УДК 378.091.26:519.237.7(045)
DOI 10.5281/zenodo.14567115

С. О. Касярум
ORCID ID 0009-0002-1518-5151
Національний університет цивільного захисту України

ЗНАХОДЖЕННЯ ВЛАСНИХ ЗНАЧЕНЬ ТА ВЛАСНИХ ВЕКТОРІВ КОРЕЛЯЦІЙНОЇ МАТРИЦІ ФАКТОРНОГО АНАЛІЗУ (НА ПРИКЛАДІ ФАКТОРНОГО АНАЛІЗУ АКАДЕМІЧНОЇ УСПІШНОСТІ ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ)

У статті порушується питання академічної успішності здобувачів вищої освіти й чинників, які мають на неї вплив. Підкреслюється, що результати педагогічних досліджень різних етапів розвитку вищої освіти демонструють різні чинники впливу на освітні досягнення студентів. Актуальними на теперішній час є вивчення впливу на академічну успішність здобувачів вищої освіти організації освітнього процесу в умовах воєнного стану і дистанційного або змішаного навчання.

Для визначення факторів впливу на академічну успішність студентів дослідники використовують факторний аналіз. Застосування факторного аналізу уможливлює виявлення прихованих структур у множині змінних. Підкреслюється, що факторний аналіз є складною процедурою, що може ускладнюватися недосконалістю зібраних даних і неточним статистичним висновком. Проте така процедура уможливлює виокремити цінні для дослідника фактори з набору зібраних даних, і зосередитися у подальшій своєї науковій роботі саме них. На теперішній час дослідниками використовуються різні методи проведення процедури факторного аналізу. Значно полегшують проведення факторного аналізу програмні продукти для проведення аналізу статистичних даних (SPSS, SAS, Excel та ін.).

У статті на прикладі визначення факторів впливу на академічну успішність студентів показано алгоритм визначення власних значень та векторів кореляційної матриці для проведення факторного аналізу. Особливістю процедури визначення головних факторів було використання автором властивостей власних векторів матриць, що становить нову деталь факторних обчислень. У статті показано, як користуючись досить обмеженими можливостями електронних таблиць Excel, можна реалізувати знаходження власних чисел та власних векторів для достатньо великих за розмірами матриць. У цьому автор статті вбачає методичну привабливість пропонованого способу розв'язання цих задач. Запропонований і презентований автором спосіб не ставить за мету заміни програмних способів, а лише уможливлює вирішити проблему доступності обчислень.

***Ключові слова**: факторний аналіз, метод головних компонент, спільність, характерність, навантаження, кореляційна матриця, власний вектор, власні значення матриці, академічна успішність, здобувачі вищої освіти.*

Постановка проблеми. Питання успішності навчальної діяльності студентів у закладах вищої освіти завжди є актуальним. Зацікавленість науковців вказаною проблемою продиктовано зміною умов, в яких здійснюється освітній процес. Так, наприклад, у період пандемії COVID-2019 освітній процес повністю перейшов на дистанційний формат, що, відобразилося на успішності навчальної діяльності студентів [2]. Серед чинників, які вплинули на навчальні досягнення здобувачів освіти, дослідники називають: непідготовленість ЗВО швидко перейти на дистанційне навчання, цифрова й методична некомпетентність викладачів, працювати у такому форматі, непідготовленість самих здобувачів освіти та інші.

Наразі в умовах воєнного стану організація освітнього процесу в країні відбувається у різних форматах (дистанційному, змішаному), що також породжує багато чинників, які впливають на навчальні досягнення студентів. Результати проведених науковцями досліджень вказують на те, що для підвищення успішності студентів в умовах дистанційного і змішаного навчання необхідними є такі фактори: розробка навчально-методичого супроводу, зокрема методичні рекомендації з чіткими інструкціями для студентів, визначення «форми взаємодії» викладача зі студентами, забезпечення з боку ЗВО розподіл ресурсів та інвестицій у розвиток технологічного потенціалу закладу, налагоджування діяльності підрозділу підтримки суб'єктів освітнього процесу [3].

Для відокремлення факторів впливу на успішність студентів існує багато способів, які використовують дослідники. Наприклад, узагальнення результатів наукових розвідок, рефлексії професійної діяльності викладачів, проведеного анкетування суб'єктів освітнього процесу (здобувачів і викладачів). Серед факторів пов'язаних з успішністю студентів вітчизняні та закордонні дослідники називають такі: стать, попередня успішність, місце проживання та дохід сім'ї, соціальне оточення, шкільний середній бал, результати вступних тестів до університету, час, витрачений на навчання, здатність до навчання, місце проживання під час університетського життя, володіння мовою, швидкість адаптації до університетського життя, зокрема й до організації освітнього процесу, побудова продуктивної (непродуктивної) взаємодії між студентом і викладачем, навички до виконання самостійної роботи, психологічні фактори тощо. Найбільш суттєвими факторами вважаються: систематичне навчання і відвідування занять, виконання завдань з дисциплін, мотивація і наполегливість у досягненні поставлених цілей, свідома саморегуляція [1].

У більшості наукових досліджень для визначення факторів, які безпосередньо впливають на успішність навчання здобувачів освіти, й наукове їх підґрунтя, дослідниками застосовується факторний аналіз.

