Тому, зазвичай, використовують найбільш популярний різновид факторного аналізу – метод головних компонент.

Сполучення кластерного і дискримінантного аналізу при оцінюванні прихованої властивості економічних об'єктів уперше було використано в 60-х роках XX ст. в ході визначення імовірності банкрутства західних компаній за допомогою багатофакторних моделей. При цьому завдання кластерного аналізу полягає в розділенні сукупності досліджуваних об'єктів на однорідні групи за рівнем латентної ознаки, що вивчається. Таке угрупування відбувається на основі значень її чинників-симптомів x_1, x_2,…, x_m за допомогою відомих алгоритмів кластерного аналізу – агломеративної ієрархічної процедури, методів *k*-середніх, «Форель», оптимізаційних алгоритмів та ін. Далі, на базі методу подвійного об'єднання (одночасно за об'єктами і чинниками-симптомами) здійснюється ідентифікація виділених кластерів, тобто встановлення груп об'єктів-лідерів, об'єктів-середняків, об'єктів-аутсайдерів і т. ін.

У процесі дискримінантного аналізу вирішуються наступні основні задачі:

1) побудова адекватної дискримінантної функції за допомогою методів кореляції і регресії;

2) визначення оцінок латентної ознаки досліджуваних економічних об'єктів на базі побудованої дискримінантної функції;

3) оцінювання прихованої властивості, що вивчається, для нових об'єктів, тобто таких, котрі не входять у досліджувану сукупність.

Перша задача вирішується шляхом розрахунку і статистичного аналізу звичайного рівняння регресії, в якому роль результативного показника виконує штучна змінна, що визначає приналежність кожного економічного об'єкта до певного кластера. А в якості факторних показників виступають чинники-симптоми x_1, x_2,…, x_m. При цьому точність майбутніх оцінок прихованої

властивості об’єктів може бути визначена на основі величини коефіцієнта детермінації побудованої адекватної дискримінантної функції у вигляді знайденого рівняння регресії.

Друга задача полягає у тім, щоб за відомими значеннями чинників-симптомів x_1, x_2,…, x_m на базі дискримінантної функції розрахувати для кожного об’єкта досліджуваної сукупності шукану оцінку латентної ознаки, яка вивчається.

Третя задача представляє собою саме безпосередню задачу дискримінації нових економічних об’єктів, тобто прогнозування рівня прихованої властивості об’єктів, що не входять у досліджувану сукупність, і віднесення їх до того чи іншого кластера.

Змішані (комбіновані) методи оцінки прихованої властивості об’єктів засновані на сполученні як експертних, так і математико-статистичних процедур і алгоритмів. До них відносяться методи теорії нечіткої логіки, метричного та неметричного шкалювання та ін. Комбіновані методи, зазвичай, мають ті ж переваги й недоліки, що і експертні та математико-статистичні підходи до оцінювання латентних ознак економічних об’єктів.

Так, методи теорії нечіткої логіки по суті представляють собою подальший розвиток і узагальнення результатів експертних оцінок. Головна новизна побудованих за допомогою елементів теорії нечіткої логіки інтегральних показників, що розглядаються в якості оцінок досліджуваних прихованих властивостей економічних об’єктів, зводиться до формалізації думок фахівців-експертів в процесі рейтингового оцінювання на базі відомих матричних моделей з усіма негативними наслідками їхнього використання: певним суб’єктивізмом, проблемами створення компетентних груп експертів тощо.

Методи багатовимірного шкалювання, зокрема, один із їх найбільш популярних напрямків (неметричне шкалювання) складається з двох послідовних математико-статистичних процедур: кластерного і факторного аналізів. При цьому в якості вхідних початкових даних застосовується інформація, отримана в результаті експертного опитування, і яка вимірюється в порядковій (ранговій) шкалі.

Вважається, що методи багатовимірного шкалювання є найбільш гнучкими й адаптованими для адекватного відображення латентних ознак економічних об’єктів. Однак, слід мати на увазі наступне: їх застосування пред’являє найвищі вимоги до професійної підготовленості, досвіду та інтуїції дослідника, оскільки від його знань і творчих здібностей багато в чому залежить правильний вибір обчислювальних процедур даних методів, коректність інтерпретації отриманих результатів тощо. Окрім того, при їх використанні зберігаються всі негативні властивості, недоліки й ризики, які притаманні самим експертним оцінкам.

На нашу думку, комбіновані методи оцінки прихованої властивості економічних об’єктів значною мірою представляють собою спробу здійснити оцифрування результатів експертних опитувань за допомогою формальних математико-статистичних моделей.

2.2 Методи експертного оцінювання

Експертні оцінки – це комплекс логічних і математико-статистичних процедур, спрямованих на отримання від фахівців інформації, її аналіз та узагальнення з метою оцінювання прихованих властивостей економічних об’єктів. Уважається, що їх доцільно застосовувати, коли вибір і обґрунтування оцінки результату не можуть бути виконані на підставі точних вимірювань і розрахунків, тобто в метричних шкалах інтервалів і відносин. Експертна оцінка отримується шляхом опитування думок фахівців і, зазвичай, представляється у вигляді рівнів оцифрованої порядкової шкали.

Чому думки експертів природно виражати саме в порядковій шкалі? Як показали численні дослідженні, людина правильніше (і з меншими труднощами) відповідає на запитання якісного, наприклад, порівняльного, характеру, ніж кількісного. Так, йому легше сказати, яка з двох гир важче, ніж вказати їх приблизний вага в кілограмах.

Вирізняють наступні види інформації, що використовуються в рамках даних методів:

- думка у вигляді відповідного числа в запропонованих межах (оцінка латентної ознаки кожного об’єкта);
- ранжирування об’єктів (розстановка їх по місцях);
- розбивка економічних об’єктів всієї сукупності на окремі класи, групи;
- попарне порівняння об’єктів і повідомлення про те, який з них «краще» за досліджуваною прихованою ознакою.

Проведення будь-якої експертизи, яка буває індивідуальною та груповою, включає в себе наступні головні стадії:

1. Формування мети експертизи. 2. Підбір експертів. 3. Вибір методики проведення опитування. 4. Обробка отриманої інформації, у тому числі перевірка узгодженості й достовірності експертних оцінок.

Однак, слід зазначити, що процедура оцінювання латентних ознак в економіці за допомогою експертних методів має свої специфічні особливості, пов’язані зі складністю, багатогранністю таких прихованих властивостей підприємства, як конкурентоспроможність, фінансовий стан, інвестиційна привабливість, сталість розвитку, пріоритетність інвестиційних проектів тощо. Це обумовлює проходження, окрім указаних вище чотирьох стадій експертизи, ряду додаткових етапів, серед яких важливішими є:

1) створення вихідного середовища оцінювання, тобто визначення певного кола чинників-симптомів досліджуваної латентної ознаки економічних об’єктів;

2) масштабування первинних чинників-симптомів, головна мета якого полягає в уникненні хибного впливу їх одиниць виміру на кінцеві результати оцінювання прихованої властивості, що досліджується;

3) встановлення та врахування системи статистичних ваг, що відображають ступінь важливості окремих чинників-симптомів;

4) безпосереднє оцінювання прихованої властивості економічних об’єктів – розрахунок шуканого, як правило у балах чи в інших умовних одиницях, інтегрального (синтетичного) показника;

5) використання отриманих результатів – ранжирування економічних об’єктів, наприклад, підприємств, за рівнем оціненої латентної ознаки, розробка відповідних рекомендацій, стратегій розвитку і т. ін.

На першому етапі здійснюється обґрунтування множини чинників-симптомів прихованої властивості економічних об’єктів, які включаються в подальше дослідження. При цьому майбутнє середовище показників повинно відповідати наступним умовам:

- обрані чинники-симптоми відображають найважливіші з точки зору оцінюваної латентної ознаки сторони об’єктів, а їх кількість обмежена;
- при визначенні чинників-симптомів слід орієнтуватись на застосування даних публічної фінансової, бухгалтерської і статистичної звітності при мінімізації використання внутрішньої інформації.

Слід мати на увазі той факт, що визначення певного кола чинників-симптомів латентної ознаки економічних об’єктів часто перетворюється у самостійну наукову проблему, вирішення якої залежить від великого числа конкретних умов і обставин даного дослідження – галузі, до якої належить підприємство, часу спостереження, особливостей виробництва і реалізації продукції тощо. Розглянемо дану проблему на прикладі створення середовища оцінки інвестиційної привабливості підприємств як одної з важливих прихованих властивостей об’єктів інвестування.

Так, в «Методиці інтегральної оцінки інвестиційної привабливості підприємств та організацій» указується, що вона базується на використанні більш ніж 40 відповідних показників за різними напрямами господарської діяльності підприємства, які складають шість груп [58].

А.А. Садеков, Н.А. Лісова розробили модель оцінки інвестиційної привабливості підприємств на основі рейтингових співставлень певної системи індикаторів, яку автори поділяють на показники фінансово-господарської діяльності, показники оцінки цінних паперів, і доповнюють показниками інвестиційної привабливості бізнес-планів інвестиційних проектів підприємств. Ця модель має модифікацію для використання в умовах кризи неплатежів, що робить її адекватною для сучасних умов господарювання [88].

В.М. Гриньова із співавторами розглядаючи різні фактори формування інвестиційної привабливості суб’єктів господарювання, вказує, що аналіз наявних методик, особливості розвитку економіки та функціонування ринкової інфраструктури дають змогу визначити найбільш значущі чинники, що застосовуються для оцінки інвестиційної привабливості підприємства. Це показники: 1) майнового стану; 2) ліквідності; 3) фінансової стійкості; 4) ділової активності; 5) рентабельності.

На їх думку, перша група (оцінка майнового становища) дозволяє з’ясувати джерела формування майна й оцінити, як воно використовується. Показники ліквідності і платоспроможності (друга група) характеризують можливість підприємства своєчасно та повною мірою провести розрахунки за

своїми зобов'язаннями. Показники фінансової стійкості (третя група) дають можливість визначити фінансовий запас міцності підприємства, тобто стабільність його діяльності з позиції довгострокової перспективи. Ділова активність підприємства (четверта група) виявляється у динаміці його розвитку, досягненні обраних цілей, розширенні ринків збуту. Показники рентабельності (п'ята група) характеризують економічну ефективність фінансово-господарської діяльності підприємства [38].

Н.А. Русак відмічає, що кожному підприємству треба знати свої шанси залучення зовнішніх інвестицій. Тому інвестиційна привабливість аналізується як при внутрішньому, так і при зовнішньому аналізі. При цьому використовується система показників, що характеризує стан фінансів підприємства, визначається рівень кожного показника і його динаміка за ряд звітних періодів. Для оцінки інвестиційної привабливості підприємства можуть застосовувати наступні групи показників: 1. Структура джерел формування капіталу. 2. Структура складу капіталу. 3. Платоспроможність і фінансова стабільність. 4. Якість прибутку й здатність її заробляти. 5. Прибутковість і оборотність капіталу [85].

Г.В. Кошельок запропонувала систему чинників-симптомів інвестиційної привабливості підприємств зернопереробного комплексу України, яка складається з 38 чинників, об'єднаних у шість груп, що характеризують на промисловому підприємстві: 1. Майновий стан. 2. Ліквідність. 3. Фінансову стійкість. 4. Ділову активність. 5. Рентабельність. 6. Ринкову активність (рис. 2.3) [43; 44; 138].

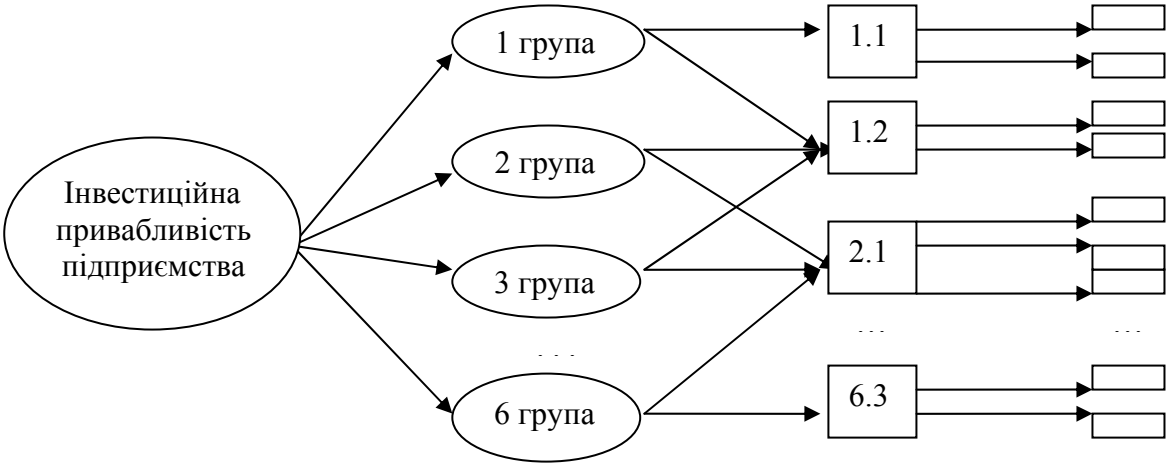


Рис. 2.3. Структура внутрішніх чинників-симптомів інвестиційної привабливості підприємства

Така ситуація наводить на думку про те, що ознака «інвестиційна привабливість» підприємства є не тільки латентною, але ще й ієрархічною, тобто такою, змістовна модель якої містить декілька взаємопов'язаних та

взаємно підпорядкованих рівнів показників. Г.В. Кошельок виділила чотири головних ієрархічних рівня даної ознаки

На першому рівні знаходиться досліджувана латентна економічна ознака – інвестиційна привабливість підприємства. Вона проявляється у вигляді чинників-симптомів – групових факторів, які знаходяться на другому рівні ієрархії і теж є латентними. Третій рівень складають чинники-симптоми, що визначають окремі групові фактори. Це коефіцієнти оновлення, зносу основних фондів, показники маневреності, покриття, ліквідності, автономності, оборотності капіталу і т. ін. І на четвертому рівні ієрархії (вихідні дані) розташовані показники бухгалтерської, статистичної та фінансової звітності, які створюють інформаційну базу для дослідження інвестиційної привабливості підприємства.

Всі латентні показники на рис. 2.3 позначені у вигляді еліпсів, на відміну від звичайних показників метричної шкали, що знаходяться на 3-му та 4-му рівнях ієрархії і позначені прямокутниками.

Отже, доволі складно надати якісь універсальні рекомендації щодо визначення певного кола чинників-симптомів досліджуваної латентної ознаки економічних об’єктів. У кожному конкретному випадку необхідно враховувати положення відповідної економічної теорії – інвестиційної привабливості (як у наведеному прикладі), конкурентоспроможності, фінансового стану і т. ін.

Другий етап оцінювання досліджуваної прихованої властивості включає масштабування первинних чинників-симптомів, тобто перехід до безрозмірних одиниць їх виміру z_i на основі певних математичних перетворень, про які йшлося при розгляданні інтервальної (абсолютної) шкали вимірювання ознак економічних об’єктів (див. розділ 1.1 даної книги).

Основною вимогою, що висувається до масштабованих чинників-симптомів, є їхня незалежність від перетворень вихідних даних типу $X_i + c$, а також cX_i, де X_i – вихідне значення i-го чинника-симптому, c – деяка константа. При цьому не слід забувати, що будь-яке масштабування даних певним чином викривляє інформацію, що міститься в досліджуваних економічних показниках.

Найбільш відомим способом масштабування змінних в математичній статистиці є їхня стандартизація, яка складається з двох послідовних операцій: центрування та нормування даних [125, с.17-18, 67].

Центрування являє собою вирахування з кожного значення даної змінної X_j за всіма спостереженнями статистичної сукупності її середнього значення $\bar{X}$. При цьому середня арифметична перетворених значень змінної стає рівною нулю. Геометрично дане перетворення рівносильне перенесенню відповідних осей координат вправо (уліво), нагору (донизу) на величину середнього значення залежно від його знака. У результаті початок координат стає «центром ваги» досліджуваної сукупності об’єктів (рис. 2.4).

Під нормуванням розуміється ділення вихідних значень показника на деяке постійне число, зазвичай на його стандартне відхилення σ_j. Геометрично нормування означає стискання (при $\sigma_j > 1$) або розтягання (при $\sigma_j < 1$) «хмари» точок спостережень у ознаковому просторі (див. рис. 2.4).

Отже, процедура стандартизації може бути виражена формулою (1.2). Стандартизація чинників-симптомів дозволяє привести всі дані до одного порядку. При нормальному розподілі змінних діапазон варіювання стандартизованих значень становить область від -3 до 3.

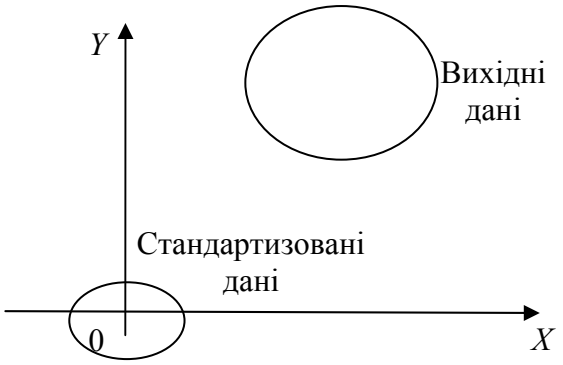


Рис. 2.4. Геометрична інтерпретація стандартизації ознак у випадку $\sigma_j > 1$

Для стандартизованих ознак справедливі наступні властивості:

1. Середня величина змінної дорівнює нулю ($\bar{Z} = 0$).

2. Дисперсія змінної співпадає зі стандартним відхиленням й складає одиницю ($\sigma^2_Z = \sigma_Z = 1$).

3. Коваріація двох стандартизованих показників збігається з їхнім коефіцієнтом кореляції.

Існують також й інші способи масштабування вихідних даних, зокрема такі, що враховують екстремальні й нормативні значення чинників-симптомів (див. формули (1.3), (1.4)). Так, формула (1.3) забезпечує приведення всіх їх значень в інтервал від 0 до 1.

Перетворення за формулою (1.4) широко застосовується при масштабуванні первинних чинників-симптомів інвестиційної привабливості підприємства в процедурах, що об'єднують різноманітні фінансові коефіцієнти, за якими відомі загальноприйняті нормативні значення. Наприклад, коефіцієнт мобільності активів ($X_{НОРМ} = 0,5$), коефіцієнт поточної ліквідності ($X_{НОРМ} = 2,0$-$2,5$), частка оборотних активів в балансі підприємства ($X_{НОРМ} = 0,33$).

Очевидно, що масштабування чинників-симптомів прихованої властивості економічних об'єктів за формулами (1.2) – (1.4) дозволяє уникнути впливу їх одиниць виміру на результати оцінювання, що, на наш погляд, є певною запорукою одержання слушних оцінок досліджуваної латентної ознаки.

На третьому етапі процедури в більшості методик експертного визначення латентної ознаки економічних об'єктів передбачається диференціація чинників-симптомів різних рівнів ієрархії за їх роллю у формуванні рівня прихованої властивості, що вивчається. Даний етап полягає у виявленні й урахуванні важливості окремих чинників-симптомів для кінцевого результату процедури оцінювання за допомогою встановленні коефіцієнтів їх вагомості.

Слід зазначити, що практично всі без винятку автори вирішують указане завдання саме на основі методу експертних оцінок. Наприклад, у вже згадуваній «Методиці» указується, що для встановлення вагомості групових та одиничних показників проводиться процедура експертної оцінки. Це мотивується тим, що у світовій і вітчизняній практиці дані методи достатньо опрацьовані і широко використовуються при вирішенні відповідних питань, тому зайве рекомендувати якусь версію методу експертної оцінки.

Зокрема, усім групам показників оцінки інвестиційної привабливості підприємств та організацій, а також показникам, розміщеним у цих групах, присвоюються, залежно від їх вагомості, відповідні числові значення. Як і граничні значення показників, вагомість може коригуватись. Указана вагомість є похідною від часу, конкретної політичної, соціальної ситуації, інших факторів, а тому ця характеристика має вірогідне походження.

Отже, для визначення інтегральної оцінки інвестиційної привабливості кожного підприємства розраховується наступна група вагових показників:

$$B_{ij} = (b_{ij} \times V_j)/100, \qquad (2.1)$$

де B_{ij} – загальна вагомість i-го показника в j-й групі (з урахуванням групової вагомості);

b_{ij} – значення вагомості i-го показника в j-й групі;

V_j – значення групової вагомості.

Значення вагомості показників і значення групової вагомості встановлюються експертним шляхом, виходячи з того, що сума рейтингу по кожній групі повинна складати 100 балів, і сума загальних групових рейтингів теж повинна дорівнювати 100 балам. Згідно з отриманими результатами найбільшу статистичну вагу в «Методиці» отримали показники ринкової активності підприємства (рентабельність акції 9,45; коефіцієнт цінності акції 9,45; дивіденд 8,10) та показники його майнового стану (коефіцієнт зносу основних засобів 8,00; коефіцієнт оновлення 6,00) [58].

Обговорюваний метод зважування чинників-симптомів латентної ознаки вирізняється відносною простотою, але для нього характерні всі недоліки, які притаманні процедурі експертної оцінки в загалі. В першу чергу, це суб'єктивізм суджень окремих фахівців-експертів щодо важливості певних чинників-симптомів та групових факторів прихованої властивості економічних об'єктів.

І. Вініченко, пропонуючи методику експрес-діагностики інвестиційної привабливості підприємств, включає в аналіз дві групи фінансових показників, що характеризують: 1) ефективність господарської діяльності (можливість одержання прибутку); 2) платоспроможність підприємств. Він пише, що значення підсумкового рейтингу інвестиційної привабливості розраховується шляхом підсумовування значень зведеного рейтингу за групою показників, що характеризують ефективність діяльності, і зведеного рейтингу за групою

показників, що характеризують платоспроможність, з ваговими коефіцієнтами 0,2 і 0,8 відповідно.

Зведений рейтинг підприємств за кожною групою визначається наступним чином. За кожним показником, що входить до групи, розраховуються значення, які підсумовуються з урахуванням вагових коефіцієнтів, що визначають ступінь впливу чинників-симптомів на приховану властивість. Отримані значення ранжирують і таким способом будуються зведені рейтинги, що характеризують ефективність діяльності й платоспроможність. Значення вагових коефіцієнтів кожного показника наведені в табл. 2.1 [16].

Нажаль, у своїй методиці автор не вказує, на яких розміркуваннях базується визначення коефіцієнтів вагомості показників інвестиційної привабливості підприємства – як групових (0,2 і 0,8), так і третього рівня ієрархії (табл. 2.1). Незрозуміло, чи то вони знайдені експертним шляхом, чи визначені на основі емпіричних спостережень?!

Таблиця 2.1

Вагові коефіцієнти факторів інвестиційної привабливості підприємств

№ п/п	Показники	Вагові коефіцієнти
1	Загальна рентабельність звітного періоду	0,30
2	Норма прибутковості капіталу	0,40
3	Рентабельність основної діяльності	0,20
4	Рентабельність активів	0,10
Разом		1,00
1	Коефіцієнт поточної ліквідності	0,40
2	Коефіцієнт автономії	0,15
3	Коефіцієнт абсолютної ліквідності	0,30
4	Коефіцієнт фінансової стійкості	0,15
Разом		1,00

В.М. Гриньова із співавторами пропонує рівень вагомості окремих показників при оцінці інвестиційної привабливості підприємства розраховувати за формулою:

$$b_i = \frac{R_i}{n}, \qquad (2.2)$$

де R_i – ранг *i*-го показника згідно з підсумковим ранжируванням.

Для визначення вагомості окремих чинників-симптомів часто використовується один із методів експертних оцінок – метод рангової кореляції. Експерти на підставі анкети розподіляють чинників-симптоми латентної ознаки підприємства за рангами (від 1 до *n)* за ступенем зменшення їх пріоритетності та важливості.

Вагомість кожної з п’яти груп визначається шляхом накопичення рангів окремих показників у цій групі, що дозволяє, на думку авторів, дати їм об’єктивну оцінку. За групами накопичені такі ранги: перша група – 31; друга

група – 231; третя група – 369; четверта група – 344; п'ята група – 60. Це свідчить про те, що на першому та другому місці за значимістю – група показників майнового стану та рентабельності, а на третьому та четвертому місцях – ліквідність і ділова активність відповідно. Останнє місце належить групі показників фінансової стійкості [38, с.80].

Як відомо, матричні моделі є одним із різновидів рейтингових оцінок. Бальний принцип зводиться до оцінки значення кожного показника у балах та визначення загальної суми балів по підприємству. Для практичної реалізації цього методу значенням досліджуваних чинників-симптомів латентної ознаки експерти присвоюють певну суму балів, а, отже, і певну статистичну вагомість.

К.А. Пріб наводить наступний приклад рейтингової оцінки інвестиційної привабливості підприємств за чотирма показниками 3-го рівня ієрархії (табл. 2.2). Вона пише, що набір оціночних показників, диференціація їх значень, та бальна оцінка вирізняється у різних методиках.

Таблиця 2.2

Оцінка значення показників у балах

Показники	Позитивна оцінка	Усереднена оцінка	Негативна оцінка
Коефіцієнт автономії	Більше 0,6	0,4-0,6	Менше 0,4
- оцінка у балах	40	30	15
Коефіцієнт покриття	Більше 2	1-2	Менше 1
- оцінка у балах	30	20	10
Тривалість фінансового циклу, днів	Менше 30	30-60	Більше 60
- оцінка у балах	40	20	10
Рентабельність активів, %	Більше 14	6-14	Менше 6
- оцінка у балах	50	35	10

На думку К.А. Пріб, метод бальної оцінки може застосовуватися не лише для сукупності підприємств, а й для окремих із них. У такому разі встановлюється шкала оцінки інвестиційної привабливості підприємства. Наприклад: понад 150 балів – інвестиційно привабливе підприємство; 130-149 балів – відносно привабливе; менше 130 балів – непривабливе [80, с.76-79].

На наш погляд, тут необхідно відмітити той факт, що хоча матричні моделі дійсно є простими й зрозумілими, але їм водночас притаманні недоліки, які серйозно ускладнюють їхнє практичне застосування і роблять незручними. Перш за все, використання бального методу ускладнюється через потребу обґрунтовано встановити бали для кожного чинника-симптому. Наприклад, не зовсім переконливо виглядає пропозиція про те, що коефіцієнт автономії у разі його значення більше 0,6 оцінюється саме у 40 балів, від 0,4 до 0,6 – 30 балів, менше 0,4 – 15 балів. Окрім того, у разі рейтингових оцінок ускладнена також диференціація показників по вертикалі табл. 2.2. Тобто проблема встановлення вагомості окремих ознак інвестиційної привабливості підприємств вирішується досить суб'єктивно з певними елементами волюнтаризму.

Очевидно, що всі без виключення розглянуті підходи до встановлення системи статистичних ваг, які відображають ступінь важливості окремих чинників-симптомів та групових факторів інвестиційної привабливості підприємства, не вирізняються оригінальністю і їм притаманні всі названі вище недоліки: недостатність обґрунтування, певний суб'єктивізм та волюнтаризм.

З цього приводу В. Захожай, М. Кіт пишуть, що обґрунтованість розрахунків рейтингу інвестиційної принадливості підвищиться, якщо надати окремим показникам диференційованої ваги згідно з їх значущістю. Цю вагу можна визначити за допомогою експертних оцінок. Проте останні здебільшого є суб'єктивними і тому ненадійними. Більшої об'єктивності можна досягти, встановивши взаємозв'язок результатів діяльності певних об'єктів і факторів, які їх визначають. Для цього доцільно скористатися методами кореляційно-регресійного аналізу [36].

На нашу думку, перша теза даних авторів не викликає сумніву, але з пропозицією використовувати для визначення вагомості чинників-симптомів кореляційно-регресійний аналіз важко погодитися. Справа в тому, що «результати діяльності певних об'єктів», тобто кількісні рівні інвестиційної привабливості підприємств, є прихованими (невідомими) ознаками, що випливає з їхнього латентного характеру. Якщо ж результативна залежна ознака невідома, то методи кореляційно-регресійного аналізу тут безсилі. Треба шукати інші засоби об'єктивного визначення системи статистичних ваг окремих чинників-симптомів та групових факторів.

Після встановлення тим чи іншим способом вагомості чинників-симптомів прихованої властивості економічних об'єктів автори більшості методик переходять до четвертого етапу процедури – розрахунку за деяким алгоритмом інтегрального показника латентної ознаки. Головна мета цього етапу є поєднання впливу окремих чинників-симптомів різних рівнів ієрархії в один інтегральний (синтетичний) показник досліджуваної прихованої властивості.

Як зазначає К.А. Пріб, інтегральна оцінка надає змогу поєднати в одному показнику багато різних за назвою, одиницями виміру, вагомістю та іншими характеристиками чинників. Це спрощує процедуру оцінки конкретної інвестиційної пропозиції, а інколи є єдино можливим варіантом її проведення і надання об'єктивних остаточних висновків. Застосування різних методик оцінки дає змогу інвестору спростити процес прийняття рішень щодо реалізації реальних інвестицій [80, с.76-80].

Найбільш розповсюдженим методом визначення інтегрального показника прихованої властивості економічних об'єктів є метод сум (простих чи зважених) або побудованих на його базі відповідних середніх арифметичних.

В його основі лежить гіпотеза про адитивний вплив на досліджувану латентну ознаку чинників-симптомів різних рівнів ієрархії. Тобто вважається, що інтегральній показник прихованої властивості кожного об'єкта I в самому загальному вигляді можна представити так:

$$I = \sum_{i=1}^{n} \sum_{j=1}^{m} B_{ij} z_{ij}, \qquad (2.3)$$

де z_{ij} – стандартизоване значення i-го чинника-симптому j-ї групи для досліджуваного економічного об'єкта.

На нашу думку, адитивна модель багатьох прихованих властивостей економічних об'єктів на рівні підприємства є цілком коректною, оскільки взаємодія більшості факторів в економіці дійсно підпорядковується принципу незалежного накопичення впливу: про рівень інтегрального показника латентної ознаки можна судити, підрахувавши внески всіх теоретично зв'язаних з ним чинників-симптомів різних рівнів ієрархії. Але тут є одне досить серйозне зауваження. Оскільки поряд з включеними до аналізу чинниками-симптомами й груповими факторами, що виділені дослідником, на рівень I_k впливають також і інші фактори, які з певних причин не були враховані в середовищі оцінювання (наприклад, деякі нефінансові, нематеріальні чинники тощо), то доцільним буде їхнє відображення в рівнянні (2.3) у вигляді додаткового доданку ε – випадкової компоненти, що відображає сумарний вплив чинників-симптомів прихованої властивості, які не потрапили до оцінювання.

Це означає, що істинна залежність між величиною шуканого інтегрального показника латентної ознаки й чинниками-симптомами різних рівнів ієрархії носить не функціональний, а стохастичний, імовірнісний характер. На поверхні економічних явищ подібні залежності проявляються у вигляді кореляційних та регресійних зв'язків між досліджуваними змінними.

Ми вважаємо, що нехтування наявності кореляційно-регресійних зв'язків між змінними моделі інвестиційної привабливості (2.3) і намагання оцінити латентний показник I шляхом визначення коефіцієнтів вагомості B_{jj} за допомогою суджень експертів не виключає можливості отримання викривлених результатів щодо істинного рівня досліджуваного економічного явища внаслідок наступних причин:

а) суб'єктивізму експертних оцінок при визначенні вагомості чинників-симптомів B_{ij};

б) неадекватності моделі (2.3) істинному взаємозв'язку між її змінними в результаті ігнорування впливу ε на величину I.

Тому слід досить обережно відноситися до результатів оцінювання інтегрального показника інвестиційної привабливості підприємства, отриманих за методом сум (як простих, так і зважених) або побудованих на їх базі відповідних середніх арифметичних.

На другому місті за популярністю знаходиться метод добутку (простого чи зваженого) або побудованих на його базі відповідних середніх геометричних. В його основі лежить гіпотеза про мультиплікативну залежність досліджуваної латентної ознаки від чинників-симптомів різних рівнів ієрархії. Тобто вважається, що інтегральній показник I можна представити наступним чином:

$$I = \prod_{j=1}^{m} \prod_{i=1}^{n} B_{ij} z_{ij}. \qquad (2.4)$$

Головна відмінність мультиплікативної моделі (2.4) від адитивної моделі (2.3) полягає в тому, що підсумовування в ній замінюється на добуток чинників-симптомів прихованої властивості об'єкта. З математико-статистичного моделювання відомо, що мультиплікативні рівняння досить легко перетворюються до адитивного вигляду шляхом так званої лінеаризації змінних, яка в даному випадку зводиться до логарифмування лівої та правої частини вираження (2.4).

Таким чином, логарифмуючи дані та здійснюючи нове позначення змінних, на базі моделі (2.4) досить легко отримати вираження типу (2.3) з усіма негативними висновками та недоліками, що були визначені вище для метода сум і зв'язаними з ним середніми величинами. Отже, обидва вказаних методи страждають одними й тими ж недоліками – викривленням отриманих оцінок в результаті суб'єктивізму експертів та неадекватністю моделі, що застосовується для визначення рівня досліджуваного латентного показника.

Ще один недолік мультиплікативної моделі (2.4) оцінювання прихованої властивості економічних об'єктів полягає в тому, що у разі нульового значення будь-якого стандартизованого чинника-симптому z_i уся величина I теж перетворюється на нуль. А така ситуація цілком можлива при $X_i \approx \bar{X}$.

На нашу думку, відмінною рисою методів експертних оцінок є той факт, що вони не забезпечують отримання шуканих метричних оцінок латентних ознак економічних об'єктів, а залишаються на рівні рангів порядкової шкали з усіма їх недосконалостями і обмеженими математичними операціями градацій даної шкали (див. розділ 1.1).

Тим не менш, існують беззаперечні позитивні прикладні аспекти знайдених оцінок – ранжирування й угрупування економічних об'єктів, наприклад, підприємств за рівнем таких латентних ознак, як інвестиційна привабливість, конкурентоспроможність, ділова активність, фінансовий і майновий стан тощо з розробкою відповідних стратегій розвитку на ближню й дальню перспективу.

2.3 Виділення репрезентативного чинника-симптому

З формально-математичної точки зору задачу оцінки латентної ознаки економічних об'єктів можна сформулювати наступним чином. Нехай є n об'єктів (підприємств, працівників тощо), кожний з яких характеризується набором із m чинників-симптомів, котрі представляють собою зовнішнє проявлення прихованої ознаки, що досліджується.

Номер об'єкта позначимо через i ($i = 1, 2, \ldots, n$), номер чинника-симптому – через k ($k = 1, 2, \ldots, m$). Тоді x_{ik} – значення k-го чинника-симптому на i-му об'єкті і вся вихідна інформація представляється у вигляді матриці $\boldsymbol{X}$, розміру $n \times m$:

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1m} \\ x_{21} & x_{22} & \ldots & x_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1} & x_{n2} & \ldots & x_{nm} \end{pmatrix} \qquad (2.5)$$

Рядки цієї матриці відповідають окремим об'єктам, а стовпці – окремим чинникам-симптомам. А сама задача оцінки латентної ознаки економічних об'єктів зводиться до повної редукції матриці ***X***, тобто до пошуку єдиного вектора x (реального або умовного, штучного), який би відповідав заздалегідь установленим вимогам і умовам.

На нашу думку, одним із простіших методів вирішення поставленої задачі є виділення найбільш репрезентативного вектора матриці ***X***. Його ідея полягає в наступному: всі чинники-симптоми (вектори-стовпці матриці ***X***) взаємозв'язані між собою лінійною залежністю різного ступеня щільності. Це явище в математиці називається *мультиколінеарністю* [121; 131, с.110-112].

Чим щільніше зв'язаний даний вектор-стовпець x_s з усіма іншими, тим більше підстав розглядати його як лінійну комбінацію векторів x_1, x_2, …, x_{s-1}, x_{s+1}, …, x_m і, відповідно, носієм інформації всієї множини чинників-симптомів. Питання лише в ступені мультиколінеарності, яка, зазвичай, вимірюється величиною коефіцієнта детермінації R_s^2 у рівняннях регресії наступного виду:

$$\hat{x}_s = a_0 + a_1x_1 + a_2x_2 + \ldots + a_{s-1}x_{s-1} + a_{s+1}x_{s+1} + \ldots + a_mx_m . \qquad (2.6)$$

Очевидно, що поставлена задача буде успішно вирішена, якщо для вектора-стовпця x_s виконуються умови:

1) $R_s^2 = \max(R_k^2)$;

2) величина R_s^2 є статистично значущою із заздалегідь заданою вірогідністю помилки [129, с.127-131].

Для вимірювання всіх значень R_k^2 можна скористатися стандартними програмами регресійного аналізу і розрахувати m рівнянь виду (2.6) з відповідними величинами коефіцієнтів детермінації. Але існує більш простий спосіб знаходження R_k^2. Доведено, що величина R_k^2 зв'язана з відповідним діагональним елементом c_{kk} матриці, оберненої до кореляційної матриці ***r***, наступним чином:

$$R_k^2 = 1 - 1/c_{kk} . \qquad (2.7)$$

Отже, спочатку треба на базі матриці (2.5) розрахувати матрицю коефіцієнтів парної кореляції ***r***, наприклад, за допомогою стандартних програм кореляційного аналізу в редакторі Excel, а потім знайти обернену матрицю

$$\boldsymbol{r}^{-1} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1m} \\ c_{21} & c_{22} & \ldots & c_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1} & c_{m2} & \ldots & c_{mm} \end{pmatrix} \qquad (2.8)$$

Діагональні елементи c_{kk} матриці (2.8) змінюються в межах від 1 до $+\infty$. Якщо вектори-стовпці (чинники-симптоми) матриці $\boldsymbol{X}$ лінійно незалежні (ортогональні), то матриці $\boldsymbol{r}$ і $\boldsymbol{r}^{-1}$ є одиничними матрицями й усі $c_{kk} = 1$, а $R_k^2 = 0$. Ймовірність такого випадку дуже низька, оскільки між будь-якими двома рядами чисел спостерігається кореляція (іноді досить щільна). Тому справедливі нерівності $1 < c_{kk} < +\infty$.

У разі високої мультиколінеарності чинників-симптомів латентної ознаки, що оцінюється, $c_{kk} \to +\infty$ і $R_k^2 \to 1$. Однак, слід мати на увазі, що лінійна залежність векторів $x_1, x_2, \ldots, x_m$ це не питання наявності, а питання ступеню. Тому вирішення поставленої задачі зводиться, по суті, до перевірки статистичної значущості максимальної величини R_k^2 (c_{kk}).

Д. Феррар, Р. Глаубер визначили мультиколінеарність як ступінь відхилення від ортогональності множини змінних $x_1, x_2, \ldots, x_m$ [155]. Такий підхід дозволив їм побудувати статистичний критерій для перевірки значущості лінійної залежності досліджуваних векторів. Його статистика має вигляд:

$$F_k = (c_{kk} - 1)(n - m)/(m - 1). \qquad (2.9)$$

В умовах справедливості нульової гіпотези $H_0 : c_{kk} - 1 = 0$ (k-й чинник-симптом є лінійно незалежним від інших) проти альтернативи $H_a : c_{kk} - 1 > 0$ величина (2.9) підпорядковується F-розподілу Фішера з рівнем значущості α і $df_1 = m - 1$, $df_2 = n - m$ числами ступені вільності (детальніше див. додаток Д).

Розрахункове значення статистики критерію отримується шляхом підстановки величини $c_{ss} = \max(c_{kk})$, n, m у формулу (2.9). Його порівнюють з α-квантилем (критичним значенням) $F_{\alpha;df_1;df_2}$, заздалегідь задавшись рівнем значущості α, який, зазвичай, обирається серед малих значень (α = 0,01; 0,05). Це дозволяє мінімізувати ймовірність помилки першого роду (відхилення H_0, коли вона вірна) й виявити найбільш репрезентативний реальний чинник-симптом у якості оцінки шуканої латентної ознаки.

α-квантиль $F_{\alpha;df_1;df_2}$ знаходиться за допомогою стандартних програм, наприклад, в редакторі Excel. Для цього необхідно відкрити робочу книгу і виконати наступні дії та команди:

- Натиснути на панелі кнопку *fx* «Майстер функцій»;
- В лівій частині вікна, яка називається «Категорії», вибрати «Статистичні»;
- Потім в правій частині вікна, яка називається «Функції», обрати «FРАСПОБР»;
- Задати значення α, df_1, df_2 для зворотної функції – Enter.

Надалі виконують порівняння розрахункового значення статистики критерію з α-квантилем. У випадку $F > F_{\alpha;df_1;df_2}$ нульова гіпотеза $H_0 : c_{ss} - 1 = 0$ відхиляється і з вірогідністю 1 - α приймається альтернатива H_a, тобто можна стверджувати, що s-й чинник-симптом є лінійно залежним від усіх інших. Це означає, що він представляється у вигляді лінійної комбінації векторів $x_1, x_2, \ldots, x_{s-1}, x_{s+1}, \ldots, x_m$ й може розглядатися в ролі оцінки прихованої властивості об’єктів, що вивчаються.

У протилежному випадку ($F \leq F_{\alpha;df_1;df_2}$) нульова гіпотеза H_0 : c_{ss} - 1 = 0 не відхиляється і з вірогідністю, що дорівнює потужності критерію [129, с.131], робиться висновок, що *s*-й чинник-симптом є лінійно незалежним від усіх інших і поставлену задачу не можливо вирішити даним методом.

Останнім часом часто користуються диференційною функцією «FРАСП», яка дозволяє за заданими значеннями F, df_1, df_2 визначити ймовірність можливої похибки p при відхиленні нульової гіпотези. Для цього виконуються наступні дії та команди:

- Натиснути на панелі кнопку *fx* «Майстер функцій»;
- В лівій частині вікна, яка називається «Категорії», вибрати «Статистичні»;
- Потім в правій частині вікна, яка називається «Функції», обрати «FРАСП»;
- Задати значення F, df_1, df_2 для диференціальної функції розподілу – Enter.

Знайдена величина p порівнюється зі заздалегідь заданим рівнем значущості α. У разі $p < \alpha$ нульова гіпотеза H_0 : c_{ss} - 1 = 0 відхиляється і з вірогідністю 1 - α приймається альтернатива H_a. У цьому випадку можна стверджувати, що *s*-й чинник-симптом є лінійно залежним від усіх інших, тобто репрезентативним, який можна розглядати в ролі оцінки прихованої властивості об'єктів, що вивчається.

У протилежному випадку ($p \geq \alpha$) нульова гіпотеза H_0 : c_{ss} - 1 = 0 не відхиляється і з вірогідністю, що дорівнює потужності критерію, робиться висновок, що *s*-й чинник-симптом є лінійно незалежним від усіх інших і поставлену задачу не можливо вирішити даним методом.

2.4 Таксономічний аналіз

Одним із порівняно простих математико-статистичних методів оцінки прихованих властивостей економічних об'єктів є таксономічний аналіз або метод таксономії. Згідно сучасному тлумачному словнику таксономія (від грецьких слів *táxis* – розташування, порядок і *nómos* – закон) – це теорія класифікації та систематизації складно організованих областей дійсності, які, зазвичай, мають ієрархічну структуру [98]. Поняття «таксономія», «таксономічний аналіз» виникли вперше в біології: швейцарський ботанік О. Декандоль у 1813 р. запропонував використовувати їх при розробці класифікації рослин.

У першій половині XX ст. проблеми теорії класифікації та таксономічного аналізу починають відігравати помітну роль не тільки в біології, але й у ряді інших наук, які мали справу з множинами ієрархічно організованих дискретних об'єктів. Це відображало загальну для науки того часу тенденцію до підвищення ролі типології в науковому мисленні й були пов'язані з дослідженнями польських учених Й. Чекановського, З. Хельвіга, В. Плюти [75; 76] та ін., які заклали фундамент сучасного порівняльного багатовимірного аналізу в економіці.

Оцінка латентних ознак економічних об’єктів на основі методів таксономії базується на використанні понять їх відстані та схожості в просторі визначених дослідником чинників-симптомів. Детальне висвітлення вказаних понять, а також математико-статистичні формули їх розрахунку наводяться в роботах Ю.А. Єгупова [28; 89], О.Г. Янкового [122, с.64-66; 125, с.51-77; 130; 133]. Уведемо наступні поняття:

1. Стимулятор – чинник-симптом, зростання якого є бажаним з точки зору оцінюваної латентної ознаки. Наприклад, стимуляторами при вивченні інвестиційної привабливості підприємств є такі показники господарської діяльності, як прибуток, рентабельність, капіталовіддача, продуктивність праці, об’єм продукції, що користується попитом.

2. Дестимулятор – чинник-симптом, зростання якого є небажаним з точки зору оцінюваної прихованої властивості. Так, дестимуляторами при вивченні інвестиційної привабливості підприємств є такі показники господарської діяльності, як собівартість продукції, втрати робочого часу, кредиторська заборгованість тощо.

3. Номінатор – чинник-симптом номінальної шкали, що характеризує атрибутивні ознаки досліджуваних об’єктів: назва підприємства, економічне призначення продукції, приналежність до галузі, об'єднання.

4. Еталон – реальна або умовна точка багатовимірного простору, чинники-симптоми якої відповідають уявленням економічної науки про найкращий рівень стимуляторів і дестимуляторов з позиції досліджуваної латентної ознаки. Наприклад, в якості еталону при оцінці інвестиційної привабливості економічних об’єктів можна розглядати підприємство прибуток, рентабельність, капіталовіддача, продуктивність праці, об’єм продукції, що користується попитом, якого максимальні, а собівартість продукції, втрати робочого часу, дебіторська заборгованість – мінімальні на сукупності всіх досліджуваних об’єктів.

5. Антиеталон – поняття, протилежне поняттю еталона, тобто реальна або умовна точка багатовимірного простору, чинники-симптоми якої відповідають уявленням економічної науки про найгірший рівень стимуляторів і дестимуляторів з позиції оцінюваної прихованої властивості.

Якщо еталон – це точка верхнього полюса, до якої повинні прагнути всі об’єкти сукупності для досягнення максимального рівня латентної ознаки, то антиеталон, навпаки – точка нижнього полюса. Від неї слід триматися подалі, щоб бути в лідерах. При цьому можлива цілком точна кількісна характеристика досліджуваної прихованої властивості об’єктів на базі функцій відстаней (метрик) табл. 2.3, в якій використано наступні позначення:

z_p, z_s – стандартизовані координати об’єктів p і s, відстань між якими вимірюється;

i – номер досліджуваного об’єкта ($i = 1, 2, \ldots, p, \ldots, s, \ldots, n$);

k – номер чинника-симптому досліджуваного об’єкта ($k = 1, 2, \ldots, m$);

N, P, g – параметри степеневих метрик, що визначають їх конкретний вид;

$\boldsymbol{r}^{-1}$ – матриця, обернена до кореляційної матриці чинників-симптомів;

^{T} – знак транспонування матриці.

Таблиця 2.3

Функції відстані (метрики)

Назва відстані	Формула
1. Лінійна (міських кварталів, манхеттенська)	$d_1(z_p, z_s) = \sum_{k=1}^{m} \lvert z_{pk} - z_{sk} \rvert$
2. Евклідова	$d_2(z_p, z_s) = [\sum_{k=1}^{m} (z_{pk} - z_{sk})^2]^{1/2}$
3. Супремум-норма (Чебишева)	$d_\infty(z_p, z_s) = max \lvert z_{pk} - z_{sk} \rvert$
4. Мінковського	$d_N(z_p, z_s) = [\sum_{k=1}^{m} \lvert z_{pk} - z_{sk} \rvert^N]^{1/N}$
5. Квадрат евклідової відстані	$d^2_2(z_p, z_s) = \sum_{k=1}^{m} (z_{pk} - z_{sk})^2$
6. Степенева	$d_{P,g}(z_p, z_s) = [\sum_{k=1}^{m} \lvert z_{pk} - z_{sk} \rvert^P]^{1/g}$
7. Коефіцієнт неузгодженості	$d_p(z_p, z_s) = (\text{число } z_{pk} \neq z_{sk})/m$
8. Махаланобіса	$d_M(z_p, z_s) = (z_p - z_s)\boldsymbol{r}^{-1}(z_p - z_s)^T$

Для відстаней Мінковського із збільшенням показника ступеня N величина відстані для заданих точок z_p, z_s не зростає. Тому можна записати:

$$d_1(z_p, z_s) \geq d_2(z_p, z_s) \geq \ldots \geq d_\infty(z_p, z_s). \qquad (2.10)$$

П’ята метрика являє собою евклідову відстань, зведену в квадрат, тобто частинний випадок метрики 2.

Степенева відстань узагальнює перші п’ять показників: при $P = g$ отримуємо відстань Мінковського, при $P = 2$, $g = 1$ – квадрат евклідової відстані, при $P = g = 2$ – евклідову відстань і т.д.

