МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

На правах рукопису

Скотаренко Федір Миколайович

УДК 519.6:004.93

РОЗРОБКА МАТЕМАТИЧНИХ МЕТОДІВ ГРУПУВАННЯ ІНФОРМАЦІЇ З МАТРИЧНИМИ ОЗНАКАМИ

01.05.04 — системний аналіз і теорія оптимальних рішень

дисертація на здобуття наукового ступеня

кандидата фізико-математичних наук

Науковий керівник

Донченко Володимир Степанович

д.ф.-м.н., проф.

Київ – 2016

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

$E,\langle,\rangle$ абстрактний евклідів простір

$\langle,\rangle$ скалярний добуток

$\|.\|$ норма евклідового простору $(E,\langle,\rangle)$

R поле дійсних чисел

R^n евклідів простір числових векторів розмірності n

$R^{m\times n}$ евклідів простір матриць розмірності $m\times n$

A матриця ($m\times n$) лінійного оператора з R^n в R^m

A^* спряжений до A оператор, $\mathrm{A}^+ : R^n \to R^m$

E_m одинична матриця розмірності m, тотожний оператор в R^m

A^T транспонована до матриці A матриця

L лінійний підпростір $L\subseteq R^n, L\subseteq R^{m\times n}$

(e,λ) сингулярність лінійного оператора $\mathrm{A} : R^n \to R^m$ e - власний вектор, λ - ненульове власне число, $\|e\|=1$

$\Gamma\ a,L$ гіперплощина: $\Gamma\ a,L\ = a+L$

P_L ортогональний проектор (ОП) на лінійний підпростір

SVD абревіатура для “Single Value Decomposition” – сингулярне подання

$(u_i, \lambda_i^2), i = \overline{1, r}$ набір ненульових сингулярностей матриць

AA^T, $A \in R^{m \times n}$, $r = rankA$, $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$

$(v_i, \lambda_i^2), i = \overline{1, r}$ набір ненульових сингулярностей матриць

$A^T A$, $A \in R^{m \times n}$, $r = rankA$, $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$

ПдО псевдообернення

SVD A сингулярне подання лінійної матриці A, лінійного оператора A

A^+ - ПдО A псевдообернення лінійної матриці A, лінійного оператора A

$Z(A^*)$ ортогональний проектор на ядро A^*

$Z(A)$ ортогональний проектор на ядро A

$KerA$ множина нулів (ядро) лінійного оператора A

$R(A) = \mathrm{A}^+ A^{+T}$ оператор групування за матрицею A, оператором A

$R(A^T) = A^{+T} \mathrm{A}^+$ оператор групування за матрицею A^T, оператором A^T

СЛАР системи лінійних алгебраїчних рівнянь $Ax = b, A \in R^{m \times n}, b \in R^n, x \in R^n$

МаСЛАР матрична системи лінійних алгебраїчних рівнянь

$AX = B, A \in R^{m \times n}, \mathrm{B} \in R^{m \times p}, x \in R^{n \times p}$

ВСТУП

Актуальність теми. У багатьох прикладних задачах виникає завдання оперування об’єктами з матричним представленням. До таких задач відносять представлення фунцій за їх спостереженями, класифікацію, кластеризацію, розпізнавання образів. Розвиток та впровадження інформаційних технологій, систем штучного інтелекту неможливе без адекватного розвитку та вдосконалення методів математичного моделювання. Актуальність обумовлена необхідністю подальшого розвитку математичних методів дослідження в прикладних задачах, що характеризуються умовами невизначеності. Так, для прикладних задач є актуальним створення математичного апарату дослідження невизначеності у зв’язку із матричними об’єктами, зокрема, базовими структурами евклідових просторів. Такий стан речей визначається тим, що для R^n існує розвинений апарат оперування з базовими структурами евклідових просторів: підпростір, гіперплощина, еліпси групування. Вони пов’язані великим арсеналом засобів опису зв’язків в евклідових просторах: сингулярного подання матриці та псевдообернення. Розвиток математичного апарату з матричними представниками потребує засобів перенесення з R^n на матричні простори. Тому дослідження, пов’язані з вивченням природи невизначеності та розвитком засобів її математичного моделювання, набувають особливої актуальності.

На сьогоднішній день задачі групування інформації чи у вигляді задачі відновлення функції, представленої своїми значеннями, чи у вигляді задач кластеризації, класифікації чи розпізнавання образів, представляють важливий клас прикладних задач із широким колом застосування. Особливого значення вони набувають у зв’язку із проблемою створення технологій природного інтерфейсу у спілкуванні людини з комп’ютером. Розпізнавання мови і рукописного тексту значно спрощує взаємодію людини з комп'ютером, розпізнавання друкованого тексту використовується для перекладу документів в електронну форму, що породжує спеціальну увагу до розвитку технологій обробки мовних сигналів та зображень. Створення систем розпізнавання та синтезу мови, системи визначення особи та її емоційного стану, системи компресії мовного сигналу, системи очищення від шумів, підвищення розбірливості є важливими задачами створення природного інтерфейсу. Проблеми, що стосуються обробки зображень, включають системи виділення однорідних (у тому чи іншому сенсі) областей на зображеннях, виділення та «захват» об'єктів рухомих зображень, системи автоматичного зчитування та розпізнавання тексту різноманітного призначення, систем ідентифікації особи за зображенням і т. ін. Досягнення у згаданих областях є вагомими, але ще далекими від завершення. Подальше просування в розв'язанні задач групування інформації вимагає як ефективного використання вже наявних можливостей математичних методів, так і розвитку їх з метою застосування саме в задачах групування інформації. Це повною мірою стосується задач обробки мовних сигналів та зображень з метою розпізнавання, класифікації, кластеризації. Розвиток та вдосконалення адекватного для прикладних задач математичного апарату, втілення його в ефективних алгоритмах обробки сигналів є актуальним завданням у прикладних математичних дослідженнях.

Розвиток сучасних технологій визначається все більш високими вимогами до них. Наприклад, це стосується задач розпізнавання мови, в тому числі дактильної. Цей напрямок є одним з актуальних та перспективних у зв`язку зі створенням природних інтерфейсів, які забезпечують спілкування людини з комп`ютером природним чином. Не менш важливим напрямком є розвиток технологій аналізу сцен, зокрема, виділення та супроводження об`єктів на зображенні та їх класифікація. Виділення облич людей на фотографіях, аналіз емоцій за мімікою губ, пошук на зображенні автомобілів та їх номерних знаків – це лише деякі приклади найбільш відомих загалу задач, які знайшли практичне застосування у відомих інтернет–порталах та системах різного напрямку. Як результат, досить поширеними є програми, які дозволяють керувати автомобілем за допомогою голосових команд або жестів.

Ефективне розв`язання, зокрема, згаданих задач базується на методах класифікації, кластеризації, розпізнавання образів. Принциповими є роботи іноземних науковців, зокрема, самоорганізаційна карта Кохонена Т.К., приховані марківські моделі Баума Л.Е., інваріантні моменти Ху М.К. та інші. Важливий внесок здійснили і вітчизняні науковці. Ще півстоліття тому Глушков В.М. створив в Києві наукову школу аналізу зображень. Цей напрямок досліджень продовжив розвивати Шлезінгер М.І. Результати досліджень Шлезінгера М.І. знайшли своє практичне застосування у проекті Viewdle, який пов`язаний із комп`ютерним зором, а саме розпізнаванням облич та об`єктів. Статистичні методи розвивали Вапник В.Н. та Червоненкіс А.Я. В області розпізнавання мовної інформації важливими є роботи Вінцюка Т.К. Видатний внесок у розвиток алгебраїчних методів, зокрема, псевдообернення та його застосування в задачах групування інформації, був здійснений Кириченком М.Ф., зокрема, у співавторстві з Кривоносом Ю.Г. Розв'язання задачі лінійної дискримінації, як правило, пов'язується з роботами Фішера Р.А. та Маклоклен Г. Розв`язання задач кластеризації на основі занурення у відповідні множини (еліпсоїдальні циліндри) запропонували Куц Ю.В. та Кириченко М.Ф. В роботах Кириченка М.Ф. та Донченка В.С. [89,90] запропонована та розвинута концепція відстаней відповідності в задачах кластеризації та класифікації інформації через асоціацію (занурення) у слушні алгебраїчні структури. Принципові результати у сфері моделювання емоційних станів обличчя людини, жестової мови, автоматизації стенографування, синтезу мови отримані Краком Ю.В.

Задачі кластеризації, класифікації, розпізнавання в різних варіантах прояву та постановки є принциповими прикладними задачами, які є предметом уваги багатьох дослідників. Важливими в цьому контексті є монографія T. Kohonen T. [27], класична робота Richard O. Duda, Peter E. Hart, David G. [41], Christopher M. Bishop [4,5], роботи Вапніка В.Н. [49,50,59] та Шлезінгера М.І. [129,130,131], Яценка В.О., M. Pardalos [36]. Вагомий внесок в розвиток математичних засобів аналізу та синтезу в задачах кластеризації, класифікації, розпізнавання образів на основі розвитку апарату псевдообернення здійснено в роботах Кириченка М.Ф., Донченка В.С. [89,90], Крака Ю.В. [100,101,102,103], Кривоноса Ю.Г. [69,88,92,93,105,118], Кудіна Г.І. Принциповим у розвитку засобів аналізу та розпізнавання мовних сигналів є робот А Вінцюка Т.К. [61], Лобанова Б.М., Карпова О.М.[78], А.В . Фролова. Задача розпізнавання мовних сигналів є предметом досліджень в роботах Клетта Д. Х., Барнета Дж. А., Бернстейна М. І., Аграновського А.В., І.О. Стелі. Розвитку методів аналізу мовних сигналів на основі вейвлет-перетворення присвячені, зокрема, роботи Addison P.S., Добеші І., Polikar R., Данилова В.Я., Малла С., Смоленцева Н.К., Сорокіна В.М. [45]. Важливим засобом аналізу мовних сигналів є апарат прихованих марківських процесів, див., наприклад роботи Rabiner L.R. [24,40] і B.H. Juang. Широким в задачах кластеризації, класифікації та розпізнавання є застосування методів штучних нейронних мереж (див., наприклад, Bengio Y., DeMori R., Flammia G., and Kompe R. і Bishop C.M.) та нечітких множин (Кієдзі Асаї, Дзюндзо Ватада, Сокуке Іваї та ін. [79]). Розв’язанню практичних задач розпізнавання зображень присвячені роботи Шлезінгера М.І. [8,9,10], Бармака О.В., Муригіна К.В., Гонсалеса Р., Путятіна Є.П. [114,115], Зайченка Ю.П. [75,76], Бідюка П.І. Важливим класом задач класифікації та розпізнавання є задача розпізнавання з учителем, коли кожен із класів представлений певним набором «представників». Одним із засобів розв’язання проблеми може вважатися дихотомічна класифікація. Стандартний недолік нестійкості дихотомії при великих об’ємах навчальних вибірок в рамках задачі кластеризації, коли важко розрізнити групи класів, перевага у швидкості розпізнавання за проходженням відповідного дерева, надає їй конкурентних переваг, особливо, якщо комбінувати її із іншими методами, що можуть компенсувати згаданий недолік. Зазвичай задачі кластеризації розв’язуються через побудову певної рекурентної процедури, важливим елементом якої є процедура рекурентної перевірки елементів набору, що аналізується на предмет виділення структурних особливостей та відмінностей, на належність групі-кластеру через визначення «найкращої відповідності» класу. Така відповідність класу-кластеру визначається на основі «відстані відповідності», якою, як правило, є евклідова відстань до певного «типового» представника класу-кластеру. В роботах [64,69,71,82] обґрунтовано та запропоновано використання специфічних для класу-кластеру відстаней відповідності. Така побудова здійснюється через «занурення» відповідної навчальної вибірки у слушне геометричне утворення в евклідовому просторі числових векторів, в якому лежать відповідні «вектори ознак» – числові вектори. Такими геометричними утвореннями виступають в одному варіанті лінійні підпростори чи гіперплощини, в іншому – еліпси чи мінімальні еліпси групування. У роботі пропонується розвиток концепції використання неоднорідних за класами – кластерами відстаней відповідності на евклідові простори матриць фіксованої розмірності за слушними еліпсами та мінімальними еліпсами групування в евклідовому просторі «матричних векторів» ознак. Доведені теореми, які є основою розвитку концепції групування на матричні простори. Можливість такого розвитку забезпечується розвитком теорії псевдообернення за Муром-Пенроузом на спеціальні класи операторів між стандартними евклідовими просторами числових векторів та матриць фіксованої розмірності.

Довгий час класичними методами, які використовувалися для опису невизначеності в математичному моделюванні об'єктів, були статистичні (теоретико-імовірнісні) методи та детерміновані в тому числі у вигляді методів розв'язку „обернених" задач. Важливим у розвитку засобів опису невизначеності та побудовою адекватних засобів розв'язання практичних задач були 50–60-ті роки XX століття, коли бурхливий розвиток техніки, промислових технологій та широке впровадження обчислювальної техніки в математичному моделюванні призвів до появи й формування майже одночасно декількох нових напрямків опису та врахування невизначеності. Серед них виділимо:

- техніка псевдообернення, створена йще на початку 20-го сторіччя за Penrose'ом як засіб розв'язання багатьох задач, у тому числі в наближеному вигляді, та наступний бурхливий розвиток запропонованого напрямку; власне, це було „оптимізаційне” представлення псевдообернення після його появи в „алгебраїчному” варіанті, запропонованому Moore’ом в 1920 р.;

- теорія оцінок з гарантованою точністю (теорія мінімаксного оцінювання), що сформувалася в 60-х роках та набула розвитку в роботах F.Schweppe, А.М. Куржанського, Н.Н. Красовського, М.Ф. Кириченка, О.Г. Наконечного, В.М. Кунцевича, М.М. Личака, Г.М. Бакана, Ф.Л. Черноуська;

- теорія нечітких за L. Zadeh підмножин (1965 р.) та подальший розвиток цієї теорії в роботах A.Kuafmann’a, N. Kasabov’a, Р. Беллмана та Л. Заде;

- інженерний засіб обробки зображень, запропонований Hough’ом (перетворення Гока) (1962 р.) та оформлений у вигляді патенту, з подальшим його розвитком в роботах Rosenfeld’a, Duda&Hart’a, Ballard’a, Merlin &.Faber’a, Cohen& Toissaint’a, Tsuji & F.Matsumoto, Xu & Oja.

Зазначимо, що саме з останнім із згаданих напрямків пов’язана поява терміну „множинні моделі невизначеності”, зокрема розвинений в роботі Донченко В.С. [74], як характеризації невизначеності та джерела її появи. Значення цього терміну виходить за рамки області, в якій він з’явився, тому що може бути основою погляду на джерела та характеризацію невизначеності загалом.

Важливість засобів опису невизначеності загалом та в рамках конкретних методів у математичному моделюванні визначає актуальність досліджень дисертаційної роботи.

У напрямку математичного моделювання, загалом, завдяки Георгію Кантору базові математичні структури представляють собою певну множину та “набір зв’язків”, які існують між елементами цієї множини. Арсенал опису зв’язків не такий вже великий. Такі “зв’язки” можуть описуватися відношеннями, операціями, функціями, наборами підмножин чи їх комбінацією.

Математична модель об’єкту представляє собою передачу математичними засобами (засобами математичного структурування) його структури. Використання тих чи інших математичних структур в математичному моделюванні визначається як специфікою самого об'єкту, так і набором можливостей передачі зв’язків між елементами в математичному об’єкті як такому.

Лінійні та евклідові простори є математичними структурами з багатим арсеналом опису “зв’язків” між елементами відповідних множин. До цих “зв’язків” відносяться: ортогональність (кут між векторами), відстань, лінійні оператори, збіжність. З іншого боку, в описі реального об’єкту природним чином з’являється набір його характеристик. Такий набір може розглядатися як елемент Евклідового простору відповідної розмірності.

Ці два фактори визначають той факт, що Евклідові простори є математичною структурою, що найчастіше використовується в математичному моделюванні.

Нагадаємо, що Евклідовим простором R^n називається множина всіх числових наборів фіксованої довжини з покоординатним додаванням, покоординатним множенням на скаляр і сумою покоординатних добутків як скалярного добутку.

Загалом, скалярний добуток позначається $\langle .,. \rangle$ і є базою для визначення норми $\| . \|^2$: $\| x \|^2 = (x, x)$, що в свою чергу є основою для визначення відстані $\rho^2(x, y) = \| x - y \|^2$. Відповідно, відстань дає можливість говорити про збіжність в Евклідовому просторі, тобто визначати топологію.

Принциповими з точки зору математичного моделювання є використання в Евклідовому просторі структур, які найчастіше використовуються в описі реальних об’єктів. До таких базових структур відносяться:

- лінійні структури: підпростори, гіперплощини та лінійні оператори;
- нелінійні структури: квадратичні форми (симетричні невід’ємно визначені), еліпси та еліпсоїди, що побудовані за ними.

Нескладно помітити, що серед перерахованих вище базових структур виділяють:

- множинні: лінійні підпростори, гіперплощини, а також еліпсоїди;

- “синглові” (одиничні): матриці лінійних операторів та матриці квадратичних форм.

Ефективність використання відповідних базових структур евклідового простору визначається конструктивністю їхнього опису, побудови та обчислення пов’язаних з ними характеристик, а також можливістю переходу від одного типу структури до іншого: від множинних до сінглових та навпаки. Це передбачає конструктивність опису лінійних підпросторів та гіперплощин чи еліпсоїдів за тими чи іншими наборами векторів – навчальними вибірками в тих чи інших задачах та засобах обчислення «ступеню близькості» до побудованих об’єктів.

Визначна роль у забезпеченні зазначеної вище конструктивності належить уже згаданому апарату псевдообернення (ПдО), започаткованого Муром та Пенроузом, і особливо розвитку цього апарату в фундаментальних роботах М.Ф. Кириченка, серед яких особливо виділяють, [84].

Зазначимо, що ПдО за Муром-Пенроузом забезпечує зв’язок між сінгловими та множинними структурами евклідового простору, що суттєво розширює можливості використання структур евклідового простору в математичному моделюванні і, зокрема, в задачах групування інформації. Під цією задачею матимемо на увазі з одного боку, задачу відновлення функції, представленої своїми спостереженнями, а з іншого – задачі класифікації, кластеризації чи розпізнавання образів.

Загалом математичні методи для задач групування інформації і відновлення функції, класифікації, кластеризації, розпізнавання образів (classification, clusterisation, pattern recognition) використовують розвинений математичний апарат, який пов’язаний з евклідовим простором як числових векторів так і евклідів простір матриць та багатий на структурні зв’язки арсенал засобів оперування, що існують в евклідовому просторі.

В задачах керування чи аналітичних задачах регресії використовують скаляр векторного аргументу. В задачах керування вектор функції, рекурентного співвідношення функції.

В той же час у багатьох важливих прикладних задачах принциповим є поява представника інформації – об’єкту; зведення обєкту до математичної моделі дає можливість вичерпно аналізувати його властивості. Такими представниками можуть бути:

- матриця спектрограми;
- матриця зображення.

Досі обєктами оперували наступним чином: формували вектори ознак, які потім перетворювались чи використовувались у розв’язанні задачі групування інформації.

Наразі, сформувалася нагальна потреба перенесення базових властивостей чи властивостей опису базових лінійних і нелінійних структур, які діють із евклідового простору числових векторів на евклідів простір матриць фіксованої розмірності та скалярним добутком, з покомпонентним додаванням і множенням на скаляр, та скаляр добутку. Перенесення апаратів SVD, ПдО і пов’язаних з ними структур в матричному випадку вимагає відповідних математичних визначень і встановлення можливсті аналогів SVD, ПдО для матричних просторів базових структур. Це можуть бути еліпси групування породжених матрицями, що вимагає розв’язання питання ПдО у зв’язку з матричними евклідовими просторами.

Зв’язок роботи з науковими програмами, планами, темами. Робота виконана у відповідності до плану наукових досліджень кафедри системного аналізу та теорії прийняття рішень факультету комп’ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка Міністерства освіти і науки України у рамках науково-дослідницької теми: №11БФ015-06 “Проблеми теорії прийняття рішень та системного аналізу стохастичних мереж” (державний номер реєстрації 0111U006680, термін виконання 2011-2015р.р. за програмою "Інформатизація суспільства").

Обґрунтованість та достовірність отриманих наукових результатів підтверджується:

– узгодженістю отриманих результатів з класично відомими;

– об’єктивним характером використаної в процесі дослідження інформації, застосуванням фундаментальних положень теорії оптимальних рішень, коректністю постановок задач, строгим математичним доведенням результатів.

Мета і задачі дослідження. Метою дослідження є розвиток математичних засобів групування інформації, розробка моделей реалізації систем розпізнавання мови та знаків дактильної мови жестів.

Для досягнення мети поставлені наступні задачі:

– реалізувати концепцію кортежних операторів для матричних кортежів, що дасть можливість побудувати конструктивну теорію SVD, ПдО (псевдообернення) для таких операторів,

– поставити та розв’язати задачу лінійної робастної дискримінації в матричних евклідових просторах,

– розробити конструктивний алгоритм реалізації задачі лінійної дискримінації,

– побудувати рекурентну формулу псевдообернення для кортежних операторів: узагальнену формулу Ґревіля-Кириченка для кортежних операторів.

Предметом дослідження є базові лінійні та нелінійні структури об’єктів в матричному евклідовому просторі.

Об’єктом дослідження є засоби оперування такими структурами в матричному евклідовому просторі.

Реалізація задач дослідження пов’язана із засобами оперування з базовими структурами, лінійними і нелінійними в матричних евклідових просторах, що запропоновані в контексті кортежних операторів.

Методи дослідження. Методологічну основу дослідження становлять алгебраїчні методи, включаючи методи сингулярного розкладу та псевдообернення, методи кластеризації та класифікації, розпізнавання образів.

Наукова новизна дисертаційного дослідження в теоретико-методологічному обґрунтуванні та реалізації концепції кортежних операторів за матричними кортежами матриць фіксованої розмірності, що дало можливість:

вперше:

- поставити та розробити конструктивний алгоритм реалізації задачі лінійної дискримінації в матричних евклідових просторах;
- вичерпно та конструктивно дослідити систему лінійних алгебраїчних рівнянь для кортежних операторів та спряжених до них;
- поставити та розв'язати задачу рекурентного обчислення псевдообернення для кортежних операторів та спряжених до них за розширенням кортежів новими матричним елементами; побудувати аналоги прямих формул Ґревіля-Кириченка для матриць.

удосконалити:

- засоби дослідження лінійних та нелінійних базових структур в матричних евклідових просторах;

набули подальшого розвитку:

- розв'язання задачі побудови сингулярного подання як самих кортежних операторів, так і спряжених до них, зведенням до задачі на власні значення для класичних матричних операторів;
- конструктивна побудова псевдообернених операторів до самих кортежних та спряжених до них, що уможливило створення основи до побудови низки методів розв'язання задач групування інформації для представників з матричних евклідових просторів;
- алгоритми розв’язання задач лінійної дискримінації у їхньому застосуванні до матричних евклідових просторів;

базові структури евклідового простору: лінійний підпростір, гіперплощина, лінійний оператор, квадратичні форми, зокрема, еліпсоїди чи їх внутрішність.