Термін «факторний аналіз» в навчальній та науковій літературі часто використовується для двох різних математичних процедур. Одна з них використовується у задачах регресійного аналізу для підтвердження впливу визначених відомих величин на досліджувану величину. Їх називають чинниками регресії. Засоби розв'язку подібних задач достатньо відомі, часто вони не потребують складних програмних продуктів.

Інша проблема, а отже, більш цікава з точки зору математики, пов'язана із потребою знаходження невідомої кількості невідомих факторів, які за припущенням впливають на систему досліджувані величини. Масив вхідних експериментальних результатів записують у вигляді матриці вхідних даних. Далі, перетворення цієї матриці приводить до появи вихідного результату – кореляційної матриці факторного аналізу. Аналіз результуючої матриці визначає ту мінімальну кількість факторів, які пояснюють досліджувані закономірності. Разом із цим кожному фактору ставиться у відповідність факторна вага, – число, яке визначає міру впливу на досліджувані величини. Задача факторного аналізу полягає у тому, що потрібно знайти мінімальну кількість факторів разом із факторною вагою, які пояснюють досліджувані закономірності. Такі величини прийнято називати головними компонентами факторного аналізу. Результати обчислень цих компонент в решті решт виявляють їх природу. Вони виявляються власними значеннями та власними векторами результуючої кореляційної матриці. Іншими словами, факторний аналіз допомагає спростити складні дані та виявити основні причини, що впливають на спостережувані явища.

З огляду на вказане вище, виникає питання застосування факторного аналізу для виокремлення факторів, які мають безпосередній вплив на академічну успішність здобувачів вищої освіти.

Аналіз актуальних досліджень. Активний розвиток факторного аналізу відбувся у першій половині XX століття. Він з успіхом використовується в психології та соціології до нині. Для цих галузей науки факторний аналіз залишається одним з головних методів досліджень. Принципи та наукове обґрунтування факторного аналізу можна знайти у класичній тепер монографії Harry H. Harman [7].

На сьогодні існує значна кількість наукових праць, в яких надаються рекомендації щодо проведення процедури факторного аналізу і проблем [4; 5; 6; 8; 9], з якими може стикатися дослідник, методологія, опис поетапності факторного аналізу, зокрема за допомогою сучасних програмних продуктів для статистичного аналізу даних (SPSS, SAS, Stata, Excel, AMOS, R, Python та ін.

У наш час цікавість до факторного аналізу знову зросла у зв'язку із його використанням у цьому аналізі великої кількості інформації, яка для традиційних задач, так і для нових задач, що з'явилися відносно недавно. Ось деякі з цих задач:

- вимірювання латентних (прихованих) величин та побудова нових узагальнених показників (традиційно психологія та соціологія);
- скорочення числа змінних. Необхідність зменшення розмірності змінних при умові збереження втрати потрібної інформації;
- виявлення груп взаємозалежних змінних та структури взаємної залежності між ними. Факторний аналіз забезпечує лаконічнішу та точнішу модель структури залежності між змінними;
- подолання мультиколінеарності змінних у регресійному аналізі в деяких економічних задачах;
- доповнення пропущених значень розріджених матриць для використання у рекомендаційних системах;
- латентний семантичний аналіз, який дозволяє вести інформаційний пошук та аналізу великої кількості документів з метою їх індексації, класифікації і таке інше, там де є потреба виявлення головних факторів з масиву інформаційних даних.

Це далеко не повний список задач, які потребують використання факторного аналізу, а лише задачі більш-менш відомі нам.

Мета статті полягає у презентації процедури виокремлення та оцінювання факторів впливу на успішність навчання студентів за допомогою факторного аналізу. Для знаходження головних компонент факторного аналізу застосовано електронні таблиці Excel, що є підходящім, доступним та зрозумілим інструментом обчислень для вирішення проблеми обчислень при появі матриць великого порядку часто десяти і більше. Особливість процедури визначення головних факторів полягає у використанні властивостей власних векторів матриць, що становить нову деталь факторних обчислень.

Виклад основного матеріалу. З точки зору математики, вибраний нами метод обчислень головних компонент факторного аналізу представляє собою відому задачу матричної алгебри, а саме, знаходження власних значень та власних векторів квадратної матриці [7]. Як правило, в освітньому процесі використовують приклади знаходження цих величин лише для матриць другого та третього порядків. Оскільки відповідна задача розв'язується методом характеристичних рівнянь, які є степеневими, то розв'язок рівнянь із степенем більше ніж третій досить проблемний.

Для матриць високих порядків часто використовують програмі засоби, серед яких відомим є метод Данилевського, але при цьому виникає проблема їх доступності. Ідея проведеного дослідження полягає у тому, як користуючись досить обмеженими можливостями електронних таблиць Excel реалізувати знаходження власних чисел та власних векторів для достатньо великих за розмірами матриць. У цьому вбачаємо методичну привабливість пропонованого способу розв'язання цих задач. Хоча маємо запевнити, що не ставимо за мету заміни програмних способів, а лише вирішуємо проблему доступності обчислень за запропонованим методом.

Орієнтуючись на читача, який можливо детально не знайомий з положеннями факторного аналізу, а також, виходячи з того, що спосіб обчислень за пропонованим методикою тісно зв'язаний з цією теорією, дозволимо собі її короткий виклад.

Параметри статистичного дослідження групи об'єктів. Як правило, у статистичному дослідженні розглядають групу об'єктів, що мають певні спільні для них ознаки. Кількісні оцінки цих ознак називають значеннями їх параметрів. Таким чином кожний об'єкт характеризують набором певної кількості значень параметрів z_j. Така характеристика об'єкта є багатовимірною випадковою величиною з компонентами z_{ij}.