Сьома метрика (коефіцієнтом неузгодженості) застосовується для розрахунку відстаней між об’єктами, які описуються атрибутивними ознаками. Вона використовується в соціологічних, політологічних, біологічних дослідженнях і являє собою частку розбіжностей значень цих ознак для двох об’єктів у загальному числі досліджуваних чинників-симптомів m.

Відстань Махаланобіса є особливою метрикою з точки зору її побудови, оскільки включає не тільки вектори точок у багатовимірному просторі, а й матрицю $\boldsymbol{r}$, що характеризує парний кореляційний зв’язок між всіма чинниками-симптомами. Дана функція відстані узагальнює метрики 5 і 2 табл. 2.3: доведено, що $d_M(z_p, z_s) \geq d_2^2(z_p, z_s)$ і рівність має місце тільки на множині лінійно незалежних (ортогональних) змінних.

Поняття схожості (близькості) об’єктів у багатовимірному ознаковому просторі є протилежним по відношенню до поняття відстані. Існує можливість простого переходу від метрик відстаней $d(z_p, z_s)$ до метрик схожості μ_{ps}. Досить застосувати, наприклад, наступне обернене перетворення:

$$\mu_{ps} = 1/[1 + d(z_p, z_s)], \qquad (2.11)$$

де $d(z_p, z_s)$ – деяка функція відстані, наприклад, з табл. 2.3.

Очевидно, що показник схожості (2.6) змінюється в межах від нуля до одиниці. При $d(z_p, z_s) \to \infty$ $\mu_{ps} \to 0$, а при $d(z_p, z_s) \to 0$ $\mu_{ps} \to 1$. Іншими словами, об'єкти повністю різні, тобто мають нульову схожість при нескінченно великій відстані між ними, і одиничну (повну) схожість, коли $z_p = z_s$. Тоді цілком доцільною виглядає наступна ідея: величина метрики подібності кожного об'єкта з еталоном або метрики відстані від кожного об'єкта до антиеталона може розглядатися як певна інтегральна (синтетична) оцінка рівня досліджуваної латентної ознаки. Насправді, чим ближче даний об'єкт до еталона, тим вище рівень прихованої досліджуваної ознаки, і навпаки. Справедливо також твердження про те, що чим далі об'єкт від антиеталона, тим він більш розвинений відносно латентної ознаки і, навпаки.

Розглянемо наступний приклад. Нехай відомі дані за чотирма підприємствами *A*, *B*, *C*, *D* ($i = 4$) про величину прибутку X_1 та собівартості продукції X_2 ($m = 2$). За антиеталон (*AE*) приймемо умовну точку, координати якої відповідають найгіршим досягнень на даній сукупності об'єктів (рис. 2.5). Тоді відстані від кожної точки (підприємства) до антиеталона будуть характеризувати рівень інвестиційної привабливості розглянутих об'єктів.

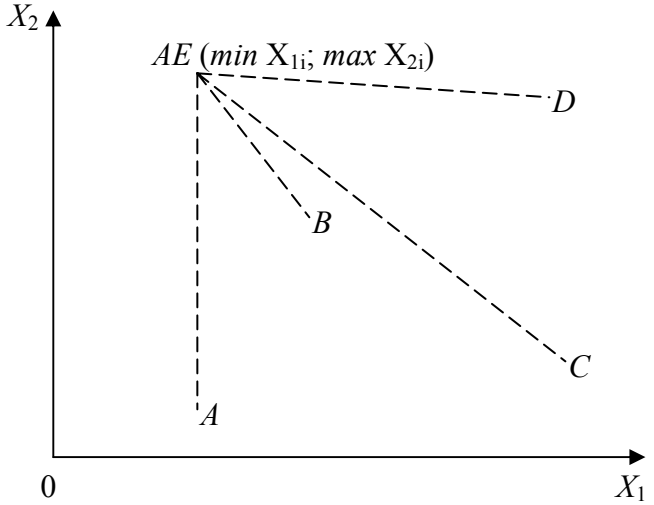


Рис. 2.5. Відстані між точками (об'єктами) і антиеталоном (*AE*) у двомірному ознаковому просторі

Рис. 2.5 наочно показує, що найбільш інвестиційно привабливим є підприємство *C*, оскільки воно далі за всіх інших розташоване від антиеталона (має порівняно високий прибуток і низьку собівартість продукції). Воно є лідером на даній сукупності об'єктів з точки зору досліджуваної латентної ознаки. Підприємство *B*, навпаки, слід віднести до аутсайдерів, оскільки величина його чинників-симптомів не надто вирізняється від координат найгіршого умовного об'єкта *AE*. Виділяють два основні алгоритми оцінки

латентної ознаки за допомогою методів таксономії – класичний та модифікований.

У класичному алгоритмі передбачається завдання еталона у вигляді точки верхнього полюса, розрахунок відстаней від нього до всіх об'єктів досліджуваної сукупності (d_i) і визначення міри схожості об'єктів з еталоном (μ_i). Остання величина розглядається як шукана інтегральна синтетична оцінка латентної ознаки. У модифікованому алгоритмі передбачається завдання антиеталона у вигляді точки нижнього полюса та розрахунок відстаней від нього до всіх об'єктів досліджуваної сукупності (d^*_i).

Класичний алгоритм оцінювання латентної ознаки на основі методу таксономічного аналізу складається з наступних основних етапів (рис. 2.6).

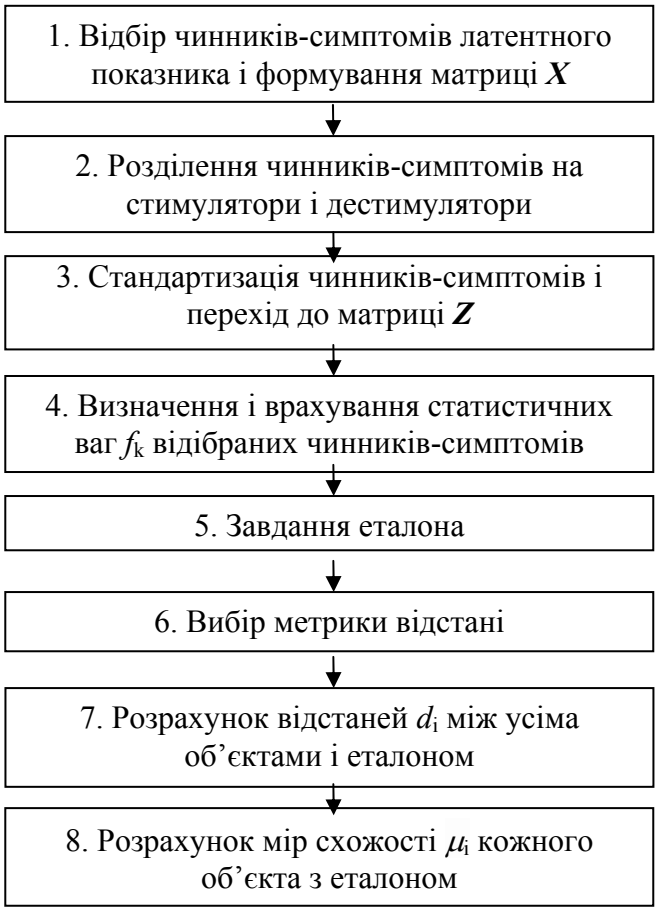


Рис. 2.6. Блок-схема класичного алгоритму оцінки латентної ознаки на базі функцій відстаней і схожості

1. Відбір чинників-симптомів латентної ознаки і формування матриці ***X***

Для забезпечення точності і обґрунтованості оцінювання необхідно застосовувати найбільш вагомі змінні, що характеризують досліджувану латентну ознаку об'єктів. Тут необхідно враховувати той факт, що абсолютні (об'ємні) чинники-симптоми, безсумнівно, важливі при індивідуальному дослідженні об'єктів, повинні бути обов'язково доповнені відносними показниками структури, координації, інтенсивності і т.п.

2. Розділення чинників-симптомів на стимулятори і дестимулятори

Даний етап є необхідним для правильного завдання еталона для порівняння. На цьому етапі на базі економічної теорії слід чітко визначитися, зростання яких чинників-симптомів бажане, позитивне, а яких – небажане, негативне з погляду оцінюваної латентної ознаки.

3. Стандартизація чинників-симптомів і перехід до матриці $\boldsymbol{Z}$

Перехід до матриці стандартизованих ознак $\boldsymbol{Z}$ виконується за формулою (1.2). Як відомо, стандартизація дозволяє привести економічні дані до одного порядку (від -3 до 3), тобто позбавитися впливу одиниць вимірювання чинників-симптомів на результати таксономічного аналізу.

4.Визначення і врахування статистичних ваг f_k відібраних чинників-симптомів

Даний етап дослідження припускає диференціювання змінних за їх впливом на латентну ознаку. На базі економічної теорії досить часто з'ясовується неоднакова роль окремих чинників-симптомів у формування її рівня. Тому змінні доцільно диференціювати за допомогою системи статистичних ваг f_k, які забезпечили б посилення важливості одних чинників-симптомів, наприклад, шляхом множення відповідних стовпців матриці $\boldsymbol{Z}$ на f_k й послаблення інших відповідно до їх реального впливу на латентну ознаку. При цьому слід дотримуватись вимоги $\sum_{k=1}^{m} f_k = 1$.

5. Завдання еталона

У якості еталона приймається реальна або умовна точка в багатовимірному просторі чинників-симптомів, координати якої характеризують якнайкращі (з урахуванням розділення змінних на стимулятори і дестимулятори) досягнення об'єктів, що вивчаються. Еталон відображає максимально можливий, потенційний рівень латентної ознаки і служить своєрідним орієнтиром (базою порівняння) для всіх об'єктів досліджуваної сукупності. Можливі різні методи завдання еталона:

1) на основі значень чинників-симптомів даної сукупності об'єктів;

2) на основі значень чинників-симптомів інших сукупностей, наприклад, з урахуванням досвіду інших підприємств, регіонів, країн тощо.

У першому випадку для стимуляторів еталонні значення (z_{0k}) визначаються так:

$$z_{0k} = \max_i z_{ik}, \qquad (2.12)$$

а для дестимуляторів:

$$z_{0k} = \min_i z_{ik}, \text{ або } z_{0k} = 0. \qquad (2.13)$$

У другому випадку в якості еталонних значень приймаються, наприклад, світові досягнення за даними чинниками-симптомами.

6. Вибір метрики відстані

У літературі не існують конкретні рекомендації щодо вибору метрик, які містить табл. 2.3. Виходячи з власного досвіду, ми вважаємо, що слід, у першу чергу, використовувати степеневі метрики – лінійну, евклідову відстань, супремум-норму, враховуючи співвідношення (2.10) між ними.

7. Розрахунок відстаней d_i між усіма об'єктами і еталоном здійснюється на основі формул табл. 2.3 у відповідності до обраній на попередньому етапі метриці. Знайдені відстані до еталону d_i тлумачаться як інтегральні резерви підвищення латентної ознаки кожного об'єкту. Для еталонної точки $d_0 = 0$.

8. Розрахунок мір схожості μ_i кожного об'єкта з еталоном

Після знаходження відстаней до еталона визначаються міри схожості μ_i кожного об'єкта з еталоном за формулою (2.11), які й розглядаються у якості шуканих оцінок латентної ознаки економічних об'єктів, що вивчається.

На рис. 2.7 наведено блок-схему модифікованого алгоритму оцінки латентної ознаки на базі функцій відстаней.

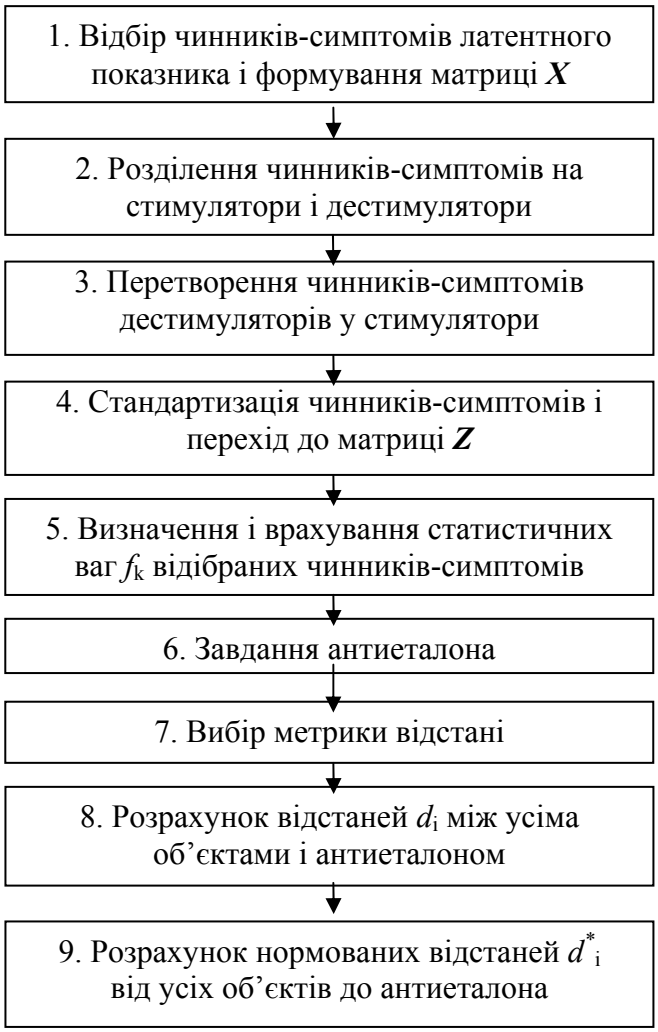


Рис. 2.7. Блок-схема модифікованого алгоритму оцінки латентної ознаки на базі функцій відстаней до антиеталона

Порівняння її з блок-схемою класичного алгоритму (рис. 2.6) показує, що вони розрізняються лише в етапах, пов'язаних з перетворенням чинників-симптомів дестимуляторів у стимулятори (етап 3), завданням антиеталона (етап 6) і нормуванням відстаней до точки нижнього полюса (етап 9).

Проте, слід мати на увазі, що вказані відмінності можуть привести до різних значень оцінок латентної ознаки, що вивчається, одержаних в результаті застосування кожного методу для одних і тих же статистичних даних.

У модифікованому алгоритмі передбачається завдання антиеталона у вигляді точки нижнього полюса z_0 з координатами

$$z_0 (a, a,..., a), \tag{2.14}$$

де a – довільне від'ємне число, що задовольняє умові:

$$a \le \min_{i,k} z_{ik} . \tag{2.15}$$

Зазвичай, a приймає значення -2, -3, -4. Вихід за ці межі практично неможливий унаслідок того, що згідно правилу трьох сигм для нормального розподілу однорідної змінної усі її стандартизовані значення з імовірністю 99,7 % знаходяться в межах від -3 до 3.

Нижній полюс (2.14) є початковою, нульовою точкою відліку рівня латентної ознаки в просторі стандартизованих чинників-симптомів, які є стимуляторами. Очевидно, що відстань від об'єкта до антиеталона в цьому просторі можна трактувати як безпосередню кількісну оцінку шуканої прихованої ознаки, яка не вимагає переходу до міри схожості. Тому на самому початку таксономічного аналізу всі чинники-симптоми, які є дестимуляторами, необхідно перетворити в стимулятори наступним чином:

$$x_{ik} = 1 / x'_{ik} , \quad \text{або} \quad x_{ik} = 1 - x'_{ik} , \tag{2.16}$$

де x'_{ik} – початкова реалізація чинника-симптому дестимулятора.

Для того, щоб усі значення оцінки латентної ознаки d_i знаходились у межах від 0 до 1 застосовується нормування:

$$d_i^* = d_i /(-2a\sqrt{n}). \tag{2.17}$$

Величина d^*_i, визначена за формулою (2.17), тлумачиться наступним чином: високі значення d^*_i свідчать про високий рівень шуканої латентної ознаки і, навпаки. Зупинимося докладніше на прикладних аспектах знайдених оцінок μ_i та d^*_i.

По-перше, і цей напрямок ми вважаємо головнішим, вони можуть бути використані в цілях кількісного оцінювання досліджуваної прихованої властивості економічних об'єктів, тобто переводу її з класу якісних латентних ознак до класу кількісних масштабованих ознак (див. рис. 1.2).

По-друге, величини μ_i та d^*_i дозволяють складати рейтинг досліджуваних об'єктів на основі ранжирування (впорядкування) точок багатовимірного

ознакового простору за величиною оціненої латентної ознаки шляхом присвоєння кожній із них відповідного рангу (1, 2, ..., *n*).

По-третє, на їх основі стає можливою багатовимірна класифікація об'єктів з виділенням, наприклад, груп лідерів, середняків і аутсайдерів. Це набуває особливої актуальності у зв'язку з постійною необхідністю розширення кола ділових партнерів шляхом визначення перспективніших партнерів в господарській діяльності, дозволяє виявити становище конкурентів, намітити стратегію і тактику розвитку даного підприємства в мінливому зовнішньому середовищі. В однорідних групах стає можливим проведення поглиблених досліджень взаємозв'язків між змінними із застосуванням апарату статистичного моделювання та прогнозування.

Щодо другого практичного спрямування результатів таксономічного аналізу, то тут треба відмітити, що під рейтингом, зазвичай, розуміють впорядкований список досліджуваних об'єктів, побудований за зменшенням досліджуваної латентної ознаки цих об'єктів. Рейтинг називається експертним, якщо впорядкування визначається експертом або групою експертів. Аналітичним називається рейтинг, в якому впорядкування об'єктів визначається за математико-статистичною моделлю.

Не дивлячись на те, що будь-яке рейтингування не дозволяє вийти за межі класу ознак, виміряних у порядковій шкалі, ми, тим не менш, вважаємо, що аналітичний рейтинг має суттєві переваги у порівнянні з експертним рейтингом внаслідок певного суб'єктивізму останнього. Тому питання складання рейтингів об'єктів за досліджуваною латентною ознакою на основі методів таксономії представляє самостійний науковий інтерес з точки зору прикладного використання математико-статистичних моделей в економіці.

Аналіз літератури стосовно даного питання дозволяє зробити наступні висновки [3; 110; 111]: чим більше відстань від еталона (антиеталона), тим більша імовірність помилки алгоритму й похибки при побудові рейтингу. Тобто методика ранжирування за класичним алгоритмом таксономії точніше визначає провідні місця і досить може помилятися при ранжируванні останніх позицій. А методика впорядкування із застосуванням модифікованого алгоритму таксономії, навпаки, більш точно визначає останні місця і загрожує помилками при визначенні лідерів.

Тому в останній час при складанні рейтингів пропонується використовувати змішаний (об'єднаний) підхід, заснований на використанні результатів двох різних методів ранжирування об'єктів – відносно еталона, та відносно антиеталона.

На підставі ранжирувань, виконаних за класичним та модифікованим алгоритмами, будується агрегована рейтингова оцінка методом строкових сум наступним чином:

$$r_{\text{i}} = R_{\text{ei}}K_{\text{ei}} + R_{\text{ai}}K_{\text{ai}}\,, \qquad (2.18)$$

де r_{i} – загальне рейтингове число *i*-го об'єкта;

R_{ei} – рейтинг *i*-го об'єкта, визначений за допомогою класичного

алгоритму;

R_{ai} – рейтинг *i*-го об'єкта, знайдений за допомогою модифікованого алгоритму;

K_{ei}, K_{ai} – коефіцієнти вагомості відповідно рейтингу, визначеного за допомогою класичного та модифікованого алгоритму.

Коефіцієнти K_{ei} та K_{ai} знаходяться з урахуванням наступних міркувань:

а) для класичного алгоритму – чим вище R_{ei} (тобто, чим більша відстань до еталона), тим меншим повинен бути коефіцієнт вагомості K_{ei};

б) для модифікованого алгоритму, навпаки, чим менше R_{ai} (тобто, чим ближче об'єкт до антиеталона), тим більшим повинен бути коефіцієнт вагомості K_{ai}.

При цьому повинні виконуватись наступні вимоги:

$$\sum_{i=1}^{n} K_{ai} = 1 , \quad \sum_{i=1}^{n} K_{ai} = 1 . \tag{2.19}$$

Коефіцієнти вагомості розраховуються на основі формул Фішберна [101]:

$$K_{ei} = \frac{2(n - R_{ei} + 1)}{(n+1)n} ; K_{ai} = \frac{2R_{ai}}{(n+1)n} . \tag{2.20}$$

.

Легко показати, що для коефіцієнтів вагомості виконуються вимоги (2.19):

1) $$\sum K_{ei} = \sum \frac{2(n - R_{ei} + 1)}{n(n+1)} = 2\frac{n \cdot n - (1 + 2 + \text{K} + (n-1) + n) + n \cdot 1}{n(n+1)} =$$

$$= 2\frac{n(n+1) - \frac{1+n}{2} n}{n(n+1)} = \frac{2n(n+1) - n(n+1)}{n(n+1)} = \frac{n(n+1)}{n(n+1)} = 1.$$

2) $$\sum K_{ai} = \sum \frac{2R_{ai}}{n(n+1)} = \frac{2(1 + 2 + \text{K} + (n-1) + n)}{n(n+1)} = \frac{2\frac{1+n}{2} n}{n(n+1)} =$$

$$= \frac{n(n+1)}{n(n+1)} = 1. \tag{2.21}$$

Тут межі підсумовування по *i* від 1 до *n* для простоти опущені.

Далі, після розрахунку об'єднаних рейтингових оцінок r_i кожному об'єкту присвоюється кінцевий ранг R_i наступним чином: об'єкту з мінімальним значенням r_i присвоюється ранг 1; об'єкту з мінімальним значенням r_i серед тих, що залишилися, присвоюється ранг 2; а об'єкту з максимальним значенням r_i – ранг *n*.

В останні роки деякі дослідники намагаються застосувати ідеї змішаного рейтингування безпосередньо до самих оцінок латентної ознаки μ_i і d_i [62; 139; 140]. Так, Т.Б. Надтока, А.Г. Виноградов пропонують об'єднаний підхід до оцінювання рівня соціально-економічного розвитку підприємства, заснований

на використанні результатів двох різних методів ранжирування об'єктів – відносно еталона, та відносно антиеталона. Вони вважають, що нормування знайдених оцінок латентної ознаки μ_i і d_i на основі максимальних значень $\max\mu_i$, $\max d_i$ забезпечує їх зіставний вигляд й можливість розрахунку осередненого показника:

$$D_i = (\mu'_i + d'_i)/2, \qquad (2.22)$$

де $\mu'_i = \mu_i/\max\mu_i$, $d_i{}'_i = d_i/\max d_i$ – нормовані значення оцінок латентної ознаки *i*-го об'єкта, отримані відповідно за алгоритмами класичної та модифікованої таксономії.

Т.Б. Надтока, А.Г. Виноградов указують на перевага такого прийому: об'єднані оцінки D_i стають більш диференційованими, оскільки їх стандартне відхилення вище [62, с.187-188].

На завершення відмітимо, що всі розрахунки, пов'язані з стандартизацією чинників-симптомів, визначенням відстаней d_i та мір схожості μ_i кожного об'єкта з еталоном, доцільно проводити в системі STATISTICA, в модулі «Кластерний аналіз» (Cluster Analysis) [11; 12]. Приклади таких розрахунків з детальним математико-статистичним та економічним тлумаченням знайдених в процесі таксономічного аналізу параметрів можна знайти в роботах [30; 125, с.51-77; 130; 143; 149; 150].

2.5 Факторний (компонентний) аналіз

Другим важливим напрямком багатовимірних математико-статистичних методів, орієнтованих на вирішення завдання оцінювання латентних ознак, є алгоритми факторного аналізу. Джерела факторного аналізу, як і інших багатомірних методів, виходять із робіт англійського психолога і статистика Ф. Гальтона. Запропоновані ним в кінці XIX ст. ідеї й підходи до обробки й інтерпретації експериментальних даних послужили основою таких сучасних дисциплін, як антропометрія, диференціальна психологія, багатовимірний статистичний аналіз.

З 1890 р. у дослідженнях Гальтона активну участь приймав англійський математик і біолог К. Пірсон, який у серії своїх робіт математизував й істотно розвив статистичні ідеї Гальтона. Пірсон, працюючи з антропометричними даними, в 1901 р. запропонував метод головних осей, котрий зараз широко відомий як метод головних компонент.

Факторний аналіз своїми коріннями тісно пов'язаний з науковими поглядами в області психології, біології, фізіології, медицини. Провідні психологи, математики і статистики кінця XIX – початку XX ст. активно обговорювали теорію існування загальної обдарованості й специфічних здатностей людини. У той час починає свою наукову кар'єру загальновизнаний основоположник факторного аналізу Ч. Спірмен. В 1903 р. він захищає

дисертацію з психології й примикає до статистичної групи Оксфордського університету, що розробляла тести інтелекту за проектом Гальтона для обстеження здатностей учнів англійських шкіл.

Починаючи з XX ст. й до теперішнього часу, на розвиток факторного аналізу сильний вплив роблять математична статистика й використання електронно-обчислювальної техніки. Д. Лоули в 1940 р. запропонував оцінювати факторні навантаження методом максимальної правдоподібності, зосередивши свої зусилля на тім, щоб дати опис факторного аналізу в термінах математичної статистики й теорії ймовірностей.

У 50-60-х роках XX ст. одержали подальший розвиток математико-статистичні основи моделей з латентними показниками, які були підсумовані в роботах американського соціолога П. Лазарсфельда, який привніс математичні методи в суспільні науки [47; 48]. Однак, спроби застосувати конкретні латентно-структурні моделі до реальних даних зіштовхнулися зі значними обчислювальними й методичними труднощами. Спочатку оптимістичні надії, пов'язані з декларуванням широких потенційних можливостей методу, змінилися більш стриманим до нього відношенням.

У другій половині минулого сторіччя в зв'язку з бурхливим розвитком комп'ютерної техніки знову відродився інтерес до моделей з латентними показниками. Зокрема, було запропоновано ряд ефективних алгоритмів, у яких успішно переборені обчислювальні й прикладні труднощі. Латентно-структурні моделі факторного аналізу стали доступними і прийнятними для використання не тільки в психологічних, соціологічних, а й в економічних та інших дослідженнях.

Якщо методи таксономії «працюють» з об'єктами (підприємствами), то методи факторного аналізу направлені на чинники-симптоми. Їх суть полягає в знаходженні нових штучних змінних (загальних факторів), які служать оцінками досліджуваної латентного ознаки, наприклад, конкурентоспроможності, фінансового стану, інвестиційної привабливості підприємства. Загальні фактори, названі в рамках даного методу головними компонентами, являють собою лінійні комбінації спостережуваних чинників-симптомів і використовуються в подальшому аналізі для пояснення кореляційних зв'язків між ними.

Серед алгоритмів факторного аналізу, які досить детально і всебічно описані в математико-статистичній літературі [6; 7; 24; 32; 37; 39; 91; 102; 126; 133; 134], можна вказати наступні найбільш відомі методи, що розрізняються за способом визначення факторного рішення: головних компонент; мінімальних залишків; центроїдний; максимальної правдоподібності; головних осей. Усі вони надають добре узгоджені результати з досить високим коефіцієнтом парної кореляції. Тому розглянемо їх застосування на прикладі найбільш популярного різновиду факторного аналізу – методу головних компонент [25; 101; 126, с.19-61].

Головні компоненти відповідають наступним основним вимогам методу:

- лінійно незалежні (ортогональні);
- стандартизовані;

- перша головна компонента повинна пояснювати максимальну частку варіації чинників-симптомів, друга головна компонента – максимальну частку варіації чинників-симптомів, що залишилася після першої компоненти, і т.д.

Звідси випливає фундаментальна ідея методу – виділення таких штучних змінних, які описували б максимальну частку варіації вихідних чинників-симптомів або кореляційні зв'язки між ними. При цьому число головних компонент може бути істотно менше числа спостережуваних чинників-симптомів. Найскладнішим і відповідальнішим моментом методу є змістовна інтерпретація виділених і виміряних головних компонент і, зокрема, першої головної компоненти, яка, зазвичай, розглядається в якості оцінки шуканої латентної ознаки.

Отримана стандартизована змінна утворює базу для ранжирування та угрупування об'єктів за величиною прихованої властивості, що визначає кореляційні зв'язки спостережуваних чинників-симптомів. За допомогою значень виміряної першої головної компоненти для кожного об'єкта можна виділити групи лідерів, середняків та аутсайдерів.

Алгоритм оцінки латентної ознаки на базі методу головних компонент представлений на рис. 2.8.

Рис. 2.8. Блок-схема оцінювання латентної ознаки економічних об'єктів за допомогою метода головних компонент

На першому етапі відбувається формування матриці ***X*** вихідних даних про чинники-симптоми прихованої властивості досліджуваних об'єктів. Така початкова інформація базується на відповідній економічній теорії латентної ознаки, що вивчається, з урахуванням знань, досвіду та навичок дослідника.

Уведемо наступні позначення:

i – номер досліджуваного об'єкта ($i = 1, 2, ..., n$);

k – номер чинника-симптому ($k = 1, 2, ..., m$);

L – номер головної компоненти ($L = 1, 2, ..., m$);

x_{ki} – значення k-го чинника-симптому в i-го об'єкта;

a_{kL} – факторне навантаження L-ої головної компоненти F_L k-м чинником-симптомом (коефіцієнт парної кореляції між F_L і Z_k).

Тоді вихідна інформація про всі значення чинників-симптомів $x_1, x_2, \ldots, x_m$ за всіма досліджуваними об'єктами може бути представлена у вигляді матриці $\boldsymbol{X}$, розміру $m \times n$:

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix} \qquad (2.23)$$

Рядки цієї матриці відповідають окремим чинникам-симптомам, а стовпці – окремим об'єктам.

Другий етап включає стандартизацію чинників-симптомів і приведення даних до одного і того ж порядку, оскільки спостережувані змінні $x_1, x_2, \ldots, x_m$, що включені в матрицю $\boldsymbol{X}$, неоднорідні, мають різні одиниці виміру. В результаті утворюється матриця $\boldsymbol{Z}$, розміру $m \times n$:

$$\boldsymbol{Z} = \begin{pmatrix} z_{11} & z_{12} & \ldots & z_{1n} \\ z_{21} & z_{22} & \ldots & z_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ z_{m1} & z_{m2} & \ldots & z_{mn} \end{pmatrix} \qquad (2.24)$$

На третьому етапі проводиться розрахунок кореляційної матриці $\boldsymbol{r}$, яка відображає зв'язки між чинниками-симптомами. Відмітимо, що матриця $\boldsymbol{r}$, розміру $m \times m$, є симетричною матрицею з одиницями на головній осі. Її визначення здійснюється на основі матричної формули:

$$\boldsymbol{r} = \boldsymbol{Z}^{\mathrm{T}} \boldsymbol{Z} / n, \qquad (2.25)$$

де $\boldsymbol{Z}^{\mathrm{T}}$ – матриця, транспонована по відношенню до матриці стандартизованих чинників-симптомів.

Матриця $\boldsymbol{r}$, розміру $m \times m$, має вигляд:

$$\boldsymbol{r} = \begin{pmatrix} 1 & r_{11} & r_{12} & \ldots & r_{1m} \\ r_{21} & 1 & r_{23} & \ldots & r_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{m1} & r_{m2} & r_{m3} & \ldots & 1 \end{pmatrix} \qquad (2.26)$$

Вона характеризує щільність і напрямок усіх кореляційних зв'язків, що існують між обраними чинниками-симптомами.

Знаходження матриці факторних навантажень ***A*** відбувається на четвертому етапі. Визначення a_{kL} – одна з найважливіших операцій факторного аналізу латентних ознак економічних об'єктів. Завдання одержання матриці ***A*** й виділення на її основі головних компонент складаються з установлення числа й виду осей координат, необхідних для відображення взаємної кореляції вихідних чинників-симптомів.

Матриця факторних навантажень ***A*** знаходиться так:

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}^{1/2}, \tag{2.27}$$

де $\boldsymbol{\Lambda}$ – діагональна матриця характеристичний коренів λ_L матриці ***r***, розміру $m \times m$;

V – матриця нормалізованих характеристичних векторів v_L матриці ***r***, що відповідають характеристичним кореням λ_L.

Із співвідношення (2.27) вектори-стовпці матриці ***A*** визначаються наступним чином:

$$\boldsymbol{a}_{\mathbf{L}} = \boldsymbol{v}_{\mathbf{L}}(\lambda_{\mathrm{L}})^{1/2}. \tag{2.28}$$

Значення окремих факторних навантажень a_{kL} знаходяться так:

$$a_{kL} = v_{kL}(\lambda_L / \sum_{k=1}^{m} v_{Lk}^2)^{1/2}. \tag{2.29}$$

Розрахувавши за формулою (2.29) значення факторних навантажень, утворюють з них квадратну матрицю ***A***, розміру $m \times m$, яка грає фундаментальну роль як в методі головних компонент, так і в усьому факторному аналізі:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{pmatrix} \tag{2.30}$$

На останньому етапі проводиться виділення, вимір та інтерпретація головних компонент, зокрема, першої головної компоненти F_1. Справа в тім, що в загальному випадку компонентного (факторного) аналізу треба визначити, які штучні змінні із m знайдених доцільно виділяти, тобто залишати для подальшого дослідження, а які можна відкинути як слабо інформативні. Для вирішення цієї проблеми використовується формула:

$$V_L = \lambda_L. \tag{2.31}$$

Вираження (2.31) визначає абсолютний вклад L-ї головної компоненти V_L у пояснення загальної варіації чинників-симптомів $x_1, x_2, \ldots, x_m$. Отже, величина характеристичних коренів матриці $\boldsymbol{r}$ ($\lambda_1, \lambda_2, \ldots, \lambda_m$) як раз і характеризує важливість кожної головної компоненти $F_1, F_2, \ldots, F_m$. Причому, виходячи з вимог методу, вони ранжирувані таким чином:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m , \tag{2.32}$$

тобто перша головна компонента пояснює максимальну долю варіації чинників-симптомів, друга головна компонента – максимальну долю варіації чинників-симптомів, що залишилася після першої компоненти, і т.д. Саме таке ранжирування й визначає роль першої головної компоненти F_1 в оцінюванні шуканої латентної ознаки економічних об'єктів.

Відносний вклад першої головної компоненти в пояснення загальної варіації чинників-симптомів обчислюється за формулою:

$$d_1 = 100(\lambda_1/m) \tag{2.33}$$

і вимірюється у відсотках. Величина m тут виражає загальну дисперсію всіх стандартизованих чинників-симптомів. Як показує досвід, поставлене завдання оцінювання вважається виконаним, якщо виконується умова $d_1 \geq 10$ %.

Величина d_1 показує, скільки відсотків варіації вихідних чинників-симптомів описує перша головна компонента, яка розглядається в якості оцінки шуканої латентної ознаки. А величина (100 - d_1) характеризує втрати інформації при відкиданні (m - 1) останніх головних компонент.

Вимір виділеної першої головної компоненти F_1 полягає в розрахунку її значень для кожного досліджуваного об'єкта відповідно до матричної формули:

$$\boldsymbol{F} = \boldsymbol{A}^{-1}\boldsymbol{Z} . \tag{2.34}$$

На практиці для кількісної оцінки латентної ознаки використовується співвідношення:

$$\boldsymbol{F} = \boldsymbol{\Lambda}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Z} . \tag{2.35}$$

Значення першої головної компоненти для i-го об'єкта у відповідності з формулою (2.35) розраховується так:

$$F_{1i} = \sum_{k=1}^{m} (a_{k1} / \lambda_1) z_{ki} . \tag{2.36}$$

Тлумачення F_1 базується на вивченні структури матриці факторних навантажень A, зокрема, її першого стовпця. При цьому особлива увага приділяється високим за абсолютною величиною значенням факторних навантажень a_{k1} ($|a_{k1}| \geq 0{,}7$). Вони визначають ті чинники-симптоми, які

«навантажують» економічним змістом виділену штучну змінну – першу головну компоненту F_1.

У разі, коли для всіх k виконується нерівність $|a_{k1}| \geq 0{,}7$, то F_1 називається генеральною компонентою. В протилежному випадку F_1 – групова компонента. Для оцінювання латентної ознаки економічних об'єктів перша ситуація є найбільш сприятлива, оскільки в цьому випадку $d_1 >> 10$ %.

Слід мати на увазі, що в реальних економічних дослідженнях досить часто зустрічаються ситуації, коли одну і ту ж головну компоненту навантажують різні за змістом чинники-симптоми, які досить трудно об'єднати. У цьому випадку тлумачення побудованої штучної змінної сильно ускладнено, а її практична цінність зводиться нанівець.

У простішому випадку, коли розглядається тільки два вихідних чинника-симптоми x_1, x_2 всю процедуру компонентного аналізу можна зобразити графічно на рис. 2.9, на якому досліджувані багатовимірні об'єкти представлені у вигляді точок кореляційного поля (еліпса) в системі координат $X_1 0 X_2$. Стандартизація вихідних змінних (другий етап методу головних компонент) полягає в перенесенні початку координат у центр кореляційного еліпса, що відповідає центруванню даних, і деякому його стиску або розтягу, тобто їх нормуванню. В результаті отримаємо вираження x_1, x_2 в системі координат $Z_1 0 Z_2$.

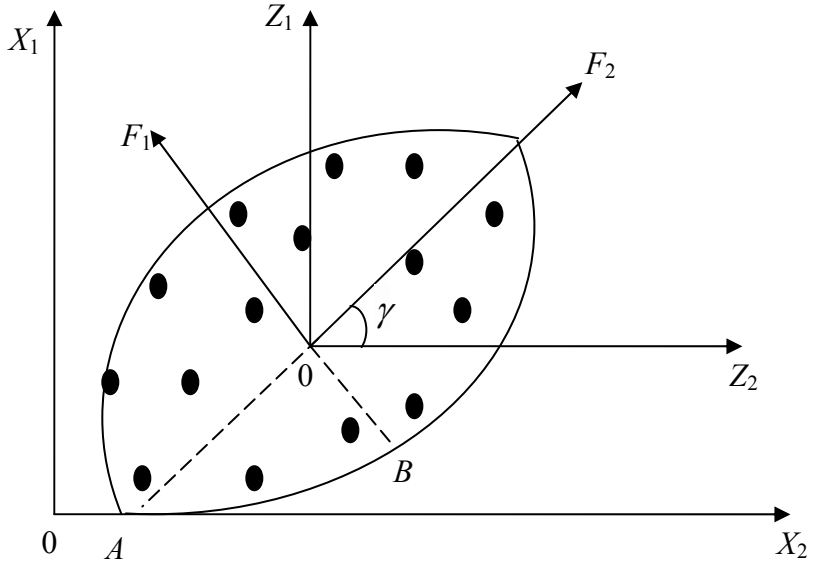


Рис. 2.9. Геометричне тлумачення метода головних компонент при $m = 2$

Зворотне лінійне перетворення стандартизованих чинників-симптомів у головні компоненти (третій і четвертий етапи процедури) зводиться до повороту координатних осей $Z_1 0 Z_2$ на такий кут γ, при якому нова координатна вісь $0F_1$, відповідна першій головній компоненті, буде зорієнтована по першій головній осі еліпса AF_1. А нова координатна вісь $0F_2$, відповідна другій головній компоненті, буде зорієнтована по другій головній осі еліпса BF_2.

Очевидно, що максимальний розкид точок, вимірюваний загальною дисперсією вихідних ознак x_1, x_2, спостерігається саме по першій головної осі

еліпса AF_1, а мінімальний – по другій головній осі еліпса BF_2 відповідно з вимогою $\lambda_1 \geq \lambda_2$. Самі осі $0F_1$, $0F_2$, тобто головні компоненти F_1, F_2 взаємно перпендикулярні і їх вимірювання (п'ятий етап методу) являє собою вираження координат всіх точок (об'єктів) у новій системі координат $F_1 0 F_2$.

Геометричне тлумачення компонентного аналізу, наведене на рис. 2.9, дає ключ до розуміння сутності обговорюваного методу, пояснює походження багатьох наукових термінів факторного аналізу. Зокрема назву методу «головних осей» бере свій початок від головних осей кореляційного еліпса, по яких зорієнтовані головні компоненти. Центроїдний метод факторного аналізу отримав свою назву тому, що нові осі $0F_1$, $0F_2$ повинні обов'язково проходити через «центр ваги» еліпса на рис. 2.9, тобто через точку 0 в системі координат $Z_1 0 Z_2$.

Виконання громіздких обчислень на комп'ютерах відкрило можливість обробляти великі кореляційні матриці, а також дозволило застосовувати методи Монте-Карло й множинного регресійного моделювання в процедурах факторного аналізу. Тепер завдяки таким комп'ютерним системам, як STATISTICA відкрився шлях до нових технологій статистичної обробки даних, що максимально скорочує рутинні технічні процедури [11; 12].

Система STATISTICA і її модуль «Факторний аналіз» (Factor Analysis) повністю вивільняють дослідника від складних та трудомістких обчислювань значень загальних факторів (головних компонент), максимально зорієнтовані на візуалізацію досліджуваної інформації за допомогою потужних графічних засобів, коли на користувача лягають тільки творчі операції прийняття рішень у процесі статистичного аналізу даних, зокрема, в ході оцінювання латентних ознак економічних об'єктів. Приклади отримання таких оцінок за допомогою системи STATISTICA наведені в роботах [30; 126, с.43-61; 136; 138].

2.6 Оцінка латентних ознак на основі сполучення кластерного і дискримінантного аналізу

Ще один підхід до оцінювання прихованих властивостей заснований на поєднанні двох напрямків багатовимірних математико-статистичних методів – кластеризації сукупності економічних об'єктів та побудові дискримінантної функції. Основні ідеї даного підходу сходять до методів експрес-діагностики, зокрема, до моделей Е. Альтмана (1969, 1983), Р. Таффлера і Г. Тішоу (1977), Г. Спрінгейта (1978) та ін. [99; 147; 148; 152-154].

У відповідності до названих напрямків багатовимірного аналізу весь процес оцінки досліджуваної латентної ознаки розділяється на три послідовні стадії [125, с.78-146, 159-180]:

1) багатовимірне угрупування сукупності об'єктів на основі їх кластеризації;

2) побудова дискримінантної функції;

3) визначення оцінок латентної ознаки об'єктів за допомогою побудованої дискримінантної функції.

Розглянемо стисло їх математико-статистичну сутність. Почнемо з першої стадії, яка завжди передує дискримінації і безпосередньому оцінюванню шуканої латентної ознаки.

Завдання кластерного аналізу може бути сформульовано, виходячи з наступних розміркувань. Нехай сукупність, що складається з n об'єктів, кожний з яких описується за допомогою m чинників-симптомів латентної ознаки, задана у вигляді матриці вихідних $\boldsymbol{X}$ або стандартизованих даних $\boldsymbol{Z}$, розміру $n \times m$.

Тоді кластером називається така компактна група об'єктів із всієї вихідної сукупності, для якої середній квадрат внутрігрупової відстані від об'єктів групи до її центра ваги менше середнього квадрата відстані від всіх об'єктів до центра ваги всієї вихідної сукупності. Чим більше серед виділених груп кластерів, тим більш успішною можна вважати отриману розбивку досліджуваної сукупності об'єктів.

Тоді завдання кластерного аналізу зводяться до пошуку й виділення у вихідній сукупності об'єктів максимального числа кластерів, які розглядаються як кількісно однорідні групи одночасно за всіма чинниками-симптомами. Іншими словами, у процесі кластеризації треба прагнути одержати таку розбивку сукупності, щоб кожний об'єкт належав до однієї й тільки до однієї групи, та відстані між об'єктами однієї групи були істотно меншими у порівнянні з відстанями між об'єктами різних груп.

Для поставленої задачі це буде означати, що вдалося виділити групи об'єктів, наприклад, підприємств з близькими значеннями чинників-симптомів, що відображають шукану латентну ознаку, наприклад, конкурентоздатність підприємств.

Як і при застосуванні методів таксономії, одним із важливіших моментів кластерного аналізу є питання визначення певного кола чинників-симптомів, що характеризують латентну економічну ознаку об'єктів. Слід мати на увазі, що це проблема не математико-статистична, а економічна, яка повинна вирішуватися в ході якісного дослідження з урахуванням положень відповідної науки (наприклад, теорії конкуренції та конкурентоспроможності підприємств).

Після визначення матриці вихідних даних $\boldsymbol{X}$, стандартизації чинників-симптомів і утворення матриці $\boldsymbol{Z}$, розрахунку матриці відстаней $\boldsymbol{D}$ між усіма об'єктами переходять до безпосереднього застосування алгоритмів кластерного аналізу, яких у теперішній час в літературі налічується більше ста [26; 55; 125]. Всі вони можуть бути згруповані у три основні напрямки:

а) процедури прямої класифікації;

б) оптимізаційні алгоритми;

в) апроксимаційні підходи.

Процедури прямої класифікації – це історично найбільш ранній напрямок кластерного аналізу, пов'язаний з іменами німецького біолога Ф. Гейнке, польського антрополога К. Чекановського, які на початку XX ст. висунули ідеї пошуку компактних груп об'єктів у ознаковому просторі.

Суть першого напрямку укладається в чіткому формулюванні поняття кластера й утворенні груп об'єктів, що відповідають даному формулюванню.

Найбільшого поширення серед процедур прямої класифікації одержали ієрархічні алгоритми, які базуються на наступному визначенні кластера: всі відстані між об'єктами усередині групи повинні бути менше будь-якої відстані між об'єктами групи й іншою частиною множини об'єктів.

Ієрархічні алгоритми дозволяють одержувати послідовну розбивку сукупності об'єктів за певним правилом. Вони підрозділяються на дивізімні й агломеративні.

Дивізімні алгоритми починають роботу з розгляду вихідної сукупності як одного кластера й послідовно розділяють її на більш дрібні групи, аж до розбивки, коли кожний об'єкт вважається окремим кластером. Агломеративні алгоритми працюють у зворотному порядку.

Ієрархічні алгоритми характеризуються рядом переваг у порівнянні з іншими процедурами кластерного аналізу. Відзначимо важливіші з них:

- відносна простота й змістовна ясність;
- допустимість втручання в роботу алгоритму;
- можливість графічного подання процесу класифікації у вигляді дерева об'єднання (дендрограми);
- порівняно невисока трудомісткість розрахунків.

Зупинимося докладніше на найбільш популярних процедурах прямої класифікації – ієрархічних агломеративних і деяких інших алгоритмах кластерного аналізу.

На першому кроці даних процедур кожний об'єкт вважається окремим кластером і провадиться об'єднання (агломерація) кластерів відповідно до деякого правила, що визначає послідовність (ієрархію) такого об'єднання. Алгоритми зазначеного типу розрізняються між собою, головним чином, критеріями, які використовуються при об'єднанні кластерів. Головні з них наступні:

1. Критерій «ближнього сусіда». В англомовній літературі даний критерій відомий як простий (одиночний) зв'язок (Single linkage). На кожному кроці поєднуються кластери K_p і K_s, відстань між найближчими об'єктами p і s яких мінімальна. При його використанні на першому кроці поєднуються два найближчих між собою об'єкта, на другому – кластери за мінімальною відстанню між двома ближніми сусідами й т.д. Звідси й назва критерію: потрібний тільки один мінімальний зв'язок, щоб приєднати об'єкт до кластера, оскільки враховується лише одиночний, простий зв'язок з однією точкою кластера.

2. Критерій «далекого сусіда». В англомовній літературі даний критерій відомий як повний зв'язок (Complete linkage). На кожному кроці поєднуються кластери K_p і K_s, відстань між найбільш віддаленими об'єктами p і s яких мінімальна.

3. Критерій «середнього зв'язку» (середньої відстані). На кожному кроці поєднуються кластери K_p і K_s, середня відстань між всіма парами об'єктів яких мінімальна. Даний критерій має дві модифікації залежно від способу розрахунку середніх відстаней між об'єктами кожного кластера:

1) критерій середньої відстані, розрахований за формулою простої середньої арифметичної (Unweighted pair-group average);

2) критерій середньої відстані, розрахований по формулі зваженої середньої арифметичної (Weighted pair-group average). У першому випадку не враховується число об'єктів у кожному кластері, тобто їхня статистична вага, а в другому – враховується.

4. Критерій «середнього сусіда» (центроїда). На кожному кроці поєднуються кластери K_p і K_s, відстань між центрами ваги яких мінімальна.

Даний критерій також має дві модифікації залежно від способу обліку чисельності кожного кластера:

1) критерій центроїда, розрахований без урахування числа об'єктів (статистичної ваги) поєднуваних груп (Unweighted pair-group centroid);

2) критерій центроїда, розрахований з урахуванням числа об'єктів (статистичної ваги) поєднуваних груп (Weighted pair-group centroid).

5. Критерій Уорда (Ward's method). Цей метод агломерації вирізняється від попередніх тим, що він ґрунтується на аналізі збільшень внутрігрупової варіації чинників-симптомів для всіх можливих варіантів об'єднання кластерів. Помічено, що метод Уорда приводить до утворення кластерів приблизно рівних розмірів у формі гіперсфер.