Особистий внесок. Дисертація є самостійною науковою працею, в якій висвітлені власні ідеї і розробки автора, що дозволили розв`язати поставлені задачі. Робота містить теоретичні та методичні положення, розроблені алгоритми та відповідні висновки, сформульовані дисертантом особисто. Використані в дисертації ідеї, положення чи гіпотези інших авторів мають відповідні посилання і використані лише для підкріплення ідей здобувача.

У роботах, що написані у співавторстві, внесок автора полягає у наступному: у роботі [9] опис методів розв'язання задачі побудови сингулярного подання як самих кортежних операторів, так і спряжених до них, зведенням до задачі на власні значення для класичних матричних операторів; у роботі [12] опис задачі лінійної дискримінації для евклідових просторів матриць; у роботі [10] формулювання та доведення прямих формул Ґревіля-Кириченка для кортежних операторів, пов'язаних з евклідовими просторами матриць; у роботі [118] формулювання задачі та доведення теорем щодо розв'язків задачі лінійної дискримінації в евклідових просторах матриць; у роботі [6] побудова апарату псевдобернененя для кортежних операторів, визначених матричними кортежами, а також для спряжених до них; у роботі [11] постановка задачі та доведення теорем про умови існування розв'язків задачі лінійної дискримінації; у роботі [8] формулювання алгоритму розв'язання задачі лінійної дискримінації, [69] опис множини розв'язків чи псевдорозв'язків СЛАР в концепції кортежних операторів за матричними кортежами.

Особистий внесок у рамках проведених досліджень полягає в розробці методів та алгоритмів реалізації концепції кортежних операторів, що дозволила перенести засоби аналізу і використання базових структур евклідового простору R^n на матричні евклідові простори. Сформульована та розв'язана задача робастної лінійної дискримінації для матричних евклідових просторів. Побудований алгоритм розв'язання такої задачі. Побудовані рекурентні формули псевдообернення за розширенням кортежу додатковими матричними елементами, що є аналогами формули Гревіля-Кириченка для матричних евклідових просторів.

Реалізована концепція, запропонована для евклідових матричних просторів, що дозволило перенести створення засобів аналізу та базові структури евклідового простору від векторного представлення до матричного.

Розв'язані такі задачі:

1) реалізація концепції кортежних операторів: зведення задачі SVD, ПДО для кортежних операторів задачі на власні значення для звичайних матриць;

2) постановка та розв'язання задачі робастної лінійної дискримінації для матричних евклідових просторів;

3) побудова рекурентних формул псевдообернення для кортежних операторів: узагальнення формули Гревіля-Кириченка для кортежних операторів;

4) розвиток можливості розв'язання задач групування в класичній задачі визначення інформативних та неінформативних ознак для задач жестової мови.

Практична цінність і впровадження результатів роботи. Отримані в дисертації наукові результати є внеском у розвиток математичних методів розв'язання прикладних задач, явищ, процесів в умовах невизначеності, коли досліджувані об'єкти представляються матрицями. Побудовані алгоритми використання отриманих результатів у розв'язанні задач групування інформації. Результати наукового дослідження методів та алгоритмів псевдообернення в евклідовому просторі матриць для розв'язання прикладних задач у сучасних інформаційних системах, отримані в рамках підготовки кандидатської дисертації "Розробка математичних методів групування інформації з матричними ознаками", впроваджені у 2013-2014 н.р. у навчальний процес кафедри системного аналізу та теорії прийняття рішень факультету кібернетики Київського національного університету імені Тараса Шевченка при викладанні курсу "Обробка інформації в умовах невизначеності" для студентів 1 курсу магістратури за спеціальністю "Системи і методи прийняття рішень". Застосовано при виконанні науково-дослідної теми №11БФ015-06 "Проблеми теорії прийняття рішень та системного аналізу стохастичних мереж" (державний номер реєстрації 0111U006680, термін виконання 2011-2015р.р. за програмою "Інформатизація суспільства").

Апробація роботи. Результати дисертаційної роботи доповідались на семінарах кафедри моделювання складних систем та кафедри системного аналізу та теорії прийняття рішень факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка, представлені на наступних конференціях: на міжнародній конференції International conference "Knowledge Dialogue- Solution" KDS-2012, Kyiv, Ukraine; на міжнародній конференції ITHEA International Conference - 2014 Varna, Bulgaria; 2-га міжнародна науково-практична конференція 'Інформаційні технології та взаємодія' (IT&I) (Київ, Україна, 2015), 18-th International Conference SAIT 2016 (System Analysis and Information Technology) (Київ, Україна, 2016), 5-та міжнародна науково-практична конференція «Проблеми інформатики та комп'ютерної техніки» ПІКТ-2016. (Чернівці, Україна, 2016), PDMU-2016 XXVII International Conference Problems Of Decision Making Under Uncertainties (Tbilisi-Batumi, Georgia, 2016).

Публікації. За результатами дослідження опубліковано 12 наукових праць – 8 наукових статей у фахових виданнях, з них 3 в іноземних журналах, що входять до міжнародних наукометричних баз, та 4 тези доповідей на міжнародних наукових конференціях.

РОЗДІЛ 1
ОГЛЯД МАТЕМАТИЧНИХ РЕЗУЛЬТАТІВ ТЕОРІЇ ПСЕВДООБЕРЕННЯ В ЕВКЛІДОВИХ ПРОСТОРАХ

У цьому розділі наведені основні визначення і твердження, що є принциповими для використання теорії ПдО (за Муром-Пенроузом) в математичному моделюванні.

1.1. Базові математичні структури в задачі групування інформації

Розв'язок таких задач пов'язаний із виділенням у сигналі, аналізованому на предмет ідентифікації, характерних інформаційних ознак, за якими і здійснюється розпізнавання. Такий набір інформаційних ознак, як правило, представляє собою набір числових характеристик, тобто числовий вектор. Таке представлення аналізованого сигналу є інтерпретаційним припущенням, що дозволяє використовувати в математичній моделі задачі розпізнавання потужний набір структур евклідового простору числових векторів фіксованої розмірності – того чи іншого варіанту R^n.

Зазвичай, таке використання обмежується структурою метричного простору – відстані в евклідовому просторі, хоча такий запас структур (базових структур евклідового простору) є значно ширшим. Він включає в себе як множинні структури: лінійні підпростори та гіперплощини, а також слушні еліпсоїди, (точніше еліпсоїдальні циліндри), так і структури «синглового» (від single – одиничний) характеру, якими є лінійні оператори та невід'ємно визначені квадратичні форми.

Ефективність використання структур (зв'язків між елементами) в евклідовому просторі R^n визначається ефективністю математичного апарату, який забезпечує використання відповідних структур для їхнього конструктивного опису та побудови, а також обчисленням «відстаней відповідності» для таких структур. Такими «відстанями відповідності» можуть бути відстані від відповідних множинних утворень: підпросторів, гіперплощин, еліпсоїдів.

Зауважимо, що такі «відстані відповідності» ефективно обчислюються за «сингловими» структурами: лінійними чи нелінійними. Тому принципової ваги набуває питання встановлення взаємної відповідності між множинними та «сингловими» базовими структурами евклідового простору R^n.

Як виявляється, адекватним математичним апаратом, що забезпечує встановлення такого зв'язку, є апарат псевдообернення (в подальшому – ПдО) за Муром-Пенроузом [37] та його розвиток в принциповій роботі М.Ф.Кириченка [86]. Детальнішу інформацію про ПдО та його роль для опису базових структур R^n, а також використання згаданих структур в розв'язанні задач групування інформації, можна знайти в роботах [85],[86].

Принципова роль фундаментальної математики в прикладних дослідженнях полягає в тому, що вона надає засоби отримання верифікованих тверджень про реальний об'єкт з метою їх використання для передбачення поведінки цього об'єкту.

Прикладна математика – це сукупність методів опису реальних об'єктів засобами математичного структурування з метою встановлення їх безумовно вірних властивостей.

Математична модель – це опис математичними засобами реального об'єкту.

Перехід від структури об'єкту до її математичного опису вимагає засобів інтерпретування елементів структури об'єкту в математичних термінах і навпаки.

Так, основні складові математичного моделювання можуть бути представлені наступною схемою (діаграмою) (рис 1.1).

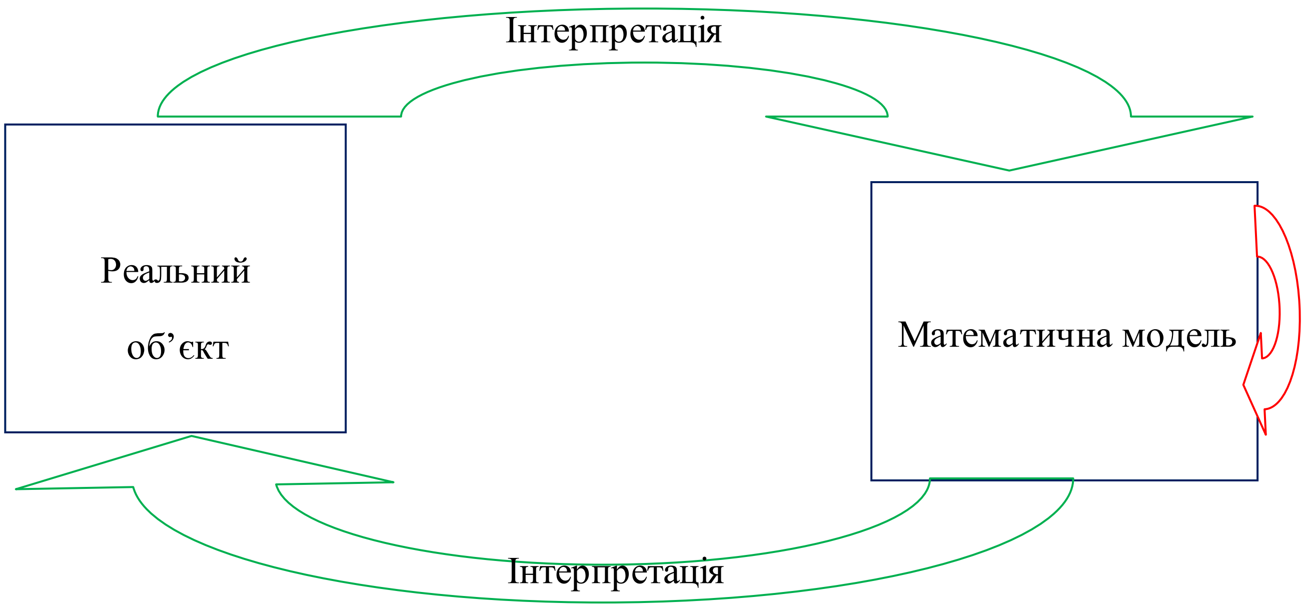

Рис. 1.1. Зв'язок між прикладною та фундаментальною математикою

Реальний об'єкт $\Longrightarrow$ Інтерпретація $\Longrightarrow$ Математична модель $\Longrightarrow$ Інтерпретація в зворотній бік $\Longrightarrow$ Результат .

Цикл всередині математичної моделі відповідає отриманню верифікованих тверджень в рамках математичної моделі.

Отже, змістом схеми є:

1) перехід від об'єкту до його математичного опису;

2) отримання безумовно вірних тверджень в рамках математичної моделі;

3) інтерпретація їх в термінах елементів структури досліджуваного об'єкту.

Так, наприклад, якщо математичною моделлю об'єкту є система лінійних рівнянь, всередині математичної моделі стрілці може відповідати твердження про розв'язок СЛАР, а обернена стрілка - інтерпретацію розв'язку СЛАР, як системи оптимальних характеристик досліджуваного об'єкту.

Опис зв'язків в математичних структурах будемо розглядати як комбінацію способів структурування. Все, що ми маємо у математиці, - це множини з тою чи іншою комбінацією способів структурування. Структурування, тобто опис певних зв'язків, може бути здійсненим одним із 4-ох способів:

1) операціями

2) функціями (наприклад скалярний добуток, норма)

3) відношеннями ≤, ≥, =…

4) граничними переходами (граничними точками) – топологією. Вона визначається наборами множин.

Звісно, у структуруванні можливе використання комбінації вище згаданих зв'язків.

Говорячи про евклідів простір як про математичну структуру, слід зазначити, що зв'язки між елементами задаються:

- операціями;
- функціями;
- топологією.

Структурування тої чи іншої множини – це опис зв'язків, які можуть існувати між елементами досліджуваних множин. Найчастіше з об'єктом асоціюється певний набір характеристик. Взяті разом, вони утворюють упорядкований набір чисел, тобто є елементами R^n.

1.2. Евклідів простір R^n та "вектори ознак"

Відзначимо, що «занурення» моделі в евклідів простір R^n є найпростішим варіантом використання, як правило, найпростіших (лінійних операцій та відстані) структур евклідового простору R^n. У багатьох прикладних задачах слушним є використання складніших модельних відповідників реальних об'єктів, складніших «векторів ознак». Матричним поданням мовного сигналу є спектрограма, яка є матрицею. Так само зображення того чи іншого об'єкту теж є матрицею. Отже, доцільним було б аналізувати саме матрицю, як модельне втілення аналізованого об'єкту, хоча наразі такі матриці піддаються подальшій обробці, щоб побудувати класичний варіант модельного представника – класичний вектор ознак у вигляді набору числових характеристик (числового вектора), що і фігурує у всіх подальших розглядах, представляючи первинний об'єкт.

У роботі [68] відзначено, що базові структури евклідового простору, які є принциповими в задачах групування інформації, можуть бути лінійні та нелінійні. Обидва типи структур евклідового простору можуть бути множинними: підпростори і гіперплощини, еліпси групування, або сингловими (одиничними): лінійні оператори, квадратичні форми. Відзначимо, що до базових нелінійних структур віднесені квадратичні форми (в роботі – невід'ємно визначені) та множинні відповідники – еліпси групування в різних формах. Важливою особливістю еліпсів групування є те, що вони містять всі вектори із набору векторів і що для них можна вибрати «мінімальні» еліпси. У згаданій роботі наведено конструктивні методи генерації та опису базових структур, що породжуються сукупністю векторів, наведені також формули взаємного переходу від одних типів структур до інших: від лінійних підпросторів і гіперплощин до матриць і навпаки, а також від набору векторів до матриць квадратичних форм і еліпсів групування. Також розглянуті конструктивні способи породження таких структур та ортогональних проекторів, пов'язаних із вказаними об'єктами. У тому ж руслі конструктивності лежить розгляд концепції групування операторів і мінімальних еліпсів групувань. Концепція групування операторів дозволяє використовувати в прикладних дослідженнях зв'язок того чи іншого набору векторів з важливим видом нелінійних структур евклідових просторів – мінімальними еліпсоїдами групувань. І в лінійному, і в нелінійному випадку згадана конструктивність забезпечується застосуванням псевдообернення за Муром-Пенроузом (ПдО), а також новими результатами в цій області, що беруть свій початок і спираються на фундаментальну роботу М.Ф Кириченка [86]. Таким чином, для розширення можливості апарату запропоновано розвиток теорії псевдообернення на матричні відображення між евклідовими просторами матриць фіксованих розмірностей із слідовими скалярними добутками. Крім того, на випадок відображень між евклідовими матричними просторами перенесена теорема про сингулярний розклад матриці оператора, доведена теорема згортки про співвідношення між сингулярним розкладом однієї і тієї ж матриці у векторному і матричному випадках.

Важливість та ефективність можливого розширення застосування наведених результатів визначається можливістю їх використання для розв'язання широкого класу прикладних задач, включаючи лінійну регресію (в тому числі векторну), задачі оптимального керування, кластеризацію, прогнозування і функціональні мережі, які є узагальненням штучних нейронних мереж. Структура лінійного простору займає провідне місце серед найважливіших математичних структур. Це, з одного боку, об'єкти на основі матриць: лінійні оператори та квадратичні форми, з іншого – структури, які можна назвати множинними. До них можна віднести підпростори і гіперплощини, а також еліпсоїди, що породжені невід'ємно визначеними квадратичними формами. Важливу роль у прикладних дослідженнях відіграють конструктивні методи опису відповідних об'єктів. Важливим в цьому відношенні стосовно до базових лінійних структур є використання ортогональних проекторів, для нелінійних – використання групуючих операторів. В обох випадках апарат псевдообернення дозволяє конструктивно будувати як ті, так і інші. Саме застосування цього апарату є основою для розв'язання основних завдань, пов'язаних з побудовою, описом і використанням базових структур лінійних, точніше, евклідових просторів. Відзначимо, що ортогональні проектори відіграють важливу роль у вичерпному дослідженні систем лінійних алгебраїчних рівнянь (СЛАР). Принципово важливі вони також і в постановці й розв'язанні важливих оптимізаційних задач з квадратичними функціоналами якості в евклідових просторах. Важлива роль ортогональних проекторів у розв'язанні прикладних задач повною мірою зберігається і в зв'язку із задачею побудови найкращих квадратичних наближень правої частини СЛАР значеннями лівої, коли СЛАР не має точних розв'язків. Такі найкращі наближення називають також псевдорозв'язками. Конструктивний опис ортогональних проекторів у зв'язку з природними підпросторами лінійного оператора прямо визначається псевдооберненням за Муром-Пенроузом [33,36,37].

Слід зазначити, також, що важливу роль у застосуванні апарату псевдообернення відіграє сингулярне представлення (його називають також SVD-представленням) довільної $m \times n$ матриці в специфічній формі запису у вигляді зваженої суми тензорних добутків спеціального набору пар векторів з прозорим геометричним змістом. При цьому матриця розглядається як матриця лінійного оператора між двома просторами числових векторів R^n та R^m. Це визначає специфіку сингулярного розкладу та інтерпретацію його складових.

Зауважимо про те, що одна і та ж $m \times n$ матриця A задається як лінійний оператор $A: R^n \to R^m$ між евклідовими просторами числових векторів, так і лінійний оператор $A: R^{n \times p} \to R^{m \times p}$ для довільного натурального p. Перехід до евклідових просторів $R^{n \times p}$, $R^{m \times p}$ матриць зі слідовими скалярними добутками вимагає врахування специфіки цих евклідових просторів для передачі принципових особливостей SVD-представлень. У даній роботі наведено SVD-представлення для матриць, що розглядаються як лінійні оператори між просторами матриць, доведена теорема згортки, що зв'язує два варіанти SVD-представлення матриці: векторного і матричного [69].

Важливим в описі нелінійних структур є мінімальні еліпси групування, як і в лінійному випадку, конструктивні способи їх опису. Кажучи про мінімальні еліпси групувань, маємо на увазі «мінімальний еліпс», який неможна стиснути, оптимальний у певному сенсі, що включає заданий набір векторів. Як виявляється, матрицею квадратичної форми для такого еліпса є групуючий оператор, який конструктивно описується засобами псевдо-обернення.

1.4. Евклідові простори

Абстрактне визначення евклідового простору. Евклідовий простір E – це скінченно вимірний лінійний простір із визначеним скалярним добутком $\langle , \rangle$.

Наведемо означення основних об'єктів, які будемо використовувати.

Евклідовий простір R^n визначатимемо як множину скінченних числових послідовностей однієї і тієї ж довжини n, записаних в стовпчик з покоординатними операціями додавання та множення на скаляр та сумою покоординатних добутків в якості скалярного добутку. Саме такий варіант евклідового простору будемо визначати через R^n, а його елементи – через $a : a^T = (a_1, ..., a_n)$, $\| T \|$ - означає транспонування. Стандартні ортонормовані базиси, складені з векторів єдиною одиничною компонентою (решта – нулі) на позиції з відповідним номером будуть позначатись для R^m і R^n відповідно через $e(j) \in R^m, j = \overline{1, m}, e_{(i)} \in R^n, i = \overline{i, n}$, а для загального випадку - $e_k \in R^n, k = \overline{1, n}$. Оператор A із R^n, в R^m: $A : R^n \to R^m$, заданий в ортонормованих базисах $e(j) \in R^m, j = \overline{1, m}, e_{(i)} \in R^n, i = \overline{i, n}$, будемо ототожнювати з $m \times n$ - матрицею $A = (a_{ij})$ цього оператора в цих базисах.

Евклідовим простором всіх $m \times n$ матриць з покоординатним множенням та додаванням, «слідовим» скалярним добутком будемо позначати $R^{m \times n}$. «Слідовий» скалярний добуток $(\cdot, \cdot)_{tr}$ двох матриць визначається відношенням:

$$(C\cdot, D\cdot)_{tr} = \sum_{i=1}^{n} \sum_{k=1}^{m} c_{ki} d_{ki} = trC^T D, C = (c_{ij}), D = (d_{ij}) \in R^{m \times n} \tag{1.1}$$

Лінійним оператором A , що задається $m \times n$ матрицею є відображення із R^n в R^m, дія якого визначається за правилами матричної алгебри.

Можна говорити про матрицю лінійного оператора в контексті абстрактного скінченного вектора лінійного простору.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{i,j});$$

Сингулярність (υ,λ) $\lambda \in R^1, \lambda \neq 0, \upsilon \in R^n$, для лінійного оператора матриці $A - n \times n$ визначається співвідношенням: $A\upsilon = \lambda\upsilon$.

Лінійний підпростір, породжений системою векторів $c_k \in R^n, k = \overline{1,K}$, буде позначатися через $L(c_k, k = \overline{1,K}) \equiv L(c_1 ..., c_K)$, а лінійний підпростір значень лінійного оператора $A\colon R^n \to R^m$ – через L_A, відповідно, для A^T - через L_{A^T}.

Для матриці $A = (a_{ij})$ будемо використовувати також блочне представлення за стовпчиками та рядками:

$$A = \begin{pmatrix} a_{(1)}^T \\ \Lambda \\ a_{(m)}^T \end{pmatrix} = (a(1) \ldots a(n)), a_{(i)} \in R^n, i = \overline{1,m}, a(j) \in R^m, j = \overline{1,n}. \qquad (1.2)$$

1.3.1. Базові структури евклідового простору

Базові структури евклідових просторів, як зазначалося вище, поділяються з одного боку, на синглові (одинарні) та множинні, з іншого, на лінійні та нелінійні.

а) Множинні лінійні структури. До них, як зазначалося вище, відносяться:

Лінійні підпростори – це підмножини, замкнені щодо лінійних операцій, тобто застосування операцій до елементів цього підпростору не виводять результат за межі цієї сукупності. Прикладом є лінійний підпростір породжений $c_1, ...c_k \in R^n$:

$$L(c_1, ...c_k) = \{x \in R^n : x = \sum \lambda_r c_r, \lambda_r \in R, r = \overline{1,k}\}$$

Гіперплощина**,** позначувана $\Gamma(a, L)$ – це підпростір L, зміщений на вектор a.

$\Gamma(a, L) = L + a = \{y : y = a + x, x \in L\}$; $\Gamma(\bar{a}, L(\tilde{a}_1, .., \tilde{a}_n))$;

Термін «гіперплощина» може тлумачитись у двох сенсах . В одному з них - це підпростір будь-якої розмірності, зміщений на вектор, в іншому – зміщений підпростір, коли його розмірність менша на 1 ніж розмірність всього простору.