Вимірювання n параметрів для N об'єктів можна записати у вигляді таблиці або матриці даних. У прикладі значення величин довільні.

	A	B	C	D	E
1	1,153813	1,906125	1,291421	1,366405	2,247261
2	3,205298	3,565386	1,576525	2,477157	2,976501
3	1,669088	1,409253	4,087802	4,163335	4,216346
4	3,73513	1,130467	4,666799	5,982727	3,306375
5	4,346538	2,785882	6,368358	5,194159	3,076479
6	4,438002	1,480941	7,345347	4,225135	4,835444
7	5,196905	3,701895	7,751854	8,348704	7,980926
8	5,60094	4,200415	8,259987	9,091952	5,383404
9	9,565203	4,361095	9,192755	8,769219	8,961486
10	9,626179	9,781365	9,873104	9,921171	9,752525

Необхідно провести обчислення: середнє значення кожного параметра; відхилення значення параметра від середнього; вибіркову дисперсію параметра; нормоване значення параметра j для об'єкту i; коефіцієнт кореляції для будь яких довільних параметрів j та k.

Після транспонування останньої матриці обчислюють коефіцієнт кореляції для будь яких довільних параметрів j та k.

	1	0,779329	0,86301	0,84794	0,886024
	0,779329	1	0,585664	0,662613	0,746813
[R*]=	0,86301	0,585664	1	0,908483	0,840778
	0,84794	0,662613	0,908483	1	0,849095
	0,886024	0,746813	0,840778	0,849095	1

Зміст таких обчислень можна сформулювати таким чином: знайдена кореляційна матриця елементи якої становлять основу головної матриці факторного аналізу.

Лінійна модель класичного факторного аналізу. Завдання факторного аналізу полягає у тому щоб виразити кожний нормалізований параметр z_j у термінах прихованих гіпотетичних факторів. Ці фактори розуміють, як деяку нову систему координат із відповідними масштабами, а F_1, $F_2,\dots,F_m$ модулі координатних векторів цієї системи. Найпростішою та вживаною є лінійна модель класичного факторного аналізу.

$$z_j = a_{1j}F_1 + a_{2j}F_2 + \dots + a_{mj}F_m + d_jU_j \quad (1)$$

У цій моделі кожний параметр z_j лінійно залежить від m спільних факторів F_1, $F_2,\dots,F_m$ ($m<<n$) та одного характерного фактора U_j,, який відображає ту частину параметра z_j , яку не можливо пояснити спільними факторами. Спільні фактори вважають змінними, а a_{1j}, $a_{2j},\dots,a_{mj}$ – коефіцієнтами при змінних F_1, $F_2,\dots,F_m$. Те саме стосується характерного фактора U_j та коефіцієнта d_j . Лінійна модель класичного факторного аналізу схожа на рівняння регресії, але на відміну від регресії нам невідомо, як вимірювати ці змінні.

На фактори моделі накладають певні вимоги. Фактори мають бути величинами із нормованим нормальним розподілом, а, отже, мати математичне очікування рівне нулю та дисперсію рівну одиниці. Крім того, як вектори нової системи координат, фактори мають бути ортогональними, а, отже, варіації між різними факторами мають бути рівними нулю.

Коефіцієнти при факторах a_{1j}, $a_{2j},\dots,a_{mj}$ називають навантаженнями. Навантаження визначають долю впливу окремого фактора на даний параметр, а їх значення є предметом обчислень у факторному аналізі.

Для j параметра та i об'єкту відповідна лінійна модель набуває вигляду:

$$z_{ij} = a_{1j}F_{1i} + a_{2j}F_{2i} + \dots + a_{mj}F_{mi} + d_jU_{ij} = \sum_{p=1}^{m} a_{pj}F_{pi} + d_jU_{ij} \cdot \quad (2)$$

Основною проблемою факторного аналізу є визначення $m \cdot n$ навантажень спільних факторів. Така задача у решті-решт розв'язується процедурою аналізу матриці коефіцієнтів кореляції між параметрами. При такому аналізі проблема статистичної достовірності коефіцієнтів кореляції, як правило, ігнорується.

Складові дисперсії лінійної моделі класичного факторного аналізу. Відповідно лінійної моделі стандартна дисперсія (нормованої величини) параметру $\boldsymbol{z_j}$ може бути визначена як

$$s_{zj}^2 = \frac{\sum_{i=1}^{N} z_{ij}^2}{N} = \sum_{i=1}^{N} \frac{\left(\sum_{p=1}^{m} a_{pj} F_{pi} + d_j U_{ij} \right)^2}{N} = \sum_{p=1}^{m} a_{pj}^2 \left(\frac{\sum_{i=1}^{N} F_{pi}^2}{N} \right) + d_j^2 \frac{\sum_{i=1}^{N} U_{ij}^2}{N} + 2 \sum_{p<q=1}^{m} a_{pj} a_{qj} \left(\frac{\sum_{i=1}^{N} F_{ip} F_{iq}}{N} \right) + 2d_j \sum_{p=1}^{m} a_{pj} \left(\frac{\sum_{i=1}^{N} F_{ip} U_{ij}}{N} \right)$$

Вираз може бути суттєво спрощений завдячуючи двом обставинам.