Загальна схема ієрархічного агломеративного алгоритму складається з наступних основних етапів (див. рис. 2.10).

1) всі об'єкти z_i розглядаються як n самостійних кластерів $K_1, K_2, \ldots, K_n$;

2) розраховуються відстані між усіма кластерами і утворюється матриця відстаней $\boldsymbol{D}$, розміру $n \times n$;

3) на базі обраного критерію визначається пара найближчих кластерів, які поєднуються в один новий кластер. Якщо відразу кілька кластерів мають мінімальну відстань між собою, то вибирають будь-яку пару;

4) обчислюються відстані від отриманого нового кластера до всіх інших. Розмірність матриці $\boldsymbol{D}$ при цьому знижується на одиницю;

5) на наступному кроці повторюється виконання пунктів 2, 3, 4 доти, поки не вийде розбивка, що складається з одного кластера – вихідної сукупності об'єктів Очевидно, що доводити ієрархічний агломеративний алгоритм до кінця не має змісту, тому що одержаний результат кластеризації є тривіальним, а завдання багатомірного групування залишається невирішеним. Необхідна об'єктивно обґрунтована зупинка процедури агломерації. Сигналом для такої зупинки може служити різкий ріст на черговому кроці мінімальної відстані між поєднуваними кластерами. Це вказує на те, що в одну групу поєднуються вже більш різнорідні об'єкти, чим на попередніх кроках.

Як видно із загальної блок-схеми наведеного алгоритму (рис. 2.10), для його успішного здійснення необхідно:

1) розрахувати відстані від нового (об'єднаного) кластера до всіх інших;

2) вчасно зупинити процедуру, вибравши оптимальне число компактних груп об'єктів.

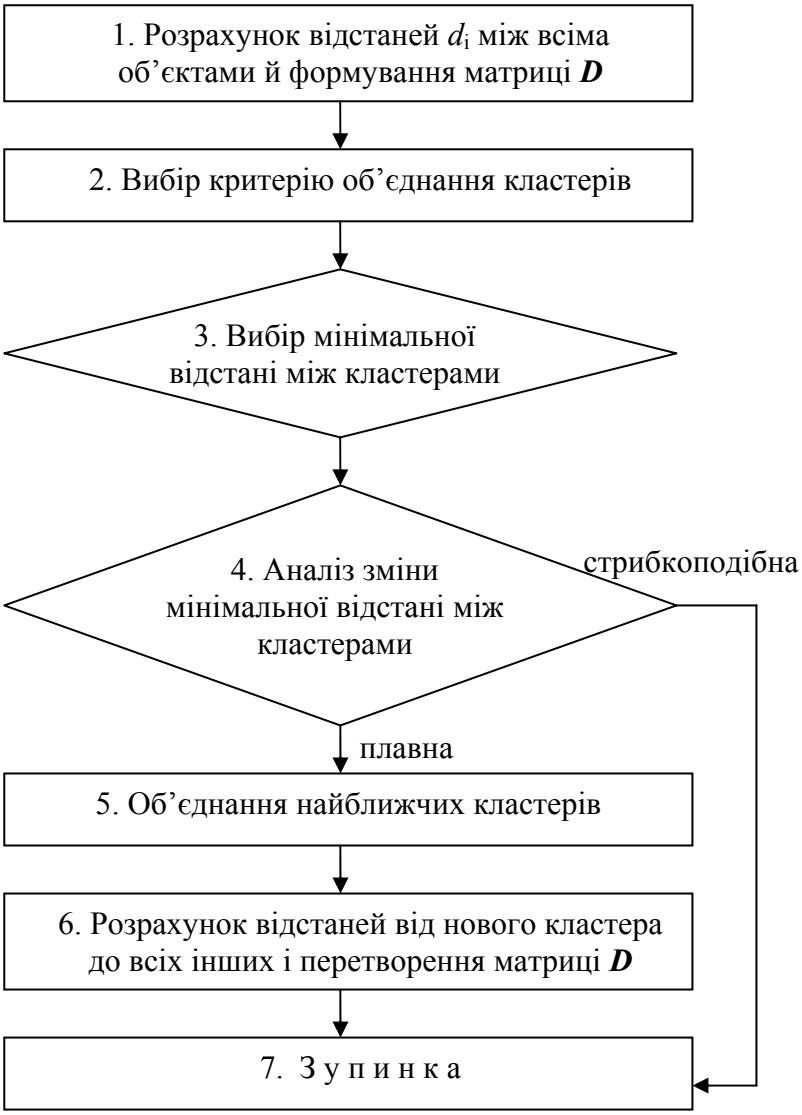


Рис. 2.10. Блок-схема ієрархічного агломеративного алгоритму

З приводу останньої проблеми можна сказати, що вона вирішується в значній мірі суб'єктивно, залежно від досвіду й інтуїції дослідника. На відміну від неї, перше завдання має цілком строге математичне рішення. Існує загальна формула для розрахунку відстані між кластером K_r, що є результатом об'єднання кластерів K_p і K_s, і кластером K_g.

Таким чином визначаються відстані від нового (об'єднаного кластера) до всіх інших. Потім відбувається перехід до третього етапу алгоритму з наступним аналізом зміни мінімальної відстані між кластерами. Зупинка багатовимірної процедури здійснюється в тому випадку, коли зазначена відстань зростає стрибкоподібно, що сигналізує про перспективу об'єднати в один кластер об'єкти, які досить віддалені один від одного.

Поряд з ієрархічними процедурами в загальному руслі методів прямої класифікації існує численна група так званих ітеративних алгоритмів кластерного аналізу. До них відноситься, алгоритм «ФОРЕЛЬ» (див. додаток

Е), метод *k*-середніх, який належить до групи ітеративних методів еталонного типу.

Назва методу і його основні ідеї були запропоновані Дж.В. Мак-Куином у 1967 р. На відміну від ієрархічних процедур, метод *k*-середніх не вимагає обчислення й збереження матриці відстаней або подібності між об'єктами. Алгоритм цього методу побудований на застосуванні не матриць відстаней або схожості, а на роботі з самими об'єктами й допускає використання тільки деяких початкових умов.

Для початку процедури багатовимірної класифікації повинні бути задані *k* довільно обраних об'єктів, які будуть служити центрами ваги майбутніх кластерів. Вважається, що алгоритми типу *k*-середніх зручні й швидкі в реалізації кластерного аналізу. Розглянемо коротко їх сутність.

Спочатку з *n* точок досліджуваної сукупності відбираються випадковим способом або задаються дослідником виходячи з будь-яких апріорних міркувань *k* точок (об'єктів). Ці точки вважаються центрами ваги (центроїдами) майбутніх кластерів на першому кроці процедури.

Кожному центроїду привласнюється порядковий номер, який одночасно є й номером кластера. На першому кроці із (*n* - *k*) об'єктів, що залишилися, вибирається точка z_i (*i* = 1, 2, … , *n*) з координатами (z_{i1}, z_{i2}, … , z_{im}) і перевіряється, до якого із центроїдів вона перебуває ближче всього. Для цього використовується одна з функцій відстаней, наприклад, евклідова відстань d_i.

Об'єкт, що перевіряється, приєднується до того центроїду, якому відповідає значення min(*d*i). Далі центр ваги заміняється новим, перерахованим з урахуванням приєднаної точки, і його статистична вага (кількість об'єктів, що входять у даний кластер) збільшується на одиницю. Якщо зустрічаються дві або більше мінімальних відстані, то *i*-й об'єкт приєднують до центроїду з найменшим порядковим номером.

На наступному кроці вибирається точка z_{i+1} і для неї вся процедура повторюється знову. Таким чином, через (*n* - *k*) кроків усі точки (об'єкти) сукупності стають віднесеними до *k* кластерів і процес багатовимірного угрупування закінчується. Слід зазначити, що процедура кластеризації об'єктів за допомогою методу *k*-середніх, зазвичай, застосовується для того, щоб переконатися у правильності результатів багатовимірного угрупування об'єктів, отриманого за допомогою ієрархічної агломеративної процедури.

Подальшим розвитком ідей і процедур агломерації є алгоритм подвійного об'єднання (Two-way joining) або блокової кластеризації. Його сутність складається в одночасному угрупуванні, як спостережень (об'єктів), так і змінних (чинників-симптомів), на відміну від звичайних ієрархічних агломеративных алгоритмів, у яких об'єднання відбувається роздільно за об'єктами, або за змінними.

Теоретична сторона методу подвійного об'єднання розроблена Дж. А. Хартиганом, який запропонував даний метод блокової кластеризації в 1975 р. Центральне місце в процедурі займає граничне значення параметра подвійного об'єднання, яке визначає, коли алгоритм одночасно врахує близькість в обох

характеристиках досліджуваної сукупності – в об'єктах та в чинниках-симптомах.

Воно задається дослідником і визначає кількість блоків: якщо граничне значення велике, то буде сформований тільки один блок у просторі «об'єкти-змінні». Якщо ж воно мале, то, навпаки, виділяється велика кількість таких блоків. Зазвичай система рекомендує величину граничного значення подвійного об'єднання на рівні половини стандартного відхилення всіх спостережень, тобто приблизно 0,5.

На наш погляд, метод блокової кластеризації може бути корисний як допоміжний інструмент багатовимірного угрупування, коли основна процедура класифікації об'єктів уже виконана (наприклад, на базі ієрархічних алгоритмів, методу *k*-середніх і т.п.) і додатково необхідно з'ясувати, як у кожному з утворених кластерів розподілені значення чинників-симптомів. Алгоритм подвійного об'єднання надає можливість ідентифікувати типи, класи об'єктів за рівнем прихованої ознаки економічних об'єктів на основі аналізу величини тих чинників-симптомів, що складають інформаційну базу дослідження.

Наприклад, якщо перший кластер буде об'єднувати підприємства з низькими значеннями показників дестимуляторів (витрат, собівартості продукції, фінансового ризику) та з високими рівнями факторів-стимуляторів (ринкової частки, рентабельності виробництва тощо), то з великою вірогідністю можна стверджувати про високу конкурентоздатність господарських об'єктів, що потрапили в дану групу. Іншими словами, виділений кластер об'єднує підприємства-лідери за рівнем конкурентоспроможності. І, навпаки, якщо деякий кластер буде об'єднувати об'єкти з високими значеннями чинників дестимуляторів та з низькими значеннями змінних-стимуляторів, то існує висока ймовірність спостереження низької конкурентоспроможності даних підприємств. Тобто даний кластер об'єднує підприємства-аутсайдери.

В ході використання алгоритму подвійного об'єднання слід орієнтуватися на число кластерів, отримане на стадії основної процедури класифікації, підбираючи за допомогою граничного значення параметра відповідне число блоків у просторі «об'єкти-змінні». Наприклад, якщо за допомогою основного алгоритму кластеризації промислових підприємств було утворено 3 кластера, то й при використанні методу подвійного об'єднання варто прагнути одержати 3 блоки.

Дана рекомендація з приводу допоміжного характеру обговорюваної процедури є тим більш раціональною, оскільки алгоритм автоматично не окреслює утворені блоки, залишаючи рішення цієї проблеми на розсуд користувача. Ясно, що, не маючи попереднього уявлення про структуру досліджуваного масиву інформації, важко інтерпретувати отримані результати блокової кластеризації в термінах «об'єкти-змінні».

Другою стадією процедури оцінювання рівня шуканої латентної ознаки об'єктів є їх дискримінантний аналіз, який виконується після здійснення багатовимірного угрупування об'єктів. Дискримінація – це класифікації нових об'єктів за деякою прихованою ознакою, яка представлена множиною відповідних чинників-симптомів, тобто розпізнавання й віднесення нових

об’єктів до вже відомих кластерів. Термін «розпізнавання» об’єктів тлумачиться як оцінювання рівня їхньої латентної ознаки, «новизна» розуміється не в абсолютному, а у відносному змісті – у порівнянні з уже вивченими й описаними класами об’єктів.

Дискримінантний аналіз – це розділ багатовимірних статистичних методів, метою якого є розробка прийомів вирішення завдань розпізнавання (дискримінації) нових об’єктів шляхом порівняння величини їх чинників-симптомів з аналогічними показниками вже досліджених кластерів. Таке порівняння дозволяє класифікувати нові об’єкти, віднести їх у ті або інші групи. Отже, кластерний аналіз, багатовимірне угрупування об’єктів, їх ранжирування, зазвичай, передують дискримінантному аналізу, становлять його основу – навчальну вибірку.

Оцінювання шуканої латентної ознаки за допомогою дискримінації укладається в побудові такої інтегральної моделі за інформацією чинників-симптомів, яка, по-перше, дозволяла б за конкретними значеннями чинників-симптомів розрахувати кількісну оцінку прихованої властивості досліджуваних об’єктів й, по-друге, з досить високою вірогідністю дозволяла б віднести новий об’єкт до того або іншого кластера. Наприклад, до першого кластера, в складі якого знаходяться підприємства-лідери за рівнем своєї конкурентоспроможності, або до другого, що містить об’єкти з середньою конкурентоздатністю, чи до третього, який складається з явних аутсайдерів за даною прихованою властивістю.

Апарат дискримінантного аналізу розроблявся багатьма вченими, починаючи із середини 30-х років XX ст. Питаннями дискримінації, як і іншими багатомірними статистичними методами, займалися Р. Фішер, П. Махаланобіс, Г. Хотеллінг.

Вирішальне правило, за допомогою якого здійснюється багатовимірна класифікація нових об’єктів, будується на основі спеціальних математичних функцій, що називаються дискримінантними. Останні визначаються таким чином, щоб мінімізувати загальні втрати при помилковому розпізнаванні нових об’єктів.

У математико-статистичній формі завдання дискримінантного аналізу формулюється в такий спосіб. Нехай кожний з n об’єктів, як звичайно, описується m чинниками-симптомами і вся інформація представлена у вигляді матриці $\boldsymbol{X}$ вихідних або $\boldsymbol{Z}$ стандартизованих даних. Окрім того, самим дослідником задається додаткова змінна S (S = 1, 2, …, p), яка вказує на приналежність даного об’єкта до певного кластера. Величина S може бути встановлена тільки на основі результатів попередньої стадії – кластерного аналізу (виявлення однорідних по відношенню до шуканої латентної ознаки груп та їх ідентифікації).

Сукупність об’єктів з відомими параметрами x_1, x_2, …, x_m, S використовується для виявлення зв’язку між значеннями цих характеристик і називається навчальною вибіркою, або вчителем. Тоді завдання полягає в побудові за допомогою вчителя дискримінантної функції $f(x)$ і віднесенні на її основі нового об’єкта до одного з p кластерів W_1, W_2, … , W_P.

Очевидно, що проблема дискримінації вирізняється від завдання кластерного аналізу саме наявністю навчальної вибірки – вчителя. Тому її іноді називають завданням розпізнавання образів з учителем, на відміну від завдання кластеризації, що звичайно іменують завданням розпізнавання образів без учителя.

У процесі проведення дискримінантного аналізу виникає, принаймні, дві головні теоретичні проблеми:

1) вибір виду функції $f(x)$;

2) визначення невідомих коефіцієнтів дискримінантної функції $f(x)$.

У математико-статистичній літературі найбільш розробленим і всебічно описаним є випадок використання лінійної дискримінантної функції як простішої й найчастіше застосовуваної в практичних дослідженнях [7, с.220-228; 125, с.159-171]. Тому надалі буде розглядатися саме цей клас дискримінантних функцій.

Що стосується визначення невідомих коефіцієнтів функції $f(x)$, то в даний час широке використання одержав шлях рішення зазначеної проблеми, заснований на використанні апарату множинного кореляційно-регресійного аналізу. При цьому в ролі результативного показника виступає змінна S, а в ролі факторних ознак – чинники-симптоми латентної ознаки $x_1, x_2, \ldots, x_m$.

Дискримінантна функція в цьому випадку має вигляд звичайного рівняння регресії:

$$\hat{S} = a_0 + a_1x_1 + a_2x_2 + \ldots + a_mx_m , \qquad (2.37)$$

де $\hat{S}$ – розрахункове значення змінної, що визначає приналежність об'єкта до певного кластеру;

$a_0, a_1, \ldots, a_m$ – невідомі коефіцієнти регресії.

Знаходження коефіцієнтів регресії $a_0, a_1, \ldots, a_m$, зазвичай, здійснюється на основі методу найменших квадратів за допомогою стандартних програм кореляційно-регресійного аналізу, наприклад, в редакторі Excel.

Після побудови дискримінантної функції у вигляді рівняння регресії (2.37) проводиться його ретельний якісний та статистичний аналіз, який складається з наступних етапів [123; 127, с. 34-45; 129, с.74-88]:

1. Перевірка знаків коефіцієнтів $a_0, a_1, \ldots, a_m$ на відповідність теоретичних уявлень про напрямок зв'язку між чинниками-симптомами й рівнем латентної ознаки досліджуваних економічних об'єктів.

2. Визначення щільності множинної кореляційної залежності між змінною S і відібраними чинниками-симптомами на основі дослідження величини коефіцієнта множинної кореляції R.

3. Аналіз точності побудованої моделі за допомогою коефіцієнта детермінації R^2, стандартної помилки регресії тощо.

4. Перевірка статистичної надійності всього отриманого рівняння регресії (2.37) на базі F-критерію Фішера, та окремих його коефіцієнтів $a_0, a_1, \ldots, a_m$ за допомогою t-статистики Стьюдента. У разі виявлення незначущих коефіцієнтів

регресії відповідні чинники-симптоми шуканої латентної ознаки вилучаються з моделі шляхом покрокового регресійного аналізу.

5. Визначення адекватності побудованого рівняння регресії економічному явищу, що моделюється, за допомогою графічного та статистичного дослідження залишків рівняння (2.37).

Тільки у випадку позитивного результату якісного та статистичного аналізу моделі, або після відповідного корегування рівняння (2.37) може бути використане в якості інструменту оцінювання шуканої латентної ознаки для всіх економічних об'єктів з навчальної вибірки, або дискримінантної функції для розпізнавання нового об'єкта. При цьому багатовимірна класифікація здійснюється за наступним вирішальним правилом: будь-який новий об'єкт $\boldsymbol{x}^*$, відноситься до кластеру W_S, якщо при підстановці його координат x^*_1, x^*_2, …, x^*_m у дискримінантну функцію (2.37) виконується співвідношення

$$\hat{S} \approx S \pm 0{,}5. \qquad (2.38)$$

Наприклад, якщо $1{,}5 < \hat{S} < 2{,}5$, то з досить високою імовірністю, яка дорівнює достовірності отриманої регресійної моделі, можна стверджувати про приналежність нового об'єкту до кластеру W_2, а у випадку $0{,}5 < \hat{S} < 1{,}5$ – до кластеру W_1. При $\hat{S} < 0{,}5$ і $\hat{S} > p + 0{,}5$ краще відмовитися від розпізнавання у зв'язку з тим, що досліджуваний об'єкт розташований у зоні підвищеного ризику помилкової класифікації.

У сучасних обчислювальних програмах дискримінантного аналізу, наприклад, в системі STATISTICA в модулі «Дискримінантний аналіз» (Discriminate Analysis), розпізнавання нових об'єктів здійснюється автоматично за допомогою наступних підходів, що передбачають розрахунок:

- квадратів відстаней Махаланобіса (Squared Mahalanobis distances);
- апостеріорних ймовірностей (Posterior probabilities).

Перший підхід базується на розрахунку квадратів відстаней Махаланобіса від центра ваги кожного кластера x_{0S} до нового об'єкта $\boldsymbol{x}^*$. Новий об'єкт відноситься до того кластера S, який відповідає мінімальній величині квадратів відстаней Махаланобіса. Дійсно, повністю логічним є ототожнювання ще не досліджуваного об'єкта з тими членами кластеру, до яких він ближче всього розташований в просторі чинників-симптомів.

Другий підхід базується на розрахунку апостеріорних ймовірностей, які характеризують частості появи даних об'єктів у певному кластері після проведення процедури багатовимірної класифікації. Вони знаходяться з урахуванням апріорних ймовірностей на основі деяких статистичних параметрів розташування кожного об'єкта в m-мірному просторі чинників-симптомів.

Апостеріорні ймовірності знаходяться за допомогою величин схожості кожного об'єкта із центрами ваги всіх кластерів, розрахованих на основі метрики Махаланобіса. Чим вище вказане розрахункове значення, тим більше є підстав очікувати появи досліджуваного об'єкта в даному конкретному кластері і, навпаки.

Таким чином, розрахувавши за формулою (2.37) для всіх об'єктів навчальної вибірки, а також для «нового» об'єкта величину $\hat{S}_i$ й віднісши його до вже відомих кластерів, наприклад, лідерів, середняків або аутсайдерів, утворених за досліджуваною латентною ознакою, можна отримати її кількісну оцінку.

Слід зазначити, що процедури кластеризації об'єктів на основі ієрархічного агломеративного алгоритму, методу *k*-середніх і подвійного об'єднання реалізовані в пакеті прикладних обчислювальних програм STATISTICA в модулі «Кластерний аналіз» (Cluster Analysis) [11; 12]. Так, для отримання багатовимірної класифікації економічних об'єктів за методом *k*-середніх дослідник повинен задати число майбутніх кластерів *k*, яке, зазвичай, визначають за результатами ієрархічної агломеративної процедури. А алгоритм подвійного об'єднання в модулі «Кластерний аналіз» надає підсумки даної процедури в графічній формі у вигляді мапи, що нагадує географічну, де коричневим та червоним кольорами позначені високі значення чинників-симптомів, жовтим – середні, а зеленим – низькі. Це надає досліднику (з урахуванням розділення чинників-симптомів латентної ознаки на стимулятори і дестимулятори) ідентифікувати виділені кластери економічних об'єктів на лідерів, середняків і аутсайдерів.

Зазначимо, що при застосуванні модуля «Дискримінантний аналіз» (Discriminate Analysis) в системі STATISTICA при розрахунку квадратів відстаней Махаланобіса та апостеріорних ймовірностей здійснюється перевірка правильності результатів багатовимірної класифікації всіх об'єктів, які були отримані на попередній стадії, тобто при кластеризації. Помилково класифіковані об'єкти позначаються зірочкою, що надає додаткову змогу досліднику корегувати результати кластерного аналізу об'єктів, рівень латентної ознаки яких досліджується.

Приклади застосування сполучення кластерного і дискримінантного аналізів при оцінці латентних ознак економічних об'єктів, зокрема, фінансового стану підприємств, наводяться в роботах [147; 148].

2.7 Методи теорії нечіткої логіки

В останні роки при здійсненні оцінки латентних ознак економічних об'єктів широку популярність набули методи комбіновані методи, що поєднують експертні прийоми і математико-статистичний апарат. До них, зокрема, відносяться методи теорії нечіткої логіки [1; 64; 106], основи якої були закладені в 60-і роки XX ст. в роботах американського математика Л.А. Заде.

Дослідження такого роду було викликано зростаючим незадоволенням науковців і практиків експертними системами, зокрема, системами «штучного інтелекту». Для створення дійсно інтелектуальних систем, здатних адекватно взаємодіяти з людиною, був необхідний новий математичний апарат, який переводить неоднозначні життєві твердження в мову чітких і формальних математичних формул. Першим серйозним кроком у цьому напрямку стала теорія нечітких множин, розроблена Заде. Його робота «Fuzzy Sets»,

опублікована в 1965 р., заклала основи моделювання інтелектуальної діяльності людини і стала початковим поштовхом до розвитку нової математичної теорії. Він же дав і назву для нової галузі науки – «fuzzy logic» (fuzzy – від англ. нечіткий, розмитий, м'який) [33; 34].

Апарат теорії нечіткої логіки, продемонструвавши ряд багатообіцяючих можливостей застосування – від систем управління літальними апаратами до прогнозування підсумків виборів, виявився разом з тим складним для втілення. Враховуючи наявний рівень технології, нечітка логіка зайняла своє місце серед інших спеціальних наукових дисциплін – десь посередині між експертними системами і нейронними мережами.

Своє друге народження теорія нечіткої логіки пережила на початку 80-х років XX ст., коли кілька груп дослідників (в основному в США і Японії) всерйоз зайнялися створенням електронних систем різного застосування, використовують нечіткі керуючі алгоритми. Сучасний період розвитку даної теорії розпочався з 90-х років минулого сторіччя і продовжується дотепер. Він характеризується бумом практичного застосування теорії нечіткої логіки в різних сферах науки і техніки. Сьогодні приклади її практичного застосування можна знайти в ході аналізі нових ринків, біржової гри, оцінки політичних рейтингів, виборі оптимальної цінової стратегії тощо.

Розглянемо коротко математичні основи теорії нечіткої логіки стосовно проблеми оцінки латентної ознаки економічних об'єктів [9; 10; 41; 79; 114]. Їх сутність полягає в тому, що в процесі аналізу використовуються певні чинники-симптоми 2-го рівня ієрархії X_j (j = 1, 2, …, m), які визначаються фахівцями в області даної прихованої властивості досліджуваних об'єктів. Для кожного з цих чинників-симптомів уводиться лінгвістична змінна «Рівень чинника-симптому X_j» з терм-множиною (від англ. term – назва) значень P_{kj}, (k = 1, 2, …, 5), кожне з яких є нечіткою множиною і описане природною мовою: P_{1j} – «Дуже низький рівень X_j», P_{2j} – «Низький рівень X_j», P_{3j} – «Середній рівень X_j», P_{4j} – «Високий рівень X_j», P_{5j} – «Дуже високий рівень X_j».

Далі визначається нечітка відповідність поточного значення показника X_j кожному лінгвістичному значенню відповідної терм-множини, тобто задаються функції належності значень показника X_j нечітким множинам. У даному випадку функції належності мають задовольняти наступним умовам:

1. Носії будь-яких трьох лінгвістичних значень показника X_j не перетинаються, а носії двох послідовних лінгвістичних значень X_j перетинаються.

2. Якщо поточне значення X_j належить носію одного лінгвістичного значення, то функція належності дорівнює одиниці; якщо поточне значення показника X_j належить перетину носіїв двох послідовних лінгвістичних значень, то сума функцій належності відповідним нечітким множинам теж дорівнює одиниці.

Таким властивостям задовольняє, наприклад, система трапецевидних функцій $Y = \mu(X)$, які зазвичай використовуються в практичних дослідженнях з застосуванням елементів теорії нечіткої логіки (рис. 2.11).

Побудувавши таким чином нечітку класифікацію показників X_j, переходять до їх ранжирування за рівнями на основі експертних оцінок. В якості рангу фахівці-експерти використовують міру належності значення показника X_j тій чи іншій нечіткій множині, що відповідає значенню лінгвістичної змінної «Рівень чинника-симптому X_j».

Оскільки побудовані функції належності задовольняють наведеним вище двом умовам, то для кожного показника сума рангів за всіма нечіткими множинами дорівнює одиниці. Причому кількість ненульових рангів може бути або один – у випадку коли значення показника гарантовано належить відповідному лінгвістичному значенню змінної, або два – у випадку, якщо не можна гарантовано встановити належність одному значенню лінгвістичної змінної «Рівень чинника-симптому X_j».

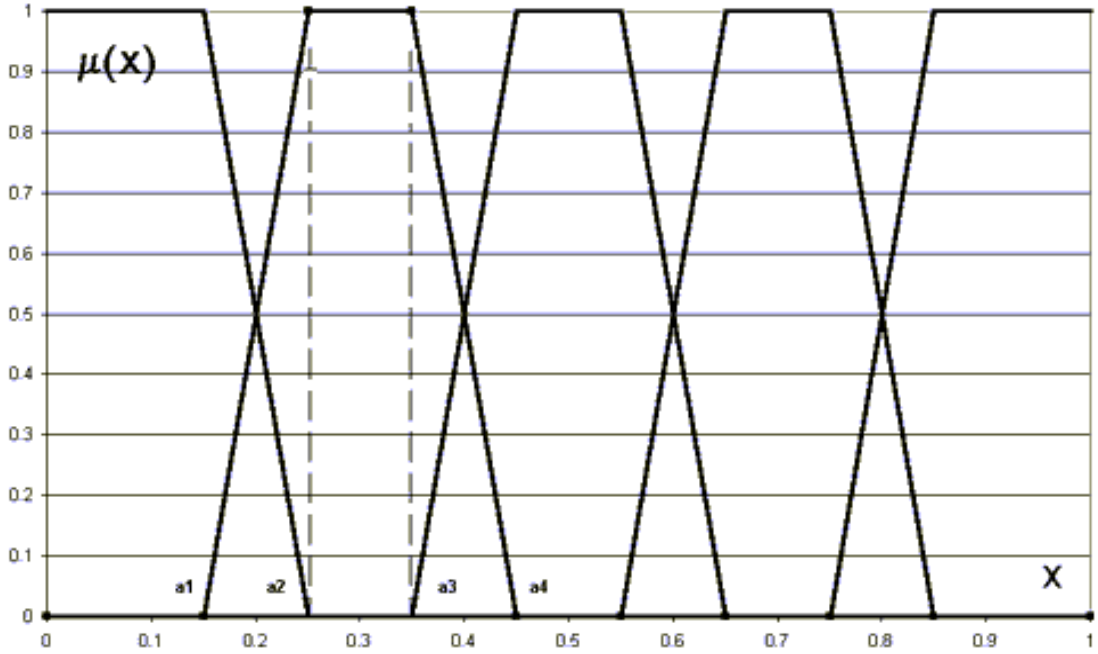


Рис. 2.11. Функції належності у вигляді трапецевидних функцій

Обчислені таким чином значення рангів λ_{ikj} ($i = 1, 2, \ldots, n$) задовольняють наступним умовам:

$$0 \le \lambda_{ikj} \le 1; \sum_{k=1}^{5} \lambda_{ikj} = 1 \qquad (2.39)$$

для будь-яких значень i, k.

Безпосередня побудова інтегральної оцінки досліджуваної прихованої властивості починається з уведення лінгвістичної змінної «Рівень латентної ознаки об’єктів» з терм-множиною значень (множиною лінгвістичних значень) «Дуже високий (ДВ)», «Високий (В)», «Середній (С)», «Низький (Н)», «Дуже низький (ДН)». Визначається носій її терм-множини – дійсна змінна U на інтервалі від нуля до одиниці. Задається система трапецевидних функцій

належності відповідних нечітких множин, областю визначення яких буде інтервал від нуля до одиниці (табл. 2.4).

Наступним кроком визначаються вагомості чинників-симптомів X_j, тобто обчислюються вагові коефіцієнти w_j. Це можна зробити, наприклад, за допомогою експертних методів (див. розділ 2.2 даної книги).

Використавши отримані коефіцієнти w_j, знаходять для кожного об'єкта середньозважені за чинниками-симптомами X_j ранги:

$$\lambda_{ik} = \sum_{j=1}^{m} w_j \lambda_{ikj}, \tag{2.40}$$

де λ_{ik} – середньозважений ранг k-го рівня для n-го об'єкта.

Обчислені таким чином значення рангів λ_{ik}, задовольняють умовам (2.39) для будь-яких значень i. Їх також представляють в таблицях для кожного чинника-симптому X_j.

Таблиця 2.4

Значення функції належності для лінгвістичної змінної «Рівень латентної ознаки об'єктів»

Значення U	Дуже низький	Низький	Середній	Високий	Дуже високий
0-0,15	1	0	0	0	0
0,15-0,25	(0,25-U)/0,1	(U-0,15)/0,1	0	0	0
0,25-0,35	0	1	0	0	0
0,35-0,45	0	(0,45-U)/0,1	(U-0,35)/0,1	0	0
0,45-0,55	0	0	1	0	0
0,55-0,65	0	0	(0,65-U)/0,1	(U-0,55)/0,1	0
0,65-0,75	0	0	0	1	0
0,75-0,85	0	0	0	(0,85-U)/0,1	(U-0,75)/0,1
0,85-1	0	0	0	0	1

Інтегральна оцінка латентної ознаки U розраховується за наступною формулою:

$$U = \sum_{k=1}^{5} \alpha_k \lambda_{ik}, \tag{2.41}$$

де α_k – абсциси максимумів функцій належності терм-множин лінгвістичної змінної.

Для трапецевидних функцій належності значення α_k знаходяться за формулою:

$$\alpha_k = 0{,}2 \times k - 0{,}1 \tag{2.42}$$

за умови, що відповідні трапеції є рівнобічними. Далі згідно даним табл. 2.4 встановлюється рівень шуканої латентної ознаки кожного *i*-го об'єкта із сукупності всіх об'єктів.

На нашу думку, обговорюваний підхід по суті представляє собою подальший розвиток і формалізацію ідей методу матричних моделей як одного із різновидів експертних рейтингових оцінок. Його удосконалення полягає в тому, що оцінки, надані експертами, для всіх латентних ознак (як факторних, так і результативної) строго масштабовані і належать інтервалу [0; 1]. Причому деякі з них можуть перетинатися, утворюючи нечітку множину. Отже, головна новизна побудованих за допомогою елементів теорії нечіткої логіки інтегральних показників, що розглядаються в якості оцінок досліджуваних латентних ознак, зводиться до формалізації думок фахівців-експертів у процесі рейтингового оцінювання на базі відомих матричних моделей з усіма негативними наслідками їхнього використання (певним суб'єктивізмом, проблемами створення компетентних груп експертів тощо).

Г.І. Великоіваненко, К.М. Мамонова з цього приводу відмічають, що значним недоліком, який проявляється кожен раз при застосуванні елементів теорії нечіткої логіки, є те, що під час проведення обчислень з використанням нечітко-множинного підходу результати розрахунків в значній мірі залежать від правильного поділу нечіткої множини на рівні, що часом досить важко зробити за обмеженої і досить неоднорідної статистики [15].

При вирішенні конкретних практичних задач оцінювання латентної ознаки економічних об'єктів з використанням елементів теорії нечіткої логіки рекомендується застосовувати математичне забезпечення, з яким можна ознайомитись в роботах [20; 50; 55]

2.8 Багатовимірне шкалювання

2.8.1 Сутність і основні поняття багатовимірного шкалювання

Методи багатовимірного шкалювання застосовуються на практиці для дослідження економічних об'єктів з метою поглиблення уявлень про їх взаємне розташування в деякій штучно побудованої системі координат (шкал). В основі теорії багатовимірного шкалювання лежить ідея про можливість розгортання об'єктів у гіпотетичному просторі латентних ознак, який адекватно відображає реальну економічну дійсність.

На відміну від інших багатовимірних статистичних методів пошук координатного простору при багатовимірному шкалюванні здійснюється не за спостережуваними чинниками-симптомами, які вимірюються в метричній шкалі і характеризують досліджувані об'єкти, а за даними, що описують відмінності, чи, навпаки, схожості цих об'єктів. Іншими словами, вихідними матрицями для застосування методів шкалювання є матриці відстаней або схожості між досліджуваними об'єктами (див. табл. 2.3, формулу (2.11)).

При цьому основне джерело інформації – це результати спостереження, реєстрації, статистичні дані про стан і поведінку економічних об'єктів у

просторі і часі, а також дані експертних опитувань, що відображають суб'єктивне сприйняття і оцінку фахівців відносного розташування об'єктів за тими чи іншими чинниками-симптомами. Два джерела інформації для подальшого аналізу визначають два основних напрями в багатовимірному шкалюванні – метричне і неметричне шкалювання.

Другий (неметричний) напрямок у багатовимірному шкалюванні є в даний час домінуючим. Його мета полягає у визначенні місцезнаходження економічного об'єкта в просторі «сприйняття суб'єктів» і створення його образу. Передбачається, що безпосередньо про сам об'єкт навіть за значеннями повного набору чинників-симптомів не можна судити цілком надійно. У той же час експерти або просто спостерігачі ще до проведення багатовимірного аналізу бачать, інтуїтивно відчувають відмінності між досліджуваними об'єктами. Неусвідомлені, нечіткі уявлення про об'єкти повинні бути конкретизовані в гіпотетичному просторі «сприйняття суб'єктів», побудованому за даними експертного опитування. Цей гіпотетичний простір має латентні осі і дозволяє виміряти і графічно візуалізувати відстані між об'єктами дослідження. Тому розмірність шуканого штучного простору латентних ознак найчастіше не перевищує трьох. Остання вимога необов'язкова, але бажана, так як до неї в багатовимірному шкалюванні вдаються досить часто саме для візуалізації досліджуваних економічних об'єктів.

Думка про можливості шкалювання спостережуваних явищ і процесів реальної дійсності з'явилася в науці досить давно. Ще І. Ньютон у 1704 р. показав, що кольори спектра можна умовно представити на колі, а їх відмінності – відстанями між точками цього кола. В XIX – початку XX ст. були зроблені спроби наочного відображення відмінностей звукових сигналів і відчуттів (винахід світломузики в 1916 р.).

Перші теоретичні обґрунтування методів і алгоритмів багатовимірного шкалювання представлені в 20–30-ті роки минулого сторіччя в роботах М. Річардсона, Л. Терстоуна, які були присвячені проблемам застосування факторного аналізу в психології людини. У 50–60-ті роки XX ст. розробка теорії багатовимірного шкалювання велася особливо активно. До неї живий інтерес був виявлений з боку представників різних наук: математики, соціології, політології, економіки та ін. Приблизно в цей же час сформувалися два основних напрямки – метричне і неметричне шкалювання.

Метричне багатовимірне шкалювання розвивалося в дослідженнях американського статистика У. Торгерсона і базувалося на використанні даних про спостережувані чинники-симптоми, що характеризують досліджувані об'єкти, і вимірюються в метричній шкалі.

Неметричне багатомірне шкалювання призначалося для обробки інформації, отриманої в результаті експертного опитування, і виміряної в порядковій шкалі. Заслуга в розробці методів цього напрямку належить Р. Шепарду, Дж. Краскалу, Л. Гуттману, Ф. Янгу, Дж. Керролу, Дж. Лінгосу, Л. Такеру та ін. Це в основному представники американської школи багатовимірного статистичного аналізу [157; 158; 161].

У 70-х роках минулого століття з'являються перші наукові праці з багатомірного шкалювання російськомовних авторів: А.Ю. Терехіної, З.А. Клігера, В.С. Кам'янського, Р.А. Сатарова [91, с.401-462].

У даний час методи багатовимірного шкалювання продовжують бурхливо розвиватися в рамках багатовимірного статистичного аналізу, прикладної математики, теорії аналізу даних. Вони вважаються найбільш гнучкими й адаптованими для адекватного відображення реальних явищ, процесів і об'єктів природного і суспільного життя. Однак, застосування багатовимірного шкалювання пред'являє найвищі вимоги до професійної підготовленості, досвіду та інтуїції дослідника, оскільки від його знань і творчих здібностей багато в чому залежить правильний вибір обчислювальних процедур методу, коректність інтерпретації отриманих результатів.

Традиційно в багатовимірному шкалюванні використовуються поняття і терміни, які рідко зустрічаються в класичному багатовимірному статистичному аналізі. Сюди відносяться такі основні поняття.

Стимул – властивість об'єкта, його ознака, латентна характеристика. Можна сказати, що стимул має саме широке тлумачення, якому найкраще відповідає термін «образ об'єкта». В практиці багатовимірного шкалювання під стимулом зазвичай розуміють об'єкт, що володіє певним набором властивих йому характеристик (як безпосередньо спостережуваних, тобто метричних, так і латентних – виміряних у порядковій шкалі).

Шкала – вісь гіпотетичного простору, яка є носієм узагальненої латентної ознаки. Шкали визначаються на завершальному етапі дослідження найчастіше за допомогою методів факторного аналізу. Як відомо, при інтерпретації виділених загальних факторів можуть виникнути різного роду труднощі, тому при тлумаченні отриманих шкал зазвичай стикаються з проблемами пошуку правильного їх найменування (ідентифікації), визначення оптимального числа, інтерпретації розташування об'єктів у гіпотетичному шкальному просторі. Успішність вирішення зазначених проблем повністю залежить від компетенції дослідника.

Експерт – поняття, що не потребує докладного роз'яснення. Це суб'єкт, який, зазвичай, з високим професіоналізмом надає судження про об'єкти і стимули. У практиці багатовимірного шкалювання зустрічається і менш суворий підхід, коли в якості експерта розглядається простий спостерігач, що реєструє поведінку об'єкта (стимулу).

Перевага – судження про стимул з точки зору його близькості до ідеалу, який уявляється експертом. Величина переваг може з часом змінюватися.

Профіль – набір спостережуваних кількісних чинників-симптомів об'єктів або ряд динаміки цих чинників-симптомів. Профілі відображають просторову і часову варіацію показників досліджуваних об'єктів.

Відмінності – відстані між об'єктами у профільному просторі або в будь-якому іншому просторі порядкових шкал.

Стрес-формули – критерії, що застосовуються для перевірки відповідності емпіричних і теоретичних рангових даних.

Так само, як і в факторному аналізі, при використанні методів багатовимірного шкалювання дослідник стикається з необхідністю оцінки та інтерпретації знайдених осей (загальних факторів), стиснення початкового простору чинників-симптомів, візуалізації геометричної конфігурації спостережуваних стимулів у координатному просторі виділених шкал.

У багатовимірному шкалюванні вихідна інформація про досліджувані об'єкти у відповідності основним напрямкам шкалювання (метричному і неметричному) подається в двох головних формах. У першому випадку, коли набори спостережуваних кількісних чинників-симптомів об'єктів (або їх ряди динаміки), тобто профілі відображають просторово-часову варіацію економічних та інших показників, то первинна інформація представляється у вигляді матриць відстаней між досліджуваними об'єктами.

У другому випадку, коли в результаті опитування експертів отримані рангові або порядкові дані типу «більше-менше», «краще-гірше», то інформація перетворюється до матриць ймовірностей. При цьому в обох ситуаціях висуваються дві головні вимоги до матриць, що отримуються на завершальному етапі обробки даних:

1) симетричність;

2) нульова сума елементів кожного рядка (стовпця).

Така матриця називається симетричною матрицею з подвійним центруванням. Якщо з якихось причин зазначена вимога не виконується, то слід здійснити додаткове перетворення вихідних даних, яке б з мінімальними втратами інформації про реальне розташування об'єктів забезпечило симетричність матриці і одночасне центрування її рядків і стовпців, по головній діагоналі якої розташовані нулі. Розрахунок матриць відстаней (схожості) за відомими профілями не представляє особливих труднощів. При цьому, як відомо, матриця відстаней $\boldsymbol{D}$ завжди є симетричною матрицею розміру $n \times n$:

$$\boldsymbol{D} = \begin{pmatrix} 0 & d_{12} & d_{13} & \dots & d_{1n} \\ d_{21} & 0 & d_{23} & \dots & d_{2n} \\ & & & \dots & \\ d_{n1} & d_{n2} & d_{n3} & \dots & 0 \end{pmatrix} \qquad (2.43)$$

Тому тут ми зупинимося лише на проблемі трансформації матриці $\boldsymbol{D}$ в матрицю з подвійним центруванням $\boldsymbol{D}^*$. Проведемо наступне перетворення всіх елементів матриці $\boldsymbol{D}$ за формулою:

$$d_{ij}^* = d_{ij} - \bar{d}_i - \bar{d}_j + \bar{d}\,, \qquad (2.44)$$

де середні $\bar{d}_i$, $\bar{d}_j$, $\bar{d}$ розраховуються за даними матриці відстаней (2.43) так:

$\bar{d}_i = (\sum_{j=1}^{n} d_{ij})/n$ – середнє значення за i-м рядком;

$\bar{d}_j = (\sum_{i=1}^{n} d_{ij})/n$ – середнє значення за j-м стовпцем;

$\bar{d} = (\sum_{j=1}^{n} \sum_{i=1}^{n} d_{ij})/n^2$ – загальне середнє.

Легко показати, що для матриці ***D****, елементи якої знаходяться за формулою (2.44), справедливі співвідношення:

$$\sum_{i=1}^{n} d_{ij}^{*} = \sum_{j=1}^{n} d_{ij}^{*} = 0 \qquad (2.45)$$

і вона є симетричною матрицею. Отже, перетворення (2.44) дозволяють трансформувати симетричну матрицю ***D*** в симетричну матрицю ***D**** з подвійним центруванням. Це означає, що деякі елементи $d_{ij}^{*} < 0$, а деякі $d_{ij}^{*} \geq 0$, тобто діапазон зміни значень d_{ij}^{*} знаходиться в межах від $-\infty$ до $+\infty$.

Розглянемо наступний тривіальний випадок, коли $\bar{d}_i \approx \bar{d}_j \approx \bar{d}$. Тоді матриця ***D**** приймає вигляд:

$$\boldsymbol{D}^{*} = \begin{pmatrix} -\bar{d} & d_{12} - \bar{d} & \dots & d_{1n} - \bar{d} \\ d_{21} - \bar{d} & -\bar{d} & \dots & d_{2n} - \bar{d} \\ & & \dots & \\ d_{n1} - \bar{d} & d_{n2} - \bar{d} & \dots & -\bar{d} \end{pmatrix} \qquad (2.46)$$

У цій ситуації всі елементи вихідної матриці відстаней D зменшуються на константу $\bar{d}$. В результаті такого перетворення значення відстаней за всіма рядками (стовпцями) зсуваються вліво по числовій осі на величину $\bar{d}$.

При цьому матриця ***-D**** в значній мірі нагадує тепер матрицю дисперсій-коваріацій вихідних профілів. Остання, як відомо, є проміжним результатом при побудові матриці коефіцієнтів парної кореляції ***r***, що відображає взаємозв'язок профілів.

Матриця ***r***, в свою чергу, може розглядатися як деяка умовна матриця схожості (див. [125, с.35-36, 49-50]). Тому можна зробити наступний висновок: перетворення всіх елементів матриці ***D*** за формулою (2.44) і множення отриманої матриці ***D**** на -1 призводить матрицю відстаней до деякої симетричної матриці схожості досліджуваних профілів з подвійним центруванням.

У разі неметричного шкалювання робота щодо організації статистичного спостереження і зведення даних, одержаних на основі опитування експертів, проводиться особливим чином. Відомості, що надходять про об'єкти або стимули, узагальнюються в квадратних або прямокутних матрицях з назвами, відповідними методиці отримання інформації або її характером. Матриця *умовних ймовірностей* (матриця ідентифікацій) містить відносні дані «по пізнанню стимулів». Типовий вигляд матриці умовних ймовірностей показаний на рис. 2.12.

Її рядки представляють собою перелік стимулів (A, B, C, D), запропонованих для оцінки, стовпці – перелік стимулів, розпізнаних експертами. Елементи матриці умовних ймовірностей дозволяють, з одного

боку, виявити помилки в розпізнаванні стимулів, а з іншого, ідентифікувати їх, якщо самі стимули нечіткі, і необхідно внести визначеність у їх класифікацію. В реальних економічних дослідженнях, зазвичай, $P_{ij} \neq P_{ji}$ і матриця ідентифікацій, взагалі кажучи, не є симетричною. При цьому виконуються такі співвідношення $\Sigma P_{ij} = 1$, тобто підсумки по рядках являють собою загальні кількості експертних оцінок.

	A	B	C	D
A	P_{11}	P_{12}	P_{13}	P_{14}
B	P_{21}	P_{22}	P_{23}	P_{24}
C	P_{31}	P_{32}	P_{33}	P_{34}
D	P_{41}	P_{42}	P_{43}	P_{44}

Рис. 2.12. Матриця умовних ймовірностей

У числі матриць умовних ймовірностей можна виділити матриці *переходу* (трансформації). Таку назву вони отримали внаслідок того, що їх елементи відображають трансформаційні процеси, що протікають у часі або в просторі з точки зору різних експертів. Рядки матриці переходу дають уявлення про досліджуване явище в момент початку спостереження даного експерта, стовпці – в момент закінчення спостереження. Іншими словами, експертами здійснюється вивчення варіації певного показника в часі. Або те ж, але при переході від одного об'єкта до іншого, тобто аналіз варіації показника в просторі.

У всіх випадках несиметричні матриці умовних ймовірностей слід перетворити у симетричні. З цією метою може бути використана, наприклад, наступна формула трансформації:

$$P_{ij}^* = (P_{ij} + P_{ji})/2. \qquad (2.47)$$

Знайдені таким чином середні значення підставляються замість елементів P_{ij} і P_{ji} матриці ідентифікацій, а по головній її діагоналі проставляються нулі (рис. 2.13).

	A	B	C	D
A	0	P_{12}^*	P_{13}^*	P_{14}^*
B	P_{12}^*	0	P_{23}^*	P_{24}^*
C	P_{13}^*	P_{23}^*	0	P_{34}^*
D	P_{14}^*	P_{24}^*	P_{34}^*	0

Рис. 2.13. Трансформована матриця умовних ймовірностей

У результаті зазначених перетворень матриця умовних ймовірностей приводиться до звичайної форми матриці відстаней (2.43). При цьому необхідно мати на увазі, що трансформація матриці ідентифікацій за формулою (2.47) не є єдиною. Існує безліч інших методів приведення матриці до симетричного вигляду, і будь-які перетворення неминуче пов'язані з втратою частини інформації про досліджуване явище чи її деяким викривленням внаслідок відступу від реальних асиметричних співвідношень.