Еліпс – позначатимемо $\Im(r,B), r \in R^1, B : B = B^T, B \geq 0$: $\Im(r,B) = \{x : x^T B x = r\}$.

Еліпсоїд – позначатимемо $\Im(r,B), r \in R^1, B: B = B^T, B \geq 0$: $\Im(r,B) = \{x: x^T Bx \leq 1\}$ – внутрішність слушного еліпсу.

b) Синглові (одиничні) структури. Як зазначалося вище це:

Матриця лінійного оператора. $A = (a_{i,j})$ із застосуванням до вектора-стовпчика у відповідності до операцій матричної алгебри.

Матриця квадратичної форми. $B = (b_{i,j})$, яка визначає квадратичну форму на вектор x у відповідності до співвідношення: $x^T Bx$.

1.3.2. Зв'язки між структурами евклідового простору

Існує набір властивостей, які описують зв'язки між базовими структурами.

1. *Набори векторів, підпростори, матриця, складена із набору векторів* $a(j), j = \overline{1,n}$. Справедливе наступне твердження:

$$L_A = L(a(1),...,a(n)), A = (a(1),...a(n)) \tag{1.3}$$

Таким чином, лінійний підпростір, породжений набором векторів, співпадає з підпростором значень матриці, складеної із векторів набору, як із стовпчиків.

2 . «*Вектори із матриць*». Для елементів стовпчикового чи рядкового представлення матриці $A \in R^{m \times n}$ справедливе твердження:

$$a(j) = Ae_{(j)}, j = \overline{1,n}, a^T_{(i)} = e^T(i)A, i = \overline{1,m}. \tag{1.4}$$

3. *Добуток матриць як сума тензорних добутків.* Для добутку довільних матриць B, C зі стовпчиковими та рядковими представленнями відповідно:

$$B = (b(1)\ldots b(r)), b(j) \in R^m, \ j = \overline{1,r}, C = \begin{pmatrix} c^T_{(1)} \\ \vdots \\ c^T_{(r)} \end{pmatrix}, c_{(r)}, c(i) \in R^n, i = \overline{1,r} \tag{1.5}$$

також для діагональної матриці $\Lambda = diag(\lambda_1,..., \lambda_r)$, справедливе співвідношення:

$$B\Lambda C = \sum_{i=1}^{r} \lambda_i b(i) c_{(i)}^T \qquad (1.6)$$

1.3.3. Ортогональні проектори

Важливою складовою апарату конструктивного опису і використання лінійних структур евклідових просторів є поняття ортогонального проектора (надалі ОП). Загальною основою ефективного використання ортогональних проекторів, як і можливості їх конструктивної побудови у зв'язку з лінійними підпросторами, є наявність для них двох еквівалентних визначень проекторів, а також можливість їхнього конструктивного опису через псевдообернення.

ОП є важливим класом операторів, які дозволяють, зокрема, конструктивно обчислювати відстані до підпростору та гіперплощин.

Як зазначалося вище, використовують два еквівалентних визначення ОП “алгебраїчне ” та “ геометричне ”.

“Геометричне” визначення ОП визначається за розкладом в ортогональну суму двох підпросторів:

Нехай E розкладається в пряму суму ортогональних підпросторів $E = L + L^{\perp}$. Це означає

$$\forall x \in E, \exists x_L \in L, x_{L^{\perp}} \in L^{\perp} : x = x_L + x_{L^{\perp}},$$

Отже, це подання однозначне. Ортогональним проектором P_L на лінійний підпростір $L \subseteq E$ називається оператор, що визначається співвідношенням:

$$P_L x = P_L(x_L + x_{L^{\perp}}) = x_L .$$

Очевидним чином, оператор ортогонального проектування є лінійним оператором. Так само в силу симетричності розклад E в пряму суму визначає одночасно два ортогональні проектори P_L і $P_{L^{\perp}}$.

Для лінійного оператора $P : E_1 \to E_2$ над евклідовими просторами E_1, E_2 природні лінійні підпростори $L_A = \Re(A)$ та $L_A{}^* = \Re(A^*)$ є множинами можливих значень лінійних операторів A, A^* та їхніх L_A, L_{A^T}, відповідно для евклідових просторів числових векторів.

L_A, L_{A^*} та їхні ортогональні доповнення позначимо $L^{\perp}{}_A, L^{\perp}{}_{A^*}$.

Для A представлених матрицями: лінійних операторів над R^n- це підпростори $L_A, L_{A^*}, L^{\perp}{}_A, L^{\perp}{}_{A^*}$.

Ортогональні проектори цих підпросторів позначатимуться Z_A з відповідним індексом. Крім того, очевидним чином:

$P_{L^{\perp}_{A^*}} = I_1 - P_{L_{A^*}} \equiv Z(A)$

$P_{L^{\perp}_A} = I_2 - P_{L_A} \equiv Z(A^*)$

$P_L + P_{L^{\perp}} = I_n$

Для матриць R^n відповідно $Z(A), Z(A^T)$.

Доповнення один до одного ОП. Розклад довільного вектора $x \in E$ в силу симетричності відносно ортогональних складових визначає одночасно два ортогональні проектори: $P_L, P_{L^{\perp}}$, $P_L x = x_L; P_{L^{\perp}} x = x_{L^{\perp}}$; з очевидним відношенням :

де I_n – тотожний оператор відповідної розмірності.

Для R^n, I_n - є одиничною матрицею.

Алгебраїчне визначення ОП. Лінійний оператор $P: E \to E$, є оператором ортогонального проектування, тоді і тільки тоді, якщо він є ідемпотентним і симетричним оператором $P^2 = P; P^* = P$. Для $P: R^n \to R^n$ спряження "$*$" замінюється на транспонування.

Доповнюючі один до одного ОП і відповідні їм підпростори. Для ортогонального проектора P_L на підпростір L оператор $Z_L \equiv E - P_L$ є ортогональним проектором на ортогональне доповнення $L^{\perp}: Z_L \equiv E - P_L = P_{L^{\perp}}$.

Абстрактне визначення ОП. Для того, щоб лінійний оператор $P: E \to E$, був оператором ортогонального проектування необхідно і достатньо, щоб він був ідемпотентним та симетричним оператором. Лінійний підпростір L_P, на який відбувається ортогональне проектування відповідно до «геометричного визначення» описується одним із двох співвідношень:

$$L_P = \{x : x = Pu, u \in E\} = \{x : x = Px, x \in E\}.$$

1.3.3.1. Ортогональний проектор на ортогональне доповнення

Крім того, при використанні також ОП на ортогональних доповненнях згаданих підпросторів, тобто $E_m - P(A^T)$ та $E_n - P(A)$, вони відповідно позначатимуться $Z(A^T)$ та $Z(A)$:

$$Z(A) = E_n - P(A)$$

$$Z(A^T) = E_m - P(A^T)$$

$$P(A\ ) = P_{L_{A^T}} = AA^+;$$

$$P(A^T) = P_{L_A} = A^+A;$$

$$Z(A) = E_n - P(A) = E_n - A^+A,$$

$$Z(A^T) = E_m - P(A^T) = E_m - A^{T+}A^T = E_m - AA^+.$$

Важливість введення ОП на ортогональне доповнення, визначається тим, що вони є ОП на ядра операторів A, A^* відповідно: $Z(A) = P_{KerA}$, $Z(A^*) = P_{KerA^*}$ для матриць A, A^T відповідно $Z(A) = P_{KerA}$, $Z(A^T) = P_{KerA^T}$

Зауваження 1. Підпростір $L_{A^T}^{\perp}$ є ядром, $KerA$, (множиною нулів) оператора A:

$$L_{A^T}^{\perp} = KerA = Z(A)R^n \ , \ Z(A^T) = P_{KerA^T}.$$

1.3.4. Сингулярний розклад та псевдообернення матриці

Ефективність використання структур (зв’язків між елементами) в евклідовому просторі R^n визначається конструктивністю математичного апарату, який забезпечує використання відповідних структур для їхнього конструктивного опису та побудови, а також обчисленням «відстаней відповідності» для таких структур. Такими «відстанями відповідності» можуть бути евклідові відстані від відповідних множинних утворень: підпросторів, гіперплощин, еліпсоїдів. Зауважимо, що такі «відстані відповідності» ефективно обчислюються за «сингловими» структурами: лінійними чи нелінійними. Тому принципової ваги набуває питання встановлення взаємної відповідності між множинними та «сингловими» базовими структурами евклідового простору R^n. Як виявляється, адекватним математичним апаратом, який забезпечує встановлення такого зв'язку, є апарат псевдообернення (в подальшому – ПдО) за Муром-Пенроузом ([33,36,37], та його розвиток в принциповій роботі М.Ф.Кириченка [86]. Детальнішу інформацію про ПдО та його роль для опису базових структур R^n, а також використання згаданих структур в розв'язанні задач групування інформації, можна знайти в роботах [87],[88].

Псевдообернення за Муром-Пенроузом зазвичай вводиться і розглядається у зв'язку з матрицями довільної розмірності і може бути визначено одним із декількох еквівалентних варіантів, включаючи SVD-подання матриці. Останній є конструктивним способом задання та використання ПдО. Саме SVD-подання матриці полягає в тому чи іншому варіанті її факторизації. Використанню SVD, а, отже, і ПдО, можна надати ширшого – операторного значення, що дозволяє користуватися технікою ПдО для лінійних операторів над широким класом евклідових просторів. Нижче наводиться теорема про SVD-подання матриці саме у такому «операторному» варіанті та відповідний варіант теореми для лінійного простору над довільними евклідовими просторами.

Теорема 1.1 (SVD-розклад для лінійного оператора, представленого матрицею ([25], див. також [6])). Для довільної матриці $A \in R^{m \times n}$ рангу $r \leq \min(m, n)$, яка розглядається як лінійне відображення між евклідовими просторами числових векторів $A: R^n \to R^m$, справедливе наступне представлення матриці у вигляді зваженої суми тензорних добутків векторів

$$A = \sum_{i=1}^{r} \lambda_i u_i v_i^T \qquad (1.7)$$

де

- $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0$ спільний набір ненульових власних чисел матриць AA^T, $A^T A$;

- $u_i \in R^m,\ i=\overline{1,r}$ – ортонормований набір власних векторів матриці AA^T, що відповідають ненульовим власним числам $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$: $AA^T u_i = \lambda_i^2, \lambda_i^2 > 0$, $i=\overline{1,r}$, $u_i^T u_j = \delta_{ij},\ i \neq j$;
- $v_i \in R^n,\ i=\overline{1,r}$ – ортонормований набір власних векторів матриці $A^T A$, що відповідають ненульовим власним числам $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$: $A^T A v_i = \lambda_i^2 v_i, \lambda_i^2 > 0$, $i=\overline{1,r}$, $v_i^T v_j = \delta_{ij},\ i \neq j$,

крім того,

$$u_i = \frac{Av_i}{\lambda_i},\ i=\overline{1,r},$$

$$v_i = \frac{A^T u_i}{\lambda_i},\ i=\overline{1,r}.$$

Дія оператора A на вектор $x \in R^n$ в SVD-розкладі, являє собою перенесення розкладу вектора по ортогональному базису або його частині $v_i \in R^n,\ i=\overline{1,r}$ з координатами $x_i = v_i^T x = (v_i, x)_{R^n},\ i=\overline{1,r}$, в одному просторі в розкладі по ортогональному базису $u_i \in R^m,\ i=\overline{1,r}$, або його частини, в іншому просторі із множенням координат на додатні числа, відповідно, $\lambda_i \in R^m,\ i=\overline{1,r}$. Тому варіантом запису SVD-розкладу (1.7), може бути наступний [68]:

$$Ax = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_{R^n} = \sum_{i=1}^{r} \lambda_i u_i l_{v_i} x,\ x \in R^n \qquad (1.8),$$

де $l_v, v \in R^n$ – лінійний функціонал, породжений v у відповідності до співвідношення: $l_v x = (v, x),\ x \in R^n$.

Зауваження 1. Зазначимо, що $v_i^T, i=\overline{1,R}$ у SVD-розкладі матриць - це подання $l_{v_i},\ i=\overline{1,r}$ засобом матричної алгебри.

Зауваження 2. Пару (власне число - власний вектор): (v,λ^2) чи (u,λ^2), називатимемо сингулярністю оператора – у цьому випадку $A^T A$ чи AA^T, відповідно.

Зауважимо, що при сингулярному поданні матриці (1.8) "операторні" об’єкти реалізовані засобами матричної алгебри:

- $(x, y)_{R^n} = x^T y$;

- $1_v x = x^T y$;
- $A^* = A^T$, $A: R^n \to R^m$, $A - m \times n$-матриця.

Саме ці зміни наведені в «операторному» поданні (1.8) SVD-матриці (1.7).

Для лінійних операторів між довільним евклідовими просторами є справедливим повний аналог (1.8).

Теорема 1.2 (про SVD-подання лінійного оператора над довільними евклідовими просторами)

Для довільної лінійного оператора A між евклідовими просторами $E_1, E_2 : A: E_1 \to E_2$ – справедливе наступне його подання у вигляді суми за елементами ненульових сингулярностей – сингулярностей з ненульовими власними числами – операторів AA^*, A^*A

$$A = \sum_{i=1}^{r} \lambda_i u_i 1_{v_i} = \sum_{i=1}^{r} \lambda_i u_i (v_i, \cdot), \qquad (1.9),$$

де

- $r = \dim R(A) = \dim L_A$ – розмірність підпростору можливих значень;
- $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0$ – спільний набір ненульових власних чисел операторів AA^*, A^*A;
- $u_i \in E_2, i = \overline{1,r}$ – ортонормований набір власних векторів оператора AA^*, що відповідають ненульовим власним числам $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0$: $AA^* u_i = \lambda_i^2,\ \lambda_i^2 > 0,$ $i = \overline{1,r},\ u_i^T u_j = \delta_{ij}, i \neq j$;
- $v_i \in E_1,\ i = \overline{1,r}$ – ортонормований набір власних векторів оператора A^*A, що відповідають ненульовим власним числам $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0$: $A^*Av_i = \lambda_i^2 v_i,\ \lambda_i^2 > 0,$ $i = \overline{1,r},\ v_i^T v_j = \delta_{ij}, i \neq j$,

крім того,

$$u_i = \frac{Av_i}{\lambda_i}, i = \overline{1,r}, \quad v_i = \frac{A^* u_i}{\lambda_i}, \ i = \overline{1,r}.$$

Так само, як і у матричному SVD, співвідношення (1.9) може скласти основу визначення ПдО для лінійного оператора між двома абстрактними евклідовими просторами.

Доведення. Доведення здійснюється так само, як для матричних операторів. Простори можливих значень оператора та спряженого співпадають із лінійними оболонками відповідних наборів: $v_i \in E_1,\ i=\overline{1,r}$ та $u_i \in E_2, i=\overline{1,r}$. Включення в один бік є очевидними, а співпадіння визначається однаковими розмірностями відповідних просторів. Далі перевіряється співпадіння вихідного оператора на $v_i \in E_1,\ i=\overline{1,r}$, а, отже і на $L(v_1,,...,v_r)$. Останній підпростір є рейнджем – $\Re(A^*)$ (множина можливих значень) для спряженого оператора, а його ортогональне доповнення $L^{\perp}(v_1,,...,v_r)$ – ядром оператора A . Отже, на векторах цього підпростору оператор A приймає нульові значення, а оператор, що визначається сумою теж, оскільки всі вектори ортонормованого набору $v_i \in E_1,\ i=\overline{1,r}$ є ортогональними до векторів $L^{\perp}(v_1,,...,v_r)$. Оскільки $E_1 = L(v_1,,...,v_r) + L^{\perp}(v_1,,...,v_r)$, то за лінійністю операторів в лівій та правій частині отримуємо співпадіння їхніх значень на будь-яких векторах-елементах області можливих значень.

Теорема 1.3. Класичне (або векторне) сингулярне (SVD) представлення довільної $m \times n$ матриці. Для довільної матриці $A \in R^{m\times n}$ рангу $r \le \min(m,n)$, яка розглядається як лінійне відображення між евклідовими просторами числових векторів $A: R^n \to R^m$, справедливе наступне представлення матриці у вигляді зваженої суми тензорних добутків векторів

$$A = \sum_{i=1}^{r} \lambda_i u_i v_i^T \, , \qquad (1.10),$$

де

- $\lambda_1^2 \ge ... \ge \lambda_r^2 > 0$ – спільний набір ненульових власних чисел матриць AA^T, $A^T A$;
- $u_i \in R^m, i=\overline{1,r}$ – ортонормований набір власних векторів матриці AA^T, що відповідають ненульовим власним числам $\lambda_1^2 \ge ... \ge \lambda_r^2 > 0$: $AA^T u_i = \lambda_i^2, \lambda_i^2 > 0,$ $i=\overline{1,r},\ u_i^T u_j = \delta_{ij}, i \ne j$;

- $v_i \in R^n,\ i = \overline{1, r}$ – ортонормований набір власних векторів матриці $A^T A$, що відповідають ненульовим власним числам $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0$: $A^T A v_i = \lambda_i^2 v_i, \lambda_i^2 > 0$, $i = \overline{1, r},\ v_i^T v_j = \delta_{ij}, i \neq j$,

крім того,

$$u_i = \frac{A v_i}{\lambda_i},\ i = \overline{1, r},$$

$$v_i = \frac{A^T u_i}{\lambda_i},\ i = \overline{1, r}.$$

Дія оператора A на вектор $x \in R^n$ в SVD-розкладі, являє собою перенесення розкладу вектора по ортогональному базису або його частині $v_i \in R^n, i = \overline{1, r}$ з координатами $x_i = v_i^T x = (v_i, x)_{R^n}, i = \overline{1, r}$, в одному просторі в розкладі по ортогональному базису $u_i \in R^m, i = \overline{1, r}$, або його частини, в іншому просторі із множенням координат на додатні числа, відповідно, $\lambda_i \in R^m, i = \overline{1, r}$. Тому варіантом запису SVD-розкладу (1.10), може бути наступний [68]:

$$Ax = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_{R^n},\ x \in R^n.$$

Важливою є наступна теорема, яка має "обернений" характер до попередньої.

Теорема 1.4. Нехай в абстрактних евклідових просторах E_1, E_2 задані:

1) набір додатних чисел $\lambda_1^2 \geq \ldots \geq \lambda_r^2 > 0,\ r > 0,\ i = 1, r,$

$r \leq min \quad dim\, E_1,\ dim\, E_2$;

2) ортонормований набір векторів $v_i \in E_1, i = 1, r;$

3) ортонормований набір векторів $u_i \in E_2, i = 1, r;$

Тоді для оператора $A_{\lambda,u,v}$, , визначеного за згаданими вище елементами співвідношенням

$$A_{\lambda,u,v} = \sum_{i=1}^{r} \lambda_i u_i \mathrm{l}\, v_i \tag{1.11}$$

його сингулярне подання співпадає із співвідношенням (1.10), що його задає.

Доведення. Доведення випливає з того, що, як свідчить безпосередня перевірка, v_i, λ_i^2, u_i, λ_i^2, $i = \overline{1, r}$ є повними ортонормованими наборами сингулярностей для операторів $A^*_{\lambda,u,v}A_{\lambda,u,v}$, $A_{\lambda,u,v}A^*_{\lambda,u,v}$ відповідно. Що й доводить теорему.

1.3.5 Псевдообернення: класичний варіант

Класичний варіант псевдообернення: псевдообернення за Муром-Пенроузом (як зазначалося вище, позначатиметься абревіатурою ПдО в подальшому) визначається для матриць A довільної розмірності і є важливим алгебраїчним засобом в математичному моделюванні, зокрема, в обробці сигналів. ПдО може бути визначене багатьма різними еквівалентними способами, як це власне окремо та незалежно один від одного зробили Мур та Пенроуз [33,36,37], але визначення через сингулярний розклад є чи не найзручнішим з технічної точки зору. Саме воно наводиться нижче.

Визначення псевдообернення (ПдО) довільної $m \times n$-матриці A, що задає лінійний оператор між евклідовими просторами числових векторів: $A: R^n \to R^m$ (векторне SVD-представлення матриці). Для довільної $m \times n$-матриці A псевдообернена до неї є $m \times n$ матриця, що позначатиметься як A^+ і визначатиметься співвідношенням

$$A^+ = \sum_{i=1}^{r} \lambda_i^{-1} v_i u_i^T : R^m \to R^n. \tag{1.12}$$

1.3.6. Сингулярний розклад та псевдообернення матриці як лінійного оператора над матричними просторами

Загалом, ПдО лінійного оператора може бути введене більш ніж десятьма варіантами. Наведене представлення є найбільш привабливим технічно – з точки зору обчислення.

Матрицю $A \in R^{m \times n}$ розглянемо як лінійний оператор з матричного евклідового простору $R^{n \times p}$ в $R^{m \times p}$. Довільна $m \times n$ матриця A може розглядатися як матриця лінійного оператора між двома евклідовими просторами матриць $R^{n \times p}$ і $R^{m \times p}$ зі слідовими скалярними добутками. Цей оператор описується стандартним чином: для довільної матриці $X \in R^{n \times p}$ результатом дії оператора є матриця $AX \in R^{m \times p}$, обчислена за стандартним визначенням добутку в матричній алгебрі.

Твердження 1.1 (Матричний сингулярний розклад $m \times n$ матриці $A \in R^{m \times n}$ (M-SVD)) [68]. Справедливе наступне представлення матриці у вигляді суми, яке реалізує перенесення розкладу по ортонормованим власним елементам (власним матрицям $A^T A$) матриці на ортонормовані власні елементи матриці $A^T A$

$$AX = \sum_{i=1}^{r} \sum_{k=1}^{p} \lambda_i u_i e_k^T (v_i e_k^T, X)_{tr} \qquad (1.13),$$

де

- $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$ загальний набір ненульових власних чисел матриць AA^T, $A^T A$;
- $u_i e_k^T \in R^{m \times p}, i = \overline{1, r}, k = \overline{1, p}$ – ортонормований набір власних $m \times p$ матриць матриці AA^T, що відповідають ненульовим власним числам: $AA^T u_i e_k^T = \lambda_i^2 u_i e_k^T$, $\lambda_i^2 > 0, i = \overline{1, r}, k = \overline{1, p}, (u_i e_k^T, u_j e_l^T)_{tr} = \delta_{ij} \delta_{kl}, i \neq j, k \neq l$;
- $v_i e_k^T \in R^{n \times p}$, $i = \overline{1, r}, k = \overline{1, p}$ – ортонормований набір власних $n \times p$ матриць матриці $A^T A$, що відповідають ненульовим власним числам: $A^T A v_i e_k^T = \lambda_i^2 v_i e_k^T, \lambda_i^2 > 0$, $i = \overline{1, r}$, $k = \overline{1, p}$, $(v_i e_k^T, v_j e_l^T)_{tr} = \delta_{ij} \delta_{kl}, i \neq j, k \neq l$.

Співпадіння векторного і матричного варіанту SVD.