Перша, стандартна дисперсія дорівнює одиниці, а всі інші величини виразу також записані у стандартній формі. Внаслідок цього маємо

$$s_{zj}^2 = 1 = \sum_{p=1}^{m} a_{pj}^2 + d_j^2 + 2 \sum_{p<q=1}^{m} a_{pj} a_{qj} r_{FpFq} + 2d_j \sum_{p=1}^{m} a_{pj} r_{FpUj}$$

Друга, приймається припущення, що всі фактори незалежні один від одного, а отже не корелюють між собою. Отже у решті-решт маємо:

$$s_{zj}^2 = 1 = \sum_{p=1}^{m} a_{pj}^2 + d_j^2 = a_{1j}^2 + a_{2j}^2 + \ldots + a_{mj}^2 + d_j^2 \quad (3)$$

Очевидно, що квадрати факторних навантажень виражають вклади відповідних факторів у дисперсію певного параметру. Суму $h_j^2 = \sum_{p=1}^{m} a_{pj}^2 = a_{1j}^2 + a_{2j}^2 + \ldots + a_{mj}^2$ називають спільністю h_j^2. Величину d_j^2 називають характерністю.

Отже, дисперсія параметра $\boldsymbol{z_j}$ складається із двох частин. Перша частина пояснюється впливом на параметр факторів, а друга частина не може бути пояснена впливом факторів, вона існує сама по собі та індивідуальна для кожного окремого параметра.

Факторне відображення та факторна структура. Користуючись лінійною моделлю класичного факторного аналізу (1), яка зв'язує параметр $\boldsymbol{z_j}$ із $\boldsymbol{m}$ спільними факторами, запишемо вирази для усіх $\boldsymbol{n}$ параметрів у вигляді:

$$\begin{cases} z_1 = a_{11}F_1 + a_{21}F_2 + \ldots + a_{m1}F_m + d_1U_1 \\ z_j = a_{12}F_1 + a_{22}F_2 + \ldots + a_{m2}F_m + d_2U_2 \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ z_n = a_{1n}F_1 + a_{2n}F_2 + \ldots + a_{mn}F_m + d_nU_n \end{cases}$$

Така система рівнянь називається факторним відображенням або просто відображенням. Її можна записати у матричній формі:

$$Z = A \cdot F + D,$$

$$Z = \begin{vmatrix} z_1 \\ z_2 \\ \ldots \\ z_n \end{vmatrix}; \; A = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{vmatrix}; \; F = \begin{vmatrix} F_1 \\ F_2 \\ \\ F_m \end{vmatrix}; \; D = \begin{vmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_n \end{vmatrix}$$

Окрім факторного відображення у факторному аналізі розглядають факторну структуру, необхідну для подальшої побудови теорії. Для знаходження факторної структури достатньо помножити кожне з рівнянь (2) на відповідні фактори, додати по усім $\boldsymbol{N}$ спостереженням для кожного параметра, розділити на кількість значень $\boldsymbol{N}$ і отримати коефіцієнти кореляції параметрів з факторами. Ця процедура враховує, що параметри та фактори – нормалізовані величини. А якщо врахувати також відсутність кореляції між факторами, то виявиться , що відповідні коефіцієнти кореляції параметрів та факторів дорівнюють навантаженням (елементам матриці $\boldsymbol{A}$).

$$r_{Zj;Fp} = a_{jp}$$

Як підсумок: елементи матриці $\boldsymbol{A}$ – це коефіцієнти кореляції між параметрами та факторами. Матрицю $\boldsymbol{A}$ називають факторною структурою.

Фундаментальна факторна теорема. На підставі рівнянь (2), які дають значення параметрів, можна обчислити коефіцієнти кореляції між ними, як добуток правих частин рівнянь, враховуючи нормованість та некорельованість факторів між собою. Нехтуючи деталями отримаємо вирази виду:

$$r_{ij} = a_{1i}a_{1j} + a_{2i}a_{2j} + \ldots + a_{mi}a_{mj}\ , \quad (i, j = 1,2,\ldots,n)$$

Перейдемо до матричної форми цього результату

$$Z = A \cdot F \Rightarrow R = \frac{Z \cdot Z'}{N} \Rightarrow R = \frac{1}{N}\left(A \cdot F \cdot F' \cdot A'\right) \Rightarrow R = A \cdot A'$$

Оскільки, фактори нормовані та некорельовані між собою, то:

$$\frac{F \cdot F'}{N} = I \ \Rightarrow I = \begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{vmatrix} \ \Rightarrow A \cdot I = A$$

Таким чином, маємо співвідношення $R = A \cdot A'$ яке виражає фундаментальну факторну теорему. Знайдене співвідношення зв’язує між собою величини, які доступні для вимірювання $\boldsymbol{R}$ і величини $A \cdot A'$, які безпосередньо не вимірюються.

Метод головних факторів. Основна задача факторного аналізу знаходження навантажень $\boldsymbol{a_{ij}}$ факторів. Існує багато методів [7] розв’язку цієї задачі (деякі з них досить архаїчні). Скористаємося методом головних факторів, розробленим Г. Томсоном.

Вище, виведена формула (3) $s_{zj}^2 = 1 = \sum_{p=1}^{m} a_{pj}^2 + d_j^2 = a_{1j}^2 + a_{2j}^2 + \ldots + a_{mj}^2) + d_j^2$, яка показує структуру дисперсії показника, що має дві складові: $h_j^2 = \sum_{p=1}^{m} a_{pj}^2 = a_{1j}^2 + a_{2j}^2 + \ldots + a_{mj}^2$ та d_j^2 . Перша, спільність h_j^2 – показує яка частина дисперсії показника пояснюється факторами, а друга, характерність d_j^2 – показує яку частину дисперсії показника не можна пояснити факторами. Чим більша перша частина та менша друга, тим є кращим результат факторного аналізу. Отже умова, за якою послідовно визначають фактори, вимагає максимуму вкладів кожного з факторів у сумарну спільність.

Для певного фактора $\boldsymbol{F_j}$ ця сума: $V_j = \sum_{i=1}^{n} a_{ij}^2 = a_{1j}^2 + a_{2j}^2 + \ldots + a_{nj}^2$

На першій стадії обчислень шукають навантаження при першому факторі таким чином, щоб була максимальною сума вкладів цього фактора у сумарну спільність, при цьому має бути забезпечена умова, що $r_{jk} = \sum_{p=1}^{m} a_{jk} \cdot a_{kp}$.

Це задача на умовний екстремум функції багатьох змінних, яку розв’язують методом множників Лагранжа. Так, для першого фактора маємо:

$2L = \sum_{i=1}^{n} a_{ij}^2 - \sum_{i=1}^{n} \lambda_{jk} r_{jk} = \sum_{i=1}^{n} a_{ij}^2 - \sum_{i=1}^{n}\sum_{p=1}^{m} \lambda_{jk} a_{jk} a_{kh}$ – функція Лагранжа,

$\frac{\partial L}{\partial a_{i1}} = a_{i1} - \sum_{i=1}^{n} \lambda_{1i} a_{1i} = 0$ – частинні похідні функції Лагранжа рівні нулю.

Знаходження умовного екстремуму функції Лагранжа краще вести у матричному форматі і тоді об'єднана система рівнянь із частинним похідними набуде виду характеристичного рівняння:

$$\begin{vmatrix} h_1^2-\lambda & r_{12} & \dots & r_{12} \\ r_{21} & h_2^2-\lambda & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & h_m^2-\lambda \end{vmatrix} = 0 \Rightarrow R\cdot\vec{a} = \lambda\cdot\vec{a}$$

Отже, задача зводиться до знаходження власних векторів $\vec{a}$ та власних значень, так званої, редукованої кореляційної матриці $\boldsymbol{R}$. У редукованої кореляційної матриці на головній діагоналі розміщені значення спільності $\boldsymbol{h_j}$ замість одиниць звичайної кореляційної матриці. Власні вектори $\vec{a}$ – матриця стовпчик спільностей відповідного фактора. Власні значення $\boldsymbol{\lambda}$ (коефіцієнти Лагранжа), – сума дисперсій параметрів або значення функції $\boldsymbol{V_j}$, яка підлягала максимізації.

У підсумку зауважимо наступне: нарешті знайдене матричне рівняння $R\cdot\vec{a} = \lambda\cdot\vec{a}$, що зв'язує кореляційну матрицю показників (реальних величин) із величинами (прихованими) та дозволяє їх визначати.

Саме з цим рівнянням пов'язана практична проблема визначення способу розв'язку рівняння з метою обчислень недоступних до вимірюванню факторних параметрів. Також маємо задачу, як знайти реальну редуковану кореляційну матрицю. Основною проблемою тут – є визначення спільностей $\boldsymbol{h_j}$.

Схема обчислень по методу головних факторів. Порядок обчислень наступний:

Для знаходження параметрів першого головного компонента $\boldsymbol{F_1}$ розв'язуємо рівняння $R\cdot\vec{a}_1 = \lambda_1\cdot\vec{a}_1$, що максимізує суму $V_1 = \sum_{i=1}^{n} a_{i1}^2 = a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2$, знаходимо перший корінь характеристичного рівняння $\boldsymbol{\lambda_1}$ та один із можливих розв'язків вектор $\overrightarrow{b_1}$ =($\boldsymbol{b}_{11}, \boldsymbol{b}_{21},\dots, \boldsymbol{b}_{n1}$). Щоб задовольнити співвідношення для $\boldsymbol{V_1}$ обчислюємо a_{i1}:

$$a_{i1} = \frac{b_{i1}\sqrt{\lambda_1}}{\sqrt{b_{11}^2 + b_{21}^2 + \dots + b_{n1}^2}} \qquad (4)$$

Отримаємо навантаження для першого фактора, яке запишемо у вигляді матриці-стовпця $\vec{a}_1$.

Щоб знайти розв'язок для другого за величиною компонента $\boldsymbol{F_2}$ потрібно вирахувати із редукованої кореляційної матриці внесок від компонента $\boldsymbol{F_1}$. Для цього знаходимо добуток матриці $\vec{a}_1$ на транспоновану матрицю $\vec{a}_1'$ та отримаємо матрицю $\tilde{R} = \vec{a}\cdot\vec{a}'$, після чого знайдемо наступну редуковану матрицю R_1:

$$R_1 = R - \tilde{R}$$

Далі процедура знаходження навантажень та наступного кореню характеристичного рівняння повторюється. Розв'язуємо рівняння $R_1\cdot\vec{a}_2 = \lambda_2\cdot\vec{a}_2$, що максимізує суму $V_2 == a_{12}^2 + a_{22}^2 + \dots + a_{n2}^2$, знаходимо другий корінь характеристичного рівняння $\boldsymbol{\lambda_2}$ та матрицю-стовпець $\vec{a}_2$.

Оскільки стандартна дисперсія окремого нормованого параметра дорівнює одиниці, то знаходження наступних компонент припиняють коли $\boldsymbol{\lambda}_j$ стає менше 1.

Отже, схема обчислень визначена – лишається доповнити схему алгоритмом обчислень.

Формування редукованої матриці. Виберемо масив даних для проведення факторного аналізу. Наприклад, нами взяті оцінки у стобальній шкалі сесійного контролю навчання 28 студентів по 11 освітнім компонентам.

Відповідно цьому, оцінки розміщені у матриці розміром 28x11 (28 рядків та 11 стовпців). Для дослідження результатів навчання студентів по окремим дисциплінам можна побудувати матрицю кореляцій оцінок між студентами. Це буде квадратна кореляційна матриця порядку 28. Тоді можемо шукати фактори, що впливають на якість навчання студентів. Якщо ми хочемо порівняти якість викладання дисциплін, то можна побудувати матрицю кореляцій оцінок студентів по дисциплінам. Така матриця буде порядку 11.

Приклад наведено лише для того, щоб показати, що перший етап роботи це здобуття реальних даних. Далі говоритимемо лише про математичні процедури.

Вибираємо масив даних, нормуємо їх та записуємо прямокутну матрицю $\boldsymbol{X}$ з $\boldsymbol{m}$ параметрами та з $\boldsymbol{n}$ вимірюваннями кожної.

Транспонуємо матрицю $\boldsymbol{X}$ у $\boldsymbol{X'}$, та обчислюємо кореляційну матрицю $\boldsymbol{R}$ як добуток $\boldsymbol{R}=\boldsymbol{X'X}$. Діагональ цієї матриці складається з одиниць. Тут виникає проблема визначення спільності $\boldsymbol{h_j}^2$, значеннями якої потрібно замінити одиниці діагоналі звичайної кореляційної матриці. Як пише Г. Харман [7], існує декілька методів оцінки спільності і жоден не є надійним. Але при великих розмірах кореляційної матриці похибка в обчисленнях, пов'язаних із цією обставиною, не є суттєвою. Чим більша матриця – тим менша похибка результатів аналізу.

Нас влаштовує оцінка спільності $\boldsymbol{h_j}^2$ через обчислення коефіцієнта множинної кореляції $\boldsymbol{r_j}$ (КМК) кожного окремого параметра з усіма іншими. У цьому є логіка, бо $\boldsymbol{r_j}$ є мірою того, що спільного у параметра $\boldsymbol{j}$ з усіма іншими параметрами (у літературі вважають, що КМК – лише нижня оцінка спільності $\boldsymbol{h_j}^2$).

Отже, знаходимо обернену матрицю $\boldsymbol{C}=(\boldsymbol{X'X})^{-1}$, вибираємо діагональний елемент $\boldsymbol{c_{jj}}$ та обчислюємо КМК як $\boldsymbol{r_j}=1-1/\boldsymbol{c_{jj}}$.

Замінюємо діагональні одиниці кореляційної матриці $\boldsymbol{R}$ коефіцієнтом множинної кореляції $\boldsymbol{r_j}$ та отримаємо редуковану кореляційну матрицю, для якої збережемо те саме позначення $\boldsymbol{R}$.

Таким чином маємо редуковану кореляційну матрицю $\boldsymbol{R}$, з якої починаємо процедуру обчислень.

	Редукована матриця (у діагоналях КМК)				
	0,874161	0,779329	0,86301	0,84794	0,886024
	0,779329	0,684417	0,585664	0,662613	0,746813
[R]=X'X=	0,86301	0,585664	0,87989	0,908483	0,840778
	0,84794	0,662613	0,908483	0,857475	0,849095
	0,886024	0,746813	0,840778	0,849095	0,832888

Обчислення головних факторів. Подальша процедура починається з виділення першого фактора, яка полягає у послідовності ітераційних обчислень. Ітерації припиняємо, коли результат обчислень перестає мінятися у межах допустимих похибок (нехай буде 5-6 знак після коми). Основу обчислень становить співвідношення: $R \cdot \vec{a}_1 = \lambda_1 \cdot \vec{a}_1$. Після виділення чергового фактора ітераційні обчислення повторюються для наступного.

Опишемо процедуру обчислень. Знаходимо добуток матриці $\boldsymbol{R}$ на матрицю стовпчик $\vec{a}_1$. У якості першого наближення стовпчика $\vec{a}_1$ виберемо матрицю, утворену сумою рядочків матриці $\boldsymbol{R}$. В цьому є певна логіка, бо результатом дії буде матриця стовпчик: $\lambda_1 \cdot \vec{a}_1$, а максимальний елемент цієї матриці буде близький до факторного коефіцієнта λ_1 (власного значення).

Вибираємо найбільший елемент стовпчика $\lambda_1 \cdot \vec{a}_1$ та ділимо на нього всі інші елементи. Отримаємо перше наближення власного вектора $\vec{a}_1$.

Знову знаходимо добуток матриці $\boldsymbol{R}$ на наближення власного вектора $\vec{a}_1$. Повторюємо подальшу процедуру, та обчислюємо друге наближення $\vec{a}_1$. Після деякої кількості ітерацій результат дії перестає змінюватися.

			Знаходження головних компонент													
						сер.знач.	мумнож						λ	bi1	bi1^2	a
	0,874161	0,779329	0,86301	0,84794	0,886024	0,850093	3,422119	1	4,033788	1	4,033928	1	**4,033943**	1	1	**0,946847**
	0,779329	0,684417	0,585664	0,662613	0,746813	0,691767	2,781033	0,812664	3,277257	0,812451	3,277334	0,812442	3,277343	0,812442	0,660061	**0,769258**
[R1]=	0,86301	0,585664	0,87989	0,908483	0,840778	0,815565	3,304786	0,965713	3,896256	0,965905	3,89645	0,965919	3,896467	0,96592	0,933002	**0,914578**
	0,84794	0,662613	0,908483	0,857475	0,849095	0,825121	3,333349	0,97406	3,929785	0,974217	3,92996	0,974226	3,929976	0,974227	0,949118	**0,922444**
	0,886024	0,746813	0,840778	0,849095	0,832888	0,83112	3,348369	0,978449	3,946891	0,978458	3,947034	0,978459	3,947048	0,978459	0,957382	**0,926451**
															4,499564	**4,479577**

Формулюємо результат виявлення першої головної компоненти факторного аналізу. Найбільший елемент отриманого стовпчика $\lambda_1 \cdot \vec{a}_1$ дорівнює факторній вазі першого фактора λ_1. Навантаження першого фактора обчислимо за формулою (4) використавши стовпчик $\boldsymbol{b}_1$ з результатами ділення елементів $\lambda_1 \cdot \vec{a}_1$ на λ_1.

Знаходження 1 фактора			
λ	bi1	bi1^2	a
4,033943	1	1	**0,946847**
3,277343	0,812442	0,660061	**0,769258**
3,896467	0,96592	0,933002	**0,914578**
3,929976	0,974227	0,949118	**0,922444**
3,947048	0,978459	0,957382	**0,926451**
		4,499564	**4,479577**

Для виявлення другого фактора $\boldsymbol{F}_2$ із результатів вимірювання параметрів потрібно виключити інформацію зв'язану із фактором $\boldsymbol{F}_1$. Для цього матрицю-стовпчик навантажень $\vec{a}_1$ першого фактора транспонуємо і отримаємо матрицю рядок $\overrightarrow{a'}_1$. Потім знаходимо матрицю $\widetilde{R} = \vec{a}_1 \cdot \overrightarrow{a'}_1$ та віднімаємо її від редукованої матриці $\boldsymbol{R}$. Маємо $R_1 = R - \widetilde{R}$ – наступну матрицю для якої повторюємо всю попередню процедуру. Знаходимо суму рядків $\boldsymbol{R}_1$, повторюємо ітерації та обчислення до знаходження $\boldsymbol{R}_2$.

Розрахунки для наступних компонент продовжуємо поки не отримаємо $\lambda_{m+1} < 1$, тоді останнім компонентом аналізу буде фактор номер $\boldsymbol{m}$.

Результати обчислень можна оформити у вигляді таблиці.

Отже, описана процедура визначення головних компонентів у факторному аналізі кореляційної матриці засобами Excel, яка відрізняється використанням методу ітерації від традиційних програмних методів. Це надає запропонованому методу певної привабливості над традиційними.

Висновки та перспективи подальших наукових розвідок. Зацікавленість науковців проблемою з'ясування факторів впливу на академічну успішність студентів обґрунтовується зміною умов організації освітнього процесу у закладах вищої освіти. На теперішній час у науковій літературі можна відслідкувати різні підходи до виокремлення таких факторів. Проте, для багатьох дослідників переважає такий спосіб як проведення факторного аналізу.

Факторний аналіз є складною процедурою, що може ускладнюватися недосконалістю зібраних даних і неточним статистичним висновком. Проте така процедура уможливлює виокремити цінні для дослідника фактори з набору зібраних даних, і зосередитися у подальшій своєї науковій роботі саме на них.

Попри існуючі способи проведення факторного аналізу запропонована і описана процедура визначення головних компонентів у факторному аналізі кореляційної матриці засобами Excel. Наведено пояснення цієї процедури з позицій теоретичних положень факторного аналізу і наведено приклад її застосування для знаходження факторів впливу на академічну успішність студентів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ / REFERENCES

1. Дужич, Н. В., Мялюк, О. П., Марущак, М. І. (2023). Аналіз чинників, які впливають на академічну успішність. Медична освіта, 3, 54–61. (Duzhych, N. V., Mialiuk, O. P., Marushchak, M. I. (2023). Analysis of factors affecting academic success. Medical education, 3, 54–61)
2. Гуменюк, Н. (2023). Дослідження впливу дистанційної освіти на успішність навчання та психологічний стан студентів в умовах воєнного періоду. Перспективи та інновації науки, 10(28), 561–569. (Gumenyuk, N. Y. (2023). Investigation of the influence of distance education on study success and psychological state of students in the conditions of marital state using information technologies. Prospects and innovations of science", 10(28), 561–569).
3. Квятковська, А. (2022). Факторний аналіз дистанційного та змішаного навчання закладів фахової передвищої освіти. Вісник КрНУ імені Михайла Остроградського, 2(133), 145–150. (Kviatkovska, A. Factor analysis of distance and blended learning in institutionsof professional prehigher education. Transactions of Kremenchuk Mykhailo Ostrohradskyi National University, 2(133), 145–150.)
4. Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., Esquivel, S. L. (2019). Practical considerations for using exploratory factor analysis in educational research. Practical Assessment, Research and Evaluation, 18(1), 6.
5. Costello, A. B., Osborne, J. (2019). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical assessment, research and evaluation, 10(1), 7.
6. Goretzko, D., Pham, T. T. H., Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. Current psychology, 40, 3510–3521.
7. Harman, H. H. (1976). Modern Factor Analysis; University of Chicago Press: Chicago, IL, USA.
8. Shrestha, N. (2021). Factor analysis as a tool for survey analysis. American journal of Applied Mathematics and statistics, 9(1), 4–11.
9. Tavakol, M., Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. International journal of medical education, 11, 245.

Kasiarum S. Finding eigenvalues and eigenvectors of factor analysis correlation matrix (a study of higher education students' academic success factor analysis).

The issue of academic success of higher education students and the factors that influence it is raised in the paper. It is emphasized that the results of pedagogical research on different stages of development of higher education demonstrate different factors of influence on students' educational achievements. Currently, the study of the impact of the organization of the educational process under martial law and distance or blended learning on the academic success of higher education students is relevant.

Researchers use factor analysis to determine the factors influencing students' academic success. The use of factor analysis makes it possible to identify hidden structures in a set of variables. It is emphasized that factor analysis is a complex procedure that can be complicated by the imperfection of the collected data and inaccurate statistical conclusions. However, such a procedure makes it possible to isolate factors valuable to the researcher from the set of collected data and to focus on them in their further scientific work. Currently, researchers use various methods for conducting the factor analysis procedure. Software products for analyzing statistical data (SPSS, SAS, Excel, etc.) greatly facilitate the conduct of factor analysis.

The article uses the example of determining the factors influencing students' academic success to show an algorithm for determining the eigenvalues and vectors of the correlation matrix for factor analysis. A feature of the procedure for determining the main factors was the author's use of the properties of the eigenvectors of matrices, which is a new detail of factor calculations. The article shows how, using the rather limited capabilities of Excel spreadsheets, it is possible to implement the finding of eigenvalues and eigenvectors for matrices of sufficiently large sizes. The author of the article sees in this the methodological attractiveness of the proposed method for solving these problems. The

method proposed and presented by the author does not aim to replace software methods, but only makes it possible to solve the problem of accessibility of calculations.

Key words**:** *factor analysis, principal components method, commonality, characteristic, load, correlation matrix, eigenvector, eigenvalues of the matrix, academic success, higher education students.*

УДК 51:37.091.3
DOI 10.5281/zenodo.14566729

С. А. Кирилащук
ORCID ID 0000-0002-8972-3541
З. В. Бондаренко
ORCID ID 0009-0009-1032-1469
В. І. Клочко
ORCID ID 0000-0003-4100-1838
Вінницький національний технічний університет

ПРОЕКТУВАННЯ ТЕХНОЛОГІЇ ВІЗУАЛІЗАЦІЇ ВИВЧЕННЯ КУРСУ ВИЩОЇ МАТЕМАТИКИ

У статті досліджено проблему проектування технології візуалізації вивчення курсу вищої математики студентами інженерних спеціальностей. Виділено основні компоненти моделі технології візуального навчання: формування базових навчальних елементів; моделі образів об'єктів навчання; підвищення якості аналітичної підготовки студентів, формування знань, умінь та навичок застосування науково-методологічного апарату. Особливістю запропонованої педагогічної технології є те, що у ході її використання поряд із візуалізацією застосовуються сучасні інформаційні технології.

Проектування візуалізації навчального матеріалу з вищої математики пропонується реалізовувати у декілька етапів: відбір навчального матеріалу, структурно-логічний аналіз і побудова структурно-логічної схеми навчальної інформації; виділення методологічних і прикладних аспектів теми; розміщення візуального навчального матеріалу з урахуванням логіки формування навчальних понять; добір опорних образів (ключових понять, символів, фрагментів схем); пошук внутрішніх логічних взаємозв'язків та міжпредметних зв'язків; складання первинного варіанта, компонування матеріалу в блоки; критичне осмислення первинного варіанта, перекомпонування, перебудова, спрощення; заключна корекція опорного конспекту, схеми чи іншого візуального образу.

При доборі змісту враховується технологія професійної спрямованості вивчення курсу вищої математики, що реалізується в трьох напрямках: методологічні знання, уміння для здійснення дослідницьких функцій; міждисциплінарна взаємодія навчальних дисциплін, що дозволяє розкрити роль математики у якісному засвоєнні спеціальних дисциплін, а також в майбутній професійної діяльності; використання комплексу навчальних професійно-орієнтованих завдань, розв'язування яких істотно впливає на мотивацію вивчення курсу вищої математики.

Дослідження підтвердило, що використання візуального контенту значно покращує ефективність подання та засвоєння понять курсу вищої математики.

Ключові слова: *курс вищої математики, проектування, візуалізація, сучасні інформаційні технології, структурно-логічної схеми, педагогічні технології, професійно-орієнтовані завдання, знаково-символьні засоби.*

Постановка проблеми. Згідно з європейськими принципами та напрямами розвитку вищої технічної освіти слід впроваджувати методики навчання, основами яких є принцип «створити» замість принципу «повторити». Ця методика особливо ефективна саме у технічній освіті, де компонента знання є лише основою для здійснення компоненти уміння, де створення