Інший вид матриць для узагальнення вихідних даних, широко застосовуваних у неметричному шкалюванні, – це матриця спільних ймовірностей. У ній елементи P_{ij} характеризують питому вагу взаємодій стимулів *i* та *j*. Рядки і стовпці матриці спільних ймовірностей містять узгоджені значення, а сама матриця є симетричною. Множина прийомів, що приводять до побудови матриці спільних ймовірностей, дуже велика і відповідає розмаїттю дослідницьких цілей та індивідуальних професійних підходів до вирішення конкретних завдань оцінки латентних ознак економічних об'єктів. Серед них найбільш поширені такі:

- метод рандомізації, сутність якого полягає в послідовному поділі експертами сукупності *n* об'єктів на групи; при першому поділі утворюються дві групи – пари схожих і пари несхожих об'єктів; потім у кожній групі відповідно знаходяться пари з найбільш і найменш схожими об'єктами і т. д.; на заключному етапі отримають $n(n - 1)/2$ пар об'єктів (по числу елементів над головною діагоналлю симетричної матриці ***D***), ранжированих за ступенем схожості або несхожості;
- метод вихідної точки, коли із загального числа об'єктів на першому кроці відбирають один і його положення в сукупності приймається за якірне щодо інших; ступінь схожості (відмінності) всіх інших об'єктів з якірним оцінюється експертами з присвоєнням рангу; на наступному кроці якірним стає інший об'єкт з подальшим присвоєнням рангів усім іншим об'єктам і т.д.; у результаті виходить $n(n - 1)$ рангових оцінок парних схожості (відмінностей), з яких легко перейти до матриці відстаней ***D***;
- метод рейтингових оцінок, коли експертам пропонується шкала з деяким числом поділок (зазвичай, 7-9), що дозволяють ранжирувати кожну пару об'єктів за ступенем їх схожості (відмінності), наприклад, як це показано на рис. 2.14.

Знайдені у такий спосіб ранги усереднюють і в результаті утворюється симетрична матриця відстаней ***D***, що включає $n(n - 1)/2$ рейтингових оцінок. Дані, отримані будь-яким з трьох викладених вище методів, можна залишити в первісному вигляді або перетворити їх в кількісні, тобто оцифрувати.

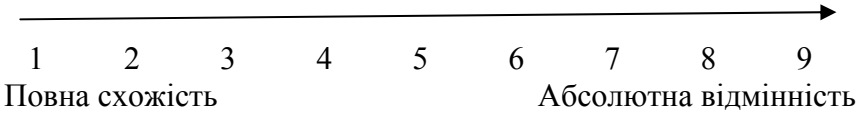


Рис. 2.14. Шкала оцінки експертами схожості (відмінності) об'єктів

Існує досить простий і надійний спосіб переходу до кількісних характеристик відмінності. Спочатку будується матриця схожості ***r***, яка являє собою звичайну матрицю коефіцієнтів парної кореляції між усіма об'єктами, обчислених за ранговим даними. Потім відстані (відмінності) між усіма об'єктами знаходиться так:

$$d_{ij} = \sqrt{1 - r_{ij}^2}. \tag{2.48}$$

2.8.2 Класична модель метричного шкалювання

У. Торгерсон на початку 50-х років XX ст. запропонував один з перших алгоритмів багатовимірного шкалювання, який в даний час відомий як класичний метод метричного шкалювання [160]. Він являє собою поєднання ідей багатовимірного представлення об'єктів у просторі чинників-симптомів і підходів факторного аналізу.

Теоретична концепція цього методу базується на таких постулатах:

1) в деякому шкальному просторі міри схожості δ_{ij} між досліджуваними об'єктами відповідають величинам μ_{ij}, що характеризує їх схожість в просторі профілів, тобто $\delta_{ij} = \mu_{ij}$;

2) у профільному і шкальному просторах середні значення координат стимулів по кожній осі дорівнюють нулю, що еквівалентно виконанню вимоги:

$$\sum_{k=1}^{m} z_{pk} = \sum_{k=1}^{m} x_{pk} = 0; \tag{2.49}$$

3) відстані між об'єктами в просторі профілів достатньо точно вимірюються квадратом евклідової метрики (див. табл. 2.3);

4) подвійне центрування матриці відстаней ***D*** відбувається за формулою

$$d_{ij}^* = -(d_{ij} - \bar{d}_i - \bar{d}_j + \bar{d})/2, \tag{2.50}$$

тобто по суті поряд з центруванням здійснюється також перехід від матриці відстаней до матриці схожості $-\boldsymbol{D}^*/2$ шляхом нормування (ділення центрованих значень на -2);

5) елементи симетричної матриці схожості з подвійним центруванням, знайдені за формулою (2.50), співпадають з елементами матриці дисперсій-коваріацій: $d_{ij}^* = \sum_{k=1}^{m} x_{pk} x_{sk}$, або в матричному вигляді

$$-\boldsymbol{D}^*/2 = \boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}. \tag{2.51}$$

Останнє припущення випливає з першого, якщо покласти $\delta_{ij} = \sum_{k=1}^{m} x_{pk} x_{sk}$ і $\mu_{ij} = d_{ij}^*$.

В класичному методі метричного шкалювання мінімізується функція:

$$f = \sum_{i,j} (\delta_{ij} - \mu_{ij})^2, \qquad (2.52)$$

яка характеризує міру відповідності близькості об'єктів в шкальному і профільному просторах. Вона виконує роль, аналогічну вимозі методу найменших квадратів в кореляційно-регресійному аналізі. Завдяки критерію (2.52) значення x_{pk}, x_{sk} знаходяться таким чином, щоб залишки (d_{ij}* - $\sum_{k=1}^{m} x_{pk}x_{sk}$) були б мінімальними і схожість об'єктів у шкальному просторі повністю відповідала схожості стимулів у профільному просторі.

При цьому сам алгоритм метричного шкалювання складається з наступних етапів (рис. 2.15).

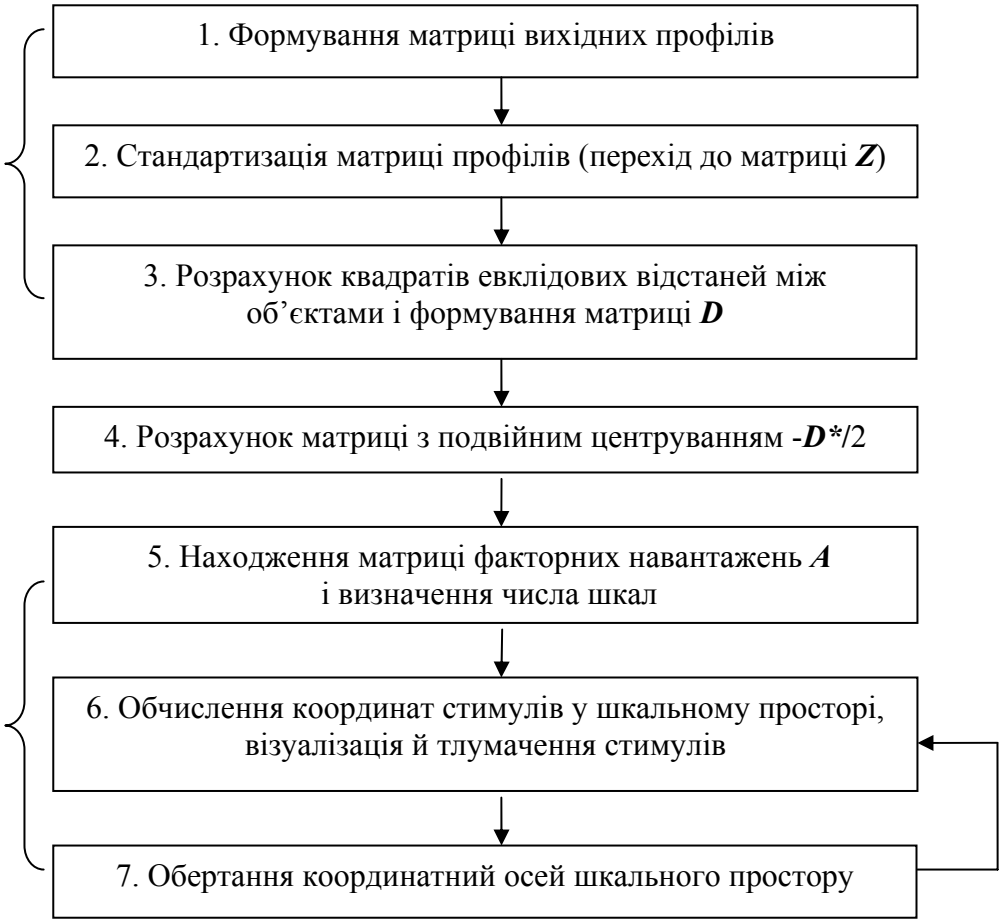


Рис. 2.15. Блок-схема алгоритму метричного шкалювання

Перші три етапи наведеної блок-схеми (позначені фігурними дужками) повністю збігаються з початковими етапами багатовимірної класифікації об'єктів, які обговорювалися в розділі 2.6 даної книги. Отже, їх можна об'єднати під умовною назвою «Етапи вимірювання відмінностей профілів» і виконувати за допомогою алгоритмів кластерного аналізу, наприклад, на базі ієрархічної агломеративної процедури. Основна їх мета – побудова матриці ***D*** в

просторі профілів за допомогою метрики квадратів евклідових відстаней з табл. 2.3.

На четвертому етапі методу реалізується подвійне центрування матриці відстаней і перехід до матриці схожості $-\boldsymbol{D}^*/2$. Він виконується на етапі подання та первинної обробки даних для метричного шкалювання на основі перетворення (2.51).

П'ятий етап процедури полягає у визначенні невідомої матриці $\boldsymbol{X}$. Він будується з урахуванням наступного математичного положення: будь-яка дійсна симетрична матриця (зокрема кореляційна матриця $-\boldsymbol{D}^*/2$) допускає так звану спектральну декомпозицію

$$-\boldsymbol{D}^*/2 = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}, \tag{2.53}$$

де $\boldsymbol{\Lambda}$ – діагональна матриця, розміру $m \times m$, характеристичних коренів λ_{L} матриці $-\boldsymbol{D}^*/2$;

$\boldsymbol{V}$ – матриця, розміру $m \times m$, нормалізованих характеристичних векторів матриці $-\boldsymbol{D}^*/2$, які відповідають характеристичним кореням λ_{L}.

Матриця $\boldsymbol{V}$ є ортогональною і для неї виконуються співвідношення $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} = \mathbf{I}$ ($\mathbf{I}$ – одинична матриця). З урахуванням вимоги (2.52), формула (2.53) перетворюється до наступного вигляду:

$$\boldsymbol{A}\boldsymbol{A}^{\mathbf{T}} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}. \tag{2.54}$$

Звідси матрицю $\boldsymbol{A}$ можна розглядати як матрицю факторних навантажень при побудові спеціальної моделі факторного аналізу. Вона знаходиться так:

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}^{1/2}\ . \tag{2.55}$$

Із співвідношення (2.55) вектори-стовбці матриці $\boldsymbol{A}$ визначаються наступним чином:

$$\boldsymbol{a}_{\mathbf{L}} = \boldsymbol{v}_{\mathbf{L}}(\lambda_{\mathrm{L}})^{1/2}, \tag{2.56}$$

де λ_{L} – L-й характеристичний корінь матриці $-\boldsymbol{D}^*/2$;

$\boldsymbol{v}_{\mathbf{L}}$ – L-й нормалізований характеристичний вектор матриці $-\boldsymbol{D}^*/2$, який відповідає характеристичному кореню λ_{L}.

Таким чином, п'ятий етап процедури метричного шкалювання, як видно з наведених співвідношень, тотожний задачі пошуку характеристичних коренів і характеристичних векторів матриці $-\boldsymbol{D}^*/2$. Вирішення проблеми розмірності шкального простору (визначення числа шкал, що виділяються) здійснюється аналогічно завданню визначення кількості шуканих загальних факторів у факторному аналізі з тією лише різницею, що в процесі багатовимірного шкалювання, зазвичай, передбачається візуалізація стимулів (шостий етап блок-схеми на рис. 2.15). Тому число шкал найчастіше обмежується 2-3.

Якщо на шостому етапі методу тлумачення результатів метричного шкалювання (яке повністю подібне тлумаченню загальних факторів) викликає труднощі, то вдаються до обертання осей шкального простору (сьомий етап блок-схеми на рис. 2.15) за відомими прийомами: квартимакс, варімакс – при ортогональному факторному рішенні; квартимін, облімакс – при косокутному рішенні (див. [126, с.105-128]). Отже, етапи 5-7 процедури метричного шкалювання також можна об'єднати (позначені фігурними дужками на рис. 2.15) під умовною назвою «Етапи факторного аналізу».

Таким чином, підтверджується висновок про те, що метричне шкалювання являє собою синтез ідей і підходів багатовимірної класифікації об'єктів у просторі чинників-симптомів і принципів оцінки латентних стимулів у шкальному просторі за допомогою факторного аналізу. По суті лише четвертий етап «Розрахунок матриці з подвійним центруванням **$-D^*/2$**» безпосередньо відображає специфіку багатовимірного шкалювання і є в цьому сенсі новим, особливим. Інші етапи (три перших і три останніх) покликані реалізувати відомі методи кластерного і факторного аналізу.

2.8.3 Неметричне шкалювання

Методи неметричного шкалювання застосовуються для обробки інформації, отриманої в результаті експертного опитування, і яка вимірюється в порядковій (ранговій) шкалі. Вони розроблялися в 60-80-ті роки XX ст. американськими вченими Р. Шепардом, Дж. Краскалом, Ф. Янгом, Л. Гуттманном, Дж. Лінгосом, Л. Такером та ін.

Головною умовою, що забезпечує адекватність побудованих статистичних моделей, тут стає відповідність монотонних співвідношень емпіричних і теоретичних даних. Якщо в реальній дійсності існує порядкова залежність відстаней між об'єктами типу $d_{ij} < d_{sp}$, то в шкальному просторі відповідно повинно виконуватися співвідношення між стимулами $\delta_{ij} < \delta_{sp}$ (тут δ_{ij} – міра відстані між відповідними стимулами).

При цьому вид монотонності (у формі монотонної функції f)

$$\delta_{ij} = f(d_{ij}) \qquad (2.57)$$

апріорі невідомий. Він підбирається шляхом послідовного перебору різних математичних функцій (лінійної, степеневої, логарифмічної та ін.). У соціально-економічних дослідженнях найчастіше використовується простіша лінійна функція $\delta_{ij} = a_0 + a_1 d_{ij}$. В якості міри відстані між об'єктами в просторі емпіричних (порядкових) даних, зазвичай, приймається універсальна метрика Мінковського (див. табл. 2.3):

$$d_{ij} = [\sum_{k=1}^{m} \left| x_{ik} - x_{jk} \right|^N]^{1/N}. \qquad (2.58)$$

Якщо, наприклад, f лінійна функція, то вираз (2.57) приймає вигляд:

$$\delta_{ij} = f(d_{ij}) = a_0 + a_1[\sum_{k=1}^{m} |x_{ik} - x_{jk})|^N]^{1/N} . \quad (2.59)$$

Універсальна лінійна модель неметричного шкалювання (2.59), побудована на основі метрики Мінковського, дозволяє легко перейти до інших лінійних моделей. Так, при $N = 1$ виходить модель неметричного шкалювання, створена на базі лінійного відстані, при $N = 2$ – на основі евклідової відстані, при $N = \infty$ – на базі супремум-норми і т. ін.

Незалежно від конкретного вибору показника ступеня N процес побудови лінійної моделі неметричного шкалювання складається з наступних основних етапів (рис. 2.16).

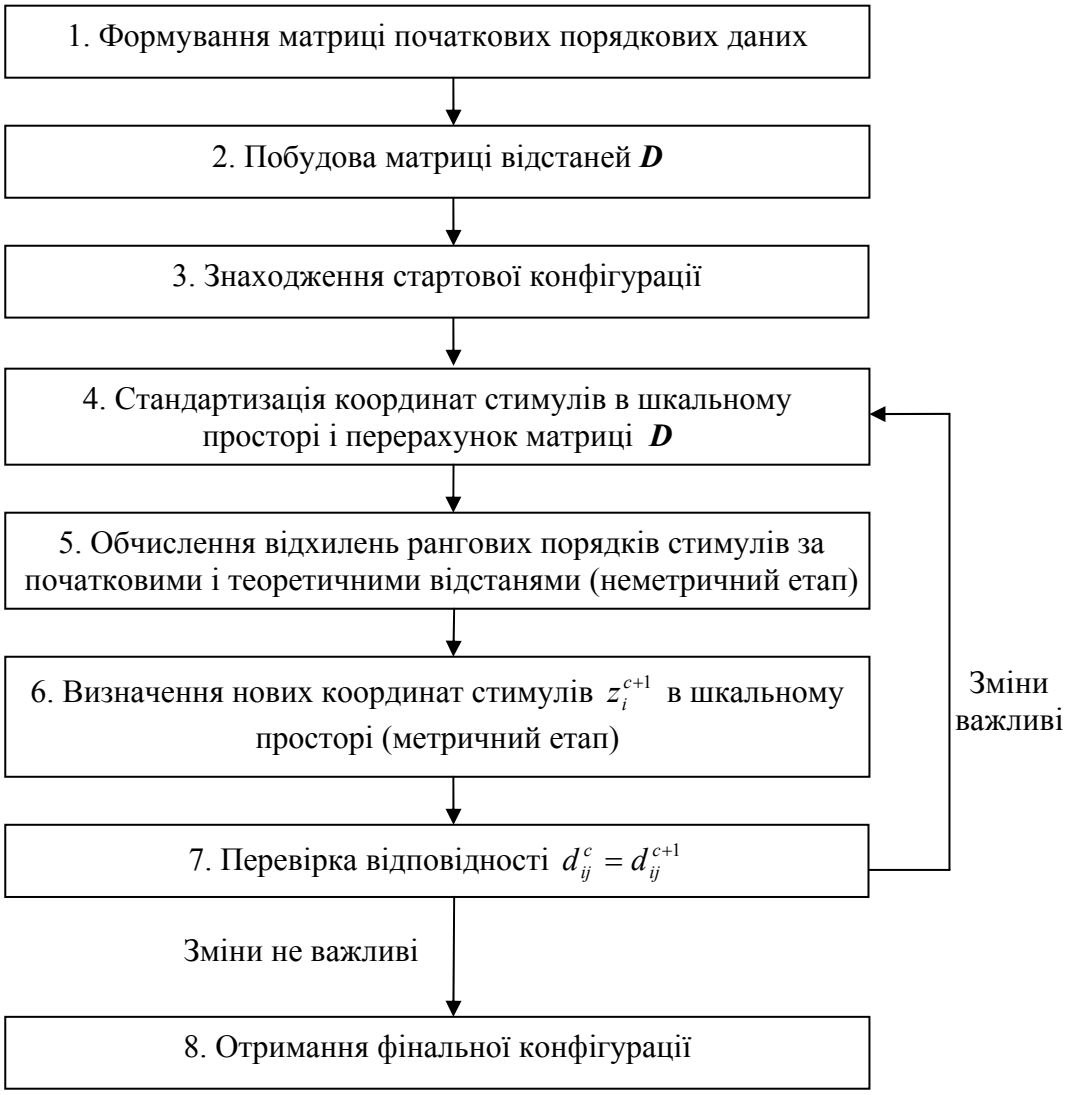


Рис. 2.16. Блок-схема алгоритму неметричного шкалювання

Як видно з рис. 2.16, процедура неметричного шкалювання являє собою ітеративний циклічний процес. На сьомому етапі перевіряється якість результатів оцінок, отриманих на черговий (c+1) ітерації, і виноситься судження про необхідність продовжувати цикл (перехід в четвертий блок) або завершити ітерації, отримавши фінальне рішення (восьмий блок).

Перші два етапи здійснюються на стадії подання та первинної обробки кількісних даних. Тут основною метою є побудова матриці відстаней між об'єктами $\boldsymbol{D}$.

На третьому етапі знаходиться первісне рішення задачі неметричного шкалювання у вигляді стартової конфігурації. Пошук стартової конфігурації та одержання перших приблизних оцінок координат векторів $\boldsymbol{x_1}$, $\boldsymbol{x_2}$, ... відбувається на основі різних підходів, наприклад, методом простої ординації Орлочі, з допомогою алгоритмів Торгерсона, Краскала, за аналогією з деякими заздалегідь відомими рішеннями та ін.

Однак, найбільш часто стартова конфігурація визначається на базі стандартного методу Гуттмана-Лингоса, який зводиться до застосування компонентного аналізу до матриці евклідових відстаней $\boldsymbol{D}$ і використання лінійної моделі (2.59). В результаті виходять початкові координатні оцінки стимулів у шкальному просторі, які дозволяють ідентифікувати та інтерпретувати виділені шкали (зазвичай 2-3).

Четвертий етап процедури полягає в стандартизації координат стимулів у шкальному просторі і перерахунку вихідної матриці відстаней $\boldsymbol{D}$. Стандартизація, зазвичай, здійснюється за відомою формулою центрування та нормування (1.2). Вона проводиться з метою збереження пропорцій стимульного простору і недопущення вироджених рішень, коли шкальний простір стискається до розмірів точки і багатовимірне неметричне шкалювання не дає скільки-небудь значимих результатів.

Це відбувається, наприклад, коли кілька стимулів отримують однакові або близькі оцінки. В такій ситуації вони практично зливаються в одну точку і стають невидимими для дослідника. Іншими словами, процедура стандартизації покликана тут «розрідити» вихідні оцінки векторів $\boldsymbol{x_1}$, $\boldsymbol{x_2}$, ... Потім на основі знайдених значень z_{ij} здійснюється розрахунок нової матриці відстаней між стимулами в стандартизованому шкальному просторі на базі формули евклідової метрики (див. табл. 2.3) при $N = 2$.

На п'ятому етапі проводиться обчислення відхилень рангових порядків стимулів за вихідними і теоретичними відстанями, знайденими за стандартизованими оцінками. Неметричний етап шкалювання призначений для впорядкування оцінок відстаней між стимулами до і після стандартизації. Справа в тому, що в стандартизованому шкальному просторі монотонність вихідних даних може порушитися. Коригування теоретичних величин d_{ij} здійснюється з метою відновлення загальної тенденції до зростання відмінності стимулів у ранжированих вихідних даних.

Невідповідність рангових порядків стимулів δ_{ij} і δ_{ij}*, побудованих за вихідними і теоретичними відстанями, добре простежується на графіку, якщо по осі абсцис відкладати вихідні (емпіричні) оцінки рангів δ_{ij}, а по осі ординат – теоретичні δ_{ij}*, як це показано на рис. 2.17.

Очевидно, що кожна з осей координат має $n(n - 1)/2$ градацій (по числу наддіагональних елементів матриці $\boldsymbol{D}$). Тут штриховою лінією показана лінійна монотонна функція (2.59), тобто коли рангові порядки стимулів, побудовані за

вихідним і теоретичним матриць відстаней, повністю збігаються. Ламана лінія відображає відхилення теоретичних рангових оцінок від емпіричних.

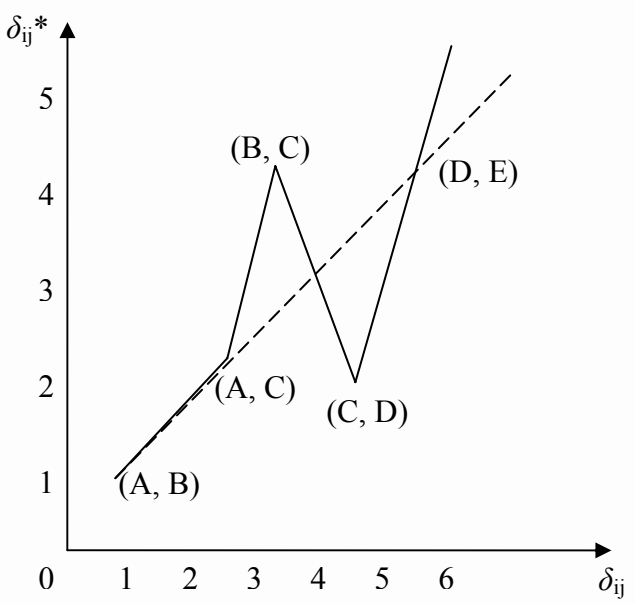


Рис. 2.17. Співвідношення рангових порядків стимулів, побудованих за вихідними і теоретичними матрицями відстаней

Графічно розбіжності зазначених рангових оцінок можна отримати також, якщо по осі абсцис замість δ_{ij} відкладати значення відстаней d_{ij} між стимулами у вихідному просторі експертних оцінок. Такий графік називається діаграмою Шепарда.

При цьому всі розрахунки, пов'язані з коригуванням теоретичних величин d_{ij} проводяться в спеціальній табл. 2.5.

Таблиця 2.5

Вихідні та скориговані оцінки відстаней між стимулами в теоретичному та шкальному просторах

Вихідний ранговий порядок (δ_{ij})	Стимул	Стимул	Стандартизовані відстані (d_{ij}^0)	Ранговий порядок у стандартизованому шкальному просторі	Скореговані оцінки відстаней (d_{ij}^1)
1	2	3	4	5	6
1	A	B	d_{AB}^0	1	d_{AB}^1
2	A	C	d_{AC}^0	2	d_{AC}^1
…	…	…	…	…	…
$n(n-1)/2$	K	L	d_{KL}^0	$n(n-1)/2$	d_{KL}^1

Слід мати на увазі, що в реальних економічних дослідженнях дані стовпців 1 і 5 табл. 2.5, зазвичай, не збігаються (див. рис. 2.17). Тому виникає необхідність відновити монотонність стандартизованих теоретичних оцінок відстаней між стимулами (тобто скоригувати їх) відповідно до даних стовпчика 1. Скориговані оцінки відстаней (d_{ij}^1) знаходяться різними методами,

наприклад, шляхом усереднення величин d_{ii}^{0} за тими парами стимулів, за якими виникли порушення монотонності у стовпцях 1 і 5.

Отримані результати записуються в стовпчик 6 табл. 2.5 і служать основою так званого метричного (шостого етапу процедури), який служить для визначення нових координат стимулів z_1^{c+1}, z_2^{c+1}... у шкальному просторі. Попередні (вихідні) координати стимулів позначаються відповідно z_1^{c}, z_2^{c}... Для перерахунку координат застосовується формула Лингоса-Роскама:

$$z_{iL}^{c+1} = z_{iL}^{c} - [\sum_{j} (1 - d_{ij}^{c+1}/d_{ij}^{c})(z_{iL}^{c} - z_{jL}^{c})]/n \,. \qquad (2.60)$$

Щоб уникнути ділення на нуль, якщо $d_{ij}^{c} = 0$, відношення d_{ii}^{c+1}/d_{ii}^{c} у формулі (2.60) довільно прирівнюється до одиниці. Значення z_{iL}^{c+1} розраховуються для всіх досліджуваних об'єктів (стимулів), потім за новими оцінками координат стимулів у шкальному просторі знову знаходяться відстані між ними і ітерація циклу закінчується. Залишається оцінити якість її результатів.

На сьомому етапі неметричного шкалювання як раз і відбувається оцінка суттєвості виконаного коригування стартової конфігурації векторів шкального простору. У разі значущості змін, отриманих на с+1 ітерації в порівнянні з ітерацією с, процедура триває (перехід по циклу в блок 4 схеми на рис. 2.16). В іншому випадку (d_{ii}^{c+1} несуттєво відрізняються від d_{ii}^{c}) цикл закінчується і знайдена конфігурація оцінок стимулів вважається фінальною.

Перевірка значущості змін, отриманих на с+1 ітерації в порівнянні з попередньою, здійснюється за допомогою спеціальних стрес-формул, а також на основі коефіцієнта відчуження. Наведемо найбільш вживані з них.

1. Стрес-формули Краскала:

$$S_1 = [\sum_{i,j} (d_{ij}^{c} - d_{ij}^{c+1})^2 / \sum_{i,j} (d_{ij}^{c+1})^2]^{1/2},$$

$$S_2 = [\sum_{i,j} (d_{ij}^{c} - d_{ij}^{c+1})^2 / \sum_{i,j} (d_{ij}^{c+1} - \bar{d}_{ij}^{c})^2]^{1/2}, \qquad (2.61)$$

де $\bar{d}_{ij}^{c} = (\sum_{i,j} d_{ij}^{c})/n^2$ – середня арифметична з усіх відстаней, знайдених на ітерації с.

2. Стрес-формули Янга:

$$SS_1 = \{\sum_{i,j} [(d_{ij}^{c})^2 - (d_{ij}^{c+1})^2]^2 / \sum_{i,j} (d_{ij}^{c+1})^4\}^{1/2},$$

$$SS_2 = \{\sum_{i,j} [(d_{ij}^{c})^2 - (d_{ij}^{c+1})^2]^2 / \sum_{i,j} [(d_{ij}^{c+1})^2 - (\bar{d}_{ij}^{C})^2]^2\}^{1/2}. \qquad (2.62)$$

3. Коефіцієнт відчуження Гуттмана

$$k = (1 - \varphi^2)^{1/2} \quad , \tag{2.63}$$

де $\varphi = [\sum_{i,j} (d_{ij}^{c} d_{ij}^{c+1})]/[\sum_{i,j} (d_{ij}^{c})^2 \sum_{i,j} (d_{ij}^{c+1})^2]^{1/2}$.

Загальний сенс виразів (2.61) – (2.63), які невід'ємні, а коефіцієнт відчуження k приймає значення від 0 до 1, полягає в наступному: чим ближче розраховані величини до нуля, тим менше змін відбулося у відстанях між стимулами в шкальному просторі на даній ітерації в порівнянні з попередньою. Тим більше є підстав для припинення процедури неметричного шкалювання та розгляду отриманої конфігурації векторів в якості фінальної.

І, навпаки, чим вище розраховані величини (чим ближче k до одиниці), тим більше змін відбулося у відстанях між стимулами в шкальному просторі на даній ітерації в порівнянні з попередньою. Тим більше є підстав для продовження процедури неметричного шкалювання з метою коригування отриманої проміжної конфігурації векторів.

Змістовна інтерпретації значень величин, знайдених за наведеними стрес-формулами, конкретизується за допомогою спеціальних таблиць. Наприклад, значення величин S_1, S_2 трактуються наступним чином (див. табл. 2.6).

Таблиця 2.6

Змістовна оцінка величин S_1 і S_2, знайдених за стрес-формулами Краскала

Ступінь відповідності d_{ij}^{c} і d_{ij}^{c+1}	Для формули	
	S_1	S_2
Низька	0,2	0,4
Задовільна	0,1	0,2
Добра	0,05	0,1
Відмінна	0,025	0,15
Ідеальна	0	0

Порівнюючи розрахункові значення S_1 і S_2, знайдені за стрес-формулами (2.61), з табличними, можна зробити висновок про необхідність нової ітерації (при низькому або задовільному ступеню відповідності d_{ii}^{c} і d_{ii}^{c+1}), або про зупинення процедури неметричного шкалювання (при гарному, відмінному або найвищому ступеню відповідності d_{ij}^{c} і d_{ij}^{c+1}) і отриманні фінальної конфігурації.

Приблизно аналогічно інтерпретується величина коефіцієнта відчуження (2.63):

- при $0 \le k < 0,3$ ступінь відповідності d_{ij}^{c} і d_{ij}^{c+1} хороша;
- при $0,3 \le k < 0,7$ ступінь відповідності d_{ij}^{c} і d_{ij}^{c+1} задовільна;
- при $k \ge 0,7$ ступінь відповідності d_{ij}^{c} і d_{ij}^{c+1} низька.

Підводячи деякі підсумки, зазначимо, що в рамках методів неметричного шкалювання вирішуються завдання, схожі з завданнями метричного шкалювання: оцінка координат стимулів і відстаней між ними, обертання осей системи координат, інтерпретація отриманих в результаті конфігурацій векторів. В той же час є й суттєві відмінності.

Неметричне шкалювання базується на більш складних алгоритмах, що включають пошук стартової конфігурації, коригування теоретичних оцінок відстаней і координат стимулів у шкальному просторі. Ітеративна процедура неметричного шкалювання будується таким чином, щоб не допустити появи вироджених рішень і суттєвих розбіжностей функціональних монотонних зв'язків емпіричних і теоретичних даних. В його алгоритмах проблемними залишаються питання вибору математичної форми монотонної функції, найбільш адекватної фактичному розподілу характеристик відмінностей стимулів, а також інтерпретація результатів аналізу.

Окрім того, існує невизначеність при встановленні конкретної метрики всередині універсальної моделі, побудованої на основі метрики Мінковського (вибір показника ступеню *N*). Успішне подолання зазначених методологічних труднощів цілком і повністю залежить від знань, досвіду та інтуїції дослідника. Дуже корисним видається попереднє і глибоке ознайомлення дослідника з іншими багатовимірними статистичними методами: основами таксономічної класифікації об'єктів, процедур кластерного і факторного аналізів.

На закінчення відзначимо, що методи побудови та використання моделей пошуку індивідуальних відмінностей джерел інформації, а також алгоритми аналізу переваг експертів за допомогою неметичного шкалювання на базі системи STATISTICA розглядаються в [126, с.239-256].

В роботі [30] розглянуто можливості оцінки конкурентоспроможності підприємств на основі методів неметричного шкалювання, а в роботі [60] наводиться візуалізація економічних об'єктів (8 молокопереробних підприємств) у просторі двох латентних осей – організаційної та економічної складових конкурентоспроможності досліджуваних підприємств.

3. ОЦІНКА ЛАТЕНТНИХ ОЗНАК В ЕКОНОМІЦІ УКРАЇНИ

3.1 Оцінка трудових латентних ознак

Перші наукові роботи в галузі економіки, присвячені вивченню латентних ознак, з'явилися в Україні в 90-х роках XX ст. в Одеському державному економічному університеті. Одне з перших таких досліджень було проведене автором даної книги у 1990-1993 рр. і спрямоване на оцінку продуктивної сили та інтенсивності праці в промисловості [124; 131]. У ньому було теоретично обґрунтовано, що на рівні підприємства ознака метричної шкали відносин «продуктивність праці» (Y) може бути виражена як добуток двох інших ознак – «продуктивна сила праці» (X) і «інтенсивність праці» (Z):

$$Y = Q/T = (Q/E)\times(E/T) = X\times Z, \qquad (3.1)$$

де Q – об'єм виробленої продукції (робіт, послуг) підприємством за певний проміжок часу у вартісному вираженні;

T – загальні витрати робочого часу всіма працівниками підприємства;

E – витрати абстрактної праці у вигляді сумарного видатку мускульної, психічної, інтелектуальної енергії людських організмів усіх працівників підприємства.

Згідно з формулою (3.1) визначення рівнів продуктивної сили праці X і інтенсивність праці Z базується на вимірюванні витрат мускульної, психічної, інтелектуальної енергії людських організмів E працівників підприємства. Основна проблема полягає в тім, що на сучасному етапі розвитку метрології, психології, наукової організації праці досить точному й безпосередньому вимірюванню підлягають тільки витрати мускульної енергії організму в процесі фізичної праці. Щодо витрат психічної, інтелектуальної енергії працівників, особливо, службовців, менеджерів та власників підприємства, то їх вимірювання в теперішній час не уявляється можливим.

Указані моменти обумовлюють невимірність ознак E і, відповідно, X, Z у метричній шкалі відносин. Про їх рівень зазвичай судять експерти за допомогою градацій порядкової шкали на основі значень чинників-симптомів – техніко-економічних показників виробництва і праці. Так, на базі дослідження темпу роботи і часу зайнятості можна дійти висновку про ступінь інтенсивності праці робітників на протязі зміни. Рівень кваліфікації і технічні параметри використаного устаткування дають певну уяву про величину продуктивної сили їх праці.

Отже, продуктивна сила і інтенсивність праці відносяться до групи латентних ознак економічних об'єктів.

Змістовну схему взаємозв'язку латентних ознак, що розглядаються, представлено на рис. 3.1.

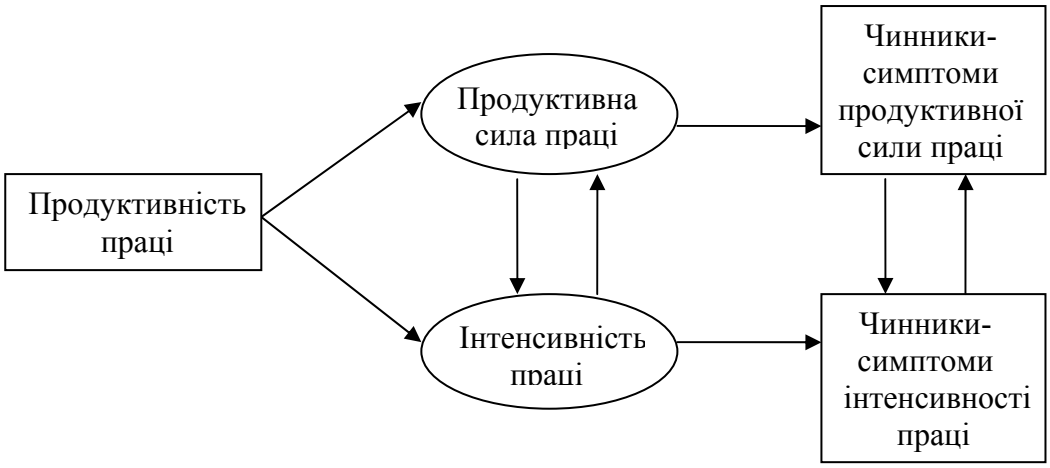


Рис. 3.1. Взаємозв'язок продуктивності, продуктивної сили і інтенсивності праці

Початковим моментом для виділення групи чинників-симптомів, які характеризують рівень продуктивної сили та інтенсивності праці на підприємстві, є основні положення теорії трудової вартості, існуючі дослідження факторів росту продуктивності праці, їх класифікація.

На базі проведеного якісного економічного аналізу в даному напрямку ми запропонували два наступних агрегованих угрупування чинників-симптомів:

А. Продуктивної сили праці в промисловості:

1. Технічний рівень виробництва.
2. Розподіл і кооперація праці.
3. Склад кадрів.
4. Природно-кліматичні умови.

Б. Інтенсивності праці в промисловості:

1. Організація виробництва.
2. Організація праці.
3. Склад кадрів.
4. Природно-кліматичні умови [124, с.16-23; 131, с.49-101].

Наявність в угрупуваннях однакових третього й четвертого пунктів свідчить про те, що існують дві великі групи чинників-симптомів, які є спільними по відношенню до обох досліджуваних латентних ознак промислових підприємств.

Оцінка продуктивної сили і інтенсивності праці здійснювалась на базі наступних математико-статистичних методів і моделей:

1) методи факторного аналізу, в тому числі метод головних компонент;
2) виділення найбільш колінеарного чинника-симптому;
3) класичний і модифікований алгоритми таксономії.

У результаті використання вказаних методів виникає можливість отримати, щонайменше, п'яти варіантів оцінок продуктивної сили і інтенсивності праці на досліджуваних підприємствах, тобто всього 25 (5×5) різних варіантів рішення поставленої задачі. Вибір найбільш точного (оптимального) варіанту рішення здійснювався на базі виконання вихідного

балансового співвідношення (3.1) чи похідних від нього співвідношень. Зокрема, з формули (3.1) випливає

$$LnY = LnX + LnZ, \qquad (3.2)$$

тобто логарифм спостережуваного показника продуктивності, який вимірюється в метричній шкалі, повинен дорівнювати сумі логарифмів знайдених оцінок продуктивної сили і інтенсивності праці. Оскільки дві останні оцінки, зазвичай, вимірюються в метричній шкалі інтервалів, тобто в масштабованому вигляді, то і показник продуктивності праці Y задля порівнянності треба стандартизувати (Y').

Співвідношення (3.2) може бути використане для побудови особливого інтегрального показника, що отримав назву критерій балансу змінних (докладніше див. додаток Ж). Для даної задачі він має вигляд:

$$B_j = \sum_{i=1}^{n} \frac{\left| LnY_i' - LnX_i' - LnZ_i' \right|}{LnY_i'}; \quad i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, m. \qquad (3.3)$$

Підсумовування по i проводиться за всіма n економічними об'єктами сукупності, що досліджується, у випадку просторової варіації латентних ознак, або за всіма точками періоду передісторії (упередження) у випадку часової варіації продуктивності, продуктивної сили і інтенсивності праці на окремому підприємстві. В результаті отримується $m = 25$ значень критерію B_j за числом різних варіантів рішення поставленої задачі.

Чим ближче знайдена величина B_i до нуля, тим точніші знайдені оцінки продуктивної сили і інтенсивності праці і, навпаки. Вона відображає сумарний відносний розбаланс ознак, що вивчаються, і реагує на всі зміни, які відбуваються на стадії відбору конкретних чинників-симптомів досліджуваних прихованих властивостей економічних об'єктів.

Таким чином, використовуючи оптимізацію отриманих рішень на основі критерію $B_i \rightarrow \min$, відбираються із множини можливих найбільш точні оцінки продуктивної сили і інтенсивності праці на аналізованих підприємствах.

Указаний підхід був апробований на 7 підприємствах приладобудування м. Одеси за даними 1991 р., а також на макрорівні при оцінці динаміки продуктивної сили і інтенсивності праці в промисловості України в 1985-1991 рр. За нашими розрахунками, продуктивна сила праці в промисловості України за даний період зросла на 23,2 % (середньорічний темп приросту 3,5 %), а інтенсивність праці знизилась на 5,5 % (середньорічний темп зниження 0,09 %).

При цьому в 1991 р. кризові явища в українській промисловості, викликані розпадом СРСР і розривом коопераційних зв'язків з колишніми постачальниками сировини, напівфабрикатів, комплектуючих виробів, з одного боку, і споживачами готової продукції, з іншого боку, негативно проявились як у зниженні продуктивної сили праці (-1,5 %), так і інтенсивності праці (-5,1 %).

Окрім того, отримані оцінки трудових латентних ознак були застосовані для комплексного прогнозування продуктивності, продуктивної сили та інтенсивності праці в промисловості України на 1992-1995 рр. Визначення форми трендів здійснювалось за допомогою статистичних методів, а також з

використанням критерію балансу змінних (3.3), що дозволило вибрати з багатьох варіантів систем прогнозних трендів найточнішу, яка забезпечила отримання найбільш збалансованих оцінок у майбутньому всіх трьох досліджуваних трудових показників.

Наступним кроком у вивченні трудових латентних ознак економічних об'єктів явилось дисертаційне дослідження Т.В. Погорєлової (1993 р., науковий керівник О.Г. Янковий) на тему «Методи оцінки ефективності праці службовців на промислових підприємствах» [77]. У цій роботі на підставі аналізу особливостей праці службовців у виробничій сфері, головних відмінностей розумової праці від фізичної розкривається основний зміст ефективності праці керівників і фахівців підприємств як латентної ознаки.

У процесі дослідження автор дійшла висновку, що вказана ознака являє собою співвідношення результатів праці службовців, виражених корисним (технічним, організаційним, економічним та ін.) ефектом, із витратами праці задля досягнення цих результатів. Але відсутність предмета праці керівників і фахівців у звичайній речовій формі робить даний ефект невимірним у метричній шкалі відносин. Тому результати праці службовців і, як наслідок, сам показник ефективності їх праці представляють собою латентну ознаку, вимірну лише в порядковій шкалі на базі експертних оцінок.

Водночас, рівень ефективності праці службовців на промислових підприємствах проявляється на поверхні у вигляді множини чинників-симптомів, що характеризують ступінь досліджуваної латентної ознаки. Т.В. Погорєловою виділено шість таких чинників-симптомів – кількісних показників метричної шкали відносин:

1) об'єм продукції на одного службовця;

2) об'єм продукції на один карбованець фонду оплати праці службовців;

3) сума прибутку (збитків) на одного службовця;

4) сума прибутку (збитків) на один карбованець фонду оплати праці службовців;

5) частка службовців у чисельності промислово-виробничого персоналу;

6) частка фонду оплати праці службовців у загальному фонді оплати праці промислово-виробничого персоналу.

Запропонований набір чинників-симптомів відповідає уявленням про те, що ефективність праці службовців – керівників і фахівців у машинобудуванні проявляється завдяки кінцевим результатам виробничої діяльності всього колективу промислових підприємств.

У процесі дослідження було побудовано сім варіантів кількісної оцінки ефективність праці керівників і фахівців за даними 8 підприємств приладобудування м. Одеси за 1990-1991 рр. за допомогою методу виділення найбільш колінеарного чинника-симптому, факторного аналізу, а також спільного застосування методу головних компонент і таксономічного аналізу. Отримані таким чином оцінки ефективності праці службовців досліджуваних підприємств представляли собою лінійні комбінації всіх шести вихідних чинників-симптомів.

Аналіз інформативності отриманих результатів показав, що найбільш точні оцінки надає спільне використання методу головних компонент і таксономічного

аналізу. Знайдені кількісні характеристики плідності праці керівників і фахівців підприємств приладобудування були використані в подальшому економічному дослідженні. Зокрема, на базі отриманих оцінок рівня ефективності праці службовців було здійснено угрупування об'єктів з виділенням груп лідерів, середняків і аутсайдерів.

Окрім того, запропонована методика була застосована на одеському заводі «Промзв'язок» при визначенні плідності праці керівників і фахівців окремих функціональних підрозділів підприємства й розподілі прибутку, призначеного для матеріального заохочення службовців за результатами виробничо-фінансової діяльності в 1991 р.

3.2 Оцінка фінансового стану підприємств

Подальшій розвиток ідей і методів вивчення латентних ознак економічних об'єктів спостерігалось на початку XXI ст. У цей період було виконане дисертаційне дослідження Г.В. Юр'євої на тему «Оцінка фінансового стану промислових підприємств – багатовимірний підхід» (2006 р., науковий керівник О.Г. Янковий) [115].

Проведений аналіз показав, що результати використання найбільш відомих в економічній літературі моделей і тестів експрес-діагностики фінансового стану вітчизняних виноробних підприємств (моделі Альтмана 1968, 1969, 1983 рр., Лісу, Таффлера-Тішоу, Спрінгейта та ін.), не завжди є узгодженими. А в деяких випадках вони суперечать один одному. Тому треба досить обережно відноситись до висновків, які будуються за допомогою цих моделей і тестів [147].

Головними причинами такого становища є:

- перекручування вихідних даних, оскільки інформаційна база (термінологія, вітчизняна звітність) часто не відповідає звітності, прийнятій в міжнародній практиці;
- в якості інструментів оцінки фінансового стану підприємств використовуються моделі, розроблені в минулому (у 60-80 роки минулого ст.) за міжнародними стандартами, які не враховують ні галузеві особливості підприємств, ні теперішні економічні умови господарювання в Україні;
- в умовах зростання змін у зовнішньому середовищі методи і моделі оцінки перспективного фінансового стану не враховують впливу зовнішніх факторів;
- більшість розглянутих вище підходів обмежена використанням лише алгоритмів кластерного і дискримінантного аналізу без врахування потенційних можливостей інших моделей з латентними показниками, побудованих, наприклад, на базі методології таксономії і факторизації чинників-симптомів фінансового стану підприємств.

На думку автора, фінансовий стан – складна, комплексна ознака, яка вивчається на базі сукупності виробничо-господарських та фінансових показників, що відображають наявність, розміщення й використання фінансових та виробничих ресурсів. Ознака «фінансовий стан» підприємства представляється

складною, прихованою властивістю виробничої системи і проявляється на поверхні у вигляді сукупності чинників-симптомів – окремих виробничо-фінансових показників та коефіцієнтів. Іншими словами, вказана ознака є латентною, оскільки не існує єдиного показника, який можна було б розглядати як характеристику рівня фінансового стану промислового підприємства. Про його рівень, зазвичай, виносять судження аудитори на основі економічного аналізу множини чинників-симптомів.

У проведеному дослідженні на основі даних фінансової та бухгалтерської звітності дев'яти виноробних підприємств України за 2004 р. здійснено оцінку їх фінансового стану за допомогою класичного і модифікованого алгоритмів таксономії, а також на базі комплексного сполучення методів кластеризації та дискримінації економічних об'єктів.

Алгоритми кластерного аналізу, зокрема, ієрархічна агломеративна процедура і метод *k*-середніх дозволили виділити дві групи підприємств за рівнем фінансового стану, а процедура подвійного об'єднання дозволила ідентифікувати отримані кластери з точки зору фінансового стану об'єктів, що їх утворюють. Отримані в результаті кластеризації досліджуваних виноробних підприємств висновки повністю тотожні тим, що були зроблені за результатами таксономічного аналізу. Це свідчить про не випадковість підсумків методів таксономії і кластерного аналізу, про наявність реальних розбіжностей у рівнях фінансового стану двох виділених груп об'єктів виноробства.