Твердження 1.2 (Твердження згортки) [68]. Векторний і матричний варіант SVD-розкладу однієї і тієї ж $m \times n$ матриці $A \in R^{m \times n}$ збігаються між собою: для довільної матриці $X \in R^{n \times p}$

$$\left(\sum_{i=1}^{r}\lambda_i u_i v_i^T\right)X=\sum_{i=1}^{r}\sum_{k=1}^{p}\lambda_i u_i e_k^T(v_i e_k^T, X)_{tr}\,. \qquad (1.14)$$

Принципову роль в описі базових структур евклідових просторів відіграє псевдообернення (ПдО) за Муром-Пенроузом [35], [37] як одномісної операції A^+ над прямокутними матрицями довільної розмірності.

Необхідно розрізняти ПдО для матриці як векторного відображення (в подальшому, за необхідності, ВПдО), властиво, класичне за Муром-Пенроузом, і ПдО для тієї ж матриці, як матричного відображення (в подальшому – МПдО).

Твердження 1.3 (Матричне ПдО матриці) [68]. Для довільної $m\times n$-матриці A МПдО, позначуване A_M^+, визначається співвідношенням

$$A_M^+X=\sum_{i=1}^{r}\sum_{k=1}^{p}\lambda_i^{-1}v_i e_k^T(u_i e_k^T, X)_{tr}: R^{m\times p}\to R^{n\times p}\,. \qquad (1.15)$$

Твердження 1.4 (Співпадіння ВПдО і МПдО) [68]. Матричний і векторний варіант ПдО для довільної $m\times n$-матриці A збігаються між собою: тобто співпадають матриці A^+, A_M^+, які визначаються співвідношеннями (1.12), (1.15) відповідно: $A_M^+=A^+$.

Властивості ПдО матриць:

1. $AA^+=\sum_{j=1}^{r}y_j y_j^T$, $A^+A=\sum_{j=1}^{r}x_j x_j^T$, - ортогональні проектори на L_A, L_{A^T}

2. $AA^+A=A$, $A^+AA^T=A^+$, $A^{+^T}A^+A=A^{+^T}$, $(A^+)^+=A$, $AA^+=A^{+^T}A^T$, $(AA^+)^T=AA^+$, $A^+A^{\;T}=A^+A$, $A^+=(A^TA)^+A^T=A^T(AA^T)^+$,

3. $Z(A^T)=I_m-AA^+$, $Z(A)=I_n-A^+A$ - проекційні матричні операції на L_A^L, $L_{A^T}^L$

I_m, I_n - одиничні матриці ортогонального проектування на L_A, L_{A^T}, перетворення в R^m, R^n.

Властивості спряжених операторів:

$A: E_1 \to E_2$

$A^*: \forall \mathrm{x} \in \mathrm{E}_1, y \in E_2$, $\langle Ax, y\rangle_2 = \langle x, A^* y\rangle$

1) $(A^*)^* = A$

2) $(A+B)^* = A^* + B^*$

3) $(AB)^* = B^* A^*$

4) $(1_a)^* = a \Rightarrow (a^*)^* = a$, де $1_a = (\mathrm{a}, \cdot)$

5) $a^* = a^T$

Доведення властивості 4: $a^* = 1_a$

$R^1 \to E, (\lambda, \mathrm{a} \in \mathrm{E})$, $(\lambda\, \mathrm{a}, \mathrm{y})_E = \lambda(\mathrm{a}, \mathrm{y})$

$(a^*) y = (a, y) \Rightarrow a^* = 1_a \Rightarrow A_a \lambda = \lambda a$.

Доведення властивості 5: $a^* = a^T$

$(\lambda\, \mathrm{a}, \mathrm{y}) = \lambda(\mathrm{a}, \mathrm{y}) = \lambda a^T y \Rightarrow a^T = a^*$).

1.3.7. Дослідження СЛАР у класичному варіанті

$Ax = y, x \in R^n, y \in R^m, A \in R^{m \times n}$.

1. Умова розв’язності системи лінійних алгебраїчних рівнянь (СЛАР). Для розв’язності СЛАР $Ax = y, y \neq 0$, необхідно і достатньо, щоб $y^T Z(A^T) y = 0$.

2. За виконання умови розв’язності $A^+ y$ розв’язок СЛАР є мінімальним за нормою.

3. Він є ортогональним до $KerA$, а множина всіх розв’язків Ω_y описується співвідношенням:

$$\Omega_y = A^+ y + KerA = A^+ y + Z(A) R^n = \{x : x = A^+ y + Z(A) v, v \in R^n\},\ Z(A) R^n = KerA \qquad (1.16)$$

Будь-який елемент Ω_y у разі нерозв'язності СЛАР називається псевдорозв'язок.

4. Псевдорозв'язок СЛАР. Якщо СЛАР $Ax = y$ нерозв'язна, тобто $y^T Z(A^T) y > 0$, то множина, визначена співвідношенням (1.16) описує сукупність всіх найкращих квадратичних наближень правої частини значеннями лівої:

$$\Omega_y = A^+ y + Z(A) R^n = \underset{x \in R^n}{Arg} \min \| Ax - y \|^2 . \qquad (1.17)$$

із збереженням властивості $A^+ y \perp KerA$, а отже $A^+ y$ - мінімальний за нормою псевдорозв'язок.

5. Значення нев'язки Δ, для довільного найкращого квадратичного наближення складає

$y^T Z(A^T) y$:

$$\Delta = \min_{x \in R^n} \left\| Ax - y \right\|^2, \; A^+ Y = \underset{x \in Arg \min\limits_{x \in R^n} \| Ax - Y \|_{tr}^2}{Arg} \min \| \mathrm{x} \| .$$

Дослідження СЛАР для матриці як лінійного оператора між матричними просторами.

1. Для того, щоб матрична СЛАР (МаСЛАР) $AX = Y, A \in R^{m \times n} X \in R^{n \times p}, Y \in R^{m \times p}$ була розв'язною необхідно і достатньо, щоб $trY^T Z(A^T) Y = 0$. У цьому випадку множина Ω_Y всіх розв'язків визначається співвідношенням:

$$\Omega_Y = \{ X : X = A^+ Y + Z(A) V, V \in R^{n \times p} \} \qquad (1.18)$$

2. У випадку, коли МаСЛАР несумісна, тобто, коли $trY^T Z(A^T) Y > 0$, множина (1.18) описує розв'язок оптимізаційної задачі найкращого квадратичного наближення правої частини Y значеннями лівої частини AX того ж рівняння:

$$\underset{X \in R^{n \times p}}{Arg} \min \| AX - Y \|_{tr}^2 = \Omega_Y = \{ X : X = A^+ Y + Z(A) V, V \in R^{n \times p} \} \qquad (1.19)$$

Величина нев'язки $\Delta = \min\limits_{X \in R^{n \times m}} \left\| AX - Y \right\|^2$ для кожного розв'язку оптимізаційної задачі складає $trY^T Z(A^T) Y$.

3. Множина розв'язків і множина псевдорозв'язків, коли МаСЛАР несумісна, описується одним і тим же співвідношенням (1.18).

4. Оскільки при сумісності МаСЛАР $\min\limits_{X \in R^{n\times p}} \| AX - Y \|_{tr}^2 = 0$ і цей мінімум досягається на розв'язках МаСЛАР, то множина Ω_Y із (1.18) описує і множину розв'язків і множину псевдорозв'язків як розв'язок однієї і тієї ж оптимізаційної задачі

$$\Omega_Y = \{X : X = A^+Y + Z(A)V, V \in R^{n\times p}\} = \underset{X \in R^{n\times p}}{Arg} \min \| AX - Y \|_{tr}^2 \tag{1.20}$$

5. Звернемо увагу, що матриця A^+Y являється розв'язком, точним або псевдорозв'язком МаСЛАР. Цей розв'язок будемо називати базовим.

6. Оптимізаційна властивість Пенроуза для матричного ПдО.

Теорема 1.5 Базовий розв'язок МаСЛАР як в точному, так і в псевдоваріанті, являється єдиним найменшим за нормою розв'язком оптимізаційної задачі пошуку найкращого квадратичного наближення правої частини МаЛР значеннями лівої частини того ж рівняння:

$$A^+Y = \underset{X \in \underset{X \in R^{n\times p}}{Arg} \min \| AX - Y \|_{tr}^2}{Arg} \min \| X \|_{tr} . \tag{1.21}$$

1.4. Відстані відповідності: проекційні та групуючі оператори.

У задачах кластеризації, класифікації представниками навчальної вибірки $a_1, ..., a_n \in R^m$ відношення до того чи іншого класу чи кластеру здійснюється за певною характеристикою ρ, яка характеризує ступінь відповідності кластеру досліджуваного з метою класифікації елементу $x \in R^m : \rho(x, Kl)$ чи $\rho^2(x, Kl)$. Ці характеристики називаються відстанями відповідності. Результати попереднього параграфу дають можливість використовувати в якості таких характеристик відстані відповідності лінійного оператора підпросторів чи гіперплощин, породжених навчальними вибірками чи еліпсами групування.

Найчастіше такою відстанню відповідності є однакова для всіх класів евклідова відстань. Такий підхід обмежує врахування геометричних характеристик класів, представлених навчальними вибірками. Таке врахування може бути досягнуте розглядом у зв’язку із навчальними вибірками класів (за навчанням із учителем) слушних геометричних об’єктів. Цими геометричними об’єктами можуть бути чи лінійні підпростори та гіперплощини, чи еліпси та мінімальні еліпси групування. Відстані відповідності, породжені цими геометричними об’єктами (об’єктами, асоційованими з навчальними вибірками) є специфічними для класів, що розширюють усталені рамки класичних варіантів. Але при цьому вони ефективно реалізуються засобами псевдообернення.

1.4.1. Відстані відповідності: лінійні структури

У випадку, коли кластер асоціюється (“занурюється”) із підпростором $L(a_1,...,a_n)$ та $L(\tilde{a}_1,...,\tilde{a}_n)$, відстань відповідності визначається співвідношенням:

$$\rho^2(x,Kl)=\left[\begin{array}{c}(x)^T Z(A^T)(x)\\ (x-\bar{a})^T Z(\tilde{A}^T)(x-\bar{a})\end{array}\right.$$

– занурення в лінійний підпростір та гіперплощину відповідно.

Теорема 1.6 (Квадрат відстані вектора до гіперплощини). Квадрат відстані $\rho^2\ b,\Gamma(a,L_A)$ вектора $b\in R^m$ від гіперплощини $\Gamma(a,L_A)=a+L_A : a\in R^m, L_A\subseteq R^m$ визначається співвідношенням

$$\rho^2(b,\Gamma(a,L_A))=\min_{y\in\Gamma(b,L_A)}\|b-y\|^2=(b-a)^T Z(A^T)(b-a)\,.$$

Відзначимо, що прив'язка підпросторів у цьому, як і в інших випадках, до множини значень оператора A, не обмежує сферу застосування результатів. Вони можуть поширюватися на ситуацію, коли підпростір породжується скінченною сукупністю векторів. Зокрема, є справедливим наступне твердження про відстань вектора від гіперплощини, породженої набором векторів $a_1,...,a_n$, зміщення якої співпадає із середнім векторів досліджуваного набору, а лінійний простір породжується векторами набору, що центруються згаданим середнім.

Теорема 1.7 [63]. Квадрат відстані $\rho^2\left(b,\Gamma(\bar{a},L(\tilde{a}_1,...,\tilde{a}_n))\right)$ вектора $a\in R^m$ від гіперплощини $\Gamma(\bar{a},L(\tilde{a}_1,...,\tilde{a}_n))=\bar{a}+L(\tilde{a}_1,...,\tilde{a}_n):\tilde{a}_i=a_i-\bar{a}\in R^n,\quad i=\overline{1,n},$ $\bar{a}=\frac{1}{n}\sum_{i=1}^{n}a_i$ визначається співвідношенням

$$\rho^2\left(b,\Gamma(\bar{a},L(\tilde{a}_1,...,\tilde{a}_n))\right)=\min_{y\in\Gamma(\bar{a},L(\tilde{a}_1,...,\tilde{a}_n))}\|b-y\|^2=(b-\bar{a})^T Z(\tilde{A})(b-\bar{a}),$$

$$\tilde{A}=(\tilde{a}_1\vdots\ldots\vdots\tilde{a}_n).$$

1.4.2. Відстані відповідності: нелінійні структури

Як зазначалося вище, нелінійними структурами евклідового простору є квадратичні форми (у роботі – невід'ємно визначені) і відповідні їм еліпсоїди або еліпсоїдальні циліндри. Серед таких нелінійних, породжених просторами структур принциповими є матриці так званих «групуючих операторів» та еліпсоїди групування на їх основі, вони природно пов'язані із груповими властивостями набору векторів.

Звичайні (центральні та нецентральні) еліпси групування відповідно матимуть вигляд:

$$\rho^2(x,Kl)=\left[\begin{array}{c} x^T\frac{1}{r}R(A^T)x \\ (x-\bar{a})^T\frac{1}{r}R(\tilde{A})(x-\bar{a}) \end{array}\right.$$

.

А мінімальні еліпси (центральні чи нецентральні) відповідно:

$$\rho^2(x,Kl)=\left[\begin{array}{c} x^T\dfrac{1}{r_{\min}}R(A^T)x \\ (x-\overline{a})^T\dfrac{1}{r_{\min}}R(\tilde{A})(x-\overline{a}) \end{array}\right.$$

Групуючі оператори виникають у зв'язку з набором векторів $a(j)\in R^m,\ j=\overline{1,n}$, і відповідної йому матриці $A=(a(1)\dots a(n))$. Як і ортогональні проектори, групуючі оператори є парними і позначаються відповідно $R(A), R(A^T)$.

Визначення групуючих операторів. Групуючі оператори $R(A), R(A^T)$ визначаються співвідношеннями

$$R(A^T)=A^{+T}A^+,\ R(A)=A^+A^{+T}.$$

Їхня важливість для практичного застосування розкривається у наведених нижче теоремах.

Твердження 1.5 Для довільного набору $a(j)\in R^m,\ j=\overline{1,n}$ з матричним представленням $A=(a(1)\dots a(n))$ і довільного нормованого вектора $u\in R^m:\|u\|=1$, справедлива рівність:

$$\|\sum_{i=1}^{n}\Pr a_i\|^2=\sum_{j=1}^{n}a^T(j)uu^Ta(j)=u^TAA^Tu. \qquad (1.22)$$

Ліва частина співвідношення (1.22) твердження 1.5 представляє собою суму квадратів проекцій векторів набору $a(j)\in R^m,\ j=\overline{1,n}$ на нормовані вектори $u\in R^m:\|u\|=1$.

Твердження 1.6 Для довільного набору $a(j)\in R^m,\ j=\overline{1,n}$, з матричним представленням $A=(a(1)\dots a(n))$ має місце співвідношення:

$$\sum_{j=1}^{n}a^T(j)u_iu_i^Ta(j)=\lambda_i^2,\ i=\overline{1,r}.$$

Твердження 1.7 Набір векторів $u_1,...,u_n$, є розв'язком оптимізаційних задач вибору одновимірних підпросторів $L_{u_1},...,L_{u_n}$ з найбільшими сумами квадратів проекцій. При чому, кожен наступний вектор вибирається в ортогональному доповненні до підпросторів, породжених побудованими векторами. Перший вектор шукається у всьому евклідовому просторі.

a) $u_1 = \underset{u\in R^n;\|u\|=1}{\arg\max} \sum_{i=1}^{n} \|\Pr_{L_u} a_j\|^2,$

$$\max_{u\in R^n;\|u\|=1} \sum_{i=1}^{n} \|\Pr_{L_u} a_j\|^2 = \lambda_1^2$$

b) $u_i = \underset{u\in R^n;\|u\|=1, u\perp L(u_1,..,u_{i-1})}{\arg\max} \sum_{i=1}^{n} \|\Pr_{L_u} a_j\|^2, i=\overline{2,r}$

$$\max_{u\in R^n;\|u\|=1, u\perp L(u_1,..,u_{i-1})} \sum_{i=1}^{n} \|\Pr_{L_u} a_j\|^2, i=\overline{2,r} = \lambda_i^2$$

Твердження 1.8 [68] (Еліпси групування набору векторів $a(j)\in R^m, j=\overline{1,n}$). Нехай $a(j)\in R^m,\ j=\overline{1,n}$ – довільний набір векторів із R^n із матричним представленням $A=\left(a(1)\vdots\ldots\vdots a(n)\right)$, $rankA=r\le\min(m,n)$. Тоді всі вектори набору належать внутрішності еліпса, точніше еліпсоїдального циліндра, який визначається рівнянням

$$x^T R(A^T)x=\sum_{i=1}^{n}\frac{(u_i^T x)^2}{\lambda_i^2}=r,\ x\in R^m, \tag{1.23}$$

де $R(A^T)$ – групуючий оператор: $R(A^T)=A^{+T}A^{+}$.

За змістом еліпси (еліпсоїди), представлені (1.23) є еліпсоїдом з осями, ортонормованим осям $u_1,...,u_r$ і довжинами напівосей $\lambda_i\sqrt{r}, i=\overline{1,r}$.

Коментар. Еліпси групування – це еліпси, де осі вибрані таким чином, щоб максимізувати квадрати проекцій. Довжини напівосей є пропорційні квадрату кореня із відповідних максимумів.

Сумування по всіх векторах набору $a(j),\ j=\overline{1,n}$, і застосування твердження 1.8 дає

$$\sum_{j=1}^{n} a^T(j)u_i u_i^T a(j) = u_i^T AA^T u_i = \lambda_i^2,\ i=\overline{1,r}\,.$$

Таким чином, після ділення обох частин останнього співвідношення на, відповідно, $\lambda_i^2,\ i=\overline{1,r}$, отримаємо

$$\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \frac{u_i^T AA^T u_i}{\lambda_i^2} = 1,\ i=\overline{1,r}\,,$$

тобто

$$\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = 1,\ i=\overline{1,r}\,.$$

Просумувавши останню рівність по $i=\overline{1,r}$, отримаємо

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \sum_{i=1}^{m} \frac{u_i^T AA^T u_i}{\lambda_i^2} = r\,.$$

Змінивши порядок сумування в подвійній сумі, отримаємо

$$\sum_{j=1}^{n}\sum_{i=1}^{m} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \sum_{j=1}^{n} a^T(j)\sum_{i=1}^{m}\frac{u_i u_i^T}{\lambda_i^2}a(j) = r\,.$$

Далі, прийнявши до уваги, що

$$\sum_{i=1}^{r}\frac{u_i u_i^T}{\lambda_i^2} = A^{+T}A^{+} = R(A^T)\,,$$

остаточно отримуємо

$$\sum_{j=1}^{n}\frac{a^T(j)u_iu_i^Ta(j)}{\lambda_i^2}=\sum_{j=1}^{n}a^T(j)\sum_{i=1}^{m}\frac{u_iu_i^T}{\lambda_i^2}a(j)=\sum_{j=1}^{n}a^T(j)A^{+T}A^{+}a(j)=$$

$$=\sum_{j=1}^{n}a^T(j)R(A^T)a(j)=r,$$

тобто

$$\sum_{j=1}^{n}a^T(j)R(A^T)a(j)=r\,. \tag{1.24}$$

Оскільки $R(A^T)$ – симетрична, невід'ємно визначена матриця, то наслідком співвідношення (1.24) є одночасне виконання нерівностей

$$a^T(j)R(A^T)a(j)\le r,\ j=\overline{1,n}\,. \tag{1.25}$$

Та ж симетричність і невід'ємна визначеність дозволяє зробити висновок, що рівняння

$$x^TR(A^T)x=r,\ x\in R^m \tag{1.26}$$

визначає еліпс, точніше – еліпсоїдальний циліндр, у R^m з довжинами $\frac{1}{\lambda_i\sqrt{r}}$, $i=\overline{1,r}$ нетривіальних півосей, при $r=rankA\le\min(m,n)$.

Таким чином, виконання нерівності (1.24) для всіх векторів набору $a(j)\in R^m$, $j=\overline{1,n}$, означає їхню одночасну приналежність внутрішності еліпсоїдального циліндра з рівнянням (1.25).

Зауваження 5. У дійсності нерівність (1.24) може давати істотне загрублення «радіуса» еліпса. Так, при векторах $a(j)\in R^m$, $j=\overline{1,n}$, близьких до ортогональних, очевидним чином, константу в правій частині (1.24) можна вибрати близькою до 1.

Твердження 1.9 (Посилення результату про еліпси групувань). Усі вектори набору $a(j) \in R^m,\ j = \overline{1,n}$, з матричною представленням $A = a(1) \vdots \ldots \vdots a(n)$ належать внутрішності еліпсоїдального циліндра

$$x^T R(A^T) x = r_{\max}^2,\ r_{\max}^2 \le r = rankA,\ x \in R^m, r_{\max}^2 = \max_{j=\overline{1,n}} a^T(j) R(A^T) a(j),$$

який будемо називати мінімальним еліпсом групувань для розглянутого набору.

У той же час групувальні оператори надають можливості знаходження групових властивостей векторів на основі мінімальних еліпсів групувань в будь-якій ситуації.

1.4.3. Застосування відстаней відповідності: кластеризація.

Визначення вищенаведених відстаней відповідності, реалізованих засобами ПдО, значно розширюють можливості кластеризації ситуації, коли вектори ознак є числовими векторами, дозволяючи ефективно занурювати класифіковані об'єкти в належні підпростори або гіперплощини, або мінімальні еліпси групувань. Теорія ПдО дає можливість зв'язувати підпростір, який породжений набором векторів, як і мінімальний еліпс групувань, з певною матрицею, яка будується конструктивно й в аналітичному вигляді. Якщо об'єкт зв'язується з гіперплощиною, то її зсув – це, як правило, середнє по векторах породженої сукупності, а підпростір – це підпростір значень матриці, яка побудована із центрованих середніх векторів породженої сукупності, як зі стовпців. У випадку групування мінімальним еліпсом групувань буде його центр – середнє по векторах сукупності, а матриця квадратичної форми – це, з урахуванням транспонування, групуючий оператор для матриці, стовпчиковим представленням якої є центровані середнім вектори вихідної сукупності. Визначення групуючих операторів дає можливість конструктивного обчислення відстаней від об'єктів (підпросторів або гіперплощин), асоційованих з породженою сукупністю в той час, як посилення результату про еліпси групувань – конструктивно будувати відстані відповідності на основі мінімальних еліпсів групувань. Детальніше про результати, що забезпечують рекурентність обчислення необхідних характеристик описано у роботах [84], [86]).

Як уже було відзначено, у якості відстаней відповідності в задачах кластеризації можуть бути використані як відстані від гіперплощин, так і відстані на основі мінімальних еліпсів групувань. Останні дають можливість будувати відстані відповідності у зв'язку з використанням базової нелінійної структури. По суті, мова йде про те, що результат твердження (1.24) можна використовувати для визначення відстані відповідності, точніше його квадрата, $\rho^2(x,Kl)$, між вектором $x \in R^m$ і кластером Kl, який породжений навчальною вибіркою $a(j) \in R^m,\ j=\overline{1,n}$.

Відстань відповідності за ортогональним проектором чи групуючим оператором класу зв’язують не з «центральним» розташуванням підпростору чи еліпсу (мінімального еліпсу групування), а пов’язують з тими, що мають певне зміщення. У випадку використання ортогональних проекторів це означає, що у зв’язку з навчальною вибіркою асоціюється гіперплощина, а у випадку еліпсів – еліпси нецентрального розташування: з центром не в нулі. Таким зміщенням в обох випадках є середнє $\bar{a}$ навчальної вибірки $a(j) \in R^m,\ j=\overline{1,n} : \bar{a} = \frac{1}{n}\sum_{j=1}^{n} a(j)$, а в якості A і у визначенні ортогонального проектора і у визначенні групуючого оператора використується матриця $\tilde{A}$, побудована за стовпчиковим поданням з векторами $\tilde{a}(j) = a(j) - \bar{a},\ j=\overline{1,n}$:

$$\bar{a} = \frac{1}{n}\sum_{j=1}^{n} a(j),$$

$$\tilde{A} = (\tilde{a}(1) \ldots \tilde{a}(n)).$$

Тоді квадрат відстані до кластеру *Kl*, представленого навчальною вибіркою $a(j),\ j=\overline{1,n}$ визначається співвідношенням

$$\rho^2(x,Kl)=\begin{cases}(x-\bar{a})^T Z(\tilde{A})(x-\bar{a})\\ \dfrac{1}{r_{\max}^2}(x-\bar{a})^T R(\tilde{A})(x-\bar{a})\end{cases}, \tag{1.27}$$

відповідно до того, який об'єкт (гіперплощина чи еліпс) асоціюється з досліджуваним кластером.

Отже, за наявності набору кластерів $Kl_l,\ l=\overline{1,L}$, відстані відповідності класам є специфічними для них, і визначаються в залежності від асоційованих геометричних об'єктів: гіперплощин чи еліпсів групування, співвідношеннями

$$\rho^2(x,Kl_l)=\begin{cases}(x-\bar{a})^T Z(\tilde{A}_l)(x-\bar{a})\\ \dfrac{1}{r_{\max,l}^2}(x-\bar{a})^T R(\tilde{A}_l)(x-\bar{a})\end{cases}, l=\overline{1,L}$$

з очевидним змістом індексованих позначень за кластерами.

Висновки до розділу 1

У першому розділі наведені основні результати, що стосуються математичних методів розв'язання задач кластеризації, класифікації та розпізнавання образів. Це відноситься до основ теорії псевдообернення за Муром-Пенроузом та результатів, що забезпечують їхнє застосування в задачах групування інформації. В практичних задачах набір інформаційних ознак, як правило, представляє собою набір числових характеристик, тобто числовий вектор, на який у прикладних дослідженнях посилаються як на вектор ознак, але є суттєва необхідність розгляду набору числових характеристик як матриць. Таке представлення, дозволяє використовувати в математичній моделі задачі розпізнавання потужний набір структур евклідового простору числових векторів фіксованої розмірності – того чи іншого варіанту R^n. Зазвичай, таке використання обмежується структурою метричного простору: відстані в евклідовому просторі, хоча такий запас структур (базових структур евклідового простору) – значно ширший. Він включає в себе як множинні структури: лінійні підпростори та гіперплощини, а також слушні еліпсоїди, точніше – еліпсоїдальні циліндри, так і структури «синглового» (від single – одиничний) характеру, якими є лінійні оператори та невід'ємно визначені квадратичні форми. Ефективність використання структур (зв'язків між елементами) в евклідовому просторі R^n визначається ефективністю математичного апарату, який забезпечує використання відповідних структур для їхнього конструктивного опису та побудови, а також обчисленням «відстаней відповідності» для таких структур. Такими «відстанями відповідності» можуть бути відстані від відповідних множинних утворень: підпросторів, гіперплощин, еліпсоїдів. Зауважимо, що такі «відстані відповідності» ефективно обчислюються за «сингловими» структурами: лінійними чи нелінійними. Тому принципової ваги набуває питання встановлення взаємної відповідності між множинними та «сингловими» базовими структурами евклідового простору R^n. Як виявляється, адекватним математичним апаратом, який забезпечує встановлення такого зв'язку, є апарат псевдообернення (в подальшому – ПдО) за Муром- Пенроузом та його розвиток в принциповій роботі М.Ф.Кириченка [83]. В багатьох прикладних задачах слушним є використання складніших модельних відповідників реальних об'єктів, в даному випадку мовних сигналів, складніших «векторів ознак». Природним поданням мовного сигналу є спектрограма, яка є матрицею. Так само зображення того чи іншого об'єкту теж є матрицею. Отже, природно було б аналізувати саме матрицю, як модельне втілення аналізованого об'єкту, хоча наразі такі матриці піддаються подальшій обробці, щоб побудувати класичний варіант модельного представника – класичний вектор ознак у вигляді набору числових характеристик (числового вектора), що і фігурує у всіх подальших розглядах, представляючи первинний об'єкт. Зауважимо, що матриці фіксованої $m \times n$ розмірності є евклідовим, що позначатиметься як $R^{m \times n}$. Відсутність уваги до використання цього варіанту евклідових просторів в математичному моделюванні визначається значною мірою недостатньою розвиненістю математичного апарату опису базових структур такого варіанту евклідових просторів. Враховуючи, що в евклідовому просторі R^n основою опису і конструктивного використання базових структур є ПдО, можна зробити висновок про важливість розвинення апарату ПдО для $R^{m\times n}$. При застосуванні базових структур у задачах кластеризації ПдО розширює можливості кластеризації, дозволяючи ефективно занурювати класифіковані об'єкти в належні підпростори, або гіперплощини, або мінімальні еліпси групувань. Теорія ПдО дає можливість зв'язувати підпростір, який породжений набором векторів, як і мінімальний еліпс групувань, з певною матрицею, яка будується конструктивно й в аналітичному вигляді. Якщо об'єкт зв'язується з гіперплощиною, то її зсув – це, як правило, середнє по векторах породженої сукупності, а підпростір – це підпростір значень матриці, яка побудована із центрованих середніх векторів породженої сукупності, як зі стовпців. У випадку групування мінімальним еліпсом групувань є його центр – середнє по векторах сукупності, а матриця квадратичної форми – це, з урахуванням транспонування, групуючий оператор для матриці, стовпчиковим представленням якої є центровані середнім вектори вихідної сукупності. Визначення групуючих операторів дає можливість конструктивного обчислення відстаней від об'єктів (підпросторів або гіперплощин), асоційованих з породженою сукупністю в той час, як посилення результату про еліпси групувань – конструктивно будувати відстані відповідності на основі мінімальних еліпсів групувань.

РОЗДІЛ 2
КОНЦЕПЦІЯ «КОРТЕЖНОГО ОПЕРАТОРА»

У цьому розділі представлена концепція «кортежних операторів», що дозволяє перенести властивості псевдообернення із евклідового простору числових векторів до евклідового простору матриць фіксованої розмірності. Таке перенесення уможливлюється введенням у розгляд спеціального класу операторів, в роботі вони названі «кортежними», – за матричними кортежами. Побудова засобів опису базових структур евклідових просторах матриць фіксованої розмірності здійснюється через побудову теорії SVD та ПдО для таких операторів. Концепція «кортежних операторів» полягає у пропозиції реалізувати засоби обробки, оперування базовими структурами евклідових просторів з $R^n \to R^{n\times m}$ на основі концепції «кортежних операторів» з матричними кортежами. Побудований новий клас операторів, які називаються «кортежами», досліджені засоби опису базових структур, побудована теорія сингулярного розкладу та псевдообернення для таких операторів, що дозволяє здійснити згадане перенесення засобів опис структур з $R^n \to R^{n\times m}$. У розділі наведені факти з теорії «кортежних операторів», необхідні для визначення SVD-подання та конструкцій на його основі: псевдообернення за Муром-Пенроузом та основних ортогональних проекторів. Важливо, що SVD-розклад та всі побудови на його основі для «кортежних операторів» зводяться до задачі на власні значення для звичайних матриць.

2.1. Абстрактна теорема про SVD-розклад

У подальшому нам знадобиться "абстрактна" теорема, під цим розумітимемо SVD-розклад довільного оператора над евклідовим простором.

Теорема 2.1 (Абстрактна SVD-теорема). Для довільного лінійного оператора $A: E_1 \to E_2$ між двома евклідовими просторами та множинами сингулярностей $(v_k, \lambda_k^2),\ (u_k, \lambda_k^2), \lambda_k^2 > 0, k = \overline{1,r},\ r = rankA$ для операторів $_{Д}$ відповідно, з спільними для обох операторів множиною власних значень $\lambda_k^2, k = \overline{1,r}:\ \lambda_{k-1} \ge \lambda_k > 0,\ \overline{k = 2,r}$ таких що $Ax = \sum_{k=1}^{r} \lambda_k u_k (v_k, x)_1, \quad A^* y = \sum_{k=1}^{r} \lambda_k v_k (u_k, y)_2$. Зокрема, мають місце наступні співвідношення:

$u_k = \lambda_k^{-1} A v_k, \mathrm{k} = \overline{1, r}$, $v_k = \lambda_k^{-1} A^* u_k, \mathrm{k} = \overline{1, r}$, A^* - спряжений до A.

2.2. «Кортежний оператор»

Кортежем за матрицями називатимемо впорядкований набір матриць $\alpha = (A_1, \ldots, A_K) \in R^{(m \times n), K}$. «Кортежним оператором» (КорО) за матричним кортежем $\alpha = (A_1, \ldots, A_K)$ називатимемо оператор, який позначатимемо $\wp_\alpha : R^K \to R^{m \times n}$ і визначається співвідношенням

$$\wp_\alpha y = \sum_{k=1}^{K} y_k A_k, \alpha = (A_1, \ldots, A_K) \in R^{(m \times n), K}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_K \end{pmatrix} \in R^K \tag{2.1}$$

Теорема 2.2 Лінійний підпростір $L(A(j), j = \overline{1, K})$, породжений набором матриць $A(j) \in R^{m \times n}, j = \overline{1, K}$ та підпростір $L_{\wp_\alpha} = \Re(\wp_\alpha)$ можливих значень «кортежного оператора» $\wp_\alpha$ з матрицею A співпадають між собою:

$\Re(\wp_\alpha) = L_{\wp_\alpha} = L(A(j), j = \overline{1, K})$.

Зауваження. Як зазначалося вище, твердження 1 для евклідових просторів числових векторів є принциповим в тому сенсі, що дозволяє зводити дослідження лінійних підпросторів, породжених елементами навчальної вибірки до дослідження матриць лінійних операторів, побудованих із елементів навчальної вибірки, як із стовпчиків.

2.3. Сингулярний розклад «кортежних операторів»

Як і у випадку евклідових числових векторів принциповим у теорії побудови ПдО «кортежного оператора» відіграють теореми про базові сингулярні асоційовані лінійні підпростори

$\wp_\alpha \wp_\alpha^*, \wp_\alpha^* \wp_\alpha$.

Важливою у визначенні SVD є подання спряженого $\wp_\alpha^*$ та теорема про сингулярності $\wp_\alpha^*\wp_\alpha$.

Теорема 2.3 [71]. Спряженим $\wp_\alpha^*$ до оператора $\wp_\alpha : R^K \to R^{m\times n}$ є лінійний оператор, що діє у зворотному до $\wp_\alpha$ напрямку: $\wp_\alpha^* : R^{m\times n} \to R^K$, і визначається співвідношенням:

$$\wp_\alpha^* Y = \begin{pmatrix} (A_1, Y) \\ \vdots \\ (A_K, Y) \end{pmatrix} = \begin{pmatrix} trY^T A_1 \\ \vdots \\ trY^T A_K \end{pmatrix} = \begin{pmatrix} \lambda_{A_1} Y \\ \vdots \\ \lambda_{A_K} Y \end{pmatrix} = \begin{pmatrix} \lambda_{A_1} \\ \vdots \\ \lambda_{A_K} \end{pmatrix} Y, Y \in R^{m\times n}.$$

де λ_A – це лінійний функціонал на $R^{m\times n}$, породжений $A \in R^{m\times n}$ у відповідності до співвідношення:

$$\lambda_A Y = (A, Y), Y \in R^{m\times n}.$$

Зауважимо, що співвідношення (2.1), визначає «кортежний оператор», що задається стовпчиковим кортежем

$$\alpha = \begin{pmatrix} \lambda_{A_1} \\ \vdots \\ \lambda_{A_K} \end{pmatrix},$$

дія якого визначається «покомпонентним» застосуванням до аргументу $Y \in R^{m\times n}$.

Доведення: доведення полягає у перевірці визначення спряженого оператора:

$$(\wp_\alpha x, Y) = (\sum_{k=1}^{K} x_k A_k, Y) =$$

$$= \sum_{k=1}^{K} x_k (A_k, Y) = (x, \begin{pmatrix} (A_1, Y) \\ (A_K, Y) \end{pmatrix})$$

Таким чином,

$$\wp_\alpha^* Y = \begin{pmatrix} (A_1, Y) \\ \vdots \\ (A_K, Y) \end{pmatrix}; \ \wp_\alpha^* = \begin{pmatrix} \lambda_{A_1} \\ \vdots \\ \lambda_{A_K} \end{pmatrix}$$

з покомпонентним застосуванням вектора стовпчика лінійних функціоналів до аргументу оператора.

Теорема 2.4 Добуток двох операторів є лінійний оператор $\wp_\alpha^* \wp_\alpha : R^K \to R^K$, що задається матрицею F - матрицею Грама, набору $A_1,...,A_K$, тобто:

$$F \equiv \wp_\alpha^* \wp = \langle A_i, A_j \rangle_{i,j=\overline{1,K}} \equiv \begin{pmatrix} trA_1^T A_1,...,trA_1^T A_n \\ \ldots \\ trA_n^T A_1,...,trA_n^T A_n \end{pmatrix}. \quad (2.2)$$

Зауважимо, що матриця F, що визначається через SVD-розклад співвідношенням (2.2) є матрицею Грама елементів $A_1,...,A_K$ матричного кортежу $\alpha = (A_1 \ldots A_K)$, що задає оператор $\wp_\alpha$.

Сингулярний розклад для матриці (2.2) є очевидним – це симетрична та невід'ємно визначена матриця. Він визначається набором ненульових сингулярностей $(v_i, \lambda_i^2), i, j = \overline{1, r}$:

$$\| v_i \| = 1,\ v_i \perp v_j,\ i \neq j;\ i, j = \overline{1,r};\ \lambda_1 > \lambda_2 > ... > \lambda_r > 0, r = rankF,$$

Це означає, що

$$\wp_\alpha^* \wp_\alpha v_i = F v_i = \lambda_i^2 v_i, i = \overline{1,r}, r = rankF$$

Таким чином, має місце теорема.

Теорема 2.5 $\wp_\alpha^* \wp_\alpha$ допускає подання через набір ненульових сингулярностей $(v_i, \lambda_i^2), i, j = \overline{1, r}$:

$$\wp_\alpha^* \wp_\alpha = \sum_{i=1}^{r} \lambda_i^2 v_i v_i^T = \sum_{i=1}^{r} \lambda_i^2 v_i (v_i, \cdot)$$

$$\wp_\alpha^* \wp_\alpha = F = \sum_{i=1}^{r} \lambda_i^2 v_i v_i^T = \sum_{i=1}^{r} \lambda_i^2 v_i (v_i, \cdot) = \sum_{i=1}^{r} \lambda_i^2 v_i 1_{v_i}$$

В подальшому кожний з вектор-рядків $v_i^T, i=\overline{1,r}$ через компоненти буде записуватися у вигляді:

$$v_i^T = (v_{1i},...,v_{Ki}),\ i=\overline{1,r}\,,$$

тобто $v_{ik},\ i=\overline{1,r},\ k=\overline{1,K}$ позначає компоненту з номером k у вектора v з номером i.

Доведення:

$$\wp_\alpha^*\wp_\alpha = \wp_\alpha^*\sum_{k=1}^{K} x_k A_k = \begin{pmatrix} \mathrm{l}_{A_1}, \sum_{k=1}^{K} x_k A_k \\ ... \\ \mathrm{l}_{A_k}, \sum_{k=1}^{K} x_k A_k \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{k=1}^{K} x_k \mathrm{l}_{A_1} A_k \\ ... \\ \mathrm{l}_{A_k}, \sum_{k=1}^{K} x_k \mathrm{l}_{Ak} A_k \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{K} x_k (A_1, A_k) \\ ... \\ \sum_{k=1}^{K} x_k (A_k, A_k) \end{pmatrix} = \begin{pmatrix} ((A_1,A_1),...,(A_1,A_K))x \\ ... \\ ((A_K,A_1),...,(A_K,A_K))x \end{pmatrix} = Fx;$$

Теорема 2.6 Матриці $U_i \in R^{m\times n}$: $U_i = \dfrac{1}{\lambda_i}\wp_\alpha v_i = \dfrac{1}{\lambda_i}\sum_{k=1}^{K} A_k v_{ik},\ i=\overline{1,r}$, що визначені за сингулярностями $(v_i, \lambda_i^2), i=\overline{1,r}$ оператора $\wp_\alpha^*\wp_\alpha$, є елементами повного набору сингулярностей $(U_i, \lambda_i^2),\ i=\overline{1,r}$ оператора $\wp_\alpha^* : R^K \to R^{m\times n}$.

Доведення: доведення випливає із співвідношень між сингулярностями двох операторів A^*A та AA^* абстрактної теореми про SVD-розклад.

Теорема 2.7 [6] (Сингулярний розклад «кортежного оператора»).

Сингулярності двох операторів: $\wp_\alpha^*\wp_\alpha, \wp_\alpha\wp_\alpha^*$ однозначно визначають сингулярний розклад операторів $\wp_\alpha$, $\wp_\alpha^*$:

$$\wp_\alpha x = \sum_{i=1}^{r} \lambda_i U_i v_i^T x = \sum_{i=1}^{r} \lambda_i U_i (v_i, x), x \in R^K ,$$

$$\wp_\alpha^* Y = \sum_{i=1}^{r} \lambda_i v_i (U_i, Y)_{tr}, Y \in R^{m\times n} ,$$

$$r = rankF .$$

Наслідок. Варіантом сингулярного подання $\wp_\alpha$ є його запис у вигляді:

$$\wp_\alpha = \sum_{k=1}^{r} \lambda_k U_k v_k^T = \sum_{k=1}^{r} \wp_\alpha v_k \ 1_{v_k}, r = rankF .$$

Зауваження: нагадаємо, що набір сингулярності $(v_i, \lambda_i^2), i = \overline{1, r}$ є сингулярностями $K \times K$ матриці F, а сингулярності $(v_i, \lambda_i^2), (U_i, \lambda_i^2), i = \overline{1, r}$ будуються за $(\gamma_i, \lambda_i^2), i = \overline{1, r}$.

Теорема 2.8 Набори ненульових сингулярностей $(v_i, \lambda_i^2), (U_i, \lambda_i^2), i = \overline{1, r}$, операторів $\wp_\alpha^* \wp_\alpha$, $\wp_\alpha \wp_\alpha^*$ відповідно пов'язані між собою «симетричними» співвідношеннями:

$$v_i = \frac{\wp_\alpha^* U_i}{\lambda_i},\ U_i = \frac{\wp_\alpha v_i}{\lambda_i},\ i = \overline{1, r}.$$

Таким чином, сингулярний розклад «кортежного оператора» має "матричний" порядок обчислювальної складності.

2.4. Пседообернення для «кортежних операторів»

Загалом у визначенні ПдО для матриць може бути застосовано до десяти еквівалентних визначень. Найбільш технічно сприйнятне – за SVD-розкладом для матриць. Теж саме можна сказати щодо «кортежних операторів». Псевдообернення буде визначене через SVD-розклад «кортежного оператора».

Означення. (*Псевдообернення через SVD-представлення*). ПдО операторів $\wp_\alpha$, $\wp_\alpha^*$ – позначатиметься верхнім індексом «$^+$» – визначається за їхнім SVD-поданням співвідношеннями, відповідно:

$$\wp_\alpha^+ Y = \sum_{k=1}^{r} \lambda^{-1} v_k \left(U_k, Y\right)_{tr} = \sum_{k=1}^{r} \lambda^{-1} v_k 1_{U_k} Y = \sum_{k=1}^{r} \lambda^{-2} v_k \left(\wp_\alpha v_k, Y\right)_{tr} = \sum_{k=1}^{r} \lambda^{-2} v_k 1_{\wp_\alpha v_k} Y \,,$$

$$\left(\wp_\alpha^*\right)^+ x = \sum_{i=1}^{r} \lambda^{-1} U_i v_i^T x = \sum_{i=1}^{r} \lambda^{-1} U_i 1_{v_k} x \,.$$

2.4.1. Властивості ПдО «кортежних операторів»

1) $(\wp^*)^+ = (\wp^+)^*$
2) $(\wp^* \wp)^+ \wp^* = \wp^+$
3) $\wp^* (\wp \wp^*)^+ = \wp^+$
4) $\wp^+ \wp \wp^+ = \wp^+$

2.4.2. Ортогональні проектори

Основними ортогональними проекторами ПдО теорії являються дві пари ортогональних проекторів. Одна пара – це пара ортогональних проекторів підпросторів, які є множинами значень $\Re$ операторів $\wp_\alpha$, $\wp_\alpha^*$ відповідно: $\Re(\wp_\alpha) \equiv L_{\wp_\alpha} \subseteq R^{m \times n}$, $\Re(\wp_\alpha^*) \equiv L_{\wp_\alpha^*} \subseteq R^K$. Ці ортогональні проектори будуть позначатися одним з двох еквівалентних способів:

$$P(\wp_\alpha^*) \equiv P_{L_{\wp_\alpha}}, L_{\wp_\alpha} \subseteq R^{m \times n}, P(\wp_\alpha) \equiv P_{L_{\wp_\alpha^*}}, L_{\wp_\alpha^*} \subseteq R^K \,.$$

Друга пара – це пара ортогональних проекторів на ортогональне доповнення $L_{\wp_\alpha}^\perp \subseteq R^{m \times n}, L_{\wp_\alpha^*}^\perp \subseteq R^K$ першої пари підпросторів. Кожен з них також буде використовуватися одним із двох еквівалентних способів:

$$Z(\wp_\alpha) \equiv P_{L_{\wp_\alpha^*}^\perp}, Z(\wp_\alpha^*) \equiv P_{L_{\wp_\alpha}^\perp} \,.$$

Очевидно, що:

$$Z(\wp_\alpha) \equiv E_K - P(\wp_\alpha), Z(\wp_\alpha^*) \equiv E_{m \times n} - P(\wp_\alpha^*) \,. \tag{2.3}$$

Відповідно до загальних властивостей ПдО

$$P(\wp_\alpha) = \wp_\alpha^+ \cdot \wp_\alpha, \quad P(\wp_\alpha^*) = \left(\wp_\alpha^*\right)^+ \cdot \wp_\alpha^* = \wp_\alpha \cdot \wp_\alpha^+ \,.$$

Відповідно,

$$Z(\wp_\alpha) \equiv E_K - \wp_\alpha^+ \cdot \wp_\alpha, \quad Z(\wp_\alpha^*) \equiv E_{m \times n} - \wp_\alpha \cdot \wp_\alpha^+ \,.$$

Як зазначалося вище, ПдО «кортежного оператора» має властивості, що фігурують у визначенні за Муром чи за Пенроузом. Зокрема, виконуються наступні співвідношення:

1. Властивості визначення ПдО за Муром для матриць:

$$\begin{cases} \wp_\alpha^+ \wp_\alpha \wp_\alpha^+ = \wp_\alpha^+, \quad \wp_\alpha \wp_\alpha^+ \wp_\alpha = \wp_\alpha \text{ – «аксіома» переваги;} \\ P_{L_{\wp_\alpha}} = \wp_\alpha^+ \wp_\alpha, P_{L_{\wp_\alpha^*}} = \wp_\alpha \wp_\alpha^+ \text{ – «аксіома» ортогональних проекторів;} \end{cases}$$

2. Властивості визначення ПдО за Пенроузом для матриць: $\wp_\alpha^+ Y = \underset{x \in Arg \min \|\wp_\alpha x - Y\|^2}{\arg\min} \|x\|^2, Y \in R^{m \times n}, Y \neq 0$ – «аксіома» оптимальності.

2.4.3. Групуючі оператори

Загалом, як і у випадку евклідових просторів числових векторів, пов'язаних із алгебраїчними структурами 2-го порядку, наведені співвідношення для групуючих операторів у випадку евклідових просторів матриць, такі оператори будемо позначати $R(\wp_\alpha)$, $R(\wp_\alpha^*)$. Групуючі оператори $R(\wp_\alpha)$, $R(\wp_\alpha^*)$, як і ортогональні проектори, є «парними» операторами.

Теорема 2.9 Для довільного лінійного оператора $\wp_E$ між парою евклідових просторів E_i, $i = 1,2$, $\wp_E : E_1 \to E_2$, існує набір сингулярностей

$$(v_i, \lambda_i^2), \ (u_i, \lambda_i^2), \ i = \overline{1,r}, \ r = rank \wp_E$$

операторів $\wp_E^* \wp : E_1 \to E_1$, $\wp \, \wp_E^* : E_2 \to E_2$, відповідно, із загальним набором власних чисел $\lambda_i^2, i = \overline{1,r}$: $\lambda_{i-1} \geq \lambda_i > 0$, $\overline{i = 2,r}$,

такі, що $\wp_E x = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_1$, $\quad \wp_E^* y = \sum_{i=1}^{r} \lambda_i v_i (u_i, y)_2$.

Крім того, виконуються співвідношення:

$u_i = \lambda_i^{-1} \wp v_i, i = \overline{1,r}$,

$v_i = \lambda_i^{-1} \wp_E^* u_i$, $i = \overline{1,r}$.

Групуючі оператори $R(\wp_\alpha)$, $R(\wp_\alpha^*)$, як і ортогональні проектори, є «парними» операторами, і визначаються співвідношеннями:

$$R(\wp_\alpha) = \wp_\alpha^+ \left(\wp_\alpha^+\right)^* = \wp_\alpha^+ \left(\wp_\alpha^*\right)^+, \; R(\wp_\alpha^*) = \left(\left(\wp_\alpha^*\right)^+ \left(\wp_\alpha^*\right)^+\right)^* = \left(\wp_\alpha^+\right)^* \wp_\alpha^+ .$$

Теорема 2.10 [71]**.** Групуючий оператор $R(\wp_\alpha^*)$ може бути представлений у вигляді:

$$R(\wp_\alpha^*)X = \sum_{k=1}^{r} \lambda_k^{-2} U_k (U_k, X)_{tr} = \sum_{k=1}^{r} \lambda_k^{-2} U_k tr U_k^T X = \sum_{k=1}^{r} \lambda_k^{-2} U_k tr X^T U_k \,,$$

Квадратична форма $(X, R(\wp_\alpha^*)X)_{tr}$, що визначається через групуючий оператор $R(\wp_\alpha)$ і описується співвідношенням:

$$(X, R(\wp_\alpha^*)X)_{tr} = \sum_{k=1}^{r} \lambda_k^{-2} (U_k, X)_{tr}^2 \,,$$

Теорема 2.11 [6]. Квадратична форма $(X, R(\wp_\alpha^*)X)_{tr}$ може бути представлена у вигляді:

$$(X, R(\wp_\alpha^*)X)_{tr} = \sum_{i=1}^{r} \lambda_i^{-4} v_i^T \begin{pmatrix} trA_1^T X trA_1^T X & trA_2^T X trA_2^T X & \cdots & trA_1^T X trA_K^T X \\ trA_2^T X trA_1^T X & trA_2^T X trA_2^T X & \cdots & trA_2^T X trA_K^T X \\ \cdots & \cdots & \cdots & \cdots \\ trA_K^T X trA_1^T X & trA_K^T X trA_1^T X & \cdots & trA_K^T X trA_1^T X \end{pmatrix} v_i =$$

$$= \sum_{i=1}^{r} \lambda_i^{-4} \left\{ v_i^T \begin{pmatrix} trA_1^T X \\ \cdots \\ trA_K^T X \end{pmatrix} \right\}^2 = \sum_{i=1}^{r} \lambda_i^{-4} \left\| v_i^T \wp_\alpha^* X \right\|^2 ..$$

Так як і у R^n вони визначають центральні чи нецентральні, звичайні чи мінімальні еліпси групування. З тими самими властивостями оптимальності, що і у випадку евклідових просторів числових векторів. Результати наведені у роботі Зінька Т.П. [6].

2.5. Дослідження СЛАР для «кортежних операторів»

Загалом, СЛАР з «кортежними операторами» досліджується так само, як і у векторному випадку, з точністю до заміни $Z(A), Z(A^*)$ на відповідно $Z(\wp), Z(\wp^*)$.

Можливість такого перенесення визначається повним співпадінням властивостей ПдО для матриць та ПдО для «кортежного оператора». Це повною мірою стосується і концепції псевдорозв'язків як найкращих квадратичних наближень правої частини СЛАР до значень лівої частини СЛАР. Отже, справджуються наступні теореми.

Теорема 2.12 Розв'язність СЛАР. Для розв'язності СЛАР $\wp_\alpha X = Y$,

$\alpha = (A_1 \ldots A_K)$, $X \in R^K, Y \in R^K$, необхідно і достатньо, щоб $(Y^T, Z(\wp_\alpha^+ Y)) = 0$;

У цьому випадку множина можливих розв'язків Ω_y визначається співвідношенням:

$\Omega_y = \wp_\alpha^+ Y + Z(\wp_\alpha) R^K$, а $\wp_\alpha^+$ є найменшим за нормою розв'язком.

Доведення:

Для розв'язності СЛАР необхідно і достатньо, щоб

$Y \in \text{Ў}(\wp_\alpha) \sim P_{L_{\wp_\alpha}} Y = Y \sim (I - P_{L_{\wp_\alpha}}) Y = 0 \sim Z(\wp_\alpha^+) Y = 0$

$\| Z(\wp_\alpha^*) Y \|^2 = 0$

$(Z(\wp_\alpha^*) Y, Z(\wp_\alpha^*) Y) = 0$

$(Y, Z(\wp_\alpha^*) Z(\wp_\alpha^*) Y) = 0 \Rightarrow (Y, Z(\wp_\alpha^*) Y) = 0$

За $Y \in \text{Ў}(\wp_\alpha)$, $\wp_\alpha^+ Y$ є розв'язком:

$\wp_\alpha(\wp_\alpha^+ Y) = \wp_\alpha \wp_\alpha^+ Y = Y$

Множина розв'язків :

$$\wp_\alpha^+ Y + Ker\wp_\alpha = \wp_\alpha^+ Y + Z(\wp_\alpha)R^K$$

Оскільки, $Ker\wp_\alpha$ - підпростір, то $\wp_\alpha^+ Y$ - розв'язок, а оскільки $\wp_\alpha^+ Y \perp Ker\wp_\alpha$ то $\wp_\alpha^+ Y$ - найменший за нормою розв'язок.

2.5.1. Псевдорозв'язок дослідження СЛАР

Як і у випадку евклідових просторів числових векторів, псевдорозв'язком називатимемо розв'язки оптимізаційної задачі

$$\arg\min_{X \in R^{m \times n}} \| \wp_\alpha X - Y \|^2$$

Як і для R^n , є справедливим наступна теорема.

Теорема 2.13 Множина всіх псевдорозв'язків кортежної СЛАР описується формлюю:

$\wp_\alpha^+ Y + Z(\wp_\alpha)R^K$, $\wp_\alpha^+ Y$ - є найменший за нормою псевдорозв'язок.

Доведення: доведення випливає з того, що розв'язки оптимізаційної задачі є розв'язками розв'язної СЛАР

$\| \wp_\alpha X - Y \| = \arg\min_{X \in R^K} \| \wp_\alpha X - Y_{L_{\wp_\alpha}} - Y_{L_{\wp_\alpha^\perp}} \| = \arg\min_{X \in R^K} \| \wp_\alpha X - Y_{L_{\wp_\alpha}} \|^2 - \| Y_{L_{\wp_\alpha^\perp}} \|^2$ оскільки

$Y_{L_{\wp_\alpha}} \in \breve{Y}(\wp_\alpha)$ то СЛАР $\wp_\alpha x = Y_{L_{\wp_\alpha}}$ є розв'язною, тому $\arg\min_{X \in R^K} \| \wp_\alpha X - Y_{L_{\wp_\alpha}} \|^2$ співпадає з множиною розв'язків СЛАР $\wp_\alpha X = Y_{L_{\wp_\alpha}} \sim \wp_\alpha X = \wp_\alpha \wp_\alpha^* Y$ у відповідності до попередньої теореми мнжина розв'язків має вигляд:

$$\wp_\alpha^+ (\wp_\alpha \wp_\alpha^+ Y) + Z(\wp_\alpha)R^K$$

За аксіомою переваги

$$\wp_\alpha^+ \wp_\alpha \wp_\alpha^+ = \wp_\alpha^+,$$

отже, остаточно, маємо запис множини псевдорозв'язків розв'язної СЛАР у вигляді:

$$\wp_\alpha^+ Y + Z(\wp_\alpha)R^K .$$

Зауваження. $\wp_{\alpha}^{+}Y$ є найменшим за нормою псевдорозв'язком. Величина нев'язки оптимізаційної задачі описується як

$$\| Z(\wp_{\alpha}) \|^2$$

Зауваження. Множина розв'язків і множина псевдорозв'язків описується одною і тою самою формулою $\wp_{\alpha}^{+}Y + Z(\wp_{\alpha})R^K$, тлумачення цієї множини як множини розв'язків має місце за умови, що

$$(Y^T, Z(\wp_{\alpha}^{+}Y)) = 0$$

У випадку якщо

$$(Y^T, Z(\wp_{\alpha}^{+}Y)) > 0,$$

то згадана множина описує псевдорозв'язки. Значення квадратичної форми у цьому випадку описує нев'язку оптимізаційної задачі, що задає псевдорозв'язки.

2.6. Еліпсоїди групування для «кортежних операторів»

Нагадаймо, що у випадку евклідового простору числових векторів, групуючі оператори задають нелінійні структури другого порядку: еліпсоїди чи мінімальні еліпсоїди групування та матриці квадратичних форм, що згадані еліпсоїди задають. Під еліпсоїдами групування: простими чи мінімальними, центральними чи нецентральними, – розуміють еліпсоїди, що оптимально – за строгого опису критерію оптимальності – «накривають» вектори із заданою послідовностю, зокрема – вектори навчальної вибірки. Вектори можуть мати центром початок координат (центральні) чи еліпсоїди, які містять елементи заданого матриць, і осі яких мають оптимальні властивості. Оптимальність полягає у тому, що осі згаданих вище еліпсоїдів групування максимізують суму квадратів проекцій елементів групування на відповідні осі.

Як і у випадку евклідових просторів числових векторів, ці ортогональні проектори є ортогональними проекторами на лінійний підпростір, породжений

$A_1,...,A_K$, що можуть представляти навчальні вибірки. Так наприклад у роботі Голіка А. [63] наведено, як за набором $A_1,...,A_K$ необхідно побудувати кортеж, за яким побудувати «кортежний оператор» $\wp_\alpha$, а $Z(\wp_\alpha)$- буде ортогональний проектор на $\wp_\alpha \wp_\alpha^*$.

Концепція кортежних операторів дозволяє повністю перенести всі твердження про «групування для векторів» на набори матриць, які можуть складати навчальну вибірку. Для послідовностей матриць виявляються всі твердження про еліпси групування: про той самий критерій оптимальності, про центральні та нецентральні, мінімальні та звичайні.

Ці твердження про групування є предметом теорем, наведених нижче.

Теорема 2.14 [72]**.** Нехай $A(j) \in R^{m\times n}$, $j=\overline{1,K}$ довільний набір матриць із $R^{m\times n}$, $\alpha = (A(1) \vdots \ldots \vdots A(K))$ – матричний кортеж, а $\wp_\alpha$– «кортежний оператор», який йому відповідає. Тоді всі матриці набору $A(j) \in R^{m\times n}$, $j=\overline{1,K}$ належать внутрішності еліпса, точніше, еліпсоїдального циліндра, який визначається рівнянням

$$X, R(\wp_\alpha^*)X_{\ tr} = trX^T R(\wp_\alpha^*)X = r,\ X \in R^{m\times n},$$

де $R(\wp_\alpha^*)$– групуючий оператор $R(\wp_\alpha^*) = \wp^{+*}\wp_\alpha^+$.

2.7. Мінімальні еліпсоїди групування для «кортежних операторів»

Теорема 2.15 [72]**.** За збереження умов попередньої теореми всі матриці набору $A(j) \in R^{m\times n}$, $j=\overline{1,K}$ належать внутрішності еліпса, точніше еліпсоїдального циліндра, який визначається рівнянням

$$X^T R(\wp_\alpha^*)X = r_{\min} \le r,\ X \in R^{m\times n},$$

в якому $r_{\min}$ визначається співвідношенням

$$r_{\min} = \max_{1\le j\le K} A^T(j)R(\wp_\alpha^*)A(j).$$

Зауваження. У практичних задачах кластеризації набори матриць $A(j)\in R^{m\times n},\ j=\overline{1,K}$ представляють собою навчальну вибірку того чи іншого кластеру. Еліпс групування в таких задачах центрується середнім $\overline{A}=\frac{1}{K}\sum_{j=1}^{K}A(j)$, а матриця квадратичної форми відповідного еліпса (групуючий оператор) будується за матрицями $\tilde{A}(j)\in R^{m\times n},\ j=\overline{1,K}$, що є центрованими середнім елементами наступної навчальної вибірки:

$$\tilde{A}(j)=A(j)-\overline{A},\ j=\overline{1,K}\ .$$

Відповідним чином «хвилька» переноситься на кортежний оператор $\tilde{\wp}$, а саме рівняння мінімального еліпсоїда групування має вигляд:

$$(X-\overline{A})^{T}R(\tilde{\wp}^{*}_{\alpha})(X-\overline{A})=r_{\max}\le r,\ \ X\in R^{m\times n}\ .$$

2.8. Відстань до множинних лінійних структур в $R^{m\times n}$ підпросторів та гіперплощини для «кортежних операторів»

Розглянемо перенесення на кортежні оператори твердження про $\rho^2\left(x,L(A,b)\right)=\left\|A^{+}(b-Ax)\right\|^{2}$

Спочатку розглянемо векторний випадок:

позначимо

$$L(A,b)=A^{+}b+Z(A)R^{n}=A^{+}b+KerA=\Gamma(A^{+}b,KerA),$$

гіперплощину, що є множиною розв'язків СЛАР

$$Ax=b:\qquad A\in R^{m\times n},x\in R^{n},b\in R^{m}$$

Теорема 2.16 Квадрат відстані $\rho^2\left(x,L(A,b)\right)$ будь якого $x\in R^{n}$, до гіперплощини, що є множиною розвязків СЛАР, визначається рівністю

$$\rho^2\left(x,L(A,b)\right)=\left\|A^{+}(b-Ax)\right\|^{2},\qquad L(A,b)=A^{+}b+\underbrace{Z(A)R^{n}}_{KerA}=\Gamma(A^{+}b,KerA)$$

Доведення. Загалом, квадрат $\rho^2 \; x,\Gamma(m,L)$ відстані будь якого $x \in R^n$, до гіперплощини $\Gamma(m,L)$ визначається співвідношенням:

$$\rho^2 \; x,\Gamma(m,L) \; = (x - m)^T P_{L^\perp}(x - m).$$

У досліджуваному варіанті

$$m = A^+b, \quad P_L = Z(A) \Rightarrow P_{L^\perp} = E_n - Z(A) = A^+A.$$

Таким чином, $\rho^2 \; x, L(A,b) \; = \rho^2 \; x,\Gamma(m,L) \; = (x - A^+b)^T \underbrace{A^+A}_{P_{KerA^\perp}} (x - A^+b) =$

$$= x^T A^+ A x - 2x^T \underbrace{A^+AA^+}_{A^+} b + b^T A^{+T} \underbrace{A^+AA^+}_{A^+} b =$$

$$= x^T A^+ A x - 2x^T A^+ b + b^T A^{+T} A^+ b. \tag{2.4}$$

З іншого боку:

$$\left\| A^+(b - Ax) \right\|^2 = \left\| A^+b - A^+Ax \right\|^2 = b^T A^{+T} A^+ b - 2x^T A^+ A A^+ b + x^T A^+ A x =$$

$$= b^T A^{+T} A^+ b - 2x^T A^+ b + x^T A^+ A x. \tag{2.5}$$

Однаковість виразів (2.4),(2.5) і дозволяє стверджувати, що

$$\rho^2 \; x, L(A,b) \; = \rho^2 \; x,\Gamma(A^+b, KerA) \; = \left\| A^+(b - Ax) \right\|^2,$$

і теорема доведена.

Перенесення на «кортежні оператори».

Тепер перенесемо міркування на кортежний випадок: позначимо

– $L(\wp_\alpha, B) = \wp_\alpha^+ B + Z(\wp_\alpha) R^K = \wp_\alpha^+ B + Ker\wp_\alpha = \Gamma(\wp_\alpha^+ B, Ker\wp_\alpha)$,

гіперплощину,

– що є множиною розв'язків СЛАР

$$\wp_\alpha x = B: \qquad B \in R^{m \times n}, x \in R^K$$

Теорема 2.17 Квадрат відстані $\rho^2\left(x, L(\wp_\alpha, B)\right)$ будь-якого $x\in R^K$ до гіперплощини, що є множиною розв'язків СЛАР, визначається рівністю

$$\rho^2\left(x, L(\wp_\alpha, B)\right) = \left\|\wp_\alpha^+(B-\wp_\alpha x)\right\|^2, \quad L(\wp_\alpha, B) = \wp_\alpha^+ B + \underbrace{Z(\wp_\alpha)R^n}_{Ker\wp_\alpha} = \Gamma(\wp_\alpha^+ B, Ker\wp_\alpha)$$

Доведення. Загалом, квадрат $\rho^2\left(x, \Gamma(m,L)\right)$ відстані будь-якого $x\in R^K$ до гіперплощини $\Gamma(m,L)\subseteq R^K$ визначається співвідношенням:

$$\rho^2\left(x, \Gamma(m,L)\right) = (x-m)^T P_{L^\perp}(x-m) = \left\langle x-m, P_{L^\perp}(x-m)\right\rangle ,$$

$P_{L^\perp}$ - ортогональний проектор на ортогональне доповнення $L\subseteq R^K$.

У досліджуваному варіанті

$$m=\wp_\alpha^+ B \quad P_L = Z(\wp_\alpha) \Rightarrow P_{L^\perp} = E_K - Z(\wp_\alpha) = \wp_\alpha^+\wp_\alpha .$$

Таким чином, $\rho^2\left(x, L(\wp_\alpha, B)\right) = \rho^2\left(x, \Gamma(\wp_\alpha^+ B, Ker\wp_\alpha)\right) = \left\langle x-\wp_\alpha^+ B, \underbrace{\wp_\alpha^+\wp_\alpha}_{P_{(Ker\wp_\alpha)^\perp}}(x-\wp_\alpha^+ B)\right\rangle =$

$$= \underbrace{\left\langle x, \wp_\alpha^+\wp_\alpha x\right\rangle}_{x^T\wp_\alpha^+\wp_\alpha x} - 2\underbrace{\left\langle x, \underbrace{\wp_\alpha^+\wp_\alpha\wp_\alpha^+}_{\wp_\alpha^+} B\right\rangle}_{x^T\wp_\alpha^+ B} + \left\langle \wp_\alpha^+ B, \underbrace{\wp_\alpha^+\wp_\alpha\wp_\alpha^+}_{\wp_\alpha^+} B\right\rangle =$$

$$= x^T\wp_\alpha^+\wp_\alpha x - 2x^T\wp_\alpha^+ b + \left\langle \wp_\alpha^+ B, \wp_\alpha^+ B\right\rangle . \qquad (2.6)$$

З іншого боку:

$$\left\|\wp_\alpha^+(B-\wp_\alpha x)\right\|^2 = \left\|\wp_\alpha^+ B - \wp_\alpha^+\wp_\alpha x\right\|^2 = \left\langle \wp_\alpha^+ B - \wp_\alpha^+\wp_\alpha x, \wp_\alpha^+ B - \wp_\alpha^+\wp_\alpha x\right\rangle =$$

$$= \left\langle \wp_\alpha^+ B, \wp_\alpha^+ B\right\rangle - 2\left\langle \underbrace{\wp_\alpha^+\wp_\alpha}_{орт.проектор=сим.+ідемп.} x, \wp_\alpha^+ B\right\rangle + \left\langle \underbrace{\wp_\alpha^+\wp_\alpha}_{орт.проектор=сим.+ідемп.} x, \underbrace{\wp_\alpha^+\wp_\alpha}_{орт.проектор=сим.+ідемп.} x)\right\rangle =$$

$$= \left\langle \wp_\alpha^+ B, \wp_\alpha^+ B\right\rangle - 2\underbrace{\left\langle x, \wp_\alpha^+\wp_\alpha\wp_\alpha^+ B\right\rangle}_{x^T\wp_\alpha^+ B} + \underbrace{\left\langle x, \wp_\alpha^+\wp_\alpha x)\right\rangle}_{x^T\wp_\alpha^+\wp_\alpha x} = \left\langle \wp_\alpha^+ B, \wp_\alpha^+ B\right\rangle - 2x^T\wp_\alpha^+ B + x^T\wp_\alpha^+\wp_\alpha x . \qquad (2.7)$$

Однаковість виразів (2.6),(2.7) і дозволяє стверджувати, що

$$\rho^2\left(x, L(\wp_\alpha, B)\right) = \Gamma(\wp_\alpha^+ B, Ker\wp_\alpha) = \left\|\wp_\alpha^+(B-\wp_\alpha x)\right\|^2$$

і теорема доведена.

Висновки до розділу 2

У цьому розділі була представлена концепція «кортежних операторів», яка дозволяє перенести властивості псевдообернення із евклідового простору числових векторів до евклідового простору матриць фіксованої розмірності. Концепція «кортежних операторів» полягає у пропозиції реалізувати засоби обробки, оперування з базовими структурами евклідових просторів з $R^n \to R^{n\times m}$ на основі концепції «кортежних операторів» з матричними кортежами. Побудований новий клас операторів, які називаються кортежами, досліджені засоби опису базових структур, побудована теорія сингулярного розкладу та псевдообернення для таких операторів, що дозволяє здійснити згадане перенесення засобів опис структур з $R^n \to R^{n\times m}$.

РОЗДІЛ 3
ЛІНІЙНА ДИСКРИМІНАЦІЯ В МАТРИЧНОМУ ЕВКЛІДОВОМУ ПРОСТОРІ

Аналіз процесу математичного моделювання в задачах класифікації, кластеризації, розпізнавання образів і в задачах групування інформації свідчить, що проблеми, які виникають у зв'язку з такими задачами часто пов'язані з обмеженнями математичного апарату, який використовується для реалізації мети математичного моделювання в таких задачах. Це стосується як класичного, для таких задач використання, так званих, «векторів ознак» – наборів числових характеристик аналізованих з метою групування об'єктів, які «представляють» об'єкти в математичній моделі задачі, так і створення та використання засобів аналізу з метою групування в ситуації, коли «вектор ознак» представляє собою матрицю.

У першому випадку мова йде про використання техніки оперування з числовими векторами: елементами евклідового простору R^n, в другому – засобів оперування з метою групування в евклідових просторах матриць фіксованої розмірності в $R^{m\times n}$. Вирішальними в алгебраїчному апараті аналізу та синтезу для задач групування інформації є сингулярний розклад та псевдообернення за Муром – Пенроузом, а також систематичне використання розширеного варіанту сингулярного подання матриці як оператора між евклідовими просторами числових векторів для визначення псевдообернення за Муром-Пенроузом, у цьому випадку дія оператора визначається стандартним використанням операцій матричної алгебри з використанням техніки оперування в евклідових просторах числових векторів. У роботах М.Ф.Кириченка та його учнів створена і удосконалена техніка опису та оперування з фундаментальними для евклідових просторів числових векторів структурами [24,25]. Таким структурами для згаданих просторів є підпростори та гіперплощини, а також еліпси та мінімальні еліпсоїди групування: центральні та нецентральні, звичайні чи мінімальні, породжені тим чи іншим набором векторів, а також відповідні цим об'єктам, ортогональні проектори та матриці квадратичних форм, що задають відповідні структури. Матриці квадратичних форм, що задають еліпси та мінімальні еліпси групування, в теорії псевдообернення називають зваженими проекційними чи групувальними. Ортогональні проектори та групуючі оператори, що описуються явними формульними виразами, і конструктивно, і в явному вигляді будуються за заданими наборами векторів, дають можливість значно розширити критерії відповідності класу чи кластеру, заданого навчальною вибіркою: набором векторів евклідового простору R^n з використанням відстаней відповідності через відстань від породженої набором гіперплощини, чи – через еліпси та мінімальні еліпси групування, що в певному сенсі оптимальним чином «накривають» навчальну вибірку класу. Так само, використання псевдообернення в R^n дає можливість поставити та розв'язати задачу розділення двох сукупностей гіперплощиною, що має певну «товщину»: розв'язати задачу «робастної» лінійної дискримінації. Крім того, використання сингулярного розкладу з його представленням матриці у вигляді суми, природним чином визначає можливість відкидання певних членів такого розкладу. Таке відкидання членів в розкладі матриці, побудованої із числових векторів ознак як із стовпчиків, дозволяє отримати стовпчики перетвореної матриці як перетворені стовпчики первісної навчальної вибірки. Таке перетворення числових векторів навчальної вибірки можна характеризувати як «алгебраїчну фільтрацію». Як вже зазначалося вище, важливі задачі групування інформації пов'язані з об'єктами, представленими не числовими векторами (елементами евклідових просторів R^n), а матрицями. Класичними прикладами таких випадків є задачі розпізнавання мовних сигналів чи задачі розпізнавання зображень. Отже, принциповим для розширення можливостей розв'язання задач групування інформації є створення техніки оперування з базовими структурами: лінійними та нелінійними – в евклідових просторах матриць, яка дозволяла б так само, як і в евклідових просторах числових векторів, конструктивний опис лінійних підпросторів та гіперплощин, а також еліпсів та мінімальних еліпсів групування, породжених наборами матриць, що є навчальними – але матричними – вибірками тих чи інших класів. У цій частині – поширення теорії ПдО на матричні простори – пропонується розвиток теорії псевдообернення на «матричні» вектори ознак, що дозволяє розширити застосування засобів кластеризації, класифікації та розпізнавання у застосуванні до представників об'єктів класифікації, якими стають матриці. Зокрема, доведена теорема про еліпси та мінімальні еліпси групування, які, як і випадку R^n, пов'язуються із групуючими операторами, але побудованими за спеціальним класом лінійних операторів. Формулюючи задачу лінійної дискримінації у згаданому вище вигляді $R^{m\times n}$, нижче будемо слідувати роботі Кириченка М.Ф для R^n.

3.1. Дихотомічна кластеризація в задачах розпізнавання

Зважаючи на важливість роботи М.Ф. Кириченка [92], наведено основні постановки задач та підходи до розв'язання задачі із згаданої роботи.

Загалом основою процедур кластеризації, наприклад, методу k -середніх, що можуть бути використані в задачах розпізнавання, є ітераційні процедури розбиття аналізованої сукупності «представників» аналізованих з метою групування об'єктів – як правило, векторів ознак, що складають навчальну вибірку, – на частини та уточнення наявного розбиття: перевірки елементів навчальної вибірки на належність наявним на кроці рекурентної процедури, що виконується, у відповідності до значення певного функціоналу, який будемо називати «відстанню відповідності». У більшості випадків такою відстанню відповідності є евклідова відстань до представників наявних класів. Стандартними кроками такого алгоритму є рекурентна процедура, кожне виконання якої складається із трьох кроків:

1. Визначення представників кластерів;
2. Перевірка кожного із елементів наявної сукупності (навчальної вибірки) на належність за відстанню до представників класів;

3. Формування нового складу класів-кластерів через віднесення вектору ознак, що піддається перевірці, до того класу, що відповідає мінімуму евклідової відстані до представників класів.

Виконання вказаних кроків завершує крок рекурентної процедури і відбувається нове її виконання з першого пункту вже у застосуванні до нового розбиття на класи.

Кількість класів, на яке здійснюється розбиття первісної сукупності, як у методі k -середніх, є параметром процедури класифікації. Хоча існують і інші процедури кластеризації, теж на основі рекурентних процедур, але об'єднання існуючих класів за мінімумом певного функціоналу, який характеризує слушність об'єднання, в задачах кластеризації з метою класифікації чи розпізнавання в подальшому використовуватимуться саме процедури першого типу, які ілюструватимуться на прикладі методу k - середніх. Дихотомне розбиття на класи: розбиття на два класи, в описаній вище процедурі, з одного боку є найпростішим варіантом застосування описаної вище процедури, а з іншого – послідовним застосуванням дихотомії можна реалізувати кластеризацію на будь-яку кількість класів. Тому в подальшому застосовуватиметься саме дихотомний варіант кластеризації з метою класифікації.

Використання послідовної дихотомічної кластеризації використовуватиметься в подальшому у роботі з метою побудови дерева розпізнавання заданого набору елементів, заданих векторами ознак, які є елементами евклідового простору R^n. Таке використання дихотомічної кластеризації полягатиме в послідовному застосуванні дихотомії до класів чергового розбиття на два класи аналізованої сукупності, починаючи з вихідної, аж до того моменту, поки кожен з класів не складатиметься з одного елементу. Отже, етап аналізу в задачі класифікації через дихотомічну кластеризацію представлятиме собою дерево, коренем якого є аналізована з метою розпізнавання кожного із своїх елементів сукупність, чергове розгалуження якого є дихотомією класу, що піддається аналізу, листям – одноелементні представники початкової сукупності, наприклад, векторів ознак набору різних слів. Застосування дихотомічної кластеризації з метою класифікації, за підходу описаного вище, має ту безперечну перевагу, що за приблизно однакових розмірів класів, що піддаються дихотомії, кількість рівнів дерева складатиме 2 log N від загальної кількості N-елементів, аналізованої з метою розпізнавання первісної сукупності. На жаль, дихотомія має і певні вади, пов'язані з нестійкістю відповідного алгоритму до збурень елементів на етапі використання дерева розпізнавання. Подолання зазначеної вади дихотомії пропонується здійснити, з одного боку, через виконання самої дихотомії: побудови дерева розпізнавання – для незбурених чи мінімально збурених векторів навчальних вибірок, що представляють аналізовані з метою розпізнавання об'єкти. Пристосування до розпізнавання «збурених векторів» пропонується здійснити через робастну лінійну дискримінацію, ефективну реалізацію якої забезпечують методи псевдообернення за Муром-Пенроузом, застосовану до класів, отриманих на етапі дихотомії, але розширених «збуреними» векторами ознак. Робастність означає, що лінійна дискримінації здійснюється не просто за знаком лінійного функціоналу, а за перевищенням відповідним модулем певного порогу $\Delta > 0$. В такому варіанті алгоритму лінійна дискримінація застосовується для кожної вершини дерева дихотомічної класифікації, тобто для розділення двох класів.

3.2. Синтез гіперплощинних кластерів засобами оптимізації

При великій розмаїтості різних методів кластеризації (короткий їхній огляд приведений у [88]) зупинимося на декількох прийнятних для використання в нашій задачі методах кластеризації в просторі ознак.

В практиці досить часто використовується алгоритм кластеризації K-*середніх* (K-means method) [88], у якому формується деяка цільова функція $\Phi(\circ)$, що виражає якість поточного розбиття множини точок $x(j),\ j=\overline{1,n}$ на K кластерів з центрами в точках $c_i,\ i=\overline{1,K}$, K – задано. У початковий момент центри кластерів вибираються довільно, далі для кожної точки множини ітеративно визначаємо його приналежність до одного з K кластерів і обчислюємо нові значення $\Phi(\circ)$ для центрів кластерів, прагнучи при цьому мінімізувати функцію $\Phi(\circ)$. У літературі можна знайти різні його модифікації [121].

Одним із класичних вважається метод сегментації інформації, заснований на розбивці і злитті кластерів (split and merge methods) [121]. У початковий момент єдиним регіоном є вся множина точок $x(j),\ j=\overline{1,n}$, і на кожній ітерації методу здійснюється розщеплення кожного регіону R_i, для якого не виконується умова $P(R_i)$ приналежності точок до i-ого регіону на деяку кількість підрегіонів. Далі зливаються ті суміжні регіони R_i і R_k, для об'єднання яких $P(R_i \cup R_k)$ ця умова виконується. Процес закінчується, якщо існуючі регіони уже неможливо ні розщепити, ні об'єднати. Як логічний предикат $P(\circ)$ можна використовувати, наприклад, критерій K-means методу, або будувати різні міри подібності [121].

Серед множини інших публікацій по цій проблемі слід зазначити фундаментальну роботу [27], у якій дається детальний аналіз і огляд різних методів, у тому числі клас методів послідовного формування кластеризуючих підпросторів (The Subspaces Methods of Classification), які в певному розумінні ідейно близькі до нижче пропонованого авторами методу кластеризації.

3.3. Лінійна дискримінація для евклідового простору числових векторів: синтез класифікаторів засобами лінійних і нелінійних оптимальних перетворень простору ознак

Для розглянутих задач розпізнавання приналежності досліджуваних об'єктів до того чи іншого класу будемо припускати відомим, до якого з класів відносяться вектори навчальної вибірки $x(j),\ j=\overline{1,n}$ в просторі ознак. Тоді доцільно розглянути послідовно такі постановки задач:

- визначити необхідні і достатні умови лінійної відокремлюваності кінцевої множини точок у багатомірному просторі;

- визначити необхідні і достатні умови лінійної полосної відділюваності точок $x(j),\ \ j=\overline{1,n}$ на два класи;

- визначити оптимальну по товщині нелінійну смугу відділюваності точок $x(j),\ \ j=\overline{1,n}$ на два класи;

- впорядкувати координати $x_i(j),\ i=\overline{1,m-1},\ \ j=\overline{1,n}$, $x_m(j)\equiv 1$ відповідно до їх впливу чи на розміри (товщину) лінійної оптимальної смуги чи на критерій лінійної полосної розділюваності;

- сформулювати правило заміни слабо впливаючих (неінформативних) координат вектора ознак на більш інформативні з множини конкуруючих ознак;

- для вибраних відповідно до прикладної доцільності класів нелінійних перетворень запропонувати засіб визначення оптимального як самого нелінійного перетворення з вибраного класу деякої координати $x_i(j)$, так і індексу i цієї координати (приклади таких класів приводяться нижче);

- згенерувати суперпозицію оптимальних перетворень координат вектора ознак з метою досягнення умов оптимальної полосної розділюваності точок у новому просторі ознак.

3.4. Умови й оптимізація полосної розділюваності образів у просторі ознак

Сформулюємо спочатку умови лінійної віддільності точок $x(j),\ \ j=\overline{1,n}$ від початку координат у просторі R^m. Використовуючи операції псевдообернення матриць, цю задачу будемо розуміти і досліджувати в сенсі існування розв\`язку системи відносно вектора $a\in R^m$

$$x^T(j)a=y_j,\quad y_j\geq\Delta,\ \ j=\overline{1,n}, \tag{3.1}$$

при деяких $\Delta>0$ і значеннях y_j.

Тоді при фіксованому $\Delta > 0$ необхідна і достатня умова існування розв`язку цієї задачі буде мати вигляд

$$\min_{y \in D(\Delta)} y^T Z(X) y = y_*^T(\Delta) Z(X) y_*(\Delta) = 0\,, \tag{3.2}$$

де

$$X = \left(x(1) \vdots \ldots \vdots x(n) \right) = \begin{pmatrix} x_{(1)}^T \\ \ldots \\ x_{(m)}^T \end{pmatrix}, \qquad Z(X) = I_n - X^+ X\,,$$

$$D(\Delta) = \left\{ y:\ y = (y_1, \ldots, y_n)^T,\ \ y_j \geq \Delta,\ \ j = \overline{1,n} \right\}, \tag{3.3}$$

при цьому шуканий вектор a приймає наступне значення

$$a(\Delta) = (X^T)^+ y_*(\Delta)\,, \tag{3.4}$$

а товщина смуги δ, що відокремлює множину точок $x(j),\ j = \overline{1,n}$ від початку координат, буде дорівнює величині

$$y_* = \delta = \frac{\Delta}{\left(y_*^T(\Delta) R(X) y_*(\Delta) \right)^{\frac{1}{2}}}, \tag{3.5}$$

де $R(X) = X^+ (X^T)^+$.

Тому що $y_*(k\Delta) = k y_*(\Delta)$, то без обмеження спільності можна покласти $\Delta = 1,\ \ y_*(1) = y_*$. Отже, максимальна товщина смуги досягається при значеннях

$$y_{opt} = \arg\min_{y \in D} y^T R(X) y, \qquad a_{opt} = (X^T)^+ y_{opt}, \tag{3.6}$$

де $D = \left\{ y : y^T Z(X) y = y_*^T Z(X) y_* = 0,\ \ e_j^T y \geq 1,\ \ j = \overline{1,n} \right\}$. Використовуючи сингулярний розклад матриць (SVD)

$$X = \sum_{j=1}^{r} u_j v_j^T \lambda_j, \quad Z(X) = I_n - \sum_{j=1}^{r} v_j v_j^T \ ,$$
$$u_i^T u_j = \delta_{ij}, \quad v_i^T v_j = \delta_{ij}, \quad XX^T u_j = \lambda_j^2 u_j, \quad X^T X v_j = \lambda_j^2 v_j,$$
$$\lambda_1^2 \geq \ldots \geq \lambda_r^2, \qquad i, j = \overline{1,n},$$

і з огляду на, що $y^T Z(X) y = 0$ для $y \in D$, можна записати наступні співвідношення

$$y = \sum_{i=1}^{r} \alpha_i v_i, \quad \sum_{i=1}^{r} \alpha_i e_j^T v_i \geq 1, \quad \forall\, j = \overline{1,n}, \quad y^T R(X) y = \sum_{j=1}^{r} \alpha_j^2 \lambda_j^{-2} \,. \tag{3.7}$$

Таким чином, задача пошуку *оптимальної лінійної віддільності точок* $x(j),\ j = \overline{1,n}$ *від початку координат* зводиться до розв’язування задачі оптимізації квадратичної функції на опуклій множині

$$\alpha_{opt} = \arg \min_{\alpha = \{\alpha :\, e_j^T (v_1 \vdots \ldots \vdots v_r)\alpha \geq 1,\ j = \overline{1,n}\}} \alpha^T diag(\lambda_1^{-2}, \ldots, \lambda_r^{-2})\alpha,$$

$$y_{opt} = \left(v_1 \vdots \ldots \vdots v_r\right) \alpha_{opt},$$

$$a_{opt} = \sum_{j=1}^{r} u_j v_j^T \lambda_j^{-1} \cdot (v_1 \vdots \ldots \vdots v_r)\alpha_{opt} = \left(\lambda_1^{-1} u_1 \vdots \ldots \vdots \lambda_r^{-1} u_r\right) \alpha_{opt}, \tag{3.8}$$

$$\delta_{opt} = \left(\alpha_{opt}^T\, diag(\lambda_1^{-2}, \ldots, \lambda_r^{-2})\alpha_{opt}\right)^{-\frac{1}{2}}.$$

Використовуючи вищеприведені умови лінійної віддільності точок у R^m для задачі лінійної полосної відокремлюваності двох класів точок у просторі ознак, одержимо як умови існування розв`язку такої задачі, так і засіб одержання самого розв`язку. Нехай відомо, що для послідовності точок $x(j)$ у просторі ознак R^m

$$x(j) \in R^m, \quad x(j) = \begin{pmatrix} x_1(j) \\ \vdots \\ x_m(j) \end{pmatrix}, \quad x_m(j) = 1, \quad j = \overline{1,n}$$

точок $x(i_k),\ k=\overline{1,n_1}$ належать першому класу, а $x(j_s),\ s=\overline{1,n_2}$ - другому класу. Тоді лінійну полосну відокремлюваність цих класів будемо розуміти в сенсі існування такого вектора $a \in R^m$, для якого

$$\begin{aligned} &a^T x(i_k) \ge 1, \quad k=\overline{1,n_1}, \\ &a^T x(j_s) \le -1, \quad s=\overline{1,n_2}. \end{aligned} \tag{3.9}$$

Тоді *умова лінійної полосної відокремлюваності* приймає наступний вигляд

$$\min_{y\in D} y^T Z(X) y = 0, \quad D = \left\{ y:\ e_{i_k}^T y \ge 1,\ e_{j_s}^T y \le -1,\ k=\overline{1,n_1}, s=\overline{1,n_2} \right\}, \tag{3.10}$$

а значення вектора a визначається оптимально в сенсі максимізації товщини розділяючої смуги

$$y_{opt} = \arg\min_{y\in D_1} y^T R(X) y, \tag{3.11}$$

де $D_1 = \left\{ y: y^T Z(X) y = 0 \right\} \cap D$,

$$a_{opt} = \left(X^T\right)^+ y_{opt}. \tag{3.12}$$

Застосовуючи представлення SVD для матриці X, ці операції можна інтерпретувати в більш зручній для використання формі

$$\alpha_{opt} = \arg \min_{\alpha \in \left\{\alpha:\ e_{i_k}^T (v_1 \ldots v_r)\alpha \ge 1,\ e_{js}^T (v_1 \ldots v_r)\alpha \le -1,\ k=\overline{1,n_1},\ s=\overline{1,n_2}\right\}} \alpha^T diag(\lambda_1^{-2},\ldots,\lambda_r^{-2})\alpha, \tag{3.13}$$

$$y_{opt} = (v_1 \ldots v_r)\, \alpha_{opt}, \tag{3.14}$$

$$a_{opt} = (u_1 \ldots u_r)\, diag\left(\lambda_1^{-1},\ldots,\lambda_r^{-1}\right) \alpha_{opt}. \tag{3.15}$$

3.5. Оптимальне нелінійне перетворення компонентів вектора ознак

Дуже часта умова лінійної полосної відокремлюваності (3.10) в обраному просторі ознак не виконується, тобто

$$\min_{y \in D} y^T Z(X) y > 0 . \qquad (3.16)$$

У цьому випадку можна використовувати наступні засоби можливого виходу з ситуації. Перш за все спробувати поліпшити вибір інформативних ознак. Другий засіб полягає у використанні каскадного полосного лінійного розділення двох класів [92] і, як нам представляється, дуже перспективно *синтезувати оптимально суперпозицію нелінійних перетворень* у просторі ознак. Цей синтез пропонується здійснювати поетапно в такий спосіб.

Розглянемо спочатку *задачу синтезу нелінійного перетворювача* при відомих вхідних і вихідних даних, відповідно $X = \left(x(1) \vdots \ldots \vdots x(n) \right)$ і $y = (y_1, \ldots, y_n)$. Спробуємо шляхом пошуку нелінійного перетворення деяких компонентів вектора $x_{(i)}$ $i = \overline{1, m-1}$ змінити їх таким чином, щоб зменшити нев'язання $y^T Z(X) y$, і визначимо серед них оптимальний компонент з індексом i_0 і відповідне їй оптимальне перетворення. Для цього виконаємо нижчезазначену послідовність дій.

1. Знайти $a = \left(X^T \right)^+ y$.

2. Обчислити

$$a(i) = \left(X_{(i)}^T \right)^+ y , \text{ де } X_{(i)} = \begin{pmatrix} x_{(1)}^T \\ \vdots \\ x_{(i-1)}^T \\ x_{(i+1)}^T \\ \vdots \\ x_{(m)}^T \end{pmatrix} , \quad i = \overline{1, m-1} . \qquad (3.17)$$

3. Утворити вектор відхилень

$$\Delta y_i = \left(\Delta y_i(1), \ldots, \Delta y_i(n) \right)^T , \quad \Delta y_i(j) = y(j) - \hat{y}_i(j), \; j = \overline{1, n} , \qquad (3.18)$$

де $\hat{y}_i = \left(\hat{y}_i(1), \ldots, \hat{y}_i(n) \right)^T = Z(X_{(i)}) y = \left(I_n - X_{(i)}^+ X_{(i)} \right) y = y - X_{(i)}^T a(i)$.

4. Замінити вилучений компонент $x_i(j)$ нелінійною функцією виду $\varphi\left(a^T x(j), \alpha(i)\right)$, невідомі параметри $\alpha(i)$ визначаються з умови мінімуму відхилення

$$\sum_{j=1}^{n} \left| \varphi\left(a^T x(j), \alpha(i)\right) - \Delta y_i(j) \right|^2 . \tag{3.19}$$

Як приклад нелінійного перетворення можна розглянути сплайни третього порядку

$$\varphi\left(a^T x(j), \alpha(i)\right) = \begin{pmatrix} 1 & z(j) & z^2(j) & z^3(j) \end{pmatrix} \alpha(i), \quad z(j) = a^T x(j), \; j = \overline{1,n}$$

$$\alpha(i) = \left(\sum_{j=1}^{n} \begin{pmatrix} 1 \\ z(j) \\ z^2(j) \\ z^3(j) \end{pmatrix} \begin{pmatrix} 1 & z(j) & z^2(j) & z^3(j) \end{pmatrix} \right)^{+} \sum_{j=1}^{n} \begin{pmatrix} 1 \\ z(j) \\ z^2(j) \\ z^3(j) \end{pmatrix} \cdot \Delta y_i(j). \tag{3.20}$$

5. Знайти компоненту з індексом $i_0 \in \left\{ i : \overline{1, m-1} \right\}$, при якій досягається мінімум величини $y^T Z\left(X_{\varphi,i}\right) y$,

де $X_{\varphi,i} = \left(x_{(1)} \ldots x_{(i-1)} \; \varphi_i \; x_{(i+1)} \ldots x_{(m)} \right)$, $\varphi_i = \left(\varphi\left(a^T x(1), \alpha(i)\right) \ldots \varphi\left(a^T x(n), \alpha(i)\right) \right)^T$.

І, таким чином, на першому етапі одержуємо найкраще перетворення оптимально обраного компонента $x_{(i_0)}$.

Після чого повторюємо для нових вхідних даних описаний вище процес нелінійного перетворення, реалізуючи, таким чином, суперпозицію нелінійних функціональних перетворень з метою мінімізації нев’язки.

У задачі синтезу нелінійної системи розпізнавання образів задача ускладнюється в зв'язку з тим, що вихідні дані (вектор y) наперед невідомі. А саме, якщо $\min\limits_{y \in D} y^T Z(X) y > 0$, то в цьому випадку, як правило, такої лінійної розділяючої смуги не існує. Виникає питання, які компоненти $x_{(i_0)}$ послідовно нелінійно перетворити таким чином, щоб розділяюча смуга вхідних даних в новому отриманому просторі існувала.

Для цього скористаємося вищеприведеним алгоритмом, у якому вихідні дані складають вектор y_*. Цей вектор y_* знаходимо, вирішуючи задачу оптимізації квадратичної функції при лінійних обмеженнях

$$y_* = Arg\min_{y\in D} y^T Z(X)\, y \quad , \tag{3.21}$$

використовуючи одну з модифікацій градієнтного методу, наприклад описану в [92].

Далі виконуємо усі вищеприведені операції п.1-п.5, думаючи $y = y_*$. Для нових векторів ознак

$$\begin{pmatrix} x_1(j) \\ \vdots \\ x_{i_0-1}(j) \\ \varphi\left(a^T x(j), \alpha(i_0)\right) \\ x_{i_0+1}(j) \\ \vdots \\ x_m(j) \end{pmatrix}, \quad j = \overline{1,n} \tag{3.22}$$

визначити нове значення $y \in D$, при якому досягається мінімум виразу $y^T Z\left(X_{\varphi,i_0}\right) y$ і відповідне нове значення

$$a = \left(X^T_{\varphi,i_0}\right)^+ y \quad . \tag{3.23}$$

Якщо це мінімальне значення $y^T Z\left(X_{\varphi,i_0}\right) y$ дорівнює нулю, тобто, виконується умова лінійної полосної роздільності в новому просторі ознак, тоді необхідно знайти

$$a_{opt} = \left(X^T_{\varphi,i_0}\right)^+ y_{opt} \quad , \quad y = \arg\min_{y\in D_1} y^T R\left((X^T_{\varphi,i_0})^+\right) y, \tag{3.24}$$

де $D_1 = \left\{ y : y^T Z\left(X_{\varphi,i_0}\right) y = 0 \right\} \cap D$. І, таким чином, побудувати оптимальну лінійну розділяючу операцію

$$a_{opt}^T x(i_k) \ge 1, \quad a_{opt}^T x(j_s) \le -1, \quad k = \overline{1, n_1}, \; s = \overline{1, n_2}$$

для нових векторів ознак, що одержуються після нелінійного перетворення компонентів.

3.6. Лінійна дискримінація в матричних евклідових просторах

Задача лінійної дискримінації (ЛД) полягає в розділенні на два класи, представлених навчальною вибіркою для відповідних гіперплощин. Для евклідових просторів ця задача була сформульована й успішно розв'язана на основі ПдО техніки в роботах [94,95]. Ця задача формулюється і розв'язується нижче із застосування апарату псевдообернення розробленої для евклідового простору матриць на основі концепції «кортежних операторів», представлених у другому розділі.

3.6.1. Постановка задачі лінійної дискримінації

Як і у наведеній вище постановці і розв'язку ЛД-задачі для R^n, нехай $X_1,...,X_K \in R^{m \times n}$ множина матриць з навчальної вибірки представлена двома классами $X_j \in Kl_1, j \in J_1, X_j \in Kl_2, j \in J_2$: $J_1 \cap J_2 = \varnothing$, $J_1 \cup J_2 = \{1, 2, .$. Необхідно знайти $\Delta > 0$ і визначити лінійний функціонал $A : R^{m \times n} \to R^1$ таким чином, щоб $(A, X_j) > \Delta, j \in J_1, (A, X_j) < -\Delta, j \in J_2$. Позначимо $\Omega(\Delta)$ як простір числових векторів $y^T = (y_1,...,y_K)$ з R^K з компонентами, які задовольняють наступним обмеженням $y_j > \Delta, j \in J_1, y_j < -\Delta, j \in J_2$.

3.6.2. Формалізація задачі лінійної дискримінації

Позначимо, як і у векторному випадку, $\Omega_\Delta \in R^N$, компоненти якої відповідають впорядкованим обмеженням та значенням лінійного функціоналу. Це означає, що вектор значень лінійного функціоналу 1_A, який треба побудувати належить Ω_Δ, яку будемо називати допустимою

$$\begin{pmatrix} l_{A,A_1} \\ \vdots \\ l_{A,A_N} \end{pmatrix} \in \Omega_\Delta \text{ чи в еквівалентному вигляді } \begin{pmatrix} A, A_1 \\ \vdots \\ A, A_N \end{pmatrix} \in \Omega_\Delta . \qquad (3.25)$$

Як випливає з теорії «кортежних операторів» розділу 2 числовий вектор

$$\begin{pmatrix} A, A_1 \\ \vdots \\ A, A_N \end{pmatrix}$$

є значенням спряженого $\wp_\alpha^*$ до «кортежного оператора» $\wp_\alpha$ побудованого за кортежем $\wp_\alpha : \alpha = (X_1, \ldots, X_K)$.

Отже, в формалізованому варіанті мова йде про побудову-відшукання матриці A, для якої її перетворення за допомогою $\wp_\alpha^*$ належало б Ω_Δ:

$$X_1, \ldots, X_K \Rightarrow \alpha = (X_1, \ldots, X_K) \Rightarrow \wp_\alpha \Rightarrow \wp_\alpha^* \Rightarrow \mathrm{A} : \wp_\alpha^* \in \Omega_\Delta$$

Має місце наступна теорема.

Теорема 3.1 Задача ЛД є рівносильною до задачі лінійного рівняння $\wp_\alpha^* A = y$ для «кортежного оператора» $\wp_\alpha, \alpha = (X_1, \ldots, X_K)$ і $(A, X_j) > \Delta, j \in J_1, (A, X_j) < -\Delta, j \in J_2$.

Доведення. Дійсно, виконання (3.25) означає, що вектор

$$y : \ (A, X_1), \ldots, (A, X_K) \ \equiv y^T \qquad (3.26)$$

належить до Ω_Δ.

Це дозволяє зробити висновок, що (3.26) рівносильне розв'язку рівняння $\wp_\alpha^* A = y, \alpha = (X_1, \ldots, X_K), y \in \Omega_\Delta$. Таким чином, доведення теореми закінчено.

Теорема 3.2 Належність розв'язків у допустимій області означає, що в цій області існує елемент для якого СЛАР є роз'язною. Отже, задача ЛД еквівалентна існуванню в області Ω_Δ вектора y, для якого $\wp_\alpha^+ A = y, y \in \Omega_\Delta$.

Теорема 3.3 Задача ЛД має розв'язок, тоді і тільки тоді коли існує $y_* \in \Omega(\Delta) \subseteq R^K$ і відповідно розв'язок визначається рівнянням (лінійною формою)

$$A = \wp_\alpha^{*+} y_* \qquad (3.27)$$

Наслідок. Задача ЛД має розв'язок тоді і тільки тоді, коли існує $y_* \in \Omega(\Delta) \subseteq R^K$ для якого виконуються такі умови $y_*^T Z(\wp_\alpha) y_* = y_*^T (E_K - \sum_{k=1}^{r} v_k v_k^T) y_* = 0$.

Доведення. Випливає із того факту, що для розв'язності СЛАР необхідо і достатньо виконання умови $y^T Z(\wp_\alpha) y = 0$ на її праву частину.

Теорема 3.4 Розв'язок матричної задачі лінійної дискримінації рівносильна задачі оптимізації в області $\Omega(\Delta) \subseteq R^K$ для квадратичної форми

$$y^T Z(\wp_\alpha) y = y^T (E_K - \sum_{k=1}^{r} v_k v_k^T) y .$$

Якщо розв'язок y_* в оптимізаційній задачі дає мінімум, який дорівнює нулю, то розв'язок задачі лінійної дискримінації визначається співвідношенням

$$A = \wp_\alpha^{*+} y_* .$$

Доведення. Зважаючи на умову розв'язності СЛАР для $y_* \in \Omega_\Delta$, мусить знайтись квадратична форма $y_*^T Z(\wp_\alpha) y_* = y_*^T (E_K - \sum_{k=1}^{r} v_k v_k^T) y_* = 0$. Зважаючи на невід'ємну визначеність, нульове значення є мінімальним для y_*. Отже, розв'язавши задачу

$y^T Z(\wp_\alpha) y = 0$ у області Ω_Δ і визначивши значення відповідного мінімуму, можна визначити і розв'язність СЛАР в Ω_Δ.

Наслідок. Розв'язання ЛД-задачі може бути досліджено у двох напрямках:

1) Квадратичної форми з визначенням точки мінімуму.

2) У разі, якщо мінімум дорівнює нулю, розв'язати СЛАР із точкою y_*, що доставляє мунумум квадратичній формі Z-оператора.

Зауваження. Неперервна функція, яка є квадратичною формою, досягає свого мінімуму у замкненій області. Замкненість задачі полягає в тому, що в обмеженому Ω_Δ при достатньо великому M включимо Δ у відповідну границю.

Теорема 3.5 Задача визначення матриці A, що є розв'язком задачі робастної лінійної дискримінації є еквівалентною існуванню розвязку для СЛАР $\wp_\alpha^* A = y$, з правою частиною, що належить допустимій області.

Доведення. Дійсно, визначення задачі лінійної дискримінації у вигляді побудови матриці A так, щоб вектор $\Omega_{(\Delta)} : \begin{pmatrix} A, X_1 \\ \vdots \\ A, X_k \end{pmatrix} \in \Omega_{(\Delta)}$ за визначенням спряженого $\wp_\alpha^*$ до $\wp_\alpha$, де $\alpha = (X_1, \ldots, X_K)$, що дає $\wp_\alpha^* A = \begin{pmatrix} A, X_1 \\ \vdots \\ A, X_k \end{pmatrix}$. Тобто, вимога $\begin{pmatrix} A, X_1 \\ \vdots \\ A, X_k \end{pmatrix} \in \Omega_{(\Delta)}$ є еквівалентною вимозі виконання умови $\wp_\alpha^* A \in \Omega_{(\Delta)}$, останнє означає, що існує $y \in \Omega_{(\Delta)}$ таке, що $\wp_\alpha^* A = y$.

Навпаки, існування $y \in \Omega_{(\Delta)}$ для якого $\wp_\alpha^* A = y$ розв'язна означає існування A такого, що $\wp_\alpha^* A = \begin{pmatrix} A, X_1 \\ \vdots \\ A, X_k \end{pmatrix} = y \in \Omega_{(\Delta)}$, і дає розв'язок робастної задачі лінійної дискримінації для такої матриці A.

Теорема 3.6 Нехай $y^* \in \Omega_{(\nabla)}$ такі, що $\wp_\alpha^* A = y$ розвязна. Тоді $A = \wp_\alpha^{*+} y^*$ є розв'язком задач лінійної дискримінації.

Доведення. Матриця $A = \wp_\alpha^{*+} y^*$ є найменшим за нормою розв'язком СЛАР $\wp_\alpha^* X = y^*$, це означає, що для неї $\wp_\alpha^* A = \begin{pmatrix} A, X_1 \\ \vdots \\ A, X_k \end{pmatrix} = y \in \Omega_{(\Delta)}$, а отже є розв'яком задачі лінійної дискримінації.

Необхідною і достатньою умовою розвязку задачі лінійної дискримінації і Z – оператора слушних кортежних операторів. У попередньому параграфі було встановлено необхідні достатні умови розв'язності задачі лінійної дискримінації в термінах СЛАР, існування розв'язків СЛАР із «кортежним оператором» $\wp_\alpha$, $\alpha = (X_1 \ldots X_K)$ і змінною правою частиною, що мусять належати допустимій області $\Omega_{(\Delta)}$.

Це означає, що в області $\Omega_{(\nabla)}$ існує елемент y^*, для якого виконується умова розв'язності СЛАР, отже справедлива є наступна теорема.

Теорема 3.7 Для розв'язності задачі лінійної дискримінації з фіксованим $\Delta > 0$ необхідно і достатньо, щоб в $\Omega_{(\nabla)}$ існував елемент y^*, для якого виконується умова розв'язності СЛАР $\wp_\alpha^* = y^*$, тобто $y_*^T Z(\wp_\alpha) y_* = 0$.

Доведення. Необхідність: розв'язність задачі лінійної дискримінації, як зазначалося вище є еквівалентною розв'язності СЛАР $\wp_\alpha^* = y^*, y \in \Omega_{(\Delta)}$, це означає, що для цього виконуються умови розв'язності $y_*^T Z(\wp_\alpha) y_* = 0$.

Навпаки, виконання для певного $y \in \Omega_{(\Delta)}$ умови $y_*^T Z(\wp_\alpha) y_* = 0$ означає, що для цього $y \in \Omega_{(\Delta)}$ СЛАР $\wp_\alpha^* X = y^*$ розв'язна.

3.7. Алгоритм задачі лінійної дискримінації

Кажучи про алгоритм для ЛД-задачі з двома класами набору матриць $X_1,\ldots,X_K \in R^{m\times n}$, виконуємо наступні кроки:

1-й крок: визначаємо матрицю Грама для набору матриць $X_1,\ldots,X_K \in R^{m\times n}$:

$$F = \begin{pmatrix} trA_1^T A_1,\ldots,trA_1^T A_K \\ \ldots \\ trA_K^T A_1,\ldots,trA_K^T A_K \end{pmatrix} = \begin{pmatrix} \langle A_1, A_1 \rangle_{tr},\ldots,\langle A_1, A_K \rangle_{tr} \\ \ldots \\ \langle A_K, A_1 \rangle_{tr},\ldots,\langle A_K, A_K \rangle_{tr} \end{pmatrix}.$$

2-й крок: визначаємо сингулярності $(v_k, \lambda_k^2), \lambda_k^2 > 0,\ k = \overline{1,r},\ \ r = rankF$;

3-й крок: визначаємо матрицю $E_K - \sum_{k=1}^{r} v_k v_k^T$ квадратичної форми ;

4-й крок: знаходимо мінімуму квадратичної форми $y^T \left(E_K - \sum_{k=1}^{r} v_k v_k^T \right) y$ в області $\Omega(\Delta) \subseteq R^K$ числовими методами і відповідний аргумент y_*;

5-й крок: виконуємо перевірку умови рівності мінімума нулю: $y_*^T\left(E_K-\sum_{k=1}^{r}v_kv_k^T\right)y_*=0$:

6-й крок:

-якщо задовольняє умові: $y_*^T\left(E_K-\sum_{k=1}^{r}v_kv_k^T\right)y_*=0$, тоді знаходження лінійної A для задачі лінійної дискримінації записується співвідношенням

$$- A=\wp_{\chi}^{*+}y_*=\sum_{i=1}^{r}\lambda_i^{-1}\frac{1}{\lambda_i}\sum_{k=1}^{K}X_kv_{ki}(v_i,y_*)= \quad =\sum_{i=1}^{r}\sum_{k=1}^{K}\lambda_i^{-2}v_{ki}X_k(v_i,y_*);$$

-якщо умова не задовольняється тоді ЗЛД не розв'язна.

3.7.1. Зауваження до алгоритму

Зважаючи на числову реалізацію:

1) Можливо потрібно буде збільшити M -границю.
2) Загалом в ЛД-задачі визначення Δ вважається параметром і може бути вибраним. Задача ЛД встановлюється як сам функціонал, так і робастна задача. Якщо не розв'язна, то можна запустити рекурентну процедуру із зменшенням кроку, що дасть змогу врахувати цей параметр ЛД-задачі.
3) Може виявитись, що ні 1, ні 2 варіанти не дадуть розв'язки, тоді можна спробувати запустити нелінійні перетворення матриць, що складають навчальну вибірку. До одного чи іншого та послідовно.

Пропонується ідея повного розв'язку лінійної дискримінації для матричних евклідових просторів, поділених на два класи і представлених навчальними вибірками.

Ми посилатимемось на ці класи як Cl_1,Cl_2 відповідно з навчальними вибірками $X(j)\in R^{m\times n},j=\overline{1,N}$ і J_1,J_2- множина індексів набору $\{1,...,N\}$ які відповідають кожному із набору навчальних вибірок для кожного з класів:

$$J_1,J_2\subseteq\{1,...,N\}:\ J_1\cap J_2=\varnothing,\ J_1\cup J_2=\{1,...,N\},$$

$$X(j) \in Cl_k \Leftrightarrow j \in J_k,\ k = 1,2,$$

$$j = \overline{1,N}$$

Розумітимемо під задачею лінійної дискримінації в $R^{m\times n}$ для двох класів Cl_1, Cl_2 представлених елементами $X(j), j \in J_1, X(j), j \in J_2$ з загальної навчальної вибірки $X(j), j = \overline{1,N}$, також як задачу побудови лінійного функціоналу $\varphi : R^{m\times n} \to R^1$ (дискримінаційної функції) для деякого $\Delta > 0$. Сама задача буде мати вигляд:

$$\begin{array}{l} y_j = \varphi(X(j)) \geq \Delta, j \in J_1, \\ y_j = \varphi(X(j)) \leq -\Delta, j \in J_2 \end{array}. \tag{3.28}$$

Лінійність для функціоналу φ означає, що функціонал може бути однозначно представлений через скалярний добуток: $\varphi(X) = (A, X)_{tr}$, де A - $m \times n$ матриця.

Теорема 3.8 Задача Лінійної Дискримінації рівносильна розв'язку задачі квадратичної оптимізації для $y^T Z(X) y$ на множині $\Omega(\Delta)$. Необхідною і достатньою умовою для існування розв'язку задачі Лінійної Дискримінації є те, щоб мінімум $y_* \in R^N$ квадратичної форми $y^T Z(X) y$ належав до $\Omega(\Delta)$:

$$y_* = \arg\min_{y \in R^N} y^T Z(X) y, y_* \in \Omega(\Delta), \tag{3.29}$$

де $Z(X) = E_N - X^+ X$, і X^+ - псевдообернення за Муром-Пенроузом для матриці X як лінійний оператор з R^N в R^N (наведено, для прикладу , [117]).

Доведення. Справді, умова $\wp^*_{\alpha_L} A = y \in \Omega(\Delta)$ вказує на те, що для довільного $y \in \Omega(\Delta)$ лінійного рівняння $\wp^*_{\alpha_L} A = y$ є розв'язною для $y \in \Omega(\Delta)$. Це значить, що y належить до області: $\wp^*_{\alpha_L}$: $y \in \Re\ \wp^*_{\alpha_L}$. Очевидно, що $\Re\ \wp^*_{\alpha_L} = \Re\ \wp_{\alpha_L} \wp^*_{\alpha_L} = \Re(X)$.

Належність до лінійного підпростору або діапазону означає, що це нерухома точка відповідного ортогонального проектора. Для $\Re\ \wp^*_{\alpha_L} = \Re(X)$ відповідні ортогональні проектори співпадають $P_{\Re\ \wp^*_{\alpha_L}} = P_{\Re(X)}$, так

$$P_{\Re(X)} y = y . \tag{3.30}$$

Отже, $y = P_{\Re(X)} y$, або $E_N - P_{\Re(X)}\ y = 0$.

Ортогональний проектор $P_{\Re(X)}$ однозначно визначається через псевдообернення для X відповідно до наступного рівняння: $P_{\Re(X)} = X^{+}X$.

Таким чином (3.30) може бути записано як $E_N - X^{+}X\ \ y = 0$, або , що теж саме зазначено наприклад в [71] ,

$$Z(X)y = 0\,. \tag{3.31}$$

У свою чергу, остання рівність еквівалентна $y^T Z(X)y = 0$.

Остання рівність означає, що абсолютний мінімум невід'ємної квадратичної форми $y^T Z(X)y, y \in R^N$ досягається в області $\Omega(\Delta)$. Це еквівалентно, існуванню $y_* \in \Omega(\Delta)$ який є мінімумом для $y^T Z(X)y, y \in \Omega(\Delta)$ і цей мінімум дорівнює нулю: $y_*^T Z(X)y_* = 0, y_g \in \Omega(\Delta)$. Кінець доведення.

Нерозв'язність задачі оптимізації з обмеженнями з теореми (3.8) означає нерозв'язність задачі Лінійної Дискримінації для моделі набору компонентів з матриць. Таким чином, компоненту потрібно спеціально змінювати.

Висновок до розділу 3

Розв'язання ЛД-задачі в матричних евклідових просторах досягається через реалізацію концепції «кортежних операторів». Задача ЛД зведена до пошуку розв'язків слушної СЛАР, з оператором, який є спряженим до «кортежного оператора» з матричним кортежем, що визначає квадратичну форму чи об'єднання двох кластерів.

Застосування техніки «кортежних операторів» зводять розв'язання задачі до дослідження розв'язності СЛАР із «кортежним оператором» і правою частиною, що мусять належати певній допустимій області. В кінцевому рахунку задача зводиться до дослідження векторів, заданих у певній області, яку можна компактувати. Таким чином, постановка і розв'язання задачі ЛД заданої в евклідовому матричному просторі дають однаковий результат, зведенням до слушної оптимізації із слушною заміною матриці на оператор, який можна вибрати спряженням до «кортежного оператора». Таким чином, складність алгоритму розв'язання задачі ЛД відповідає складності розв'язання задачі на власні значення для матриці Грама, об'єднаних навчальною вибіркою кластерів. Ця дихотомічна задача постає у роботах [94, 95]. Ця задача у варіанті дихотомії так само буде розглядