

1. Факторний аналіз

При всьому різноманітті методи факторного аналізу мають загальний алгоритм розв'язання, представлений на рис. 1.

Даний алгоритм починається побудовою матриці вихідних даних X , а завершується отриманням матриць факторного відображення і значень факторів A і F . З урахуванням прийнятих позначень, де n – число спостережень, m – число аналітичних ознак X , r – число значущих узагальнених ознак (латентних факторів). На схемі (рис.1) показана розмірність матриць даних для кожного алгоритмічного кроку. Факторний аналіз призначений для виявлення та кількісного опису узагальнюючих характеристик, які досить описують весь набір вихідних і результуючих показників, відображаючи тим самим закономірності розвитку досліджуваної системи. Отримані в результаті розрахунків моделі відображають механізм внутрішніх взаємозв'язків стану діяльності підприємства. Розпізнавання основних факторів, визначення для них назви проводиться на основі вагових коефіцієнтів a_{jr} з матриці відображення A .

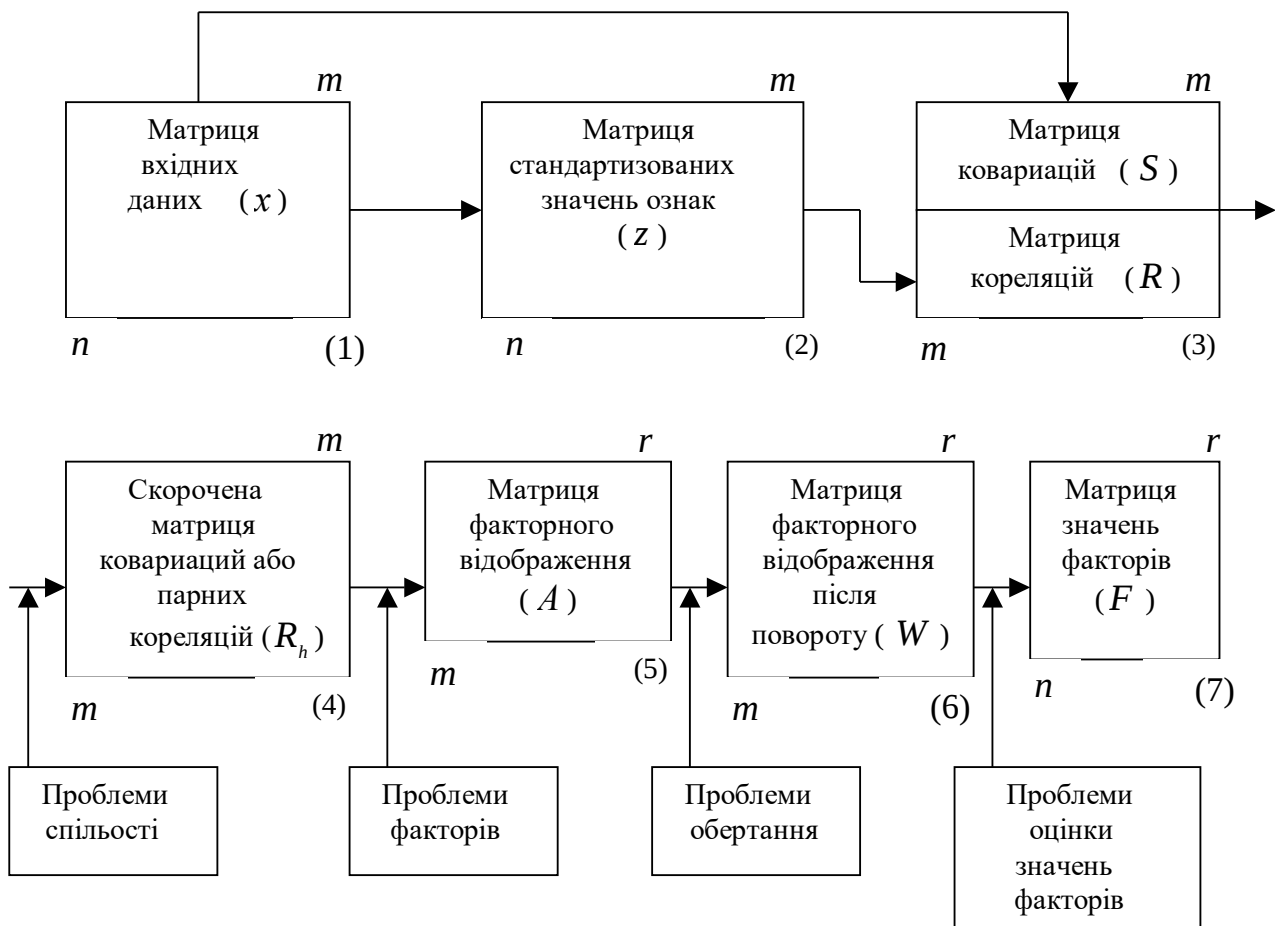


Рис. 1 Алгоритмічна схема реалізації методів факторного аналізу

Відбір значущих елементарних ознак при визначенні назви головної компоненти виробляється насамперед за абсолютною величиною вагового коефіцієнта a_{jr} . Знак коефіцієнта набуває значення в подальшому, при логічному поясненні складу і встановленні його несуперечності назвою головної компоненти.

Латентні фактори – компоненти. Коли декілька ознак пов'язані між собою кореляційними співвідношеннями, то це означає, що деякі ознаки визначаються через інші або вся така група ознак є проявами (наслідками) спільної для них причини, природа якої нам невідома, тобто не вивчена. Факторний аналіз базується саме на останньому припущенні. Ці латентні фактори, які знаходяться за лаштунками нашого уявлення про процес,

безпосередньо нами не вимірюються, але саме вони визначають всі ознаки, що ми спостерігаємо. Знаючи прояви (наслідки) латентних факторів, можна побудувати модель для вивчення цих латентних факторів, кількість яких значно менша від загальної кількості ознак. До речі, якщо можна коротко описати великий масив даних, то це означає, що знайдена певна об'єктивна закономірність, яка обумовила можливість цього короткого опису, тобто стиснення інформації.

Початкова система факторів (компонент) складається за екстремальним принципом – перша компонента має пояснювати максимум всієї змінності всіх ознак; друга компонента, що незалежна від першої, має пояснювати максимум залишкової змінності ознак, і так далі. Невелика кількість таких компонент спроможна відновити майже всю загальну змінність ознак.

Компоненту складаємо у вигляді лінійної комбінації

$$\mathbf{V} = u_1 \mathbf{Z}_1 + u_2 \mathbf{Z}_2 + \dots + u_m \mathbf{Z}_m = \mathbf{ZU},$$

де коефіцієнти u_k треба знайти з умови максимуму дисперсії

$$S_V^2 = \frac{1}{n-1} \mathbf{V}'\mathbf{V} = \frac{1}{n-1} (\mathbf{ZU})'(\mathbf{ZU}) = \frac{1}{n-1} \mathbf{U}'\mathbf{Z}'\mathbf{ZU} = \mathbf{U}'\mathbf{RU}.$$

Тут враховано, що компоненти – центровані (як лінійна комбінація стандартизованих змінних $\mathbf{Z}_k$).

Теорема. Компоненти – взаємно ортогональні, їх дисперсії дорівнюють власним числам кореляційної матриці.

Кілька перших (головних) компонент з найбільшими власними числами пояснюють майже всю загальну змінність системи ознак, а рештою компонент з нульовими і малими значеннями власних чисел можна знехтувати.

Факторні навантаження є β -коефіцієнтами регресії в розкладі вихідних ознак за системою факторів-компонент; спільності показують якість апроксимації вихідних ознак дією латентних факторів-компонент; сума

врахованих власних чисел – частку загальної змінності, що відновлена моделлю факторного аналізу.

Проблеми факторного аналізу. Факторний аналіз пройшов довгий шлях у своєму розвитку з того часу, коли ще взагалі не було сучасної електронної обчислювальної техніки. Тоді дійсно кожний етап факторного аналізу був проблемою і для їх вирішення було запропоновано багато штучних прийомів. Зараз все це не є предметом для обговорення, і наразі, проблеми факторного аналізу слід розуміти як етапи його реалізації, проте, дотримуючись класичних понять, наведемо оцінку кожної проблеми сьогодні.

Проблема факторів полягає у визначенні найменшої кількості факторів, яких досить для відновлення матриці кореляцій m змінних. Наразі ми не маємо даної проблеми – потрібно обчислити всі власні числа кореляційної матриці і, згідно з прийнятими рекомендаціями, визначити кількість головних компонент. Пригадаємо дві найпростіші й найпоширеніші рекомендації: 1) сума перших власних чисел має бути близькою до $0,8 \cdot m$; 2) всі враховані в моделі власні числа мають бути більшими за одиницю: $x_i \geq 1$. Але в минулому обчислення вручну власних чисел і власних векторів матриць великого розміру дійсно було проблемою і навіть автор методу головних компонент Хоттелінг пропонував почекати, доки не з'явиться автоматична обчислювальна техніка.

Проблема спільностей. Кореляційна матриця відтворюється точно, якщо врахувати всі компоненти, і з деякими похибками, якщо врахувати лише головні компоненти $\hat{R} = AA'$. На діагоналі відтвореної матриці тоді будуть не одиниці, а спільності $R_i^2 = \sum_{j=1}^k a_{ij}^2$. Було запропоновано редукувати кореляційну

матрицю, замінюючи одиниці на діагоналі на оцінки спільностей. У якості оцінки спільностей рекомендовано брати максимальний коефіцієнт кореляції в рядку чи в стовпчику або коефіцієнт детермінації після множинного регресійного аналізу (є й інші рекомендації). Очікувалось, що попередня редукція кореляційної матриці приведе до меншої кількості головних компонент або до більш чіткої структури факторного розв'язку після

перетворення повороту. Але чисельні розрахунки на комп'ютері не підтвердили цих очікувань – ніякої переваги від редукції кореляційної матриці не спостерігалось. Водночас були певні негаразди – після редукції окремі власні числа виявилися від'ємними, а отже, вони не можуть бути дисперсіями. Якщо дослідник впевнений у необхідності редукції, то наразі немає ніяких проблем у попередній оцінці спільностей: для цього потрібно провести на комп'ютері факторний аналіз двічі: перший раз – для визначення спільностей і другий раз – з редукованою матрицею.

Проблема повороту. Матриця повороту T конструється як сукупність поворотів у площині кожних двох факторів. Наприклад, для $k = 3$ перший цикл перетворень буде $T_{11}T_{13}T_{23}$, де T_{ij} – матриця повороту у площині факторів f_i, f_j . Ці цикли перетворень повторюються, доки збільшується сума четвертих степенів факторних навантажень. У минулому під час розрахунків вручну це, можливо, й було проблемою, але наразі всі ітераційні перетворення виконує комп'ютер і це не є вже проблемою.

Проблема вимірювання факторів полягає в обчисленні значень факторів f_i для кожного об'єкта (кожного спостереження).

Приклад застосування факторного аналізу

За допомогою спеціального статистичного пакету Statgraphics були обчислені фактори для кожного виду продукції, що експортується підприємством ПАТ «АВТРАМАТ». Так латентні фактори, що впливають на основні показники ЗЕДП ПАТ «АВТРАМАТ» по виробництву тракторних поршнів виражаються такими рівняннями:

$$F_{11} = 0,1456x_{11} - 0,0322x_{21} + 0,462x_{31} + 0,811y_{11} + 0,0696y_{21} + 0,8188y_{31} + \\ + 0,7744y_{41} + 0,2365y_{51} + 0,1819y_{61} + 0,905z_{11} + 0,8598z_{21} - 0,2644z_{31} + 0,0836z_{41};$$

$$F_{21} = 0,8674x_{11} - 0,7183x_{21} + 0,4824x_{31} - 0,1604y_{11} + 0,5521y_{21} - 0,1236y_{31} - \\ - 0,2093y_{41} + 0,1423y_{51} + 0,2069y_{61} + 0,1605z_{11} + 0,3594z_{21} + 0,6233z_{31} + 0,7237z_{41};$$

$$F_{31} = 0,1466x_{11} - 0,0171x_{21} - 0,0765x_{31} + 0,4606y_{11} - 0,588y_{21} + 0,4762y_{31} + \\ + 0,5205y_{41} + 0,9149y_{51} + 0,9196y_{61} - 0,0089z_{11} - 0,0364z_{21} - 0,2562z_{31} + 0,2965z_{41};$$

Виявлені фактори на 75,65% пояснюють змінність системи основних показників ЗЕДП по виробництву тракторних поршнів.

У всій сукупності підприємств найпершим, отже, і найбільш істотним є узагальнюючий фактор, утворений показниками $z_{11}, z_{21}, y_{31}, y_{11}, y_{41}, x_{31}, z_{31}, y_{51}, y_{61}, x_{11}, z_{41}, y_{21}, x_{21}$. Дана послідовність сформована у напрямку зниження факторних навантажень. Найбільше факторне навантаження має показник z_{11} ($a_{11} = 0,905$) – питома вага тракторних поршнів в загальному обсязі експорту. Далі, за пріоритетністю іде z_{21} ($a_{21}=0,8598$) – питома вага тракторних поршнів в загальному обсязі виробництва. Наступний показник y_{31} ($a_{31}=0,8188$) – індекс фізичного обсягу експортованих тракторних поршнів. Наступний за ним y_{11} ($a_{41} = 0,811$) – індекс вартості експортованих тракторних поршнів. Далі y_{41} ($a_{51} = 0,7744$) – індекс структури експортованих тракторних поршнів. Наступний за пріоритетом x_{31} ($a_{61}=0,462$) – ефективність реалізації тракторних поршнів на внутрішньому ринку. Аналіз послідовності дозволяє визначити цей фактор, як організаційно-виробничий, тобто якість продукції та оптимальна структура виробництва впливають на експорт тракторних поршнів.

Наступний фактор утворений показниками $x_{11}, z_{41}, x_{21}, z_{31}, y_{21}, x_{31}, z_{21}, y_{41}, y_{61}, z_{11}, y_{11}, y_{51}, y_{31}$. Ця послідовність сформована тим же чином, що і попередня. Найбільше факторне навантаження в цій послідовності має показник x_{11} ($a_{11}=0,8674$), який характеризує валютну ефективність експорту тракторних поршнів. Далі за послідовністю їдуть показники z_{41} ($a_{21}=0,7237$), x_{21} ($a_{31}=0,7183$), z_{31} ($a_{41}=0,6233$), y_{21} ($a_{51}=0,5521$) та x_{31} ($a_{61}=0,4824$). Аналіз послідовності дозволяє зробити висновок, що найбільш на ефективність зовнішньоекономічної діяльності ПАТ

«АВТРАМАТ» впливають зовнішні економічні фактори, а саме курс національної валюти.

Третій латентний фактор, впливає по пріоритетності на такі показники: $y_{61}, y_{51}, y_{21}, y_{41}, y_{31}, y_{11}, z_{41}, z_{31}, x_{11}, x_{31}, z_{21}, x_{21}, z_{11}$. Факторне навантаження в цієї сукупності формується показниками y_{61} ($a_{11}=0,9196$), y_{51} ($a_{21}=0,9149$), y_{21} ($a_{31}=0,5880$), y_{41} ($a_{41}=0,5205$), y_{31} ($a_{51}=0,4762$), y_{11} ($a_{61}=0,4606$), що відображають недоліки в організаційній структурі підприємства.

Аналіз складу такої послідовності показників дозволив припустити, що цей фактор характеризує якісні зміни в зовнішньоекономічній діяльності ПАТ «АВТРАМАТ», причому пріоритетність в факторних навантаженнях належить показникам y , що свідчить про динамічні зрушення, які відбуваються в структурі експорту підприємства за рахунок зміни ціни та попиту на продукцію, укладання нових контрактів з клієнтами.

Таким чином, схематично модель латентних факторів впливу на ЗЕД ПАТ «АВТРАМАТ» по випуску тракторних поршнів має такий вигляд (рис.2).

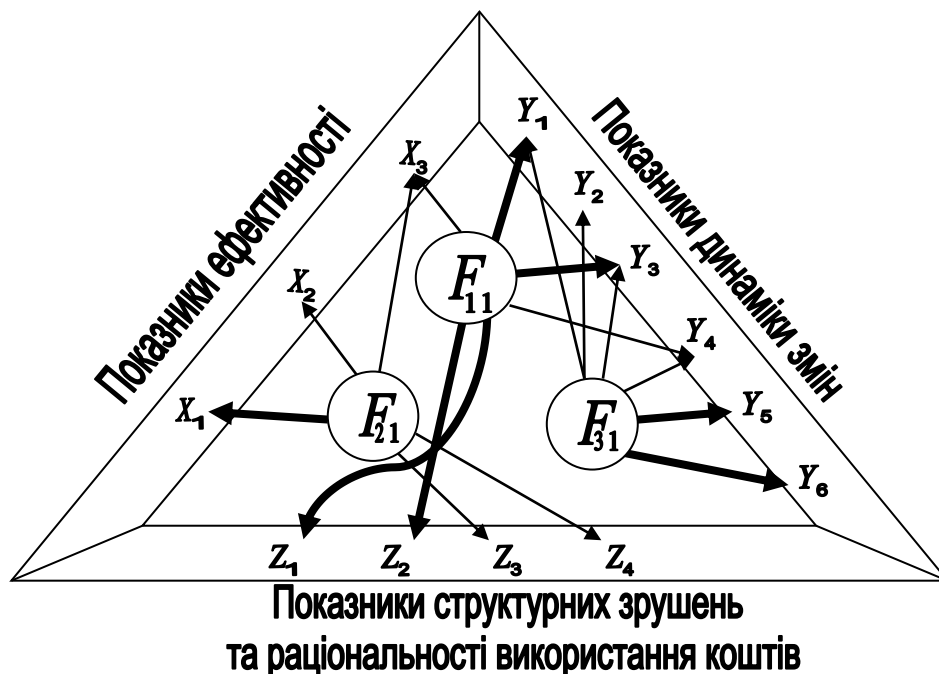


Рис. 2. Схema впливу латентних факторів на показники ЗЕД ПАТ «АВТРАМАТ» по виробництву тракторних поршнів