В контексті проблем проведеного дослідження дискримінація укладалася в побудові такої інтегральної моделі за інформацією чинників-симптомів (фінансових коефіцієнтів), яка з досить високою вірогідністю дозволяла б віднести нове виноробне підприємство до двох визначених кластерів – першого, в складі якого підприємства-лідери за своїм фінансовим станом, або другого, що містить об'єкти з потенційною неплатоспроможністю.

Проведений на основі розрахунку квадрату відстаней Махаланобіса й апостеріорних ймовірностей дискримінантний аналіз підтвердив попередні висновки, зроблені на стадії кластеризації досліджуваних виноробних підприємств. Отримана класифікація виноробних підприємств придатна для розпізнавання нових об'єктів, наприклад, підприємств цієї ж галузі за даними інших років.

Для побудови універсальної дискримінантної функції для виноробної галузі України був проведений регресійний аналіз отриманої класифікації підприємств. У якості залежної змінної виступала величина, яка характеризує приналежність підприємств до різних кластерів, а роль факторних ознак відігравали 6 чинників-симптомів фінансового стану:

1. Коефіцієнт фінансового ризику.
2. Коефіцієнт мобільності власного капіталу.
3. Коефіцієнт поточної ліквідності (покриття).
4. Коефіцієнт абсолютної ліквідності.
5. Рентабельність власного капіталу.
6. Загальний коефіцієнт оборотності капіталу.

В результаті відсіву незначущих змінних автором отримано чотирьох факторну інтегральну модель фінансового стану об'єктів виноробства, яка за своїми статистичними і апроксимаційними властивостями розглядалася як адекватна інтегральна дискримінаційна функція досліджуваної галузі. Вона допомогла визначити місце кожного підприємства в економічному середовищі, а також ефективність і ризикованість ділових відносин. Крім того, розроблена модель надавала можливість проводити аналіз як у просторі (порівняння різних підприємств між собою), так і у часі (встановлення динаміки зміни фінансового стану одного й того ж підприємства за обраний період часу).

На думку Г.В. Юр'євої, побудована галузева дискримінантна функція не є чимось застиглим і непорушним, як це трапилося з західними моделями та тестами експрес-діагностики, які протягом останніх 50 років механічно використовуються у незмінному вигляді у вітчизняних дослідженнях і часто дають суперечливі результати. Визначаючи поняття «фінансовий стан» підприємства як динамічну категорію, що по своєї суті є латентною ознакою, також і отриману дискримінантну функцію треба розглядати в діалектичному розвитку, в динаміці. Це означає, що з часом її коефіцієнти повинні перераховуватися по мірі появлення нової статистичної інформації про зміну фінансових коефіцієнтів виноробних підприємств.

У роботах [147-149] доведено, що змістовна модель ознаки «фінансовий стан» підприємства є складною і розвиненою. Вона поряд із звичайними чинниками-симптомами (коефіцієнтами ліквідності, оборотності оборотних активів, ділової активності тощо) включає також групові ознаки фінансового стану підприємства, які відображають його фінансову стійкість, ліквідність, платоспроможність, прибутковість, рентабельність, майновий стан. Дані узагальнюючі ознаки за своєю суттю теж є прихованими, так як кожна з них не може бути вичерпно охарактеризована лише якимось одним економічним показником (рис. 3.2).

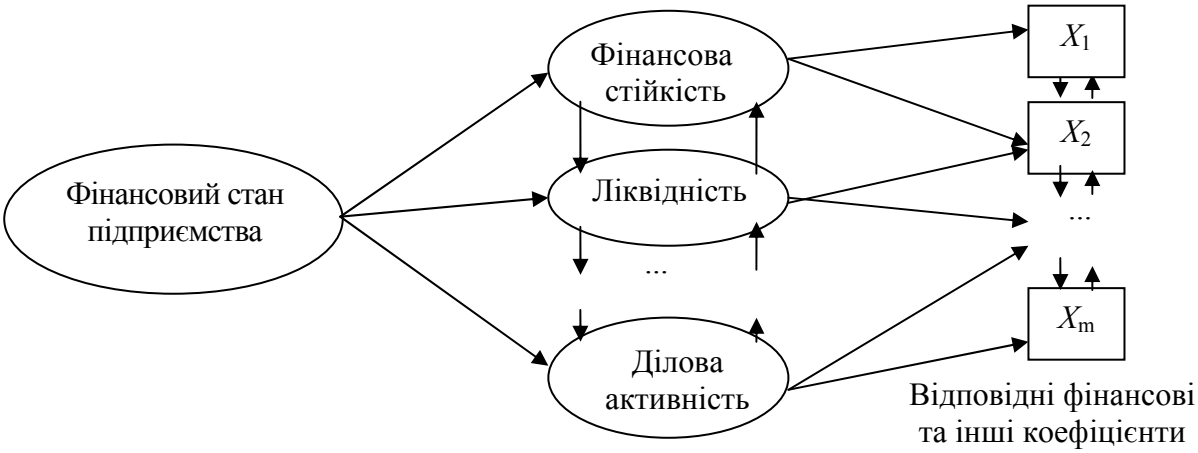


Рис. 3.2. Схема взаємодії ознак і показників фінансового стану промислового підприємства

3.3 Оцінка конкурентоспроможності підприємств

Надалі вивчення латентних ознак в економіці продовжила О.Б. Чернишова (2009 р., науковий керівник О.Г. Янковий), яка досліджувала конкурентоспроможність 9 підприємств кондитерської галузі харчової промисловості України [104; 105].

Проведений критичний аналіз показав, що конкурентоспроможність – це складна економічна ознака, що виражає результат взаємодії всіх елементів системи внутрішніх (виробничих, економічних, науково-технічних та ін.) і зовнішніх (життєутворюючих, адаптивних) відносин між промисловими підприємствами з приводу можливості реалізації даного виду продукції на конкретному ринку чи його сегменті. Це комплексне поняття, яке внаслідок щільного зв'язку з категорією «конкуренція», охоплює безліч сторін, факторів та умов суперництва підприємств за прихильність споживачів до товарів певних виробників. Рівень конкурентоспроможності підприємства – дзеркало, в якому відображаються сукупні підсумки роботи практично всіх служб підприємства, а також стан зовнішнього довкілля.

Автор дійшов висновку, що в економічній науці не існує єдиного показника, котрий можна було б розглядати в ролі індикатора рівня конкурентоспроможності суб'єкта господарювання. В формах статистичній звітності промислових підприємств сьогодні відсутній як показник потенційної (прогнозної) здатності вести суперництво у ринкових умовах господарювання, тобто показник конкурентоспроможності, так і показник успішності реальної (фактичної) конкурентної боротьби, тобто показник конкурентності.

Здійснене дослідження показало, що висока результативність роботи, визначена лише за одним показником, не є гарантом успішного продовження діяльності підприємства в майбутньому. Причина полягає в тім, що кожен показник зорієнтований лише на відображення тільки одного боку конкурентоспроможності підприємства. При цьому інші її сторони й грані можуть не висвітлюватись, що неминуче приведе до хибних висновків щодо рівня конкурентоспроможності.

Так, незважаючи на популярність вимірників конкурентоспроможності, заснованих на величині прибутку як критерію оцінки успішності економічної діяльності підприємства, їхня слабкість очевидна. Практика вітчизняної харчової промисловості показує, що досить частими є факти маніпуляції розмірами прибутку для прикриття фактичних результатів діяльності підприємств кондитерської галузі.

Оманним в оцінці конкурентоспроможності підприємства є й загальноприйнятий тест лідерства через показник частки ринку, коли він використовується ізольовано від інших характеристик даної ознаки. Справді, існує багато прикладів того, коли підприємства, володіючи найбільшою часткою ринку, за своєю прибутковістю знаходилися далеко позаду конкурентів, у яких показник частки ринку був значно нижчим. Це означає, що не їм, а вони платять за становище лідера.

Ринкова сила товаровиробника проявляється в його здатності диктувати ринку ціну, що перевищує ціни його пріоритетних конкурентів. Одним із критеріїв цієї здатності є еластичність попиту за ціною для диференційованого товару. Чим нижче за абсолютною величиною цінова еластичність попиту, тим менш чутливий ринок до збільшення ціни. Отже, попит на продукцію виробника, що володіє ринковою силою (конкурентоспроможністю) менш еластичний, ніж попит на товар підприємства, який слабо диференційований. У результаті такий виробник має більше шансів домогтися того, щоб група споживачів, чутливих до елементу диференціації, погодилася платити підвищену ціну.

Таким чином, можна говорити, принаймні, про наявність системи трьох важливіших узагальнюючих чинників-симптомів, які характеризують рівень конкурентоспроможність підприємства як латентної ознаки [144-146]:

1. Рентабельність виробництва.

2. Ринкова позиція, що виражається певною часткою ринку або його відповідного сегменту.

3. Цінова еластичність попиту споживачів на диференційовані товари.

Окрім указаних узагальнюючих чинників-симптомів конкуренто-спроможності підприємства, існують також її складові, які проявляються в величині первинних факторів латентної ознаки, що вивчається. Дійсно, кожну з таких складових, як економічний, техніко-технологічний рівень виробництва, ефективність управління, інноваційної, кадрової, інформаційної діяльності підприємства практично неможливо визначити одним-єдиним показником.

Це означає, що вказані складові конкурентоспроможності підприємства також відносяться до класу латентних ознак. Вони, в свою чергу, характеризуються відповідними значеннями економічних, техніко-технологічних, кон'юнктурних та інших чинників-симптомів метричної шкали відносин, які висвітлюють названі вище складові конкурентоспроможності підприємства.

Узагальнення проведеного дослідження дозволило автору представити змістовну схему взаємодії ієрархічних рівнів конкурентоспроможності підприємства в такий спосіб (рис. 3.3).

О.Б. Чернишова вважає, що саме така схема взаємодії ієрархічних рівнів ознак та чинників-симптомів конкурентоспроможності підприємства є найбільш обґрунтованою й повністю прийнятною для практичного оцінювання латентного економічного явища, яке аналізується [105, с.18-19]. Вона запропонувала будувати майбутню оцінку конкурентоздатності підприємств на безпосередньому використанні саме узагальнюючих чинників-симптомів – рентабельності виробництва, частки ринку або його відповідного сегменту, цінової еластичності попиту споживачів. При цьому первинні чинники-симптоми відіграють аналітичну роль щодо оціненого рівня досліджуваного латентного показника.

Методологічну основу запропонованої загальної схеми процедури оцінки конкурентоздатності підприємств склали існуючі багатовимірні статистичні методи – канонічний аналіз, таксономія об'єктів в її класичному та

модифікованому варіантах, кластерний та дискримінантний аналіз з урахуванням особливостей структури даної латентної економічної ознаки.

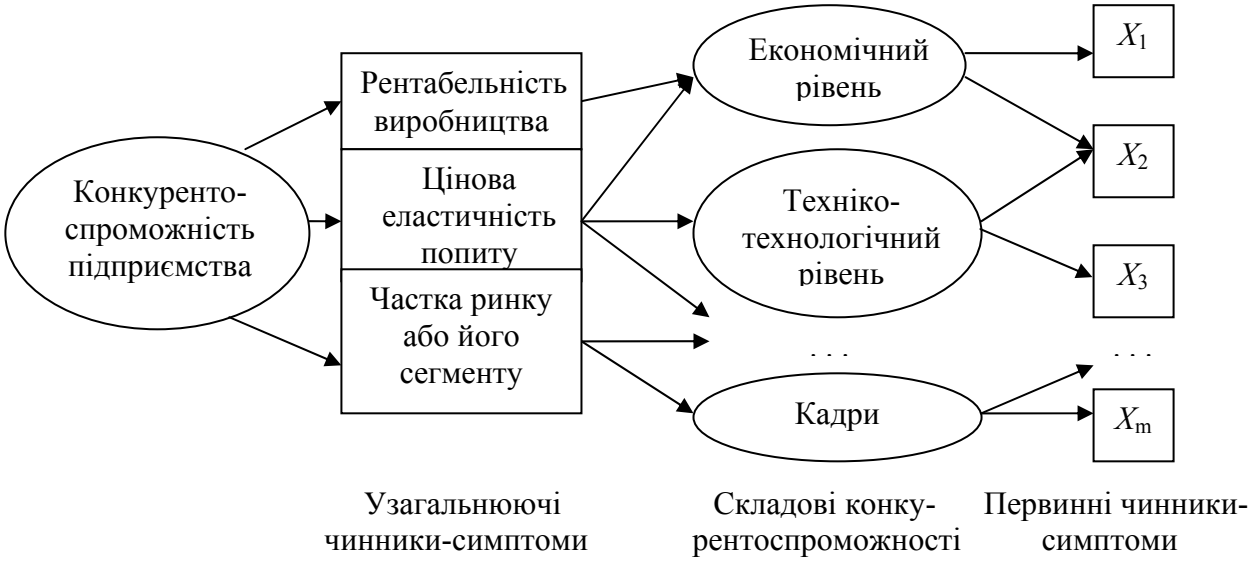


Рис. 3.3. Взаємодія ієрархічних рівнів конкурентоспроможності підприємства

Автор вважає, що канонічний аналіз двох груп економічних показників (узагальнюючих та первинних чинників-симптомів в процесі оцінювання конкурентоспроможності підприємств) доцільно проводити за аналогією з регресійним моделюванням, коли пошук найкращого рівняння базується на переборі усіх можливих сполучень вихідних змінних. Це відкриває можливість знаходження найбільш точних оцінок шуканої прихованої властивості.

Для уникнення помилкових висновків та отримання найбільш надійних оцінок рівня конкурентоспроможності досліджуваних підприємств О.Б. Чернишова запропонувала використовувати змішаний підхід при проведенні їх таксономії, заснований на застосуванні результатів обох головних алгоритмів таксономічного аналізу об'єктів – класичного та модифікованого. Його основу складає методика визначення об'єднаної оцінки латентного економічного показника на базі методів таксономії, застосовуються гнучкі статистичні ваги, які змінюються при поступовому русі по ранжированому ряду оцінок, отриманих за класичним та модифікованим методами таксономічного аналізу конкурентоздатності промислових підприємств.

Отримана за допомогою різних багатовимірних методів множина оцінок конкурентоспроможності тестується з метою вибору оптимальної оцінки на основі певного переліку первинних чинників-симптомів, який був визначений в ході канонічного аналізу. Вказане тестування ґрунтується на виборі найбільш точної регресійної моделі, яка зв'язує певну оцінку конкурентоздатності промислових підприємств з заданим колом її первинних чинників-симптомів.

Окрім О.Б. Чернишової, конкурентоспроможність як латентну ознаку розглядали також:

- Ю.А. Єгупов, В.О. Літвінова в дослідженні конкурентоздатності продукції машинобудування [29, 51];

- О.Г. Янковий, Л.В. Єргієва в процесі оцінки рівня та резервів росту конкурентоспроможності підприємств кабельної промисловості [30; 31; 136];
- О.М. Бабій, Н.О. Зайцева при вивченні виноробних підприємств [2; 35];
- А.Г. Гончарук, Т.М. Гамма, М.Г. Молла в ході аналізу діяльності молокопереробних підприємств харчової промисловості [18; 60; 61].

Автори вказаних робіт єдині в думці про прихований характер конкурентоспроможності, котра відноситься до латентних ознак сучасного підприємства, і оцінювати яку доцільно за допомогою математико-статистичних методів і моделей. Розбіжності ж в основному відносяться до переліку чинників-симптомів, що відображають рівень прихованої властивості підприємств різних галузей промислового виробництва України.

Так, В.О. Літвінова (2012 р., науковий керівник Ю.А. Єгупов) зазначає, що, оскільки конкурентоспроможність продукції є латентною відносною ознакою, то для її оцінки необхідно використовувати апарат багатовимірного статистичного аналізу, який складається з великої кількості методів – кластерного, дискримінантного, факторного і таксономічного аналізу. Серед указаних методів у контексті рішення поставленої проблеми автор особливо виділяє метод таксономії, який вирізняє простота, відсутність будь-яких вимог до сукупності досліджуваних об’єктів, зручний масштаб отриманих оцінок. Його застосування дозволило розрахувати таксономічний показник конкурентоспроможності 20 водонагрівачів 8 вітчизняних та зарубіжних фірм-виробників, а також здійснити їх ранжирування.

За результатами проведеного дослідження було визначено резерви росту конкурентоспроможності водонагрівача NovaTec (Україна) і розроблено конкретні заходи її підвищення, оцінено капітальні витрати та прогнозовано вплив заходів на зміну факторних показників і рівень конкурентоздатності даного виробу. На підставі цього розраховано показник інвестиційної вартості і відсоткового приросту аналізованої прихованої властивості водонагрівача за кожним заходом та в комплексі.

Л.В. Єргієва (2012 р., науковий керівник О.Г. Янковий) звернула увагу на методологічні проблеми дослідження конкурентоспроможності промислових підприємств, зокрема, пов’язані з наявністю безлічі різноманітних методик її оцінювання. Всі існуючі в практиці економіки і управління методи оцінки конкурентоспроможності економічних об’єктів можна класифікувати за двома основними критеріями: за ступенем об’єктивності (суб’єктивності) результатів оцінки, а також за підходом до оцінки – якісна чи кількісна.

За останній час помітно зріс інтерес до методів багатовимірного порівняльного аналізу. До них відносяться кластерний аналіз, таксономія, факторний аналіз, кореляційний аналіз, регресійний аналіз. Серед зазначених методів виділяються два магістральні напрями: метод таксономії та метод головних компонент. Тому автор пропонує провести оцінку конкуренто-спроможності 9 підприємств кабельної галузі України на основі саме цих методів, а також за допомогою неметричного багатовимірного шкалювання.

Дані методи дозволили виявити лідерів та аутсайдерів серед досліджуваних підприємств, що набуває особливої актуальності в теперішні

часи. Було порівняно показники двох лідерів – ПАТ «Завод «Південкабель» та ПАТ «Одескабель» за даними 2009-2011 рр. Відзначено, що перше підприємство лідирує за усіма показниками крім коефіцієнта автономії та витрат на 1 грн. реалізованої продукції. Отже, був зроблений висновок, що ПАТ «Завод «Південкабель» є безперечними лідером за показниками, які характеризують прибутковість серед усіх досліджуваних підприємств кабельної промисловості.

Явними аутсайдерами виявилися підприємства ВАТ «Азоакабель», ЗАТ «Кам'янець-Пділськкабель», ПАТ «Чернігівкабель». З метою кількісної оцінки резервів росту конкурентоспроможності було проведене порівняння показників господарської діяльності ПАТ «Чернігівкабель» з відповідними показниками лідера – ПАТ «Завод «Південкабель».

Н.О. Зайцева (2013 р., науковий керівник О.М. Бабій) у рамках дослідження рівня і резервів підвищення конкурентоспроможності 11 підприємств виноробної промисловості України встановила, що найбільшого поширення набули методи побудови матриць і визначення комплексного (інтегрального) показника. Проте недосконаліть указаних методів і доцільність їх застосування лише для здійснення експрес-аналізу конкурентоздатності об'єкта практично не дає можливості точного й надійного прогнозування показника, що вивчається. Це пояснюється тим, що конкурентоспроможність підприємства є латентною, прихованою ознакою, оцінити яку можна лише за допомогою аналізу взаємозв'язків чинників-симптомів.

Подолання вказаних недоліків здійснювалось у виконаному дослідженні на основі залучення багатовимірних методів, зокрема, канонічних кореляцій і множинної регресії. Виявлено, що побудова інтегрального показника конкурентоздатності підприємства і на його основі – матриці групового ранжирування є найбільш вдалим засобом визначення умов, які дозволяють об'єктам-конкурентам потрапити до вибірки, що аналізується.

Використання запропонованого підходу дозволило виділити з первинної вибірки, що містила 11 виноробних підприємств, групу з 4 об'єктів (ПрАТ «Агрофірма ім. С. Лазо», ПрАТ «Березанський», ПАТ «Цюрупінське, ЗАТ «Октябрьський вино-коньячний завод»). На базі дослідження даних відкритої звітності зазначених підприємств отримано оцінки їх конкурентоспроможності у кожний конкретний період часу за 2004-2011 рр. і здійснено ранжирування аналізованих об'єктів.

За результатами канонічного аналізу встановлено найвпливовіші чинники, які детермінують підвищення конкурентоспроможності виноробних підприємств півдня України: ріст рівня використання виробничих потужностей, прискорення оборотності капіталу, зростання темпів споживання шампанського, зменшення частки вартості ліцензії на оптову торгівлю у витратах підприємства. За допомогою проведеного кореляційно-регресійного моделювання оцінено величину резервів можливого росту конкурентоздатності кожного з 4-х виділених об'єктів.

М.Г. Молла (2014 р., науковий керівник Т.М. Гамма) конкурентоспроможність 8 молокопереробних підприємств Одеської та Миколаївської областей розглядала як латентну ознаку, що проявляється на поверхні у вигляді

сукупності чинників-симптомів – окремих групових та первинних складових. На погляд автора, до складових конкурентоспроможності підприємства відносяться наступні: організаційна, кадрова, потенційна та група економічних складових (ділова активність, фінансова стійкість, платоспроможність, прибутковість).

Оцінку виділених складових конкурентоздатності було проведено за допомогою методу об'єднаної таксономії на базі первинних чинників-симптомів за 2011 р., що дозволило здійснити ранжирування досліджуваних молокопереробних підприємств з подальшим використанням методу головних компонент. На основі розрахунків компонентного аналізу було проведене багатовимірне неметричне шкалювання, що дозволило візуалізувати отримані результати, тобто побудувати графік з положенням кожного підприємства у просторі двох осей – організаційної та економічної складових конкурентоспроможності досліджуваних об'єктів.

Запропонований метод аналізу організаційно-економічних складових конкурентоздатності надав змогу об'єктивно оцінити положення кожного молокопереробного підприємства, розробити дієві заходи не тільки задля посилення їх конкурентних позицій, але й для забезпечення життєздатності аналізованих об'єктів.

3.4 Оцінка інвестиційної привабливості підприємств

Г.В. Кошельок (2010 р., науковий керівник О.Г. Янковий) оцінювала рівень і вивчала внутрішні фактори та резерви росту інвестиційної привабливості промислових підприємств [42; 43; 137; 138].

На думку автора, інвестиційна привабливість підприємства – це складна ознака, яка відображає відносини між інвестором і реципієнтом щодо можливого вкладання коштів першого у виробничо-фінансову діяльність другого. Вона представляє собою приховану інтегральну характеристику підприємства, яка відображає доцільність вкладання в нього капіталу, ступінь задоволення вимог або інтересів інвестора, рівень фінансово-майнового стану підприємства та перспектив його розвитку, сукупність зовнішніх та внутрішніх умов його функціонування. Природа інвестиційної привабливості підприємства багатоаспектна: вона поєднує внутрішні та зовнішні, суб'єктивні та об'єктивні, потенційні та реальні начала.

При вивченні ознаки «інвестиційна привабливість» підприємства слід враховувати значення не одного фактора, а цілий комплекс, систему різноманітних чинників-симптомів, які з різних боків детально і всебічно характеризують цю складну економічну властивість. Виходячи з проведеного аналізу існуючих підходів до класифікації внутрішньовиробничих факторів інвестиційної привабливості підприємства, автор запропонувала систему з 38 чинників-симптомів, які об'єднані у 6 складових, визначаючих рівень економічної ознаки, що досліджується. Це майновий стан, ліквідність, фінансова стійкість, ділова активність, рентабельність, ринкова активність. На

рис. 2.3 наведено ієрархічну структуру та схему взаємодії внутрішніх чинників-симптомів і складових інвестиційної привабливості підприємства.

У відповідності до визначеної структури латентної ознаки «інвестиційна привабливість промислового підприємства» Г.В. Кошельок запропонувала двох стадійну процедуру її оцінювання. На першій стадії на основі первинних чинників-симптомів оцінювались складові інвестиційної привабливості, а на другій стадії на базі знайдених оцінок складових – безпосередньо сама прихована властивість підприємств, що вивчалися. Причому на обох стадіях застосовувались методи факторного аналізу.

В якості конкретного об'єкта дослідження інвестиційної привабливості виступали 12 борошномельно-круп'яних підприємств чотирьох областей півдня України, дані яких вивчались за 2004-2008 рр. В результаті проведеного факторного аналізу й відсіву змінних у фінальній факторній моделі інвестиційної привабливості досліджуваних підприємств залишилися три з шести складових 2-го рівня ієрархії: ліквідність, фінансова стійкість, ділова активність, які характеризувалися відповідно трьома, чотирма і п'ятьма первинними чинниками-симптомами 3-го рівня ієрархії (див. рис. 3.4).

Отримані результати оцінювання інвестиційної привабливості 12 борошномельно-круп'яних підприємств дозволили визначити чотири групи об'єктів: рейтинг інвестиційної привабливості досліджуваних підприємств очолював ВАТ «Кулєвчанський КХП», стандартизована оцінка якого складала 1,438. Друге місто посідав ДП «Куліндоровський КХП» (1,155), третє – ВАТ «Новокаховський КХП» (0,947). До підприємств-середняків відносились: ВАТ «Херсонський КХП» (0,676), ВАТ «Калакчацький КХП» (0,081).

Відстаючими підприємствами були: ВАТ «Бєлгород-Дністровский КХП» (-0,031), ВАТ «Миколаївський КХП» (-0,042), ВАТ «Урожайненський КХП» (-0,186), ВАТ «Берєзінський КХП» (-0,284). До групи явних аутсайдерів потрапили ДП «Одеський КХП» (-0,602), ВАТ «Симферопольський КХП» (-0,948). І найгіршим аутсайдером виявилося ВАТ «Чубівський КХП» (-2,203).

Подальший детальний аналіз чинників-симптомів і складових інвестиційної привабливості дозволив виявити, які саме фактори обумовили приналежність кожного з 12 борошномельно-круп'яних підприємств, що вивчалися, до групи лідерів, середняків, відстаючих і аутсайдерів, а також розробити конкретні виробничі заходи для підвищення досліджуваної латентної ознаки.

Окрім Г.В. Кошельок, інвестиційну привабливість підприємств як приховану властивість економічних об'єктів розглядала також І.В. Нападовська (2014 р., науковий керівник А.В. Коверда), яка запропонувала методику згортання окремих чинників-симптомів даної латентної ознаки в інтегральний показник на основі методу головних компонент [63]. За допомогою його інструментарію множину чинників-симптомів, обчислену для 5 підприємств кабельної промисловості за 2009-2013 рр., поетапно зведено до ефективності використання капіталу і ресурсів, що пояснює понад 46 % варіації факторів інвестиційної привабливості, та узагальненого майнового та фінансового стану, що детермінує майже 29 % від її залишку.

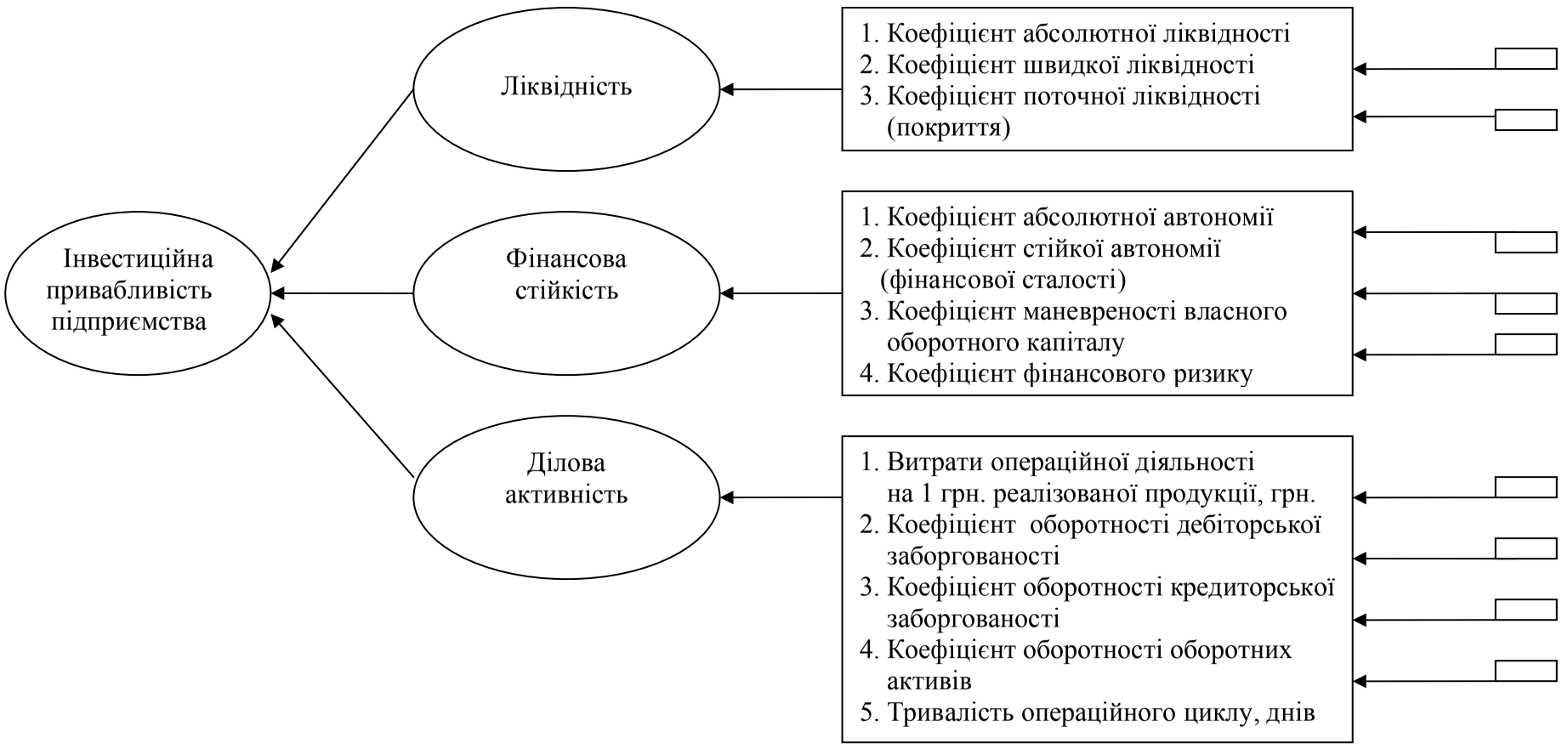


Рис. 3.4 Фінальна структура інтегрального показника інвестиційної привабливості борошномельно-круп'яних підприємств зернопереробного комплексу півдня України

Це науково спростувало орієнтацію більшості методик оцінки інвестиційної привабливості на дані балансу, кількісно підтвердивши корисність інформації звітів про фінансові результати та рух грошових коштів, а також довело перевагу дохідної концепції оцінювання в порівнянні з витратною (майновою). За результатами обчислень здійснено ранжирування найбільших підприємств кабельної промисловості України за рівнем інвестиційної привабливості.

3.5 Оцінка глобального економічного критерію при оптимізації виробничої програми підприємства

В.В. Куперман (2012 р., науковий керівник О.Г. Янковий) в процесі багатокритеріальної оптимізації виробничої програми промислового підприємства з'ясувала, що глобальний економічний критерій є, по суті, латентною ознакою й проявляється на поверхні у вигляді локальних критеріїв економічного ефекту або ефективності (чинників-симптомів) господарської діяльності [46; 141]. Їх роль виконують такі показники метричної шкали відносин, як бухгалтерський прибуток, обсяг реалізації продукції (робіт, послуг), частка певного ринку даного підприємства, рентабельність виробництва та продукції та ін. (рис. 3.5).

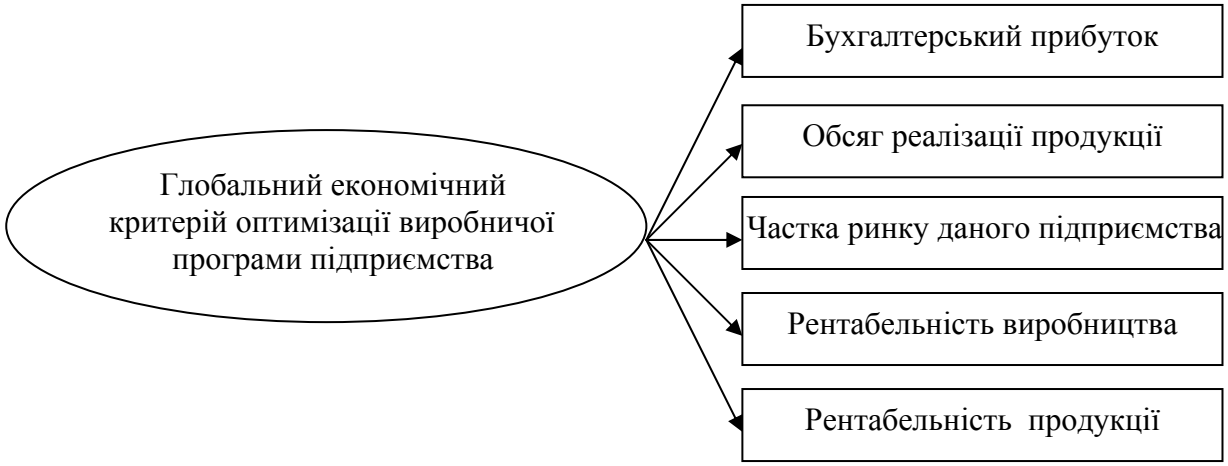


Рис. 3.5. Структура глобального економічного критерію оптимізації виробничої програми підприємства

Однак, слід мати на увазі, що на заваді багатокритеріальної оптимізації виробничої програми підприємства стоїть проблема розв'язання відповідної математичної задачі, оскільки в цьому випадку цільова функція є не скаляром, а вектором і завдання, по суті, зводиться до векторної оптимізації номенклатури та асортименту майбутнього плану виробництва. Ця проблема, зазвичай, розв'язується на основі певної схеми компромісу шляхом пошуку множини Парето-оптимальних рішень, яка є підмножиною допустимої безлічі рішень задачі векторної оптимізації.

На думку автора, одним із дієвих шляхів пошуку області Парето-оптимальних рішень є адитивний метод визначення глобального критерію K на основі локальних критеріїв $K_1, K_2, \ldots, K_S$:

$$K = A_1K_1 + A_2K_2 + \ldots + A_SK_S, \tag{3.1}$$

де $A_1, A_2, \ldots, A_S$ – додатні чи від'ємні вагові коефіцієнти локальних критеріїв.

В результаті підстановки в (3.1) виражень локальних цільових функцій глобальний критерій оптимальності представляється так:

$$\begin{aligned} K = &(A_1c_1 + A_2m_1 + \ldots + A_sp_1)X_1 + (A_1c_2 + A_2m_2 + \ldots + A_sp_2)X_2 + \ldots \\ &+ (A_1c_n + A_2m_n + \ldots + A_Sp_n)X_n \rightarrow \max \ (\min), \end{aligned} \tag{3.2}$$

де $c_i, m_i, \ldots, p_i$ – відомі коефіцієнти локальних цільових функцій.

На практиці величини $A_1, A_2, \ldots, A_S$, зазвичай, знаходяться експертним шляхом з усіма негативними наслідками даного методу. Щоб уникнути вказаних недоліків, В.В. Куперман запропонувала вирішувати проблему оцінки латентного показника «глобальний критерій діяльності підприємства», що використовується в процесі оптимізації виробничої програми суб'єкта господарювання, за допомогою сполучення кластерного і регресійного аналізу.

Запропонований підхід складається з наступних головних етапів:

- теоретичне обґрунтування важливіших локальних критеріїв діяльності підприємства, що детермінують глобальний критерій;
- визначення рішень задач однокритеріальної оптимізації на множині обраних локальних критеріїв діяльності підприємства;
- виділення декількох однорідних груп (кластерів) допустимих планів виробництва, близьких до оптимальних рішень задач однокритеріальної оптимізації;
- ідентифікація виділених груп, ранжирування їх за ступенем впливу на глобальний критерій;
- квантифікація кластерів – присвоєння кожній групі певного значення глобального критерію з урахуванням отриманих ними рангів;
- побудова і аналіз регресійної моделі, що описує залежність присвоєних значень глобального критерію від важливіших локальних критеріїв ефекту (ефективності) діяльності підприємства.

Практична реалізація вказаних етапів запропонованої процедури була успішно апробована за даними бухгалтерської звітності ПП «Гармаш», ПАТ «Одескабель» за 2011 р.

Так, дослідження місії та цілей ПАТ «Одескабель» дозволило висунути гіпотезу про пріоритетність, принаймні, трьох економічних цілей розвитку підприємства, які опосередковані наступними локальними критеріями:

K_1 – прибуток від реалізації продукції, тис. грн.; K_2 – рентабельність реалізованої продукції, %; K_3 – об'єм реалізації продукції, тис. грн.

Перший із них характеризує абсолютний ефект фінансово-економічної діяльності, другий – ефективність витрат виробничих ресурсів, третій – конкурентну позицію ПАТ «Одескабель».

Аналіз результатів проведеної процедури показав, що багатокритеріальна оптимізація виробничої програми підприємства приводить до плану, який повністю співпадає з оптимізованим варіантом випуску продукції за критерієм K_1. Прибуток від реалізації продукції за цим планом є максимальним і складає 63497 тис. грн. Рентабельність реалізованої продукції (13,4 %) близька до середньої величини із рівнів аналогічних показників виробничих програм, оптимальних за локальними критеріями K_1, K_3 (18,8 і 6,6 %). Виручка від реалізації продукції теж приблизно представляє собою середню величину по відношенню до подібних характеристик виробничих планів, знайдених за локальними критеріями K_2, K_3.

Порівняння основних показників фактичної виробничої програми ПАТ «Одескабель» за 2011 р. з показниками оптимізованої програми в розрізі головних асортиментних груп продукції дозволило констатувати, що підприємство мало можливість: 1) знизити собівартість кабельно-провідникової продукції на 18,7 %; 2) збільшити загальний прибуток від реалізації продукції на 54,3 %; 3) підвищити показник рентабельності продукції до 13,4 %.

Указані переваги реально забезпечувались структурними зрушеннями в асортименті товарів, що випускаються на ПАТ «Одескабель», зокрема, підвищенням частки найбільш прибуткових видів продукції – кабелів міських телефонних, волоконно-оптичних, радіочастотних.

3.6 Оцінка інноваційно-інвестиційних складових розвитку підприємства

М.В. Обертайло (2013 р., науковий керівник В.Г. Семенова,) аналізуючи інноваційно-інвестиційні складові розвитку машинобудівних підприємств, досліджувала їх з позиції теорії латентних ознак [65]. На думку автора, для всебічного аналізу інноваційно-інвестиційного розвитку доцільно виділити 7 основних складових латентної ознаки, які, в свою чергу, теж є прихованими властивостями підприємств і визначаються відповідними групами чинників-симптомів (див. рис. 3.6).

В роботі [90] В.Г. Семенова, М.В. Обертайло класифікували всі складові інноваційно-інвестиційного розвитку підприємства на кількісні й якісні. До кількісних вони віднесли фінансову, економічну, науково-технічну та кадрову складову. До якісних – соціально-культурну, організаційно-управлінську і маркетингову складові.

Автори навели перелік абсолютних та відносних чинників-симптомів, що характеризують кількісні латентні складові інноваційно-інвестиційного розвитку машинобудівних підприємств. Наприклад, фінансову складову характеризували коефіцієнт поточної ліквідності, фінансовий левериджс, коефіцієнт автономії. Економічну складову – продуктивність праці 1-го

робітника, рентабельність (збитковість) продукції, витрати на 1 грн. реалізованої продукції. Науково-технічну складову – наукомісткість виробництва, фондоємність, фондоозброєність. І кадрову складову – частка виконавців науково-технічних робіт.

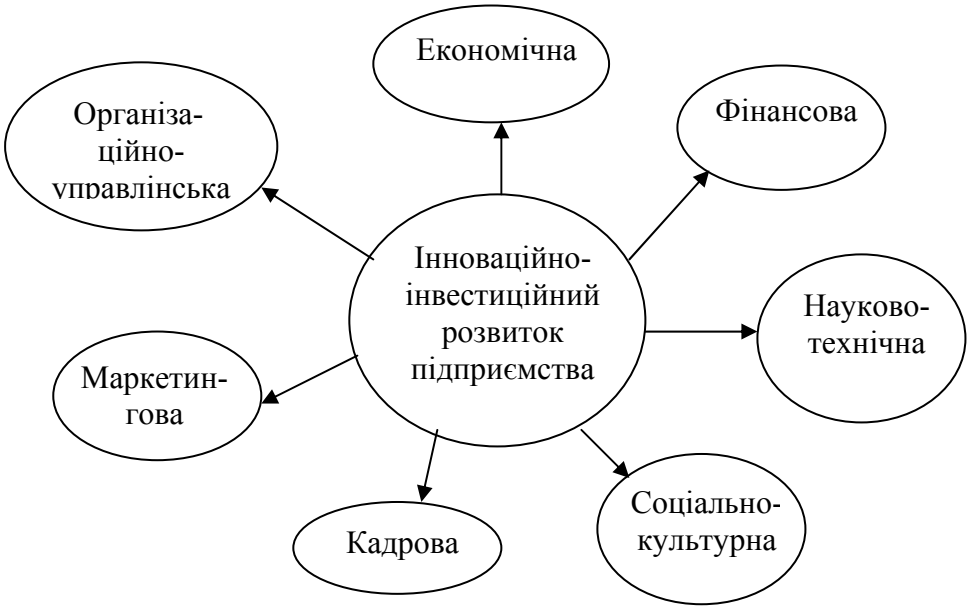


Рис. 3.6. Взаємодія інноваційно-інвестиційних складових розвитку машинобудівного підприємства

В.Г. Семенова, М.В. Обертайло зазначили, що досліджувані приховані властивості підприємств доцільно аналізувати за допомогою багатовимірних статистичних методів, серед яких важливе місце займають метод таксономічного аналізу (на основі метрик відстаней та схожості) та метод головних компонент (як різновид факторного аналізу).

Автори звернули увагу на те, що вищезазначені методи мають свої переваги та недоліки. Тому, на їх думку, для повного та всебічного аналізу й оцінки рівня інноваційно-інвестиційного розвитку групи підприємств необхідно використовувати обидва методи. Це дозволить подолати недоліки та посилити їх позитивні риси.

В результаті їх використання метод таксономії відкрив можливість провести ранжирування групи машинобудівних підприємств за рівнем розвитку кожної кількісної складової та за рівнем інноваційно-інвестиційного розвитку в цілому. Дозволив виділити серед економічних об’єктів, що вивчаються, підприємства-лідери, середняки та аутсайдери шляхом порівняння значень основних обраних чинників-симптомів з еталоном.

Метод головних компонент надав змогу оцінити приховані (латентні) ознаки, що здійснюють суттєвий вплив на рівень інноваційно-інвестиційного розвитку підприємства. За допомогою методу головних компонентів також було проведене ранжирування підприємств, яке повністю співпало з попереднім результатом.

3.7 Оцінка пріоритетності внутрішньовиробничих інвестиційних заходів та загроз на підприємстві

У роботах О.Г. Янкового, Н.В. Мельник, В.О. Янкового, котрі відображають результати економічної оцінки інвестиційних проектів, показано, що така властивість портфеля інвестицій на підприємстві, як «пріоритетність» запланованого виробничо-фінансового заходу є латентною, яка не може бути виміряною в метричних шкалах інтервалів чи відносин. У роботах [142; 143; 150] доведено, що дана ознака проявляється на поверхні у вигляді певних значень важливіших критеріїв проектного аналізу (рис. 3.7).

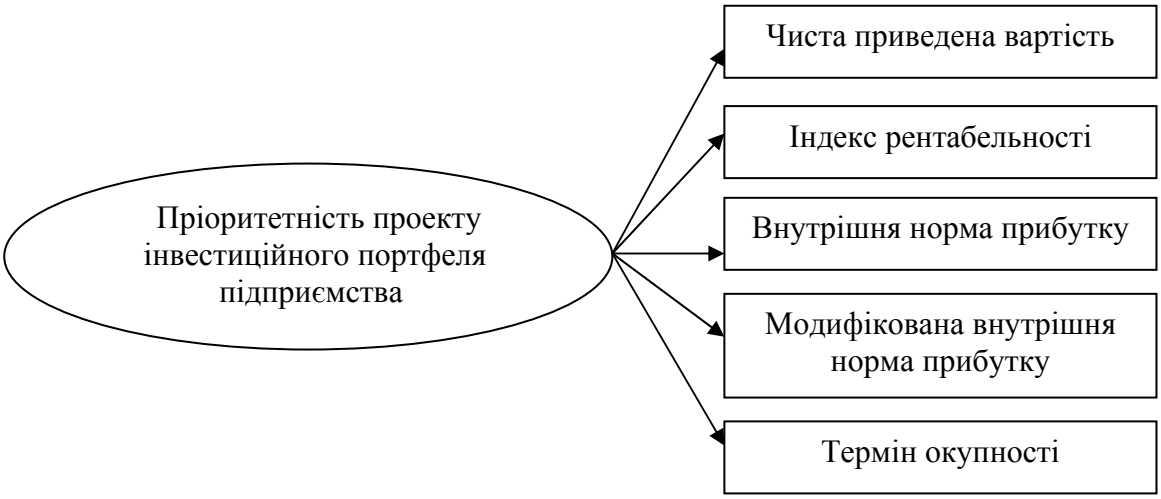


Рис. 3.7. Структура пріоритетності проекту інвестиційного портфеля підприємства

Н.В. Мельник (2013 р., науковий керівник О.Г. Янковий) запропонувала оцінювати пріоритетність проекту інвестиційного портфеля підприємства двома методами: 1) за допомогою класичного алгоритму таксономічного аналізу; 2) на базі методу головних компонент [57]. За кожним методом були знайдені оцінки досліджуваної прихованої властивості проектів і проведено ранжирування запланованих виробничих заходів керівництвом ПрАТ «Одесавинпром» на 2012 р.

Порівняння отриманих результатів показало, що ранги, отримані за методом головних компонент, повністю співпадають з їх ранжируванням за методом таксономії. Це підтвердило гіпотезу про правильність запропонованого методичного підходу до визначення пріоритетності інвестиційних проектів на основі багатовимірних статистичних методів.

Беззаперечним інвестиційним проектом-лідером виявилося облаштування автоматичної системи стабілізації проти випадання винного каменя *DT* (потужністю 400 дал/год), що використовує технології спіральних органічних мембран. На другому місці щодо пріоритетності знаходився захід зі створення інвестиційно-інноваційної науково-технічної групи, у функції якої буде входити моніторинг найкращих вітчизняних та іноземних «ноу-хау» в області

виноробства і формування пропозицій щодо їх впровадження у виробництво. Їх економічні показники характеризуються досить високими значеннями ефективності, резерву надійності та порівняно низьким терміном окупності. Більш прибуткові інвестиції з низькими значеннями вказаних критеріїв опинились усередині та в кінці ранжируваного ряду інвестиційних проектів ПрАТ «Одесавинпром».

О.Ю. Малютенко (2015 р., науковий керівник В.О. Кравченко) в процесі аналізу управління стратегічними ризиками борошномельних підприємств харчової промисловості України довів, що ранжирування ризик-факторів доцільно здійснювати за двома групами чинників: факторів погроз, тобто таких, які призводять до втрат, та факторів втрат від упущеної вигоди суб'єкта господарювання. При цьому сама пріоритетність ризик-факторів підприємства, з одного боку, є основою формування набору напрямків або сценаріїв розвитку підприємства, а, з іншого боку, представляється прихованою властивістю зазначених стратегічних чинників [55].

Автором установлено, що процедуру ранжирування кожного з визначених ризик-факторів підприємства доцільно проводити за трьома показниками: 1) значущість групи чинників; 2) значущість кожного фактора всередині групи; 3) вірогідність появи даного чинника. Він показав, що апарат теорії нечіткої логіки є дієвим інструментом не тільки у вирішенні задачі оцінювання кожного з вказаних показників, а й узгодження думок експертів щодо ризик-факторів підприємства з метою зменшення суб'єктивності майбутніх оцінок.

У ході проведеного дослідження було надано рекомендації задля обрання лінгвістичної змінної, яку формалізовано за термами: (Н) – низький, (НС) – нижче за середнього, (С) – середній, (ВС) – вище за середній, (В) – високий. Сформовано нечітку базу знань, що включає функції приналежності в графічному і математичному вигляді та матриці знань. Розроблений методичний інструментарій дозволив провести за даними ДП «Одеський КХП» вибір найбільш прийнятного сценарію розвитку підприємства, який включає впровадження поточного моніторингу, освоєння експортування борошномельної продукції тощо. Ці напрями розвитку забезпечують максимальний позитивний вплив на фінансові результати діяльності ДП «Одеський КХП» з урахуванням оцінених за допомогою апарата теорії нечіткої логіки стратегічних ризик-факторів.

3.8 Оцінка сталого розвитку підприємств

О.М. Гончаренко (2014 р., науковий консультант О.Г. Янковий) в своєму докторському дослідженні підприємств різних галузей промисловості України дійшла висновку, що рівень їх сталого розвитку представляє собою безпосередньо невимірну латентну ознаку, яка проявляється у вигляді безлічі показників (чинників-симптомів) [17]. Вона запропонувала при оцінюванні рівню сталого розвитку підприємства враховувати три приховані складові: Φ_1 – фінансовий стан; Φ_2 – виробнича діяльність; Φ_3 – інноваційно-інвестиційна

діяльність (рис. 3.8). При цьому автор детально обґрунтувала перелік економічних показників метричних шкал, що характеризують кожну з латентних складових Φ_1, Φ_2, Φ_3.

Так, О.М. Гончаренко зазначила, що фінансовий стан підприємства Φ_1 проявляється в конкретних значеннях коефіцієнтів швидкої ліквідності, рентабельності поточних активів, рентабельності власного капіталу, коефіцієнта чистого прибутку, коефіцієнта виробничої собівартості, рентабельності залученого капіталу, коефіцієнта платоспроможності та ін.

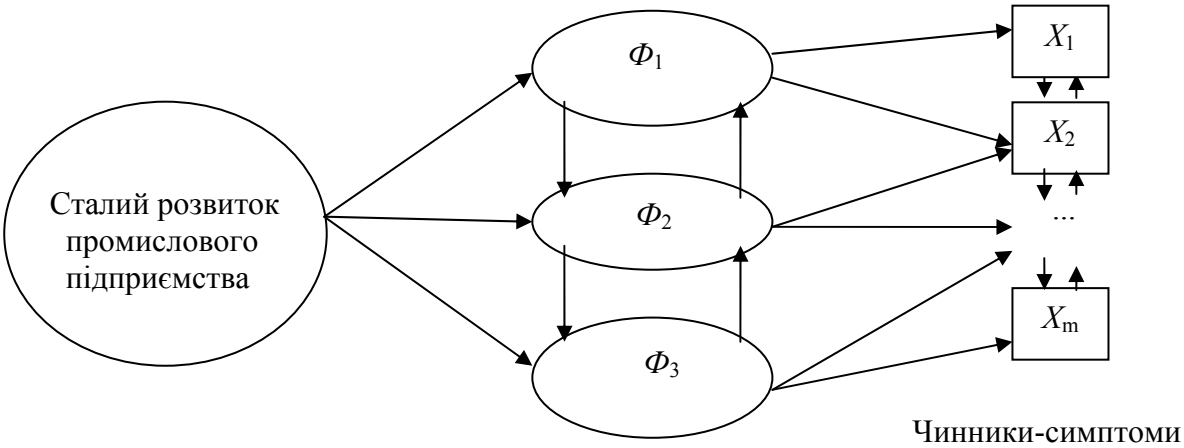


Рис. 3.8. Складові й чинники-симптоми сталого розвитку промислового підприємства

За допомогою запропонованого автором двох стадійного класичного алгоритму таксономії була проведена оцінка стійкого розвитку підприємств машинобудування Одеської області (табл. 3.1).

Для досліджуваних підприємств машинобудування знайдені в результаті таксономічного аналізу значення схожості з еталоном μ_i розглядались в якості шуканого латентного показника «рівень стійкого розвитку підприємства».

Таблиця 3.1

Результати оцінювання стійкого розвитку підприємств машинобудування Одеської області за 2013 р.

№	Підприємства	μ_i	Ранг
1	ПАТ «Одескабель»	0,3205	1
2	ТОВ «Холдингова компанія «МІКРОН»	0,2358	3
3	ПрАТ «ВО «Стальканат-силур»	0,3006	2
4	ПАТ «Одеський завод з випуску ковальсько-пресових автоматів»	0,1814	5
5	ПрАТ «Будгідравлика»	0,1825	4

Аналіз даних табл. 3.1 показав, що найближче за своїми економічними складовими до еталону знаходилось ПАТ «Одескабель», що мало максимальну з ним схожість. Друге місце за рівнем стійкого розвитку займало ПрАТ «ВО «Стальканат-Сілур». Це машинобудівні підприємства-лідери. До аутсайдерів

належали підприємства, що вирізнялись найменшою схожістю з еталоном.

Практичне застосування отриманих результатів полягає в тім, що знання лідерів і аутсайдерів у сучасних умовах набуває особливої актуальності в зв'язку з постійною необхідністю вибору найбільш перспективних партнерів, визначення оцінок розвитку самого підприємства, пошуку оптимальної кон'юнктурної позиції в бурхливому зовнішньоекономічному середовищі.

Рівень сталості розвитку енергетичного господарства та його складових (соціальної, екологічної, економічної) бюджетних та комерційних підприємств розглядався у дослідженні Г.О. Пудечевої (2015 р., науковий керівник О.О. Дегтярьова) з позиції теорії латентних ознак [81]. Задля їх оцінювання використовувалась двох етапна процедура метода головних компонентів.

Оцінки були здійснені для енергетичних господарств загальноосвітніх шкіл трьох районів Одеської області за 2010-2012 рр. Їх ранжирування за значеннями генеральної компоненти дозволило виділити групу лідерів і аутсайдерів. Автором запропоновано систему заходів, спрямованих на забезпечення сталого розвитку енергетичних господарств шкіл-аутсайдерів за основними групами чинників-симптомів – організаційними, соціальними, матеріально-технічними.

3.9 Оцінка рівня розвитку підприємств

О.В. Шикіна (2015 р., науковий керівник С.С. Галасюк) у процесі дослідження організаційно-економічного забезпечення готельних підприємств малої місткості дійшла висновку, що їх функціонування є досить складним економічним явищем, рівень розвитку якого представляє собою латентну ознаку [112]. Тому автором для її оцінки з метою прийняття раціональних управлінських рішень запропоновано імплементувати в процедуру діагностики, формування і ранжирування готельних підприємств метод таксономічного аналізу за класичним, модифікованим і об'єднаним алгоритмами.

Застосування цього методу сприяло виділення з 91 готельного підприємства малої місткості м. Одеси чотирьох груп: лідерів, послідовників лідерів, середняків, аутсайдерів. При цьому враховувались такі організаційно-економічні показники готельного бізнесу, як ціна номеру, категорія і вік готелю, кількість номерів, наявність додаткових послуг, сприйняття готелю споживачем, віддаленість від моря та ін. (табл. 3.2)

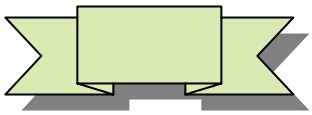
Таблиця 3.2

Структура групи готельних підприємств малої місткості м. Одеси за рівнем розвитку

№	Групи готельних підприємств	Кількість	Питома вага, %	Інтервал рівня розвитку	Ранги
1	Лідери	10	11,0	0,7497 – 0,6520	1 – 10
2	Послідовники лідерів	15	16,5	0,6434 – 0,5615	11 – 25
3	Середняки	44	48,3	0,5484 – 0,4503	26 – 69
4	Аутсайдери	22	24,2	0,4495 – 0,2923	70 – 91
Разом		91	100,0	-	-

Дані табл. 3.2 дозволили автору підготувати рекомендації щодо подальшого розвитку ринку послуг готельного бізнесу м. Одеси. Так, для групи лідерів запропонована стратегія, націлена на захист свого положення за допомогою додаткових інвестицій, високопрофесійних дій персоналу, надання якісних, у тому числі додаткових послуг. Стратегічним завданням послідовників лідерів має бути постійний моніторинг своїх слабких і сильних сторін, застосування методів боротьби за лідерство, збільшення частки на ринку готельних послуг до максимуму, організація філій.

Підприємства-середняки зацікавлені вступати в об'єднання типу асоціації, кластера, мережі, що пояснюється їх бажанням не тільки позиціонувати унікальність власного готельних послуг на ринку, але й користатися конкурентними перевагами інтегрування з подібними підприємствами. Група аутсайдерів приречена перманентно захищати свої ринкові позиції, концентруючи інвестиції у ті сегменти, де є суттєва норма прибутку і відносно невеликий ризик. Підприємства даної групи характеризуються нестабільністю й нестійкими показниками розвитку, залежністю від загальної політики та впливу держави на малий бізнес, але можуть розглядатись як потенційні учасники асоціації малих готелів міста.



Таким чином, з оглядом на перелік наукових публікацій, дисертаційних робіт, захищених за період 1993-2015 рр. під безпосереднім керівництвом професора О.Г. Янкового, під керівництвом провідних викладачів кафедри економіки підприємства, а також інших кафедр Одеського національного економічного університету, можна казати про наявність у даному вищому навчальному закладі наукової школи дослідження латентних ознак економічних об'єктів.

Як показала практика, домінуючою складовою всіх наведених розробок є оцінка прихованих властивостей підприємств, що вивчалися, без якої стає неможливими детальний аналіз факторів, резервів і прогнозів на майбутнє важливіших латентних ознак сучасного виробництва продукції, виконання робіт та надання послуг. До них відносяться, зокрема, рівні деяких трудових показників (продуктивної сили та інтенсивності праці, ефективності праці службовців), фінансового стану, конкурентоспроможності, інвестиційної привабливості, сталого та інвестиційно-інноваційного розвитку підприємств, а також конкурентоздатності продукції, глобального економічного критерію оптимальності виробничої програми, пріоритетності внутрішньовиробничих інвестиційних заходів та загроз на підприємстві та ін.

Ми сподіваємось, що даний напрямок економічної науки буде й надалі успішно розвиватись, залучаючи в ряди своїх послідовників нових молодих дослідників прихованих властивостей економічних об'єктів. А перелік останніх буде ще більше розширюватись за рахунок проникнення ідей і методів оцінювання латентних ознак у всі сфери і галузі економіки.

ЛІТЕРАТУРА

1. Астапенко Н. В., Мяздриков А. Н. Разработка экспертной системы выбора наиболее предпочтительных поставщиков на базе методов теории нечетких множеств [Электронный ресурс] – Режим доступа : http://www.rusnauka.com/13_EISN_2012/Economics/8_109465.doc.htm
2. Бабій О. М., Зайцева Н. О. Оцінка рівня та резервів підвищення конкурентоспроможності промислових підприємств / Конкурентоспроможність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 86-113.
3. Баканов М. И. , Шеремет А. Д. Теория экономического анализа: Учебник. – М.: Финансы и статистика, 1996. – 288 с.
4. Баумгартен Л. В. Анализ методов определения конкурентоспособности организаций и продукции / Л. В. Баумгартен // Маркетинг в России и за рубежом. – 2005. – № 4(48). – С. 24-29.
5. Бирман И. Я. Оптимальное программирование. – М.: Экономика, 1968. – 232 с.
6. Благуш П. Факторный анализ с обобщениями; пер. с чешск. / Благуш П. – М.: Финансы и статистика, 1989. – 248 с.
7. Болч Б., Хуань К. Дж. Многомерные статистические методы для экономики. Пер. с англ. – М.: Статистика, 1979. – 317 с.
8. Бондаренко А. Д. Мировые рынки ресурсов : [учебник] / Бондаренко А. Д. – Донецк : ДонГУ, 1998. – 378 с.
9. Борисов А. Н., Алексеев А. В., Меркурьев Г. В. и др. Обработка нечеткой информации в системах принятия решений. – М.: Радио и связь, 1989. – 304 с.
10. Борисов А. Н., Крумберг О. А., Федоров И. П. Принятие решения на основе нечетких моделей: примеры использования. – Рига «Знание», 1990. – 184 с.
11. Боровиков В. П. *STATISTICA* – Статистический анализ и обработка данных в среде *Windows* / В. П. Боровиков, И. П. Боровиков. – М.: ИИД «Филинъ», 1998. – 608 с.
12. Боровиков В. П. Популярное введение в программу *STATISTICA* / В. П. Боровиков – М. : КомпьютерПресс, 1998. – 267 с.
13. Боровиков В. П., Ивченко Г. И. Прогнозирование в системе *STATISTICA* в среде *Windows*. – М.: Финансы и статистика, 1999. – 384 с.
14. Бююль А., Цёфель П. SPSS : искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. – СПб.: ООО «ДиаСофтЮП», 2002. – 608 с.
15. Великоіваненко Г. І., Мамонова К. М. Ієрархічна модель оцінки інвестиційної привабливості підприємства / Г. І. Великоіваненко, К. М. Мамонова // Моделювання та інформаційні системи в економіці. – 2007. – Вип. 76. – С. 48-61.

16. Вініченко І. І. Удосконалення методики оцінки інвестиційної привабливості аграрних підприємств / І. І. Вініченко // “Інвестиції : практика та досвід” [Електронний ресурс]. – Режим доступу : http://www.investplan.com.ua/index.php?iid=29&operation=9
17. Гончаренко О. М. Формування механізму сталого розвитку підприємства в умовах невизначеності економічного середовища / Автореф. дис. на здобуття наук. ступеня доктора екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / О. М. Гончаренко. – Одеса, ОНЕУ, 2014. – 39 с.
18. Гончарук А. Г., Гамма Т. М., Молла М. Г. Конкурентоспроможність молокопереробних підприємств : оцінка та шляхи підвищення / Конкурентоспроможність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 114-141.
19. Горбачик А.П. Руководство пользователя OCA for Windows™ [Электронный ресурс] – Режим доступа : http://oca.com.ua/arc/oca_manual.pdf
20. Дилегенский Н. В. Нечеткое моделирование и многокритериальная оптимизация производственных систем в условиях неопределенности : технология, экономика, экология / Н. В. Дилегенский. − М.: Машиностроение-1, 2004. − 397 с.
21. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Пер. с англ. – М.: Финансы и статистика, Кн. 1, 1986. – 366 с.
22. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Пер. с англ. – М.: Финансы и статистика, Кн. 2, 1987. – 351 с.
23. Дубенко Е. А. Латентность показателя конкурентоспособности предприятия / Матер. Междунар. науч.-практ. конф. «Актуальные проблемы современных наук: теория и практика» (01.06.05 г.-08.06.05 г.) [Электронный ресурс]. – Режим доступа : http://rusnauka.com
24. Дубров А. М. Многомерные статистические методы / А. М. Дубров, А. С. Мхитарян, Л. И. Трошин. – М.: Финансы и статистика, 2000. – 352 с.
25. Дубров А. М. Обработка статистических данных методом главных компонент / Дубров А. М. – М.: Статистика, 1978. – 135 с.
26. Дюран Б. Кластерный анализ / Б. Дюран, П. Оделл; пер. с англ. – М.: Статистика, 1977. – 128 с.
27. Енюков И. С. Методы, алгоритмы, программы многомерного статистического анализа / Енюков И. С. – М.: Финансы и статистика, 1986. – 232 с.
28. Єгупов Ю. А. Таксономічний аналіз як інструмент інтерактивної оптимізації виробничої програми м’ясопереробного підприємства / Ю. А. Єгупов. – Вісник Запорізького національного університету : Збірник наукових праць. Економічні науки. – 2011. – № 4. – С. 113-118.
29. Єгупов Ю.А., Літвінова В.О. Оцінка порівняльної ефективності шляхів підвищення конкурентоспроможності продукції / Конкурентоспромож-

ність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 141-159.
30. Єргієва Л. В. Конкурентоспроможність підприємств кабельної промисловості : оцінка рівня та резервів росту / Конкурентоспроможність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 66-85.
31. Єргієва Л. В. Фактори та резерви підвищення конкурентоспроможності промислових підприємств (на прикладі кабельної промисловості) / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / Л. В. Єргієва. – Одеса, ОНЕУ, 2012. – 20 с.
32. Жуковская В. М. Факторный анализ в социально-экономических исследованиях / В. М. Жуковская, И. Б. Мучник. – М.: Статистика, 1976. – 151 с.
33. Заде Л. А. Основы нового подхода к анализу сложных систем и процессов принятия решений / Л. А. Заде // Математика сегодня. – М.: Знание, 1974. – 549 с.
34. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. А. Заде. – М.: Мир, 1976. – 165 с.
35. Зайцева Н. О. Економічне вимірювання резервів підвищення конкурентоспроможності промислових підприємств харчової промисловості / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / Н. О. Зайцева. – Одеса, ОНЕУ, 2013. – 20 с.
36. Захожай В. Статистика інвестиційної діяльності / В. Захожай, М. Кіт // Персонал. – 2007. – № 8 [Електронний ресурс]. – Режим доступу : http://personal.in.ua/article.php?ida=548
37. Иберла К. Факторный анализ / К. Иберла; пер. с нем. – М.: Статистика, 1980. – 398 с.
38. Інвестування : підручн. / [Гриньова В. М., Коюда В. О., Лепейко Т. І., Коюда О. П.]. – К.: Знання, 2008. – 452 с.
39. Ким Дж.-О. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка; пер. с англ. – М.: Финансы и статистика, 1989. – 215 с.
40. Клигер С. А. Шкалирование при сборе и анализе социологической информации [Текст] : монография / С. А. Клигер, М. С. Косолапов, Ю. Н. Толстова. – М.: Наука, 1978. – 112 с.
41. Кофман А. Введение в теорию нечетких множеств. – М.: Радио и связь, 1982. – 432 с.
42. Кошельок Г. В. Внутрішні фактори та резерви росту інвестиційної привабливості промислових підприємств / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / Г. В. Кошельок. – Одеса, ОДЕУ, 2010. – 21 с.

43. Кошельок Г. В. Система факторів, які впливають на оцінку інвестиційної привабливості підприємств харчової промисловості / Г. В. Кошельок // Вісник соціально-економічних досліджень. – ОДЕУ, 2009, № 36. – С. 98-104.
44. Кошельок Г. В. Теоретичні основи оцінки фінансового стану підприємств харчової промисловості / Янковий О. Г., Кошельок Г. В. // Наукові праці ОНАХТ. – Одеса, 2007. Вип. 30. Т. 2. – С. 367-371.
45. Кротков А. М. Конкурентоспособность предприятия : подходы к обеспечению, критерии, методы оценки / А. М. Кротков, Ю. А. Еленева // Маркетинг в России и за рубежом. – 2001. – № 6 (28). – С. 59-68.
46. Куперман В. В. Оптимізація виробничої програми промислового підприємства / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / В. В. Куперман. – Одеса, ОНЕУ, 2012. – 19 с.
47. Лазарсфельд П. Латентно-структурный анализ и теория тестов // Математические методы в социальных науках. М.: Прогресс, 1973. – С. 42-54.
48. Лазарсфельд П. Логические и математические основания латентно-структурного анализа / П. Лазарсфельд // Математические методы в современной буржуазной социологии. – М., 1966. – С. 344-401.
49. Лекции по статистике. Предмет статистики [Электронный ресурс] – Режим доступа : http://bib.convdocs.org/v9667
50. Леоненков А. В. Нечеткое моделирование в среде MATLAB и Fuzzy TECH / А. В. Леоненков. − С.-Пб.: БХВ-Петербург, 2005. − 736 с.
51. Літвінова В. О. Механізм обґрунтування ефективних шляхів підвищення конкурентоспроможності продукції / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / В. О. Літвінова. – Одеса, ОНЕУ, 2012. – 20 с.
52. Мазорчук М.С., Добряк В.С., Базилевич К.А., Хазай М.Ю. Оценка эффективности одномерных шкал при анализе данных в социологических исследованиях / Радіоелектронні і комп'ютерні системи. – 2013. – № 1 (60). – С. 99-104.
53. Максимов И. Оценка конкурентоспособности промышленного предприятия / И. Максимов // Маркетинг. – 1996. – № 3. – С. 33-39.
54. Малаева Т. Оценка конкурентоспособности фирмы / Т. Малаева // Бизнес-информ. – 1998. – № 17/18. – С. 115-117.
55. Малютенко О. Ю. Управління стратегічними ризиками борошномельних підприємств / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / О. Ю. Малютенко. – Одеса, ОНЕУ, 2015. – 20 с.
56. Мандель И. Д. Кластерный анализ / И. Д. Мандель – М.: Финансы и статистика, 1988. – 176с.

57. Мельник Н. В. Оцінка реальних інвестицій на промисловому підприємстві / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / Н. В. Мельник. – Одеса, ОНЕУ, 2013. – 20 с.
58. Методика інтегральної оцінки інвестиційної привабливості підприємств та організацій, затв. наказом Агентства з питань запобігання банкрутству від 23.02.98 № 22 // Державний інформаційний бюлетень про приватизацію. – 1998. – № 7. – С. 5-16.
59. Многомерный статистический анализ в экономике / Под ред. проф. В. Н. Тамашевича. – М.: ЮНИТИ-ДАНА, 1999. – 598 с.
60. Молла М. Г. Організаційно-економічні складові забезпечення конкурентоспроможності підприємств / Дисерт. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / М. Г. Молла. – Одеса, ОНЕУ, 2014. – 247 с.
61. Молла М. Г. Оцінка конкурентоспроможності підприємств на основі економіко-статистичних методів / М. Г. Молла // Економіка та управління національним господарством : стан, тенденції та перспективи : матер. першої міжнар. наук.-практ. конф. 29-30 трав. 3014 р. – Одеса, ОНЕУ, 2014, С. 112-114.
62. Надтока Т. Б., Виноградов А. Г. Многомерное оценивание уровня социально-экономического развития предприятия / Т. Б. Надтока, А. Г. Виноградов // БізнесІнформ. – 2014. – № 1. – С. 184-191.
63. Нападовська І. В. Економічний аналіз та аудит оцінювання потенційних об’єктів інвестування / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.09 – бухгалтерський облік, аналіз та аудит (за видами економічної діяльності) / І. В. Нападовська. – Одеса, ОНЕУ, 2014. – 19 с.
64. Недосекин А.О. Применение теории нечетких множеств к задачам управления фінансами [Электронный ресурс] – Режим доступа : http://masters.donntu.edu.ua/2005/kita/sroka/library/art8.htm
65. Обертайло М. В. Складові інноваційно-інвестиційного розвитку підприємств (на прикладі машинобудівної галузі) / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / М. В. Обертайло. – Одеса, ОНЕУ, 2013. – 20 с.
66. Общая теория статистики и математическая статистика. Курс лекций / Сост. Мизина Е. В. – Донецк : ДонГТУ, 2001. – С. 3-25.
67. Орлов А. И. Прикладная статистика / А. И. Орлов. – М.: Издательство «Экзамен», 2004. – 656 с.
68. Осипов Г. В., Андреев Э. П. Методы измерения в социологии. – М.: Наука. – 1977. – 184 с.
69. Осипов Г. В. Методы измерения в социологии. – М.: Наука, 2003 [Электронный ресурс] – Режим доступа : http://soc.univ.kiev.ua/lib/pub/o/osipov/osipov.pdf

70. Основные правила образования групп и интервалов группировки [Электронный ресурс] – Режим доступа : http://www.pravo.vuzlib.org/book_z379_page_13.html
71. Основы метрологии [Электронный ресурс] – Режим доступа : http://bib.convdocs.org/v1428
72. Паниотто В. И., Максименко В. С. Измерение и анализ распределений [Электронный ресурс] – Режим доступа : http://kiis.com.ua/materials/books/376072_C6170_paniotto_v_i_maksimenko_v_s_kolichestvennye_metody_v_sociolo.pdf
73. Панфилова Т. С. Методологические основы оценки конкуренто-способности предприятий / Т. С. Панфилова, Е. Г. Норова // Менеджер. – 2002. – № 3 (19). – С. 32-38.
74. Петренко О.О. Методичний підхід до визначення адаптивності системи професійного навчання працівників [Електронний ресурс]. – Режим доступу : http://ir.kneu.edu.ua:8080/bitstream/2010/4925/1/187%20-%20191.pdf
75. Плюта В. Сравнительный многомерный анализ в эконометрическом моделировании / В. Плюта; пер. с пол. В. В. Иванова. – Москва: Финансы и статистика, 1989. – 175 с.
76. Плюта В. Сравнительный многомерный анализ в экономических исследованиях : методы таксономии и факторного анализа / Плюта В.; пер. с пол. В. В. Иванова; под ред. В. М. Жуковской. – М.: Статистика, 1980. – 151 с.
77. Погорєлова Т. В. Методи оцінки ефективності праці службовців на промислових підприємствах / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.05 – економіка, планування, організація управління народним господарством і його галузями, 08.00.11 – статистика / Т. В. Погорєлова. – Одеса, ОДЕУ, 1993. – 17 с.
78. Попов О. А. Измерение в психологии. Шкалы измерения [Электронный ресурс] – Режим доступа : http://psystat.at.ua/publ/1-1-0-28
79. Поспелов Д. А. Нечеткие множества в моделях управления и искусственного интеллекта. – М.: Наука, 1986. – 312 с.
80. Пріб К.А. Інвестування : навч. посібник для дистанційного навчання / К. А. Пріб [за наук. ред. М. Ю. Коденської]. – К. Університет «Україна», 2007. – 259 с.
81. Пудичева Г. О. Управління енергетичним господарством бюджетних та комерційних підприємств на засадах контролінгу / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / Г. О. Пудичева. – Одеса, ОНЕУ, 2015. – 20 с.
82. Рабочая книга по прогнозированию / Под ред. И. В. Бестужева-Лады. – М.: Мысль, 1982. – 430 с.

83. Решетникова Т. П. Методика оценки конкурентоспособности предприятия / Т. П. Решетникова // Математические методы в экономике. – 2001. – № 3-4. – С. 74-82.
84. Решетникова Т. П. Методы оценки конкурентоспособности предприятия / Т. П. Решетникова // Вестник Нижегородского ун-та. – 2001. – № 9. – С. 166-173.
85. Русак Н. А. Финансовый анализ субъекта хозяйствования: справочное пособие / Русак Н. А., Русак В. А. – М.: Высшая школа, 1997. – 309 с.
86. Савчук С. И. Измерение интенсивности конкуренции на уровне предприятий / С. И. Савчук // Актуальні проблеми економіки. – 2005. – № 2 (44). – С. 39-45.
87. Савчук С. И. Основы теории конкурентоспособности / Савчук С. И. – Мариуполь, 2007. – 520 с.
88. Садеков А. А. Инвестиционная привлекательность предприятия (методология и методика оценки) : Монография / А. А. Садеков, Н. А. Лисова. – Донецк: ДонГУЭТ, 2001. – 270 с.
89. Своробович Л. Н., Єгупов Ю. А. Методические основы формирования системы материального поощрения на машиностроительных предприятиях / Л. Н. Своробович, Ю. А. Єгупов. – Науковий вісник. Од. держ. екон. ун-т. Всеукр. асоц. молод. наук. – Науки : економіка, політологія, історія. – 2010. – № 10 (111). – С. 210-223.
90. Семенова В. Г., Обертайло М. В. Інноваційно-інвестиційний розвиток – основа конкурентоспроможності промислових підприємств / Конкурентоспроможність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 211-235.
91. Сошникова Л. А. Многомерный статистический анализ в экономике : [учеб. пособ.] / [Л. А. Сошникова, В. Н. Тамашевич, Г. Уебе, М. Шефер]. – М.: ЮНИТИ, 1999. – 593 с.
92. Статистические методы анализа информации в социологических исследованиях / [Отв. ред. Г. В. Осипов и др.]. – М.: Наука, 1979. – 319 с.
93. Статистичний показник [Електронний ресурс] – Режим доступу : http://ebooktime.net/book_82_glava_5_1.3.
94. Статистический словарь / Гл. ред. М. А. Королёв. – М.: Финансы и статистика. – 1989. – 623 с.
95. Стивенс С. С. Математика, измерение и психофизика // Экспериментальная психология / Под ред. С. С. Стивенса / Пер. с англ. под ред. действ. чл. АМН СССР П. К. Анахина, докт. пед. наук В. А. Артемова. – М.: Иностранная литература, 1960, – Т. 1. – С. 19-92.
96. Суппес П., Зинес Дж. Основы теории измерений / П. Суппес, Дж. Зинес. – В сб.: Психологические измерения. – М.: Мир, 1967. - С. 9-110.
97. Суть та різновиди статистичних показників [Електронний ресурс] – Режим доступу : http://lib.chdu.edu.ua/pdf/posibnuku/7/19.pdf

98. Таксономия (теория классификации) [Электронный ресурс] – Режим доступа : http://bse.sci-lib.com/article108590.html
99. Тимчев М., Логинова Н. Модели прогнозирования кризисов в микроэкономических системах [Электронный ресурс] – Режим доступа : http://eztuir.ztu.edu.ua/770/1/370.pdf
100. Фатхутдинов Р. А. Конкурентоспособность : экономика, стратегия, управление: учеб. пособ. / Фатхутдинов Р. А. – М.: ИНФРА, 2000. – 312 с.
101. Фишберн П. Теория полезности для принятия решений / П. Фишберн. – М.: Наука, 1978. – 352 с.
102. Харман Г. Современный факторный анализ / Харман Г.; пер. с англ.– М.: Статистика, 1972. – 486 с.
103. Цилюрик С. В. Латентні ознаки формування міграційного потенціалу в сучасній Україні / С. В. Цилюрик // Міжнародний науковий форум : соціологія, психологія, педагогіка, менеджмент. – 2013. – Вип. 12. – С. 34-40.
104. Чернишова О. Б. Оцінка конкурентоспроможності промислових підприємств (на прикладі підприємств кондитерської галузі харчової промисловості України) / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / О. Б. Чернишова. – Одеса, ОДЕУ, 2009. – 21 с.
105. Чернишова О. Б. Сутність та структура поняття конкурентоспромож-ності підприємства / Конкурентоспроможність підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 6-21.
106. Чернов В. Г., Дорохов А. В., Градусов Д. А., Уланов Е. А. Использование нечетких множеств для классификации банковских ссуд физических и юридических лиц. / Управления проектами, системный анализ и логистика. – К.: НТУ – 2012. – Вып. 10. – С. 597-602.
107. Чорний А. Ю. Вітчизняний досвід моделювання латентних економічних категорій : індекс задоволеності споживачів / А. Ю. Чорний // Актуальні проблеми економіки. – 2009. – № 8. – С. 273-278.
108. Чорний А. Ю. Моделювання невідчутних активів як латентних показників / А. Ю. Чорний // Економічний аналіз. – 2011. – № 8. – С. 373-376.
109. Чорний А. Ю. Статистичне оцінювання конкурентоспроможності підприємств як латентного показника / А. Ю. Чорний // Теоретичні та прикладні аспекти економіки. Вип. 26 [Електронний ресурс]. – Режим доступу : http://tppe.econom.univ.kiev.ua/data/2011_26/zb26_28.pdf
110. Чугумбаев Р. Р. Модель комплексной сравнительной рейтинговой оценки финансового состояния предприятий-эмитентов по методу расстояний. Преимущества и недостатки методических положений

[Электронный ресурс] : – Режим доступа : http://web:http://www. socionet.ru
111. Шеремет А. Д., Сайфулин Р. С. Финансы предприятий / А. Д. Шеремет, Р. С. Сайфулин. – М.: ИНФРА-М, 1998. – 343 с.
112. Шикіна О. В. Організаційно-економічне забезпечення функціонування готельних підприємств малої місткості / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.00.04 – економіка та управління підприємствами (за видами економічної діяльності) / О. В. Шикіна. – Одеса, ОНЕУ, 2015. – 20 с.
113. Шкардун В. Д. Методика исследования конкуренции на рынке / В. Д. Шкардун, Т. М. Ахтямов // Маркетинг в России и за рубежом. – 2000. – № 4. – С. 44-54.
114. Штовба С. Д. Введение в теорию нечетких множеств и нечеткую логику [Электронный ресурс] – Режим доступа : http://matlab.exponenta.ru/fuzzylogic/book1
115. Юр'єва Г. В. Оцінка фінансового стану промислових підприємств – багатовимірний підхід / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.04.01 – фінанси, грошовий обіг та кредит / Г. В. Юр'єва. – Одеса, ОДЕУ, 2006. – 19 с.
116. Ядов В. А. Методы социологического исследования [Электронный ресурс] – Режим доступа : http://djerelo.com/sociology/47-metodu-socuologicheskogo-issledavania- iadov/2794-19-obschaya-harakteystyka-shkal
117. Ядов В. А. Стратегия социологического исследования. Описание, объяснение, понимание социальной реальности [Текст]: Университетский учебник / В. А. Ядов. – 3-е изд., испр. – М.: Омега-Л, 2007. – 567 с.
118. Янковий В. О. Класифікація ознак економічних об'єктів / В. О. Янковий / Сучасні технології управління підприємством та можливості використання інформаційних систем : стан, проблеми, перспективи : матеріали IX-ї міжнар. наук.-практ. конф. 28-99 бер. 2014 р. – Одеса : видавець Букаєв В.П., 2014. – С. 133-135.
119. Янковий В. О. Шкали вимірювання економічних ознак / В. О. Янковий / Економіка підприємства : сучасні проблеми теорії та практики : матеріали 3-ї міжнар. наук.-практ. конф. 18-19 вер. 2014 р. – Одеса : Атлант, 2014. – С. 130-132.
120. Янковий О. Г. Дрейф ознак маркетингового дослідження вдовж шкал вимірювання / О. Г. Янковий / Маркетинг і цифрові технології. – Матер. I міжнар. наук.-практ. конф. 29-30 трав. 2014 р. – Одеса, ТЕС, 2014. – С. 224-226.
121. Янковой А. Г. К вопросу измерения и проверки мультиколлинеарности / А. Г. Янковой. Вероятностно-статистические методы в экономико-математическом моделировании. – М.: Сб. ЦЕМИ, 1988, С. 49-58.
122. Янковий О. Г. Маркетинговий інструментарій статистичної оцінки конкурентоспроможності підприємств / О. Г. Янковий – Економіка : реалії часу. Науковий журнал. – 2014. – № 4 (14). – С. 63-72 [Електронний ресурс] : – Режим доступу :

http://economics.opu.ua/files /archive/2014/n1.html
123. Янковой А. Г. Математико-статистические методы и модели в управлении предприятием / А. Г. Янковой : Учебн. пособие. – Одесса : ОНЭУ, ротапринт, 2014. – 250 с.
124. Янковий О. Г. Методи оцінки факторів та резервів підвищення продуктивності праці у промисловості України / Автореф. дис. на здобуття наук. ступеня доктора екон. наук за спец. 08.00.05 – економіка, планування, організація управління народним господарством і його галузями, 08.00.11 – статистика / О. Г. Янковий. – Одеса, ОДЕУ, 1993. – 39 с.
125. Янковой А. Г. Многомерный анализ в системе STATISTICA / А. Г. Янковой. – Вып. 1. – Одесса: Оптимум, 2001.– 216 с.
126. Янковой А. Г. Многомерный анализ в системе STATISTICA / А. Г. Янковой. – Вып. 2. – Одесса : Оптимум, 2002. – 325 с.
127. Янковой О. Г. Моделювання парних зв'язків в економіці / О. Г. Янковой // Навч. посібник. – Одеса : Оптимум, 2001. – 198 с.
128. Янковой А. Г. Основы теории статистики / А. Г. Янковий : Учеб. пособие. – Одесса, ОИФ УГУФМТ, 2007. – 111 с.
129. Янковой А. Г. Основы эконометрического моделирования / А. Г. Янковой : Учеб. пособие. – Одесса, ОГЭУ, 2006. – 133 с.
130. Янковой А. Г. Оценка латентных показателей с помощью таксономического анализа / А. Г. Янковой // Ринкова економіка : сучасна теорія і практика управління, т. 11, вип. 8. – Одеса, ОНУ, 2006, С. 140-158.
131. Янковой А. Г. Производительная сила и интенсивность труда в промышленности : монография / А. Г. Янковой. – Одесса, препринт, 1993. – 227 с.
132. Янковий О. Г. Роль оцінки латентних ознак в управлінні підприємством / О. Г. Янковий / Матеріали 3-ї міжнар. наук.-практ. конф. «Економіка підприємства : сучасні проблеми теорії та практики», 18-19 вер. 2014 р. – Одеса : Атлант, 2014. – С. 121-123.
133. Янковий О. Г. Статистична оцінка латентних економічних ознак / О. Г. Янковий. – Теорія і методологія статистичного аналізу : Зб. матер. міжнар. наук.-практ. конф., присвяченої 100-річчю від народження Пасхавера Й.С. – К.: КНЕУ, 2006. – С. 327-335.
134. Янковой А. Г. Факторный анализ и метод главных компонент в исследованиях латентных показателей / А. Г. Янковой // Ринкова економіка : Сучасна теорія і практика управління, т. 6 (вип. 7). – Одеса : Астропринт, 2003, С. 81-104.
135. Янковой А. Г., Егупов Ю. А. Методы таксономии в анализе экономической эффективности деятельности предприятий / А. Г. Янковой, Ю. А. Егупов // Хозрасчет, финансы и кредит. – 1986. – Вып. 28. – С. 16-21.
136. Янковий О. Г., Єргієва Л. В. Методи оцінки конкурентоспроможності підприємств / О. Г. Янковий, Л. В. Єргієва / Конкурентоспроможність

підприємства : оцінка рівня та напрями підвищення : [монографія] / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2013. – С. 34-65.
137. Янковий О. Г., Кошельок Г. В. Теоретичні основи оцінки фінансового стану підприємств харчової промисловості / О. Г. Янковий, Г. В. Кошельок. – Одеса, Наукові праці ОНАХТ, 2007, вип. 30, т. 2. – С. 367-371.
138. Янковий О. Г., Кошельок Г. В. Фактори і резерви росту інвестиційної привабливості підприємств промисловості / О. Г. Янковий, Г. В. Кошельок / Економіка підприємства в умовах ринкових перетворень : Монографія / за заг. ред. О. Г. Янкового. – Одеса, Атлант, 2010. – С. 7-46.
139. Янковий О. Г., Кошельок Г. В., Чернишова О. Б. Об'єднана таксономія підприємств кондитерської галузі за рівнем конкурентоспроможності / О. Г. Янковий, Г. В. Кошельок, О. Б. Чернишова / Розвиток фінансових відносин в умовах трансформаційних процесів : український вимір / Матер. симпозіуму з нагоди 65-річчя Харків. інст. фінансів УДУФМТ. – Х.: ХІФ УДУФМТ, 2008. – С. 131-138.
140. Янковой А. Г., Кошелек Г. В., Чернышова О. Б. Объединённая таксономическая оценка латентных экономических показателей / А. Г. Янковой, Г. В. Кошелек, О. Б. Чернышова. // Матер. 3-ї міжнар. наук.-практ. конф. «Проблеми вдосконалення інституційних умов функціонування української економіки». – Одеса, 2008, ОНУ, ч.1, С. 45-47.
141. Янковой А. Г., Куперман В. В. Многомерные методы оптимизации производственной программы предприятия / А. Г. Янковой, В. В. Куперман // Сучасна економіка : Вип. 4. – К.: ДІПК, 2011.– С. 13-23.
142. Янковий О. Г., Мельник Н. В., Янковий В. О. Оцінка інвестиційних проектів на підприємствах харчової промисловості України / О. Г. Янковий, Н. В. Мельник, В. О. Янковий // Економіка харчової промисловості. – 2013. – № 2(18). – С. 49-53.
143. Янковий О.Г., Мельник Н.В., Янковий В.О. До проблеми визначення пріоритетності інвестиційних проектів промислового підприємства / Янковий О.Г., Мельник Н.В., Янковий В.О. // Економіка харчової промисловості. – 2015. – № 1(25). – С. 71-76.
144. Янковий О. Г., Чернишова О. Б. Оцінка конкурентоспроможності промислових підприємств у маркетинговому дослідженні / О. Г. Янковий, О. Б. Чернишова // Вісник соціально-економічних досліджень. – ОДЕУ, 2008, № 33. – С. 250-257.
145. Янковий О. Г., Чернишова О. Б. Теоретичні основи багатовимірного статистичного аналізу конкурентоспроможності продукції / О. Г. Янковий, О. Б. Чернишова // Торгівля і ринок. – ДонНУЕТ, 2007, вип. 23. – С. 272-278.
146. Янковий О. Г., Чернишова О. Б. Фактичний та потенційний аспекти поняття «конкурентоспроможність» підприємства / О. Г. Янковий, О. Б. Чернишова. – Одеса, Наукові праці ОНАХТ, 2007, вип. 30, т. 2. – С. 363-366.

147. Янковий О. Г., Юр'єва Г. В. Критичний аналіз методів експрес-діагностики фінансового стану промислових підприємств / О. Г. Янковий, Г. В. Юр'єва // Держава та регіони : Економіка та підприємництво. – Запоріжжя, 2005. – № 5. – С. 277-284.
148. Янковий О. Г., Юр'єва Г. В. Методологічні аспекти аналізу фінансового стану підприємства / О. Г. Янковий, Г. В. Юр'єва // Вісник Хмельн. нац. ун-ту : Екон. науки. т. 1, № 2. – Хмельницький, 2005. – С. 108-113.
149. Янковой А. Г., Юрьева А. В., Янковая Е. Д. Оценка латентных экономических признаков предприятий с помощью методов таксономии / А. Г. Янковой, А. В. Юрьева, Е. Д. Янковая // Зерновые продукты и комбикорма. – 2005. – № 3. – С. 11-16.
150. Янковий О. Г., Янковий В. О. Оцінка пріоритетності інвестиційних проектів / О. Г. Янковий, В. О. Янковий // Зовнішня торгівля : економіка, фінанси, право. – 2013. – № 1. – С. 19-31.
151. Яшин Н. С. Конкурентоспособность промышленных предприятий : методология, оценка, регулирование / Яшин Н. С. – Саратов : Изд. Центр СГЭА, 1994. – 248 с.
152. Altman E. I. (1968) Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy // *The Journal of Finance, September*, pp. 589-609. (USA).
153. Altman E. I. (1983) Corporate Financial Distress. – *New York, John Wiley*. (USA).
154. Altman E. I. (1984) Further Empirical Investigation of the Bankruptcy Cost Question // *Journal of Finance, September,* pp. 1067-1089. (USA).
155. Farrer D. E., Glauber F. R. (1967) Multicollinearity in Regression Analysis: The Problem Revisited. – // *The Review of Economics and Statistics*, vol. 49, № 1, pp. 92-107. (USA).
156. Coombs C.H. (1952) A Theory of psychological Scaling // *Ending Res. Bull.*, p. 34. (USA).
157. Guttman, L. A. (1968) General nonmetric technique for finding the smallest coordinate space for a configuration of points. *Pyrometrical,* pp. 33, 469-506. (USA).
158. Kruskal, J. B. (1964) Nonmetric multidimensional scaling: A numerical method. *Pyrometrical*, pp. 29, 1-27, 115-129. (USA).
159. Stevens S. S. (1946) On the Theory of Measurement Scales // *Science. No. 103.* pp. 677–680. (USA).
160. Torgerson W.S. (1958) Theory and Methods of Scaling. *N.Y.: Willey,* 455 p.
161. Young, F. W, & Hamer, R. M. (1987) Multidimensional scaling: History, theory, and applications. *Hillsdale, NJ: Erlbaum*, P. 217. (USA).

Д О Д А Т К И

Додаток А

Властивості середньої арифметичної (простої)

1. Середня постійної величини C дорівнює їй самій:

$$\overline{C} = \frac{\sum_{i=1}^{N} C}{N} = \frac{NC}{N} = C. \qquad \text{(A1)}$$

2. Зміна кожного значення показника на постійну величину C змінює середню на ту ж величину:

$$\overline{X} \pm C = \frac{\sum_{i=1}^{N} (X_i \pm C)}{N} = \frac{\sum_{i=1}^{N} X_i \pm NC}{N} = \overline{X} \pm C. \qquad \text{(A2)}$$

3. Множення кожного значення показника на постійну величину C призводить до зміни середньої в C разів:

$$C\overline{X} = \frac{\sum_{i=1}^{N} CX_i}{N} = \frac{C\sum_{i=1}^{N} X_i}{N} = C\overline{X}. \qquad \text{(A3)}$$

4. Алгебраїчна сума відхилень всіх значень показника від його середньої величини дорівнює нулю:

$$\sum_{i=1}^{N} (X_i - \overline{X}) = \sum_{i=1}^{N} X_i - N\overline{X} = \sum_{i=1}^{N} X_i - \sum_{i=1}^{N} X_i = 0. \qquad \text{(A4)}$$

5. Середня суми двох змінних дорівнює сумі середніх цих змінних:

$$(X + Y) = \frac{\sum_{i=1}^{N} (X_i + Y_i)}{N} = \frac{\sum_{i=1}^{N} X_i + \sum_{i=1}^{N} Y_i}{N} = \overline{X} + \overline{Y}. \qquad \text{(A5)}$$

6. Сума квадратів відхилень змінної від середньої арифметичної менше, ніж від будь-якої іншої величини K:

$$S = \sum_{i=1}^{N} (X_i - K)^2 = \min \text{ при } K = \overline{X}. \qquad \text{(A6)}$$

Властивості дисперсії (простої)

1. Дисперсія постійної величини C дорівнює нулю. З урахуванням першої властивості середньої арифметичної можна записати:

$$\sigma_C{}^2 = \frac{\sum_{i=1}^{N}\left(C - \overline{C}\right)^2}{N} = \frac{\sum_{i-1}^{N}(C - C)^2}{N} = 0. \qquad (Б1)$$

2. При зміні кожного значення економічного показника на постійну величину C його дисперсія не змінюється. З урахуванням п'ятої і другої властивостей середньої арифметичної можна записати:

$$\sigma^2_{X+C} = \frac{\sum_{i=1}^{N}\left[(X_i + C) - (X_i + C)\right]^2}{N} = \frac{\sum_{i=1}^{N}\left[X_i + C - \overline{X} - C\right]^2}{N} = \sigma_X{}^2. \qquad (Б2)$$

3. Множення кожного значення показника на постійну величину C призводить до зміни його дисперсії в C^2 разів. З урахуванням третьої властивості середньої арифметичної можна записати:

$$\sigma^2_{CX} = \frac{\sum_{i=1}^{N}\left(CX_i - C\overline{X}\right)^2}{N} = \frac{C^2\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}{N} = C^2\sigma_X{}^2. \qquad (Б3)$$

Отже, величина дисперсії залежить від масштабу одиниць виміру економічного показника.

4. Дисперсія суми двох змінних X і Y дорівнює сумі дисперсій цих змінних плюс подвоєна їх коваріація. З урахуванням п'ятої властивості середньої арифметичної можна записати:

$$\sigma_{(X+Y)}{}^2 = \frac{\sum_{i=1}^{N}\left[(X_i + Y_i) - (X + Y)\right]^2}{N} = \frac{\sum_{i=1}^{N}\left[(X_i - \overline{X}) + (Y_i - \overline{Y})\right]^2}{N} =$$

$$= \frac{\sum_{i=1}^{N}\left[\left(X_i - \overline{X}\right)^2 + 2\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) + \left(Y_i - \overline{Y}\right)^2\right]}{N} = \sigma_X{}^2 + \sigma_Y^2 + 2\,\mathrm{cov}(X, Y). \qquad (Б4)$$

Очевидно, якщо показники X і Y лінійно незалежні ($\mathrm{cov}(Y, X) = 0$), то дисперсія їх суми дорівнює сумі їх дисперсій: $\sigma^2_{(X+Y)} = \sigma^2_X + \sigma^2_Y$.

5. Дисперсія різниці двох змінних X і Y дорівнює сумі дисперсій цих змінних за вирахуванням подвоєною їх коваріації:

$$\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_Y - 2\mathrm{cov}(Y, X). \qquad (Б5)$$

Дану властивість легко довести, використовуючи розкладання величини $\sigma^2_{(X–Y)}$, аналогічне (Б4). Ясно, якщо економічні ознаки X і Y лінійно незалежні ($\mathrm{cov}(Y, X) = 0$), то дисперсія їх різниці дорівнює сумі їх дисперсій: $\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_Y$.

Додаток В

Властивості коваріації

1. Величина коваріації не змінюється при переміні місць показників, тобто справедливе співвідношення cov(*Y*, *X*) = cov(*X*, *Y*). Це безпосередньо випливає з визначення коваріації:

$$\operatorname{cov}(Y,X)=\frac{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)\left(X_i-\overline{X}\right)}{N}=\frac{\sum_{i=1}^{N}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{N}=\operatorname{cov}(X,Y). \quad \text{(B1)}$$

2. Коваріація змінних дорівнює нулю, якщо хоча б одна з них є константою. Нехай *X* = *C*, тоді з урахуванням першої властивості середньої арифметичної

$$\operatorname{cov}(Y,C)=\frac{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)\left(C-\overline{C}\right)}{N}=\frac{\sum_{i=1}^{N}(Y_i-\overline{Y})\times 0}{N}=0. \quad \text{(B2)}$$

3. Коваріація показників не змінюється при додаванні до будь-якого з них постійної величини. Наприклад, додамо до всіх значень показника *X* константу *C*, тоді з урахуванням п'ятої і другої властивостей середньої арифметичної можна записати:

$$\operatorname{cov}(Y,X+C)=\frac{\sum_{i=1}^{N}(Y_i-\overline{Y})\left[(X_i+C)-(X_i+C)\right]}{N}=\frac{\sum_{i=1}^{N}(Y_i-\overline{Y})\left[X_i+C-\overline{X}-C\right]}{N}=\frac{\sum_{i=1}^{N}(Y_i-\overline{Y})(X_i-\overline{X})}{N}=\operatorname{cov}(Y,X). \quad \text{(B3)}$$

4. Якщо всі значення змінної *Y* помножити на постійну величину *C*, а всі значення змінної *X* помножити на постійну величину *D*, то це призведе до зміни коваріації показників у *CD* разів. З урахуванням третьої властивості середньої арифметичної можна записати:

$$\operatorname{cov}(CY,DX)=\frac{\sum_{i=1}^{N}\left(CY_i-C\overline{Y}\right)\left(DX_i-D\overline{X}\right)}{N}=\frac{CD\sum_{i=1}^{N}(Y_i-\overline{Y})(X_i-\overline{X})}{N}=CD\operatorname{cov}(Y,X). \quad \text{(B4)}$$

Отже, величина коваріації залежить від масштабу одиниць виміру обох економічних показників.

5. Квадрат коваріції змінних не перевершує добуток їх дисперсій, тобто

$$\operatorname{cov}^2(Y,X)\le\sigma_Y^2\sigma_X^2, \quad \text{(B5)}$$

причому рівність у (В4) виконується при суворій лінійній залежності економічних показників (математичне доведення даної властивості див. у роботі [128, с.108-109]).

Слідство. З четвертої властивості коваріації змінних випливає дуже важливий наслідок: коефіцієнт парної лінійної кореляції між показниками r_{YX} за абсолютною величиною не перевищує 1. Дійсно, за визначенням r_{YX} можна записати:

$$r_{YX}^2 = \frac{\mathrm{cov}^2(Y,X)}{\sigma_Y^2 \sigma_X^2} \le 1, \tag{B6}$$

або $$-1 \le r_{YX} \le 1 ; \quad |r_{YX}| \le 1. \tag{B7}$$

6. Коваріація показника з самим собою збігається з його дисперсією. Дійсно, при $Y = X$ основна формула коваріації (B1) приймає вигляд:

$$\mathrm{cov}(X,X) = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})}{N} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N} = \sigma_X^2 \tag{B8}$$

Додаток Г

Властивості регресії

Як для парної, так і для множинної регресії виводяться такі основні властивості, які тут наведено без доказів. Математичне доведення вказаних властивостей можна знайти в роботі [129, с.48-56, 67-74].

1. Якщо до кожного значення економічних змінних Y, X_1, X_2, …, X_m додати відповідні константи c_0, c_1, c_2, …, c_m, то коефіцієнти регресії α_1, α_2, …, α_m не зміняться, а новий коефіцієнт a'_0 збільшиться в порівнянні з a_0 на величину $c_0 - \alpha_1 c_1 - \alpha_2 c_2 - \ldots - \alpha_m c_m$.

Слідство A. Якщо від кожного значення економічних змінних Y, X_1, X_2, …, X_m відняти їх середні величини, то гіперплощина регресії пройде через початок координат. Інакше кажучи, центрування вихідних статистичних даних призводить до того, що $a'_0 = 0$.

2. Якщо кожне значення економічних змінних Y, X_1, X_2, …, X_m помножити на відповідні константи c_0, c_1, c_2, …, c_m, то в порівнянні з вихідним рівнянням коефіцієнт регресії α'_j (j – номер факторного показника, $j = 1, 2, ..., m$) збільшиться в $\frac{c_0}{c_j}$ разів, а коефіцієнт a'_0 стане рівним $c_0(\bar{Y} - \bar{X}_1 - \bar{X}_2 - \ldots - \bar{X}_m)$.

Слідство B. Якщо кожне значення економічних змінних Y, X_1, X_2, …, X_m розділити на їх стандартні відхилення σ_Y, σ_1, σ_2, …, σ_m, (тобто піддати нормуванню), то в порівнянні з вихідним рівнянням коефіцієнт регресії α'_j збільшиться в $\frac{\sigma_j}{\sigma_Y}$ разів, а величина a'_0 стане рівною $\frac{\bar{Y} - \bar{X}_1 - \bar{X}_2 - \ldots - \bar{X}_m}{\sigma_Y}$.

Слідство C. Якщо з кожного значення економічних змінних Y, X_1, X_2, …, X_m відняти їх середні величини, а потім отримані різниці розділити на їх стандартні відхилення σ_Y, σ_1, σ_2, …, σ_m, тобто призвести стандартизацію вихідних статистичних даних, то нові коефіцієнти регресії приймуть значення: $a'_0 = 0$, $a'_j = a_j \frac{\sigma_j}{\sigma_Y}$. Вони називаються β-коефіцієнтами регресії.

3. Пряма регресії проходить через «центр ваги» кореляційної хмари вихідних даних, тобто через точку з координатами $Y = \bar{Y}$, $X_1 = \bar{X}_1$, $X_2 = \bar{X}_2$, …, $X_m = \bar{X}_m$. Іншими словами, справедливе співвідношення:

$$\bar{Y} = a_0 + \alpha_1 \bar{X}_1 + \alpha_2 \bar{X}_2 + \ldots + \alpha_m \bar{X}_m . \qquad (Г1)$$

4. Сума залишків рівняння регресії дорівнює нулю.

Слідство D. Середні величини фактичних і розрахункових (за рівнянням регресії) значень результативної змінної Y збігаються.

5. Залишки не корельовані зі спостереженими значеннями будь-якої факторної змінної X_j і розрахунковими значеннями результативної змінної $\hat{Y}$.

6. Середній квадрат залишків є функцією від коефіцієнтів регресії α_1, α_2, …, α_m.

7. Дисперсія результативної змінної Y дорівнює сумі двох доданків: факторної і залишкової дисперсії:

$$\sigma^2_Y = \sigma^2_F + \sigma^2_e, \qquad (Г2)$$

де $\sigma_Y^2 = \dfrac{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}{N}$ – загальна дисперсія змінної Y;

$\sigma_F^2 = \dfrac{\sum_{i=1}^{N}(\hat{Y}_i - \bar{Y})^2}{N}$ – факторна дисперсія результативної ознаки Y, обумовлена впливом усіх факторів, які увійшли в рівняння регресії;

$\sigma_e^2 = \dfrac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{N}$ – залишкова дисперсія змінної Y, обумовлена впливом усіх інших факторів, що не увійшли в рівняння регресії.

8. Розділивши обидві частини (Г2) на загальну дисперсію σ^2_Y, отримаємо наступне співвідношення:

$$1 = \frac{\sigma_F^2}{\sigma_Y^2} + \frac{\sigma_e^2}{\sigma_Y^2}. \qquad (Г3)$$

Перший дріб у правій частині формули (Г3) суть частка факторної дисперсії в загальній дисперсії результативної змінної, а другий – питома вага залишкової дисперсії в σ^2_Y. Відношення факторної дисперсії до загальної дисперсії Y називається коефіцієнтом детермінації і позначається через R^2. Коефіцієнт детермінації показує частку варіації результативної змінної Y, яка пояснюється усіма факторами, що ввійшли в рівняння регресії.

Оскільки факторна і залишкова дисперсії є невід'ємними величинами і являють собою частини загальної дисперсії змінної Y, то R^2 змінюється в інтервалі від 0 до 1.

При $R^2 = 1$ $\sigma^2_e = 0$ і це означає, що варіацію Y визначає тільки зміна факторів, які ввійшли в рівняння регресії. Всі інші фактори не діють. Отже, між досліджуваними економічними змінними існує функціональний причинно-наслідковий зв'язок.

При $R^2 = 0$ $\sigma^2_F = 0$ і це означає, що варіацію Y повністю визначає лише всі інші фактори, які не ввійшли в рівняння регресії. Тому говорять про лінійну незалежність досліджуваних показників, про відсутність між ними кореляційної

зв’язку. У практичних ситуаціях в області економіки, зазвичай, має місце співвідношення $0 < R^2 < 1$.

9. Коефіцієнт множинної кореляції пов’язаний з коефіцієнтом детермінації наступним чином: $R=+\sqrt{R^2}$.

Для парного лінійного рівняння регресії справедливі наступні співвідношення:

$$R^2 = r^2_{YX},\ R = |r_{YX}|. \qquad (Г4)$$

10. Розширення кола факторних змінних у рівнянні регресії не знижує величини його коефіцієнта детермінації, тобто можна записати:

$$R^2_1 \le R^2_2 \le ... \le R^2_m. \qquad (Г5)$$

Додаток Д

Процедура перевірки статистичних гіпотез

У ході оцінювання, моделювання і прогнозування ознак економічних об'єктів досить часто виникають завдання перевірки деяких статистичних припущень (гіпотез). Наприклад, це можуть бути гіпотези про надійність знайдених коефіцієнтів кореляції (парних, множинних), параметрів рівняння регресії тощо. При цьому, зазвичай, використовується поняття *гіпотетичної генеральної сукупності*, тобто такої, з якої могли бути отримані аналізовані статистичні дані у вигляді вибірки, незалежно від того, як вони були отримані в дійсності. Поняття гіпотетичної генеральної сукупності дозволяє здійснювати перевірку статистичних гіпотез, а математичний апарат і послідовність, тобто сама процедура перевірки гіпотез, є універсальною, придатною для переважної більшості подібних досліджень.

Розглянемо основні поняття і порядок проведення процедури перевірки статистичних гіпотез, використовуючи наступний приклад. Нехай в процесі оцінки латентної ознаки за методом виділення репрезентативного чинника-симптому отримані такі результати (див. розділ 2.3): $R_s^2 = \max(R_k^2) = 0{,}758$; $c_{ss} = 4{,}132$; $n = 20$; $m = 6$. Треба з'ясувати значущість мультиколінеарності s-го чинника-симптому за критерієм Д. Феррара, Р. Глаубера (2.9).

Статистичну гіпотезу, що перевіряється, називають нульовою і позначають через H_0. Конкуруючу гіпотезу, яка суперечить нульовій, називають альтернативною і позначають H_a. Назва «нульова» гіпотеза має історичні корені, пов'язані з тим, що найчастіше перевірялися припущення про рівність нулю деякого статистичного параметра, наприклад, коефіцієнта множинної кореляції між змінними ($H_0 : R_s^2 = 0$).

В обговорюваній задачі нульова гіпотеза записується наступним чином: $H_0 : c_{ss} - 1 = 0$. Альтернатива може бути записана так: $H_a : c_{ss} - 1 > 0$.

Простою називають гіпотезу, яка містить лише одне припущення, а складною – гіпотезу, що складається з кінцевого або нескінченного числа припущень. У даному прикладі нульова гіпотеза $H_0 : c_{ss} - 1 = 0$ є простою, а альтернативна $H_a : c_{ss} - 1 > 0$ – складною.

Зазначимо тут, що в загальному вигляді процедура перевірки статистичних гіпотез полягає у відхиленні або невідхиленні H_0 на основі деякого вирішального правила, яке називається *статистичним критерієм*. При цьому відхилення нульової гіпотези зазвичай рівносильне прийняттю альтернативи H_a, а невідхилення нульової гіпотези фактично означає її прийняття, тобто визнання H_0 справедливою.

Висунута нульова гіпотеза може бути вірною або невірною, тому завжди існує можливість помилитися – як при відхиленні, так і при невідхиленні H_0.

Помилка *першого роду* при перевірці статистичних гіпотез відбувається в тому випадку, коли відхиляється вірна нульова гіпотеза.

Помилка *другого роду* при перевірці статистичних гіпотез полягає у невідхиленні нульової гіпотези, коли вірна альтернатива.

Зазначимо одразу, що наслідки помилок першого і другого роду можуть бути дуже різними з точки зору прийняття подальших невірних рішень.

Так, помилка першого роду в обговорюваній задачі означатиме визнання чинника-симптому x_s колінеарним (репрезентативним), коли насправді, тобто в гіпотетичній генеральної сукупності, його зв'язок з іншими факторами практично відсутній ($R_s^2 = 0$; $c_{ss} = 1$). В результаті помилкового визнання виявленої множинної кореляційної залежності значущою запропоновані оцінки досліджуваної латентної ознаки (наприклад, фінансового стану промислових підприємств, їх конкурентоспроможності тощо) не будуть відображати дійсний рівень прихованої властивості економічних об'єктів.

Помилка другого роду в обговорюваній задачі означатиме визнання зв'язку даного чинника-симптому x_s з іншими незначущим, несуттєвим, коли насправді (в гіпотетичній генеральній сукупності) все навпаки: $R_s^2 > 0$; $c_{ss} > 1$. У результаті помилкового невідхилення H_0 виявиться незадіяним потужний і ефективний спосіб оцінки латентних ознак.

Ясно, що в даній ситуації більш небезпечною для оцінювання прихованих властивостей економічних об'єктів є помилка першого роду – відхилення H_0, коли вона вірна. Зрозуміло, що можна навести безліч прикладів, коли більш згубні наслідки тягне за собою помилка другого роду.

Правильне рішення в результаті процедури перевірки статистичних гіпотез буде прийнято в двох випадках:

1. Нульова гіпотеза не відхиляється (приймається), причому в дійсності вона вірна.

2. Нульова гіпотеза відхиляється, коли насправді вірна альтернатива.

Ймовірність зробити помилку першого роду називається рівнем значущості і позначається через α. Рівень значущості – це керований параметр процедури перевірки статистичних гіпотез, який задається самим дослідником виходячи з аналізу згубних наслідків помилок першого і другого роду.

Ймовірність зробити помилку другого роду позначається через β. Однак, загально прийнято розглядати вірогідність не допустити помилку другого роду, яка дорівнює 1 - β. Остання величина називається потужністю статистичного критерію і характеризує його здатність розрізняти нульову і альтернативну гіпотези. Потужність статистичного критерію – розрахункова величина, яка залежить від ряду параметрів, зокрема, від обраної величини α, від характеристики $\varphi = H_a - H_0$, яка відображає відмінності між нульовою та альтернативною гіпотезами, від числа ступенів вільності *df* (від англ. degree of freedom).

Статистичний критерій будується таким чином, щоб одночасно мінімізувати ймовірність помилок першого і другого роду або максимізувати ймовірність не допустити помилку першого роду і потужність критерію. Відразу зазначимо, що α і β пов'язані між собою зворотною залежністю: обравши дуже низький рівень значущості, ми тим самим збільшуємо ймовірність помилки другого роду і, навпаки. Проблема забезпечення проведення процедури перевірки гіпотез, що забезпечує $\min(\alpha; \beta)$ або $\max(1 - \alpha; 1 - \beta)$ є однією з центральних проблем математичної статистики.

На практиці найбільш часто α приймається на рівні 0,1; 0,05; 0,01. Однак, коли більш згубними є наслідки помилки другого роду, α вибирається серед порівняно великих значень задля мінімізації ймовірності β (α = 0,2; 0,3; 0,4). У даній задачі, виходячи з висновку про більш згубні наслідки для результатів оцінювання латентної ознаки, що вивчається, настання помилки першого роду, приймемо α = 0,01 або 1 %. Очевидно, що ймовірність не допустити помилку першого роду (відхилити нульову гіпотезу, коли насправді вірна альтернатива) дорівнює 1 - α = 1 - 0,01 = 0,99 або 99 %.

Головне місце в критерії як вирішальному правилі або інструменті перевірки гіпотез займає статистична характеристика або просто статистика, яка являє собою деяку математичну функцію від вихідних статистичних даних. Статистика підпорядковується заздалегідь відомому теоретичному закону розподілу. Зазвичай, це нормальний розподіл, *t*-розподіл Стьюдента (псевдонім англійського статистика початку XX ст. Уіл'яма Госсета), F-розподіл Фішера, χ^2-розподіл. Графіки диференціальної і інтегральної функцій указаних розподілів наведені на рис. Д1 - Д3.

Звернемо увагу на те, що при числі ступенів вільності $df \geq 30$ графік диференціальної функції (функції щільності ймовірностей) *t*-розподілу Стьюдента (рис. Д1) практично не вирізняється від нормальної кривої Гауса. Причому обидва графіки стандартизованих функцій нормального розподілу і *t*-розподілу Стьюдента є симетричними відносно вертикальної прямої, яка проходить через нульову точку.

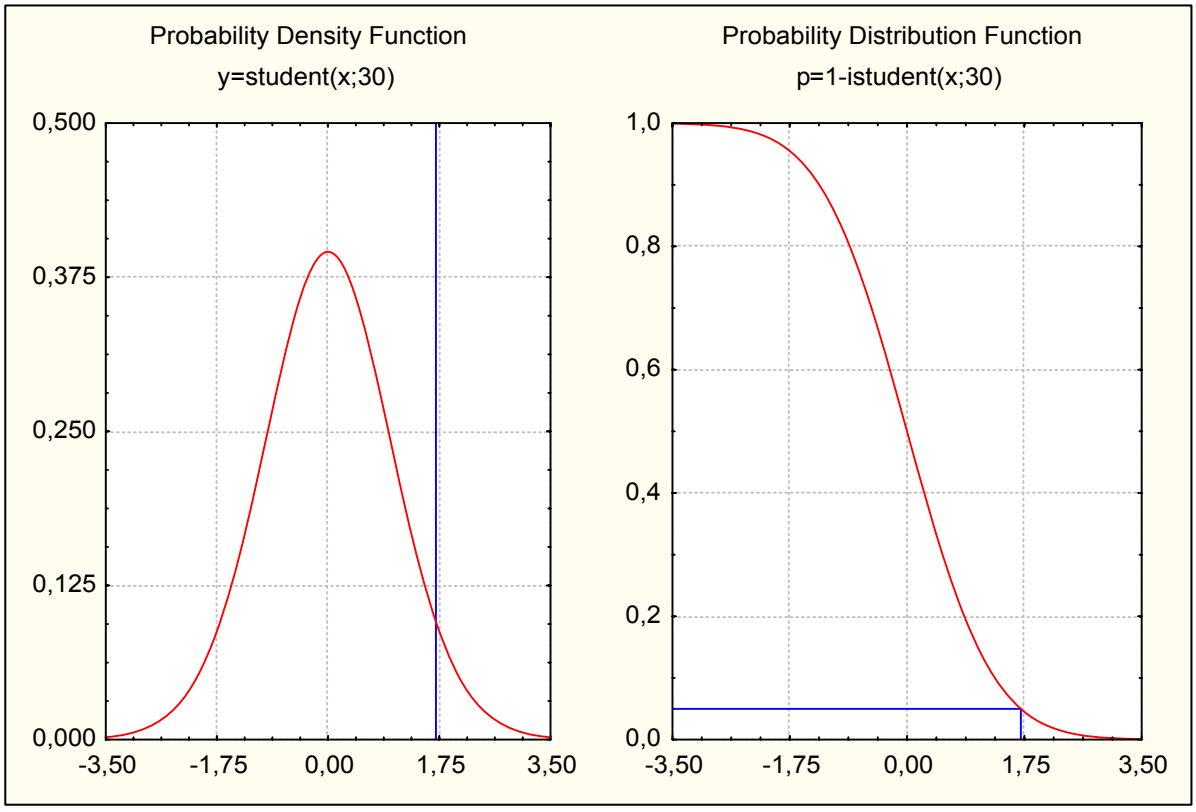


Рис. Д1. Графіки диференціальної (зліва) і інтегральної функцій t-розподілу Стьюдента

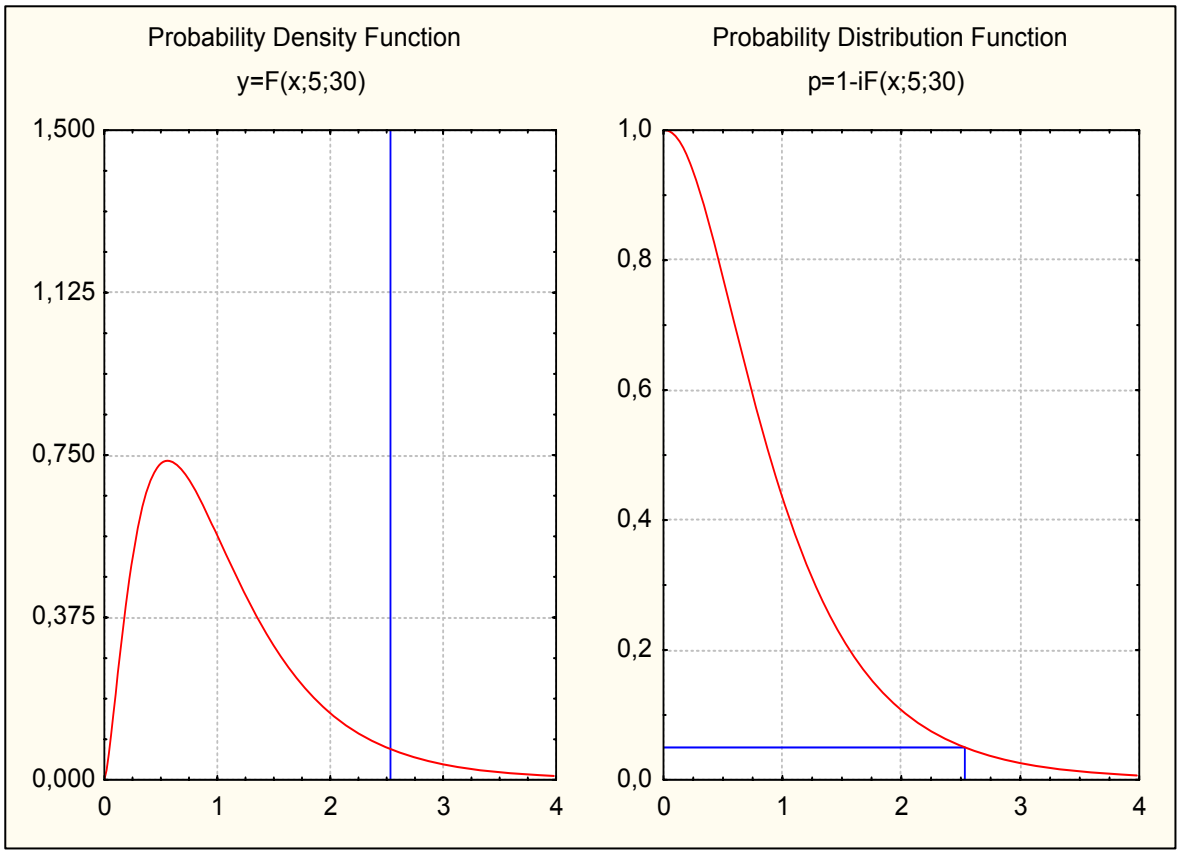


Рис. Д2. Графіки диференціальної (зліва) і інтегральної функцій F-розподілу Фішера

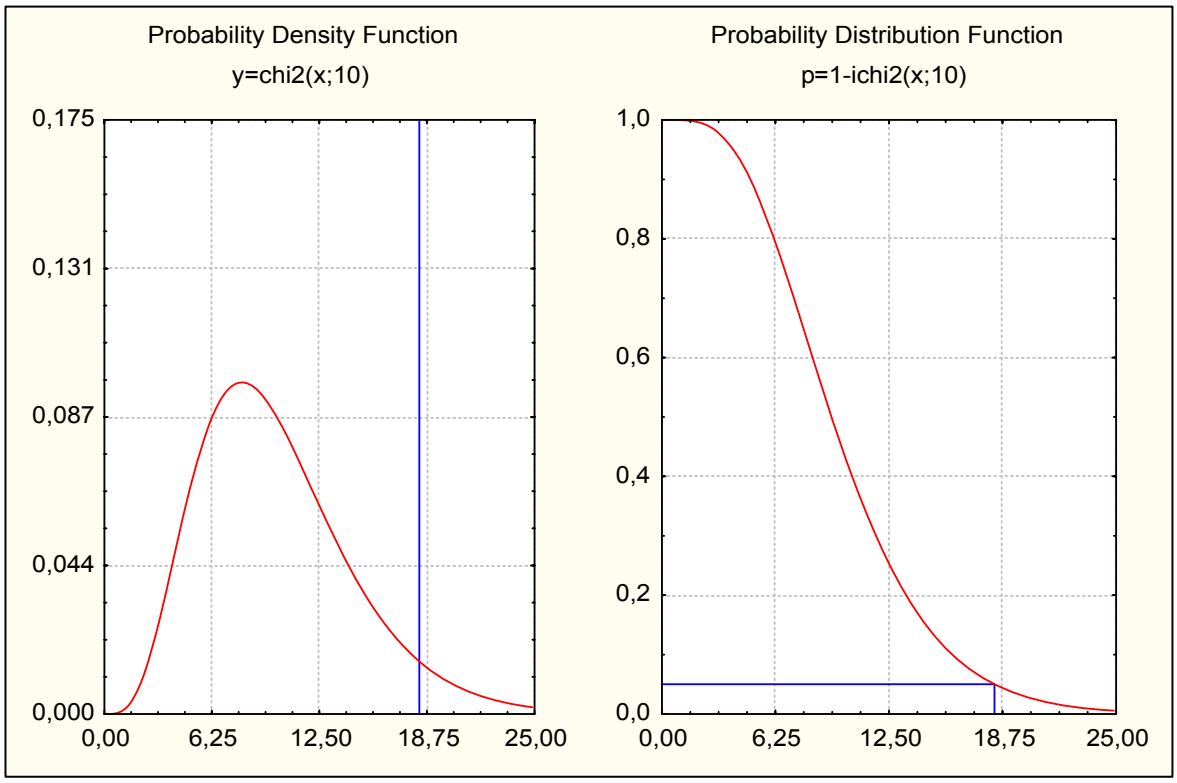


Рис. Д3. Графіки диференціальної (зліва) і інтегральної функцій χ^2-розподілу

Що стосується розподілів Фішера і χ^2, то їх графіки функцій щільності ймовірностей (рис. Д2, Д3) є несиметричними: права гілка більш поката в порівнянні з лівою гілкою, тобто можна казати про правосторонню скошеність їх розподілів.

Усі можливі значення статистик критеріїв, що підпорядковуються нормальному і *t*-розподілу Стьюдента, належать інтервалу від -∞ до +∞ (див. рис. Д1). А область визначення статистик критеріїв, що підпорядковуються *F*-розподілу Фішера і χ^2-розподілу, знаходиться від 0 до +∞ (див. рис Д2, Д3).

Критичною називається область значень статистики критерію, при потрапляння в яку її розрахункової величини нульова гіпотеза H_0 відхиляється (фактично приймається альтернатива H_a) з вірогідністю 1 - α.

Областю припустимих значень називається інтервал, при потраплянні в який розрахункового значення статистики критерію нульова гіпотеза H_0 не відхиляється (фактично приймається) з достовірністю, що дорівнює потужності статистичного критерію 1 - β.

Межею критичної області і області припустимих значень служить критична точка – α-квантиль відповідного розподілу випадкової величини *Z*. Для несиметричних розподілів, які визначені на позитивній частині осі абсцис, наприклад, для *F*-розподілу Фішера і χ^2-розподілу, α-квантиль (Z_α) задовольняє наступній умові:

$$P(Z > Z_\alpha) = \alpha. \qquad (Д1)$$

Це означає, що ймовірність прийняття випадковою величиною *Z* значення, яке перевищує α-квантиль, дорівнює α.

α-квантиль знаходиться за допомогою стандартних програм, наприклад, в редакторі Excel. Для цього необхідно відкрити робочу книгу і виконати наступні дії та команди:

- Натиснути на панелі кнопку *fx* «Майстер функцій»;
- В лівій частині вікна, яка називається «Категорії», вибрати «Статистичні»;
- Потім в правій частині вікна, яка називається «Функції», обрати необхідну функцію, наприклад, «НОРМОБР» («НОРМСТОБР»), «FРАСПОБР», «ХИ2ОБР», «СТЬЮДРАСПОБР» та ін.;
- Задати значення α і відповідні числа ступенів вільності *df* для зворотної функції – Enter.

У задачі, що розглядається, $\alpha = 0{,}01$; $df_1 = m - 1 = 6 - 1 = 5$; $df_2 = n - m = 20 - 6 = 14$. Звідси в редакторі Excel знаходимо: $F_{0,01;5;14} = 4{,}695$.

Критична область завжди будується на основі принципу практичної неможливості настання малоймовірних подій в одиничному випробуванні. Малоймовірні події наступають для статистик критеріїв, що підпорядковуються *F*-розподілу Фішера (χ^2-розподілу), коли їх розрахункові величини приймають високі значення, тобто потрапляють в праві крайні «хвости» графіків їх диференціальних функцій (див. рис. Д2, Д3).

Оскільки статистика критерію – одномірна випадкова величина, то всі її розрахункові значення належать деякому інтервалу і можуть бути графічно представлені у вигляді відрізка прямої. Тому критична область і область припустимих значень теж є інтервалами, розділеними критичними точками – α-квантилями, наприклад Z_α. Якщо припустити, що статистика деякого критерію підпорядковується F-розподілу (χ^2-розподілу), то область її всіляких значень в даній задачі може бути схематично представлена на рис. Д4.

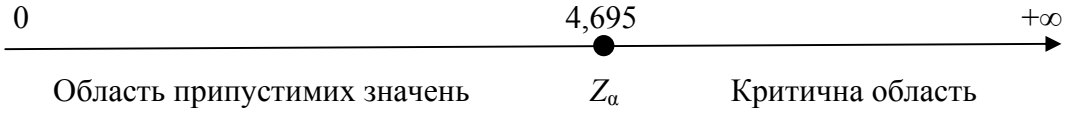


Рис. Д4. Графік всіляких значень статистики критерію, яка підпорядковується *F*-розподілу Фішера

Якщо ж малоймовірна подія сталася, тобто розрахункове значення статистики критерію виявилося в критичній області, то це можна пояснити тим, що нульова гіпотеза H_0 в гіпотетичній генеральній сукупності є помилковою і, отже, повинна бути відхилена. Справедливою в даному випадку визнається альтернатива H_a. При цьому можна зробити, принаймні, два висновки:

1) статистика критерію в умовах справедливості альтернативи H_a підпорядковується так званому нецентральному F-розподілу (нецентральному χ^2-розподілу), тобто з деяким параметром нецентральності – зрушенням праворуч за віссю абсцис (рис. Д5);

2) критична область для статистик критеріїв, що підпорядковуються F-розподілу Фішера, χ^2-розподілу (рис. Д4), є односторонньою.

Рис. Д5. Графіки диференціальних функцій розподілу статистик критерію в умовах справедливості H_0 і H_a (для F- і χ^2-розподілів)

Для симетричних розподілів, які визначені на всій осі абсцис, наприклад, для стандартизованого нормального розподілу, *t*-розподілу Стьюдента, α-квантиль позначається через $Z_{\alpha/2}$, і задовольняє наступній умові:

$$P(|Z| > Z_{\alpha/2}) = \alpha. \qquad (Д2)$$

Це означає, що ймовірність прийняття випадковою величиною Z значення, яке по модулю перевищує α-квантиль, дорівнює α.

Вимогу (Д2) запишемо наступним чином:

$$P(-\infty < Z < -Z_{\alpha/2}) + P(+Z_{\alpha/2} < Z < +\infty) = \alpha\ . \qquad (Д3)$$

Вона визначає такі дві точки на осі абсцис графіка нормальної кривої, або графіка диференціальної функції *t*-розподілу Стьюдента (рис. Д1), для яких ймовірність того, що випадкова величина Z прийме значення, менше $-Z_{\alpha/2}$ (лівий α-квантиль), і більше $+Z_{\alpha/2}$ (правий α-квантиль), дорівнює прийнятому рівню значущості α.

Оскільки нормальний і *t*-розподіл симетричні, то справедлива рівність

$$P(-\infty < Z < -Z_{\alpha/2}) = P(+Z_{\alpha/2} < Z < +\infty) = \alpha/2. \qquad (Д4)$$

Тобто ймовірність потрапляння значення випадкової величини Z в одну з областей, які визначені співвідношенням (Д4), дорівнює α/2. Окрім того, очевидно, що $|-Z_{\alpha/2}| = |+Z_{\alpha/2}|$. Це означає, що абсолютна величина лівого і правого α-квантилів стандартизованого нормального і *t*-розподілу Стьюдента збігаються.

Отже, критична область для вказаних розподілів, побудована згідно принципу практичної неможливості настання малоймовірних подій в одиничному випробуванні, має вигляд, представлений на рис. Д6

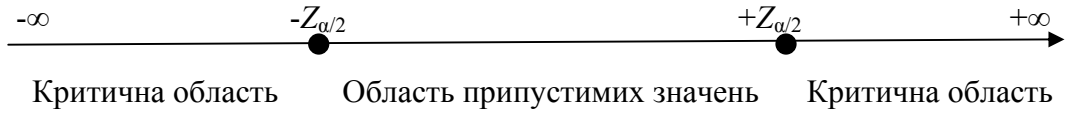


Рис. Д6. Графік всіляких значень статистики критерію, яка підпорядковується нормальному розподілу (*t*-розподілу Стьюдента)

При цьому за аналогією з несиметричними розподілами можна зробити теж два висновки:

1) статистика критерію в умовах справедливості альтернативи H_a підпорядковується так званому нецентральному нормальному розподілу (нецентральному *t*-розподілу Стьюдента), тобто з деяким параметром нецентральності – зрушенням праворуч за віссю абсцис (див. рис. Д7);

2) критична область для статистик критеріїв, що підпорядковуються *F*-розподілу Фішера, χ^2-розподілу (рис. Д6), є двосторонньою, тобто складається з

двох частин – лівосторонньої ($Z < -Z_{\alpha/2}$) і правосторонньої ($Z > +Z_{\alpha/2}$) критичних областей.

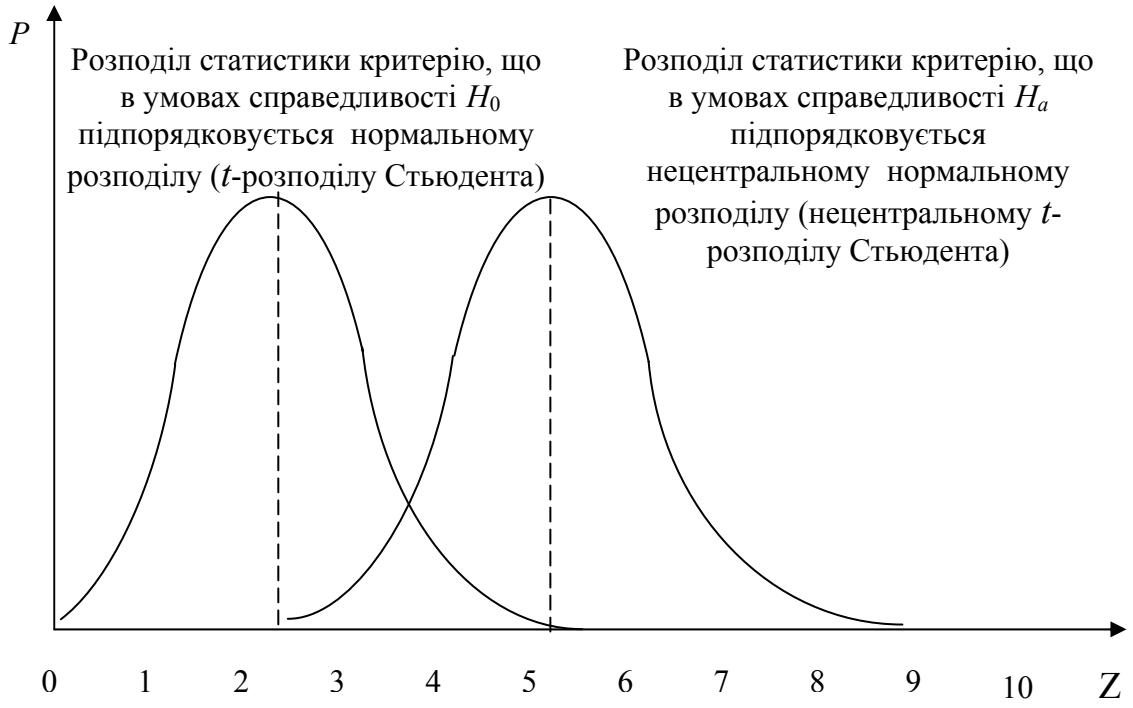


Рис. Д7. Графіки диференціальних функцій розподілу статистик критерію в умовах справедливості H_0 і H_a (для нормального і *t*-розподілів)

Застосування двосторонньої перевірки статистичних гіпотез з лівосторонньою або правосторонньою критичною областю вважається переважним, оскільки в цьому випадку в разі відхилення нульової гіпотези H_0 достовірність альтернативи H_a дорівнюю не 1 - α, а 1 - $\alpha/2$, тобто є вищою в порівнянні з односторонньою перевіркою. Тут виникає питання: за рахунок чого підвищуються вірогідність висновку? Відповідь одна: тільки за рахунок використання додаткової інформації про якісну економічну сутність статистичних параметрів, що перевіряються.

Справа в тім, що головними умовами застосування двосторонньої перевірки є наступні:

1. Статистика критерію підпорядковується симетричному закону розподілу.

2. Досліднику заздалегідь відомий (наприклад, з економічної теорії) істинний знак статистичного параметра, що досліджується.

Так, якщо перевіряється значущість коефіцієнта регресії при чиннику, про який відомо, що він прямо впливає на результативну ознаку, то лівосторонню критичну область слід відхилити як теоретично неможливу і скористатися лише правосторонньою критичною областю.

У даній задачі відомо, що статистика критерію Д. Феррара, Р. Глаубера (2.9) підпорядковується несиметричному *F*-розподілу Фішера, тому можлива лише одностороння перевірка значущості мультиколлінеарності *s*-го чинника-симптому від інших змінних.

Якщо в разі відхилення H_0 висновки щодо достовірності H_a більш-менш зрозумілі (вірогідність прийняття H_a дорівнює 1 - α при односторонній перевірці, 1 - $\alpha/2$ – при двосторонній), то при невідхиленні H_0 ситуація дещо ускладнюється, оскільки для визначення її достовірності необхідно розрахувати потужність статистичного критерію 1 - β, що не так уже і просто.

Часто потужність плутають з імовірністю 1 - α, оскільки вона пов'язана з попаданням розрахункового значення статистики критерію в область припустимих значень. Величина області припустимих значень обернено пропорційна величині α, оскільки обидві зазначені області (критична й область припустимих значень) суть частини одного цілого – множини всіляких значень статистики критерію (див. рис. Д4, Д6). Однак 1 - $\beta \neq$ 1 - α, так як на величину потужності критерію окрім α впливають також й інші фактори.

Як зазначалось вище, величина 1 - β істотно залежить від характеристики альтернативної гіпотези $\varphi = H_0 - H_a$, яка відображає відмінності між нульовою та альтернативною гіпотезами. За визначенням, потужність – це ймовірність відхилити H_0, коли вірна H_a. Потужність характеризує чутливість критерію, його здатність уловлювати відмінності між нульовою та альтернативною гіпотезами. Існує прямий зв'язок між 1 - β і величиною φ. Для φ, близьких до нуля, потужність критерію невисока: йому важко вловлювати малі відмінності між гіпотезами і помилка другого роду допускається часто. При $\varphi \rightarrow \max(\varphi)$ 1 - $\beta \rightarrow \max(1 - \beta)$.

Третім важливим чинником, що визначає величину 1 - β, є число ступенів вільності $df = N - m$, яке залежить від обсягу статистичних спостережень N і числа змінних m. У відповідності до закону великих чисел при N (df) $\rightarrow +\infty$ потужність 1 - $\beta \rightarrow 1$. При $m \rightarrow N$ $df \rightarrow 0$ і 1 - $\beta \rightarrow \alpha$.

Слід мати на увазі, що обидва розглянуті параметри (φ, df) можуть виступати як заданими, так і розрахунковими показниками статистичного критерію. Справа в тім, що взаємозв'язок параметрів критерію виражається у вигляді функції потужності. Для складної альтернативної гіпотези ця функція є безперервною, монотонно зростаючою функцією від α, φ, df. Функція потужності, зазвичай, задається у вигляді графіків, побудованих для нецентральних розподілів, тобто розподілів статистики критерію в умовах справедливості альтернативної гіпотези.

Однак, навіть не вдаючись до використання функцій і графіків потужності статистичних критеріїв можна дати наступну рекомендацію щодо оцінки 1 - β: малий об'єм вибірки ($N < 30$) зазвичай вказує на низький рівень потужності критерію і, навпаки. Цей простий, хоча і наближений метод оцінки 1 - β досить добре зарекомендував себе у практиці перевірки статистичних гіпотез.

На рис. Д8 наводиться класична блок-схема процедури перевірки статистичних гіпотез в економічному дослідженні. Щоб її реалізувати в даній задачі розрахуємо фактичне значення статистики критерію (29):

$$F_s = (c_{ss} - 1)(n - m)/(m - 1) = (4{,}132 - 1)(20 - 6)/(6 - 1) = 8{,}77.$$

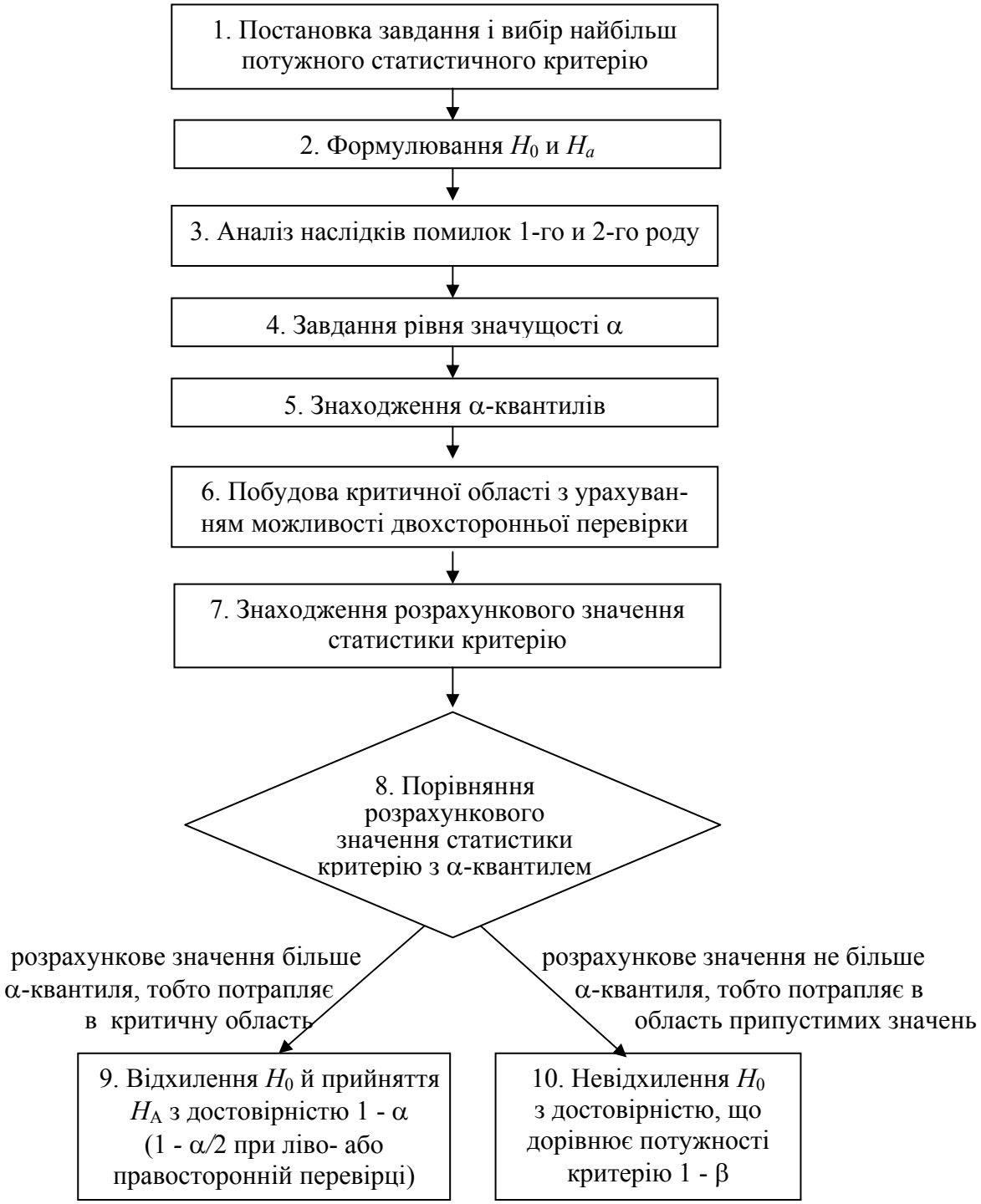


Рис. Д8. Блок-схема класичної процедури перевірки статистичних гіпотез

Отже, порівнюючи розрахункове значення статистики критерію Д. Феррара, Р. Глаубера F_s = 8,77 зі знайденим вище α-квантилем $F_{0,01;5;14}$ = 4,695, приходимо до висновку про те, що $F_s > F_{0,01;5;14}$ (8,77 > 4,695). Це означає, що розрахункове значення статистики критерію потрапляє в критичну область (див. рис. Д9) і нульова гіпотеза H_0 : c_{ss} - 1 = 0 відхиляється. З достовірністю 1 - α = 1 - 0,01 = 0,99 або 99 % визнається справедливою альтернатива H_a : c_{ss} - 1 > 0.

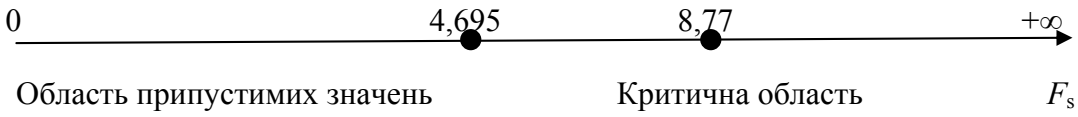


Рис. Д9. Класична перевірка гіпотези значущості мультиколінеарності *s*-го чинника-симптому за критерієм Д. Феррара, Р. Глаубера

Висновок однозначний: з достовірністю 99 % можна стверджувати, що s-й чинник-симптом є мультиколінеарним, репрезентативним, лінійно зв'язаним з рештою змінних x_1, x_2, …, x_{s-1}, x_{s+1}, …, x_m. Тому його можна розглядати як лінійну комбінацію інших векторів і, відповідно, носієм інформації всієї множини чинників-симптомів. А це, в свою чергу, надає вагомі підстави вважати його певною оцінкою латентної ознаки, що вивчається.

Вище був обговорений класичний (традиційний) підхід до перевірки статистичних гіпотез, що сформувався в першій половині XX ст., коли ЕОМ тільки розроблялися і єдиним джерелом інформації про критичні точки були статистичні таблиці α-квантилів різних теоретичних законів розподілу. В сучасних умовах персональний комп'ютер дозволяє дуже швидко розрахувати не тільки значення зворотної функції щільності ймовірностей, яка визначає відповідні α-квантилі будь-якого розподілу за заданим рівнем значущості α (α/2), але знайти і самі значення диференціальної функції (функції щільності ймовірностей).

Такий розрахунок базується на взаємно однозначній відповідності двох важливіших параметрів будь-якого розподілу – α-квантиля і α, що знаходяться між собою в оберненому зв'язку (рис. Д10).

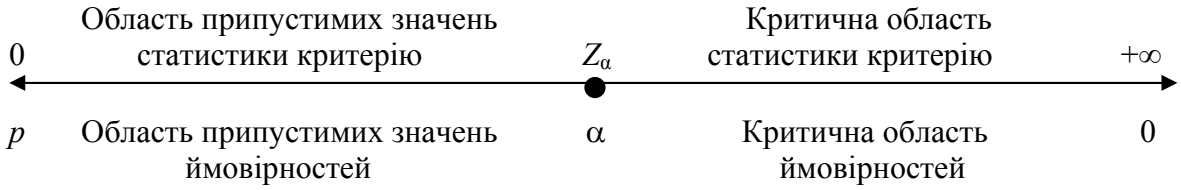


Рис. Д10. Графік взаємної відповідності значень α-квантилей і ймовірностей розподілу статистики критерію

На рис. Д10 верхня вісь всіляких значень статистики критерію зорієнтована праворуч за аналогією з рис. Д9. А нижня вісь ймовірностей настання відповідних подій, що полягають у прийнятті статистикою критерію значень, які більше певної випадкової величини Z, в тому числі й Z_α, зорієнтована ліворуч. Як видно з рис. Д10, області припустимих значень і критичні області на обох осях повністю співпадають. Це відкриває другий (сучасний) шлях використання принципу практичної неможливості настання малоймовірних подій в одиничному випробуванні в ході процедури перевірки статистичних гіпотез. Він полягає в наступному:

• задаються рівнем значущості α ($\alpha/2$ при односторонній перевірці), який і служить межею між малоймовірною й допустимою подіями, по суті, – між критичною областю та областю припустимих значень (нижня частина рис. Д10);

• знаючи теоретичний закон розподілу статистики критерію в припущенні, що H_0 справедлива, за вихідними статистичними даними визначають розрахункове значення цієї статистики, наприклад, точку F_s;

• за допомогою диференціальної функції знаходять фактичну вірогідність p настання події, що полягає в тім, що статистика критерію прийме більше значення за F_s в одиничному випробуванні (так звану p-значущість);

• якщо $p \geq \alpha$, то це означає, що настала подія, яка є припустимою, і, отже, нульова гіпотеза H_0 в дійсності (в гіпотетичній генеральної сукупності) вірна, тому вона не відхиляється (область припустимих значень у нижній частині на рис. Д10));

• якщо $p < \alpha$, то це означає, що настала подія, яка є малоймовірною, і нульова гіпотеза H_0 в гіпотетичній генеральної сукупності є помилковою, отже, вона повинна бути відхилена (критична область у нижній частині на рис. Д10).

Порівнюючи графік на рис. Д6 і графік у нижній частині рис. Д10, неважко помітити, що при сучасному підході критична область завжди одностороння, тому розрахункове значення статистик, що підпорядковуються симетричним розподілам (нормальному, t-розподілу Стьюдента), береться за абсолютною величиною, оскільки значення ймовірності p невід'ємні. Окрім того, слід пам'ятати, що при звичайній орієнтації осі ймовірностей настання відповідних подій, що полягають у прийнятті статистикою критерію значень, які більше певної величини Z, критична область знаходиться не в правій, а в лівій частині графіка, оскільки високим по модулю значенням Z відповідає низька ймовірність і, навпаки (рис. Д11).

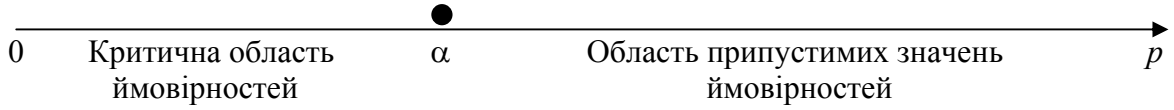


Рис. Д11. Графік *p*-значущості розподілу статистики критерію

На рис. Д12 наводиться блок-схема сучасної процедури перевірки статистичних гіпотез в економічному дослідженні, яка забезпечує отримання повністю тотожних результатів у порівнянні з класичним алгоритмом, представленим на рис. Д8.

На наш погляд, сучасний підхід представляється більш привабливим, оскільки він не потребує розрахунку α-квантилів різних теоретичних законів розподілу і припускає використання персонального комп'ютера. До того ж, більшість стандартних програм математичного забезпечення (редактор Excel, система STATISTICA та ін.) передбачає автоматичне визначення p-значущості розрахункових значень основних статистичних параметрів. Наприклад, при застосуванні програми «Регресія» лістинг містить значимість F моделі в цілому

й *p*-значущість окремих її коефіцієнтів, які, зазвичай, застосовуються при перевірці надійності указаних параметрів.

Рис. Д12. Блок-схема сучасної процедури перевірки статистичних гіпотез

Щоб реалізувати блок-схему сучасної процедури перевірки статистичних гіпотез в даній задачі знайдемо *p*-значущість розрахункового значення статистики критерію Д. Феррара, Р. Глаубера (F_s = 8,77) і побудуємо графік, подібний графіку на рис. Д11.

Для розрахункового значення статистики критерію F_s = 8,77 з числом ступенів вільності $df_1 = m - 1 = 6 - 1 = 5$; $df_2 = n - m = 20 - 6 = 14$ знайдемо ймовірність події, яка полягає в прийнятті статистикою критерію величини, що більше 8,77, тобто *p*-значущість. Для цього на персональному комп'ютері

скористуємося диференціальною функцією F-розподілу Фішера за допомогою наступних дій і команд:

- Натиснути на панелі кнопку *f*x «Майстер функцій»;
- В лівій частині вікна, яка називається «Категорії», вибрати «Статистичні»;
- Потім в правій частині вікна, яка називається «Функції», обрати «НОРМРАСП» («НОРМСТРАСП»), «FРАСП», «ХИ2РАСП», «СТЬЮДРАСП» та ін.;
- Задати розрахункове значення статистики критерію, наприклад, F_s, а також число ступенів вільності df_1, df_2 для диференціальної функції розподілу – Enter.

У задачі, що розглядається, $df_1 = m - 1 = 6 - 1 = 5$; $df_2 = n - m = 20 - 6 = 14$. Звідси в редакторі Excel знаходимо: p-значущість = 0,0006 (рис. Д13).

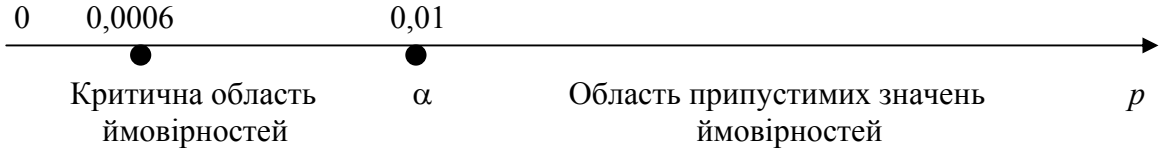


Рис. Д13. Сучасна перевірка гіпотези значущості мультиколінеарності *s*-го чинника-симптому за критерієм Д. Феррара, Р. Глаубера

Оскільки p-значущість = 0,0006 < 0,01 = α, то розрахункове значення ймовірності події $F_s > 18,2874$ потрапляє в критичну область і нульова гіпотеза $H_0 : c_{kk} - 1 = 0$ так само, як і при класичному підході до перевірки гіпотез, відхиляється. З достовірністю 1 - α = 1 - 0,01 = 0,99 або 99 % визнається справедливою альтернатива $H_a : c_{kk} - 1 > 0$. Тому з імовірністю 99 % можна стверджувати, що s-й чинник-симптом є мультиколінеарним, репрезентативним і його можна розглядати як певну оцінку прихованої властивості економічних об'єктів, що вивчаються. Дані висновки повністю тотожні тим, що були отримані вище при застосуванні класичного підходу до перевірки статистичних гіпотез.

Зазначимо тут, що малоймовірна подія може відбутися не тільки тому, що гіпотеза H_0 помилкова, а й з інших причин: малий об'єм вибірки внаслідок нестачі спостережень N, недотримання головних передумов, що лежать в основі застосовуваного критерію і т.п. У цьому випадку відхиливши вірну нульову гіпотезу, роблять помилку першого роду. Щоб цього не сталося, виникає природне прагнення дослідника зменшити ймовірність її настання α.

Проте в результаті зниження α критична область скорочується і зростає область припустимих значень, так як обидві зазначені області – частини одного цілого. В такій ситуації в результаті застосування статистичного критерію все частіше буде відбуватися помилка другого роду: хибні нульові гіпотези не будуть відхилятися, тобто відбудеться зростання β і потужність критерію 1 - β стане низькою. Єдиний спосіб одночасного зниження α і β полягає в зростанні обсягу статистичних спостережень N.

Додаток Е

Алгоритм кластеризації «ФОРЕЛЬ»

Поряд з ієрархічними алгоритмами, серед процедур прямої класифікації еталонного типу, широко застосовуваними є й інші алгоритми, зокрема методи пошуку згущувань економічних об'єктів, які не потребують попереднього завдання числа кластерів. Вони базуються на визначенні кластера типу таксон.

Якщо сукупність об'єктів, що знаходяться в гіперсфері, радіусу T, не змінює свого складу при невеликих відхиленнях центру гіперсфери від центра ваги даної групи, то такий кластер називають таксоном. Одним з найважливіших ітеративних методів кластерного аналізу, що входять в зазначений напрямок, є алгоритм «ФОРЕЛЬ».

Головна мета даного методу – згрупувати багатовимірні об'єкти на основі пошуку згущувань точок у просторі чинників-симптомів. Алгоритм використовує ідею «плаваючої» гіперсфери, яка зупиняється в процесі реалізації алгоритму в місцях локального скупчення точок.

Для здійснення даного алгоритму необхідно задати скалярну величину T – радіус гіперсфери в просторі чинників-симптомів, яка виконує в даній процедурі роль еталона. Нижньою межею T є мінімальна відстань між об'єктами вихідної сукупності, а верхньою – половина відстані між найбільш віддаленими точками. У першому випадку число виділюваних кластерів дорівнює числу об'єктів n, а в другому – єдина гіперсфера містить в собі всю сукупність досліджуваних об'єктів.

Робота алгоритму (див. рис. Е1) починається з вибору довільної точки Z_0 (z_{01}, z_{02}, … , z_{0m}), прийнятої за центр гіперсфери заданого радіуса T. Потім визначається сукупність точок, що потрапили всередину цієї гіперсфери. Далі розраховуються координати центра ваги утвореної сукупності точок: $\overline{Z}_1$ ($\overline{z}_{11}$, $\overline{z}_{12}$, … , $\overline{z}_{1m}$). У цьому випадку мають справу вже не з координатами конкретного об'єкта вихідної сукупності, а з вектором середніх значень.

Точка $\overline{Z}_1$ вважається новим центром гіперсфери і відбираються всі об'єкти, що потрапили всередину її. Як і на попередньому кроці, розраховується центр ваги $\overline{Z}_2$ нової гіперсфери за всіма чинниками-симптомами об'єктів, що потрапили в неї. Точка $\overline{Z}_2$ з новими координатами приймається за новий центр гіперсфери і т. д.

При цьому склад точок, що потрапили всередину гіперсфери може вирізнятися від сукупності точок, які опинилися всередині попередньої, що виражається у відмінності від нуля відстані між центром гіперсфери і центром ваги утвореної потім сукупності об'єктів, наприклад, $d(Z_0, \overline{Z}_1) \neq 0$.

Повтор даної процедури веде до руху центру гіперсфери в область згущення точок. «Плавання» гіперсфери продовжується до тих пір, поки черговий перерахунок координат центру гіперсфери не дасть той же самий результат, що і на попередньому кроці, тобто $d(\overline{Z}_i, \overline{Z}_{i-1}) = 0$.

При цьому точки, що потрапили в гіперсферу, приймаються за окремий кластер і з наступного аналізу виключаються. Серед решти об'єктів знову

довільно вибирається нова точка і алгоритм повторюється з самого початку, тобто з випадкового об’єкта – центру гіперсфери радіуса *T*.

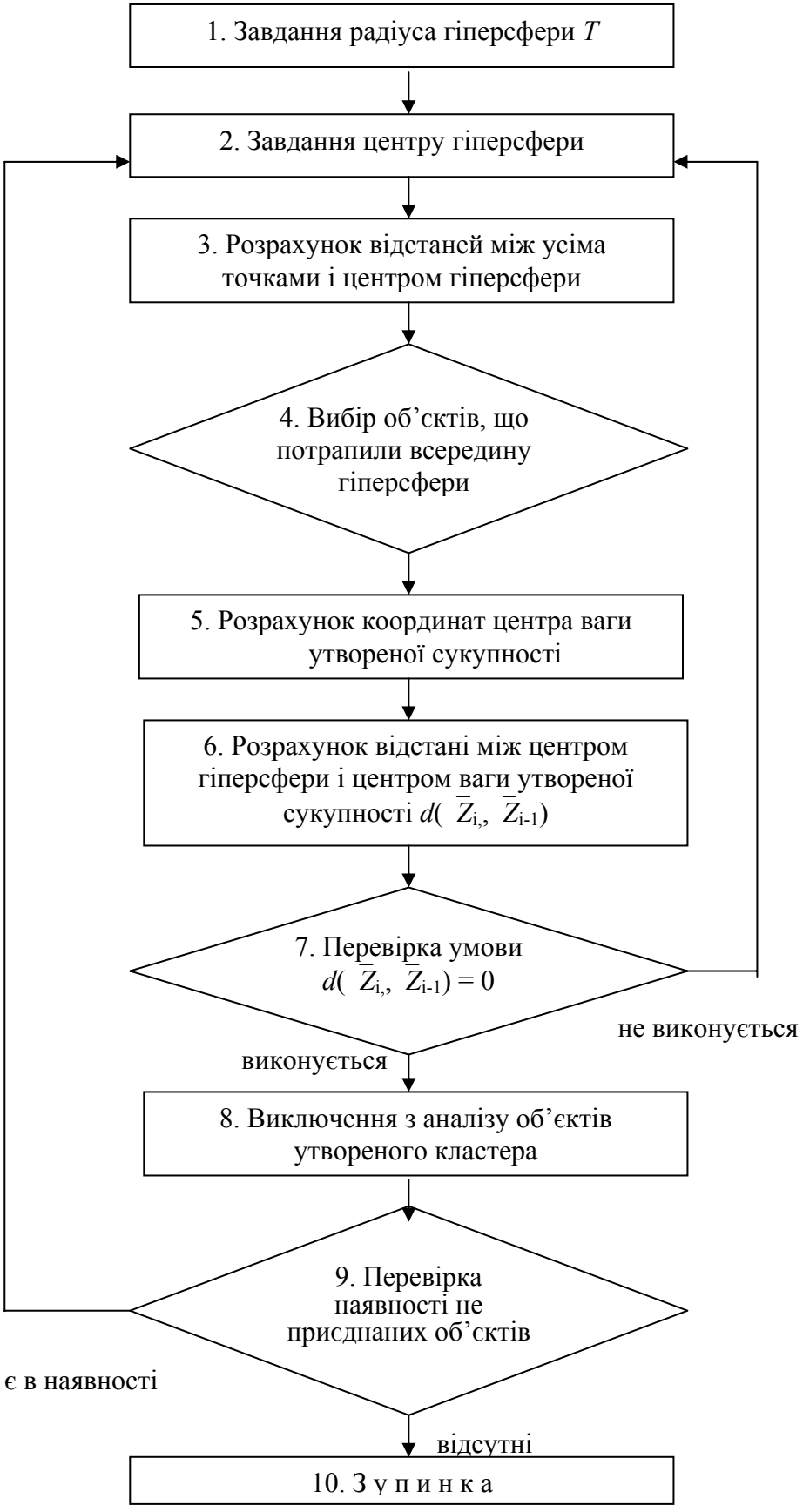


Рис. Е1. Блок-схема алгоритму «ФОРЕЛЬ»

Таким чином, робота алгоритму «ФОРЕЛЬ» завершується за кінцеве число кроків, і всі точки (об'єкти) розподіляються по кластерам. Описана процедура в двомірному просторі ознак графічно зображена на рис. Е2.

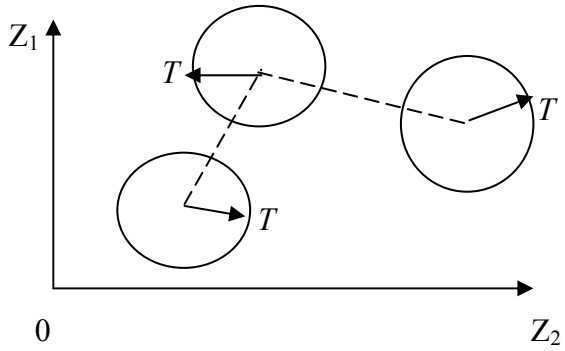


Рис. Е2. Розподіл об'єктів по кластерам на основі алгоритму «ФОРЕЛЬ»

Штриховий лінією на рис. Е2 показаний шлях, який «пропливла» гіперсфера в просторі чинників-симптомів z_1, z_2. У даному умовному прикладі спостерігаються три чітко виражених скупчення досліджуваних об'єктів, накритих гіперсферу, заданого радіуса T.

Очевидно, що результат роботи алгоритмів даного типу, зокрема число утворюваних кластерів R, багато в чому залежить від вибору величини T. Досвід проведення багатовимірного угруповання показує, що для отримання оптимального числа кластерів типу таксонів всю процедуру обчислень слід повторювати при різних значеннях T, щоразу змінюючи довжину радіуса на невелику величину.

Сигналом стійкості розбиття об'єктів є утворення постійного числа кластерів для ряду послідовних значень T і досить різка зміна R на попередньому і наступному кроках. Цієї рекомендації відповідає район «плато» T_1T_2 на рис. Е3.

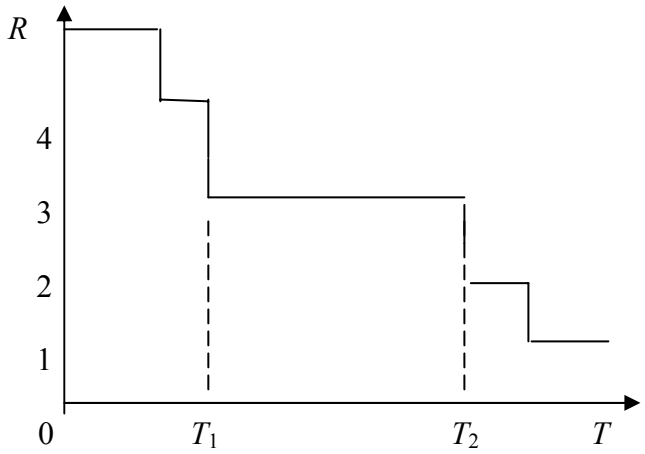


Рис. Е3. Зворотна залежність числа створюваних кластерів *R* від величини радіуса *T*

Приклад. Нехай є сукупність 10 магазинів, які за рік характеризуються такими чотирма економічними ознаками:

x_1 – виробничі витрати, тис. грн.;

x_2 – продуктивність праці, тис. грн.;

x_3 – річний обіг, млн. грн.;

x_4 – питома вага втрат товарів, %.

Статистична інформація по досліджуваним торговельним підприємствам наведена в табл. Е1.

Таблиця Е1

Вихідні економічні показники 10 магазинів

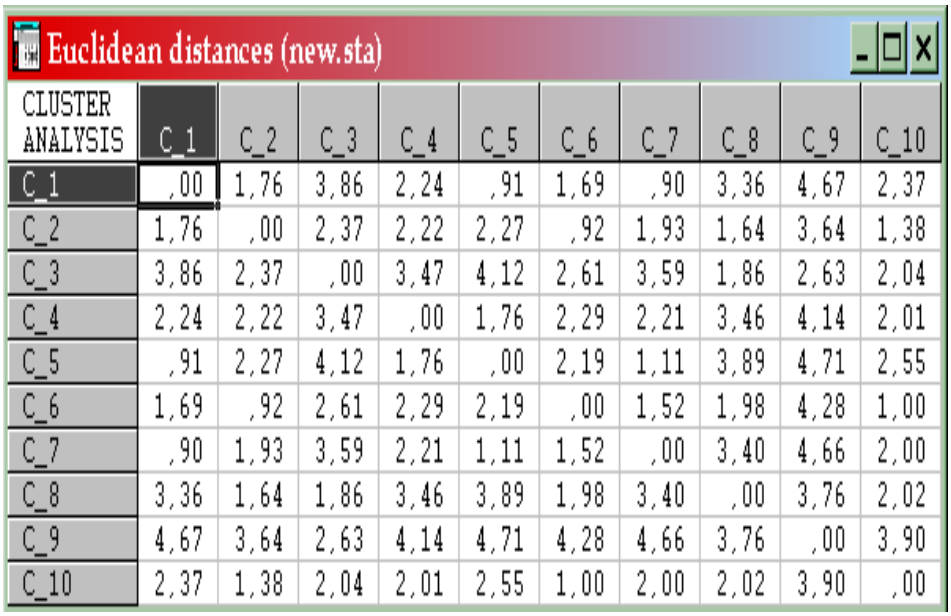
№ п/п	x_1	x_2	x_3	x_4
1	6711	1821	46	0,85
2	12578	3597	50	1,30
3	24845	5076	52	1,50
4	11262	3923	40	1,60
5	6932	1859	42	1,10
6	14406	3494	49	0,02
7	8782	1940	45	0,07
8	13472	5074	55	1,08
9	16264	4195	50	5,02
10	18583	4295	47	0,14

За наведеними економічними ознаками здійснимо с допомогою системи STATISTICA багатовимірне угруповання зазначених магазинів на основі алгоритму «ФОРЕЛЬ». Хоча система STATISTICA і не реалізує даний алгоритм, все ж скористаємося нею для знаходження деяких проміжних розрахунків, а саме, матриці стандартизованих даних (рис. Е4), вихідної матриці відстаней ***D*** між усіма об’єктами (рис. Е5), а також для знаходження відстаней до центрів тяжіння груп і центрів «плаваючої» гіперсфери в процесі реалізації алгоритму «ФОРЕЛЬ» (див. блок-схему на рис. Е1).

Data: NEW.STA 4v * 10c

NUM VAL	1 VAR1	2 VAR2	3 VAR3	4 VAR4
1	-1,358	-,985	-,352	-,290
2	,055	-,491	,527	,022
3	1,233	1,518	,967	,161
4	,315	-,541	-1,670	,230
5	-1,328	-,837	-1,231	-,116
6	-,027	-,413	,308	-,865
7	-1,264	-,306	-,571	-,830
8	1,231	-,198	1,626	-,130
9	,531	2,035	,527	2,600
10	,611	,217	-,132	-,782

Рис. Е4. Вихідна матриця стандартизованих даних *Z*



Euclidean distances (new.sta)

CLUSTER ANALYSIS	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10
C_1	,00	1,76	3,86	2,24	,91	1,69	,90	3,36	4,67	2,37
C_2	1,76	,00	2,37	2,22	2,27	,92	1,93	1,64	3,64	1,38
C_3	3,86	2,37	,00	3,47	4,12	2,61	3,59	1,86	2,63	2,04
C_4	2,24	2,22	3,47	,00	1,76	2,29	2,21	3,46	4,14	2,01
C_5	,91	2,27	4,12	1,76	,00	2,19	1,11	3,89	4,71	2,55
C_6	1,69	,92	2,61	2,29	2,19	,00	1,52	1,98	4,28	1,00
C_7	,90	1,93	3,59	2,21	1,11	1,52	,00	3,40	4,66	2,00
C_8	3,36	1,64	1,86	3,46	3,89	1,98	3,40	,00	3,76	2,02
C_9	4,67	3,64	2,63	4,14	4,71	4,28	4,66	3,76	,00	3,90
C_10	2,37	1,38	2,04	2,01	2,55	1,00	2,00	2,02	3,90	,00

Рис. Е5. Матриця *D* евклідових відстаней між усіма точками (об'єктами)

Аналіз даних матриці ***D*** на рис. Е5 показує, що min$d = d_{1,7} = 0{,}9$, а max$d = d_{5,9} = 4{,}71$. Отже, гіперсфера, радіусу $T = 0.9$, виділить рівно 10 кластерів, що містять по одному об'єкту, а гіперсфера, радіусу $T = 2{,}36$, «накриє» всю вихідну сукупність, тобто виділить один кластер, що містить усі 10 об'єктів. Ясно, що зазначені тривіальні рішення навряд чи представляють якої-небудь практичний інтерес. Приймемо за радіус гіперсфери значення T з інтервалу [0,9; 2,36]. Нехай $T = 1{,}5$.

Виберемо довільний об'єкт (наприклад, магазин № 1) в ролі вихідного центру гіперсфери Z_0(-1,358 -0,985 -0,352 -0,290).

Для знаходження точок, що потрапили всередину отриманої гіперсфери, слід проаналізувати всі відстані від них до об'єкта № 1, тобто перший рядок (стовпець) матриці ***D*** на рис. Е5. Оскільки $d_{1,7} = 0{,}9 < 1{,}5$ і $d_{1,5} = 0{,}91 < 1{,}5$, то, очевидно, що поряд з точкою № 1 всередину гіперсфери потрапили також точки № 5 і № 7.

По матриці стандартизованих ознак ***Z*** (табл. на рис. Е4) в редакторі Excel розрахуємо центр ваги утвореної групи: $\overline{Z}_1$ (-1.317 -0.709 -0.718 -0.412).

Порівняння точок Z_0 і $\overline{Z}_1$ показує, що відбувся деякий рух гіперсфери, радіуса $T = 1.5$. Приймемо точку $\overline{Z}_1$ за новий центр гіперсфери і утворюємо нову матрицю стандартизованих даних (рис. Е6).

Щоб з'ясувати, які точки потрапили всередину гіперсфери з центром в точці $\overline{Z}_1$, знайдемо в системі STATISTICA евклідові відстані від даної точки до об'єктів з №№ 2, 3, 4, 6, 8, 9, 10 (рис. Е7).

Аналіз даних першого рядка (стовпця) матриці евклідових відстаней на рис. Е7 показує, що всі $d_i > 1{,}5$. Отже, жоден новий об'єкт не потрапив усередину гіперсфери.

Data: NEW.STA 4v * 8c

NUMERIC VALUES	1 VAR1	2 VAR2	3 VAR3	4 VAR4
(1-5-7)	-1,3167	-,7091	-,7178	-,4120
2	,0554	-,4913	,5274	,0222
3	1,2328	1,5178	,9668	,1608
4	,3149	-,5406	-1,6700	,2300
6	-,0266	-,4131	,3076	-,8647
8	1,2312	-,1981	1,6260	-,1303
9	,5314	2,0352	,5274	2,5997
10	,6110	,2174	-,1318	-,7816

Рис. Е6. Перетворена матриця Z_1

Euclidean distances (new.sta)

Continue...	(1-5-7	2	3	4	6	8	9	10
(1-5-7)	,00	1,92	3,82	2,00	1,73	3,51	4,64	2,25
2	1,92	,00	2,37	2,22	,92	1,64	3,64	1,38
3	3,82	2,37	,00	3,47	2,61	1,86	2,63	2,04
4	2,00	2,22	3,47	,00	2,29	3,46	4,14	2,01
6	1,73	,92	2,61	2,29	,00	1,98	4,28	1,00
8	3,51	1,64	1,86	3,46	1,98	,00	3,76	2,02
9	4,64	3,64	2,63	4,14	4,28	3,76	,00	3,90
10	2,25	1,38	2,04	2,01	1,00	2,02	3,90	,00

Рис. Е7. Евклідові відстані від точки $\bar{Z}_1$ до інших об'єктів (рядок 1)

Тому наступний центр сукупності збігається з центром попередньої гіперсфери ($\bar{Z}_2 = \bar{Z}_1$) і відстань $d(\bar{Z}_1, \bar{Z}_2) = 0$. Це означає, що об'єкти з №№ 1, 5, 7 утворюють кластер типу таксон і з подальшого аналізу виключаються.

З решти магазинів виберемо знову довільну точку (наприклад, магазин № 2) у ролі нового вихідного центру гіперсфери Z'_0(0,0554 -0,4913 0,5274 0,0222). Для знаходження точок, що потрапили всередину отриманої гіперсфери, слід проаналізувати всі відстані від них до об'єкта № 2, тобто другий рядок (стовпець) матриці евклідових відстаней на рис. Е7. Оскільки $d_{2,6} = 0,92 < 1,5$ і $d_{2,10} = 1,38 < 1,5$, то, очевидно, що поряд з точкою № 2 всередину гіперсфери потрапили також точки № 6 та № 10.

По матриці стандартизованих ознак $\boldsymbol{Z_1}$ (рис. Е6) в редакторі Excel розрахуємо центр ваги утвореної групи: $\bar{Z}_3$(0,2133 -0,2290 0,2344 -0,5414).

Порівняння точок Z'_0 і $\overline{Z}_3$ показує, що відбувся деякий рух гіперсфери, радіусу $T = 1{,}5$. Приймемо точку $\overline{Z}_3$ за новий центр гіперсфери і утворимо нову матрицю стандартизованих даних (рис. Е8).

Data: NEW.STA 4v * 5c

NUMERIC VALUES	1 VAR1	2 VAR2	3 VAR3	4 VAR4
(2-6-10)	,213	-,229	,234	-,541
3	1,233	1,518	,967	,161
4	,315	-,541	-1,670	,230
8	1,231	-,198	1,626	-,130
9	,531	2,035	,527	2,600

Рис. Е8. Перетворена матриця Z_2

Щоб з'ясувати, які точки попали всередину гіперсфери з центром в точці $\overline{Z}_3$, в системі STATISTICA знайдемо евклідові відстані від даної точки до об'єктів з №№ 3, 4, 8, 9 (рис. Е9).

Euclidean distances (new.sta)

Continue...	(2-6-1	3	4	8	9
(2-6-10)	,00	2,26	2,08	1,77	3,90
3	2,26	,00	3,47	1,86	2,63
4	2,08	3,47	,00	3,46	4,14
8	1,77	1,86	3,46	,00	3,76
9	3,90	2,63	4,14	3,76	,00

Рис. Е9. Евклідові відстані від точки $\overline{Z}_3$ до інших об'єктів (рядок 1)

Аналіз даних першого рядка (стовпця) матриці евклідових відстаней на рис. Е9 показує, що всі $d_i > 1{,}5$, отже, жоден новий об'єкт не потрапив усередину гіперсфери. Тому наступний центр сукупності збігається з центром попередньої гіперсфери ($\overline{Z}_4 = \overline{Z}_3$) і відстань $d(\overline{Z}_4, \overline{Z}_3) = 0$. Це означає, що об'єкти з №№ 2, 6, 10 утворюють таксон і з подальшого аналізу виключаються.

З решти магазинів виберемо знову довільну точку (наприклад, магазин № 3) в ролі нового вихідного центру гіперсфери Z''_0(1,2328 1,5178 0,9668 0,1608).

Для знаходження точок, що потрапили всередину отриманої гіперсфери, слід проаналізувати всі відстані від них до об'єкта № 3, тобто другий рядок (стовпець) матриці евклідових відстаней на рис. Е9.

Оскільки всі $d_i > 1{,}5$, то жоден новий об'єкт не потрапив усередину гіперсфери. Тому наступний центр сукупності збігається з центром попередньої гіперсфери ($\overline{Z}_3 = Z''_0$) і відстань $d(\overline{Z}_3, Z''_0) = 0$. Це означає, що об'єкт № 3 утворює самостійний кластер типу таксон і з подальшого аналізу виключається.

Дослідження рядків 4, 8, 9 табл. евклідових відстаней на рис. Е9 показує, що відповідні об'єкти, прийняті за вихідні центри тяжкості гіперсфери, радіусу

T = 1,5, також утворюють окремі кластери, тому ніякі інші точки не потрапляють усередину неї (d_i > 1,5).

Таким чином, алгоритм «ФОРЕЛЬ» призводить до наступного угрупування вихідної сукупності магазинів: (3), (4), (8), (9), (1, 5, 7), (2, 6, 10).

Аналіз даних матриць евклідових відстаней (рис. Е5, Е7, Е9), отриманих у процесі виділення таксонів, показав, що знайдене рішення задачі кластеризації магазинів на основі алгоритму «ФОРЕЛЬ» є досить стійким. Воно буде виходити щоразу, коли значення радіуса гіперсфери буде вибиратися з інтервалу [0,93; 1,63]. Довжина «плато» (див. рис. Е3) з постійним числом виділених кластерів R = 6 становить 0,7 або 47,9 % всього діапазону можливих значень параметра T.

Очевидно, що алгоритм «ФОРЕЛЬ» так само, як і інші процедури кластерного аналізу, не вільний від певних елементів вольових рішень, пов’язаних з вибором величини радіуса гіперсфери T. По всій видимості, в багатовимірних методах немає такого способу кластеризації, який зміг би абсолютно об’єктивно виявити особливості структури досліджуваної статистичної сукупності об’єктів, як немає загальновизнаного визначення поняття її кількісної однорідності. Кожен автор трактує це поняття по-своєму і, відповідно, в праві відстоювати власні підходи і результати розбиття вихідної множини досліджуваних економічних об’єктів.

Додаток Ж

Принцип балансу змінних

У сучасній науці розроблено ряд методів (багатовимірний статистичний аналіз, розрахунок послідовних різниць і характеристик приросту), які володіють певними достоїнствами та недоліками і часто дають цілком задовільні результати при оцінці латентних ознак і прогнозуванні соціально-економічних показників, що розглядаються ізольовано. Однак ситуація принципово змінюється, коли здійснюється одночасна оцінка (прогнозування) декількох взаємозалежних змінних типу продуктивність, продуктивна сила і інтенсивність праці робітників; чисельність всього населення, чоловіків і жінок; рівень народжуваності, смертності та природного приросту населення.

В соціально-економічних дослідженнях такі показники зустрічаються достатньо часто, оскільки різні сторони процесу відтворення в економіці, а також і відтворення населення, тісно взаємопов'язані, утворюють єдиний комплекс, що розвивається в процесі суспільного життя. У цьому випадку наведені вище методи вибору оціночної моделі або предиктора (системи прогнозних рівнянь, трендів тощо) не виключають небезпеку отримання помилкового результату, що «розходиться», порушуючи фундаментальні якісні співвідношення і взаємозв'язки. Як би не були глибоко і ретельно розроблені прийоми екстраполяції окремих ізольованих рядів динаміки, вони не можуть застрахувати від отримання «віяла» прогнозів, що вирізняється абсурдністю передбачених значень досліджуваних ознак з точки зору їх внутрішньої залежності.

Щоб уникнути зазначених недоліків при комплексному дослідженні соціально-економічних показників, між якими спостерігаються об'єктивні взаємозв'язки, ми пропонуємо застосовувати принцип балансу змінних. Він може бути сформульований наступним чином: остаточний висновок про придатність тих чи інших методів оцінювання (систем прогнозних рівнянь) визначається ступенем виконання для оцінених (прогнозованих) значень взаємозв'язаних змінних деякого балансового співвідношення, справедливого для них апріорі, наприклад, у періоді передісторії.

Принцип балансу змінних уперше був висунутий в теорії самоорганізації як один із напрямів технічної кібернетики, очолюваний академіком О.Г.Івахненком, при моделюванні складних імовірнісних систем за допомогою метода групового обліку аргументів. Цікаві результати застосування принципу балансу змінних і побудованих на його основі критеріїв у сфері економіки та демографії наведені в роботах автора даної монографії, а також у кандидатській дисертації О.І. Яшкіної (2006 р., науковий керівник О.Г. Янковий) на тему «Методи прогнозування взаємопов'язаних показників соціально-економічного розвитку України» (див. літературу до додатку Ж).

За своїм змістом принцип балансу змінних є головним у всій сукупності принципів математико-статистичного моделювання і прогнозування, так як він базується на абсолютно достовірної інформації про оцінюваний (майбутній)

стан соціально-економічної системи: які б не були окремі значення змінних, для них має дотримуватися певний баланс, що випливає із сутності досліджуваних ознак. Тому критерії, побудовані на основі принципу балансу змінних, повинні займати пріоритетне становище в ієрархії критеріїв методів оцінювання прихованих властивостей об'єктів і відбору форм моделей при прогнозуванні. Задоволення їм є необхідною і достатньою умовою отримання найбільш точних оцінок латентних ознак і достовірних прогнозів.

Раціональність такого підходу не викликає сумніву, бо про приховані властивості об'єктів, що підлягають оцінці, а також про майбутнє ми знаємо точно тільки те, що в ньому збережуться ті ж самі взаємозв'язки між ознаками, які мали місце в минулому і в сьогоденні. Принцип балансу змінних виконує роль зовнішнього доповнення, що несе нову інформацію про досліджувані соціально-економічні процеси. Це означає, що він може і повинен застосовуватися поряд з уже розробленими математико-статистичної теорією і апробованими на практиці методами оцінювання латентних ознак і вибору форм моделей (рівнянь трендів різної складності) при визначенні прогнозних предикторів.

У даній роботі розглянемо побудову та використання критерію балансу змінних на прикладі наступного взаємозв'язку між соціально-економічними показниками:

$$Y_1 = Y_2 + Y_3. \qquad (Ж1)$$

У ролі Y_1 може виступати показник народжуваності, Y_2 – смертності, Y_3 – природного приросту населення країни (регіону, міста тощо). Окрім того, Y_1, Y_2, Y_3 можна розглядати як логарифми продуктивності, продуктивної сили та інтенсивності праці робітників підприємства (див. співвідношення (3.2)).

Нехай завдання полягає в отриманні найкращого комплексного прогнозу змінних Y_1, Y_2, Y_3 на деякий період упередження L на основі відомих досліджуваних показників за певний період передісторії N. Процес прогнозування з застосуванням критерію балансу змінних в цьому випадку складається з таких головних етапів.

1. *Визначення можливих функцій-кандидатів, які найбільш точно описують досліджувані ряди динаміки.*

На цьому етапі широко застосовуються традиційні методи статистичного моделювання: якісний теоретичний аналіз, візуальне вивчення графіків, розрахунок послідовних різниць, характеристик приросту та інших параметрів, включаючи критерії математичної статистики. У результаті на першому етапі з усього переліку опорних функцій, число яких може бути досить великим, відбираються декілька головних (n_i, i = 1, 2, 3), які найбільш адекватно описують взаємозв'язані процеси, що вивчаються.

Ряди динаміки кожного показника, взяті за період передісторії, можна представити за допомогою трендових моделей типу

$$Y_i = f_i(X) + \varepsilon_i, \qquad (Ж2)$$

де X – фактор часу;

ε_i – випадкова компонента.

За методом найменших квадратів визначаються параметри кожної з відібраних функцій (Ж2), здійснюється екстраполяція і розраховуються прогнозні значення показників на заданий період упередження $L = X1 - X2$.

2. *Послідовний перебір обмеженого числа головних функцій, виділених на попередньому етапі.*

В якості оцінки кожної комбінації виступає ступінь дотримання вихідного балансового співвідношення для всіх точок періоду упередження $L = X1 - X2$. Число всіх можливих комбінацій визначається добутком $n_1 \times n_2 \times n_3$ (n_i – кількість початкових функцій, виділених для відповідного взаємопов'язаного показника).

Найкраща комбінація відповідає мінімальному значенням критерію балансу змінних:

$$B_j = \frac{\sum_{X1}^{X2} (\hat{Y}_{1X} - \hat{Y}_{2X} - \hat{Y}_{3X})^2}{\sum_{X1}^{X2} \hat{Y}_{1X}^2}. \qquad (Ж3)$$

Тобто в якості оцінки кожної комбінації прогнозних трендів виступає ступінь дотримання вихідного балансового співвідношення для всіх точок періоду упередження.

B_j характеризує сумарний відносний розбаланс комплексного прогнозу досліджуваних показників. Чим ближче він до нуля, тим правильніше обраний предиктор і, навпаки. Отже, критерій (Ж3) дозволяє здійснити оптимізацію визначення предіктора на множині виділених вихідних функцій. Якщо на першому етапі не пропущено жодна придатна функція, яка достатньо точно описує динаміку досліджуваних показників, то множина значень B_j у міру перебору форм трендів обов'язково проходить через мінімум, який визначає оптимальну комбінацію.

3. *Вибір оптимальної довжини періоду передісторії*. Відомо, що динаміка соціально-економічних показників схильна до коливань, обумовлених дією суб'єктивних, природно-кліматичних та інших чинників, що відображені в випадковій компоненті ε_i. Тому дуже важливо правильно визначити довжину періоду передісторії, вловити сучасні тенденції, щоб прогнозоване значення не суперечило логіці сучасного розвитку досліджуваного процесу.

З цією метою застосовується ітеративний підхід. Його суть полягає в утворенні нових усічених з початку рядів динаміки шляхом виключення з вихідних рядів G ($G = 1, 2, ..., N/2$) перших рівнів. По кожному усіченому ряду будуються трендові моделі, визначаються прогнозні значення і знаходяться величини B_j. Іншими словами, досліджується поведінка критерію балансу змінних при поступовому скороченні довжини періоду передісторії N.

Такий підхід дозволяє здійснити адаптацію трендового предиктора до нових умов соціально-економічного процесу, звільнитися від впливу «старих» спостережень, посилити роль останніх точок ряду динаміки, що вкрай важливо

при короткостроковому і середньостроковому прогнозуванні досліджуваних взаємопов'язаних показників.

Тут слід відмітити наступний момент: якщо всі трендові моделі предиктора представлені однаковими математичними функціями $f_i(X)$, наприклад, лінійними або параболами 2-го ступеню, то зазвичай, спостерігається взаємозв'язок між самими коефіцієнтами трендів, який збігається з балансовим співвідношенням (Ж1). Як показали практичні розрахунки, розбіжності проявляються в четвертому і далі знаку після коми. Звідси така висока збалансованість вирівняних значень змінних і для точок періоду упередження.

Ми вважаємо, що зазначені факти – глибокий мінімум критерію $B_j(G)$ і дотримання вихідного балансового співвідношення для самих коефіцієнтів знайденого предиктора свідчать про прояв сформованої в досліджуваний період часу певної економіко-математичної закономірності розвитку взаємозв'язаних соціально-економічних показників, що вивчаються.

4. *Визначення точкових та інтервальних значень прогнозу*. Цей етап завершує процес комплексного прогнозування взаємопов'язаних соціально-економічних показників і являє собою звичайну екстраполяцію побудованих трендів і розрахунок довірчих інтервалів за відомими формулами [123; 127-129].

Екстраполяція заснована на гіпотезі про інерційність розвитку соціально-економічних систем, тобто на припущенні про те, що в недалекій перспективі виявлені в минулому зв'язки і закономірності кардинально не зміняться і будуть діяти деякий час і в майбутньому. Слід зазначити, що таке припущення досить реально, особливо для короткострокового і середнє строкового періодів, оскільки докорінні зміни показників таких систем, наприклад, підприємств, зазвичай, вимагають значних зусиль, коштів і досить тривалого часу.

Існують об'єктивні взаємозв'язки між числом оцінюваних коефіцієнтів тренда і довжиною досліджуваного ряду динаміки N, а також між величиною N і періодом упередження L. Доведено, що число спостережень N повинно в 3-4 рази перевищувати число оцінюваних коефіцієнтів тренда m. Для лінійного рівняння це число дорівнює двом (a_0, a_1), тому, щоб побудувати лінійний тренд $\hat{Y} = a_0 + a_1X$, необхідно володіти не менше 6-8 рівнями часового ряду. При цьому повинна виконуватися умова: $L \leq N/3$, тобто довжина періоду упередження не може перевищувати третини довжини періоду передісторії. Порушення зазначених співвідношень призводить до отримання ненадійних трендових моделей і, відповідно, помилкових прогнозних оцінок.

До кількісних прогнозів будь-якого соціально-економічного показника висуваються дві головні вимоги: точність і достовірність. Точність характеризується ступенем варіювання прогнозного значення $\hat{Y}_{N+L}$, яка вимірюється граничною помилкою прогнозу Δ. Чим менше Δ, тим точніше прогноз і, навпаки. Нульове значення Δ відповідає так званому точковому прогнозу у вигляді одного числа – точки на числовій осі. Точковий прогноз уважається максимально точним.

Достовірність прогнозу відображає ймовірність збігу майбутнього фактичного значення досліджуваного показника Y_{N+L} з прогнозним значенням $\hat{Y}_{N+L}$. Очевидно, що зі зростанням граничної помилки прогнозу Δ його достовірність збільшується, а зниження Δ (зростання точності прогнозу) призводить до зменшення достовірності прогнозу. Отже, точковий прогноз (Δ = 0) має мінімальну достовірність, так як точний збіг фактичного майбутнього значення Y_{N+L} з прогнозним значенням $\hat{Y}_{N+L}$ є малоймовірним. Це означає, що характеристики точності і достовірності прогнозу знаходяться між собою в обернено пропорційній залежності: зростання одного параметра викликає зменшення іншого і, навпаки.

Точкові прогнози знаходяться на базі рівнянь тренда, які відповідають найкращому предиктору за критерієм балансу змінних $B_j(G)$. З цією метою в моделі трендів підставляють замість X майбутні значення фактору часу $X = N + L$, які відповідають цільовим періодам упередження $L = X1 - X2$. Наприклад, для лінійного тренда точковий прогноз має наступний загальний вигляд:

$$\hat{Y}_{N+L} = a_0 + a_1(N + L). \qquad (Ж4)$$

Як було зазначено вище, отриманий точковий прогноз $\hat{Y}_{N+L}$ вирізняється максимальною точністю і мінімальною достовірністю. Тому в практиці прогнозування соціально-економічних показників, зазвичай, йдуть на компроміс: на основі точкового прогнозу будують інтервальний прогноз у вигляді двох чисел – нижньої і верхньої меж довірчого інтервалу із заздалегідь заданою вірогідністю потрапляння в нього фактичного значення Y_{N+L}. Межі довірчого інтервалу прогнозу розраховуються за такою загальною схемою:

$$\hat{Y}_{N+L} \pm \Delta. \qquad (Ж5)$$

Іншими словами, в центрі довірчого інтервалу знаходиться точковий прогноз $\hat{Y}_{N+L}$, а його ширина дорівнює 2Δ. В результаті отримання довірчого інтервалу за схемою (Ж5) точність прогнозу падає, оскільки розширюються його межі, але при цьому забезпечується заздалегідь задана достовірність прогнозу. Для простіших лінійного і параболічного (2-го ступеню) трендів формули розрахунку граничної помилки прогнозу Δ наводяться в роботі [123, с.23, 24].

Але, слід мати на увазі, що автоматичний розрахунок Δ в стандартних програмах кореляційно-регресійного аналізу редактора Excel не передбачений. А вручну здійснення обчислень величини Δ є дуже трудомісткою задачею. При великих m, або у випадку складних математичних функцій $f_i(X)$ подібний розрахунок зовсім не представляється можливим. Тому в реальних соціально-економічних дослідженнях з цією метою, зазвичай, користуються системою програм STATISTICA [11; 12], модуль «Множинна регресія» (Multiple Regression), в якому автоматично знаходяться величини точкового прогнозу $\hat{Y}_{N+L}$, нижньої ($\hat{Y}_{N+L}$ - Δ) та верхньої ($\hat{Y}_{N+L}$ + Δ) меж його довірчого інтервалу.

Література до додатку Ж

1. Янковой А. Г. Принцип баланса переменных в задачах долгосрочного прогнозирования взаимосвязанных социально-экономических показателей / К.В. Колузанов, А. Г. Янковой // Моделі і методи оптимального планування в сільському господарстві : Одеса, ОСГІ. – 1989. – С.35-40.

2. Янковой А. Г. Принцип баланса переменных в демографических исследованиях / А. Г. Янковой // Економіка Радянської України. – 1990. – № 6. – С. 84-86.

3. Янковой А. Г. Краткосрочное прогнозирование взаимосвязанных показателей на базе критерия баланса переменных / А. Г. Янковой. – Тезисы докл. Всесоюзной науч. конф. «Организация управления производством в новых условиях хозяйствования». – Одесса : ОПИ. – 1990. – С.148-150.

4. Янковой А. Г. Опыт комплексного прогнозирования показателей естественного движения населения / А. З. Подгорный, А. Г. Янковой. – Тезисы докл. Всесоюзной науч. конф. по исторической демографии. – Донецк, ДДУ. – 1991. – С. 18-19.

5. Янковой А. Г. Прогнозирование социально-экономических показателей на основе принципа баланса переменных / А. Г. Янковой // Экономика и математические методы, т. 29, вып. 1. – 1993. – С. 108-118.

6. Янковой А. Г. Комплексное прогнозирование показателей естественного движения населения / А. З. Подгорный, А. Г. Янковой, Е. В. Коробкина, Т. В. Погорелова // Демографічні дослідження. Вип. 17, 1993.- К.: Наукова думка. – С. 90-99.

7. Янковой А. Г. Прогнозирование демографических показателей Украины / А. Г. Янковой // Ринкова економіка : Сучасна теорія і практика управління, т. 5. – Одеса : Астропринт. – 2002. – С. 34-49.

8. Янковий О. Г., Яшкіна О. І. До проблеми вибору математичної форми трендів при прогнозуванні соціально-економічних показників / О. Г. Янковий, О. І. Яшкіна // Вісник соціально-економічних досліджень. – Одеса, ОДЕУ. – 2003. – № 14. – С. 341-346.

9. Янковий О. Г., Яшкіна О. І. Прогнозування взаємопов’заних показників соціально-економічного розвитку України / О. Г. Янковий, О. І. Яшкіна // Статистика України. – 2006. – № 3. – С. 61-65.

10. Яшкіна О. І. Методи прогнозування взаємопов’язаних показників соціально-економічного розвитку України / Автореф. дис. на здобуття наук. ступеня канд. екон. наук за спец. 08.02.03 – організація управління, планування і регулювання економікою / О. І. Яшкіна. – Одеса, ОДЕУ, 2006. – 19 с.

О.Г. Янковий

ЛАТЕНТНІ ОЗНАКИ В ЕКОНОМІЦІ

МОНОГРАФІЯ

Підписано до друку 12.05.2015 р. Формат 60×84/16.
Папір офсетний. Ум.-друк. арк. 9,77. Замовлення № 517
Тираж 300 прим.

Друкарня «Атлант» ВОІ СОІУ
65029, м. Одеса, Ольгієвський узвіз, 8
тел.:728-45-71
e-mail: ev_atlant@mail.ru
Свідоцтво:23179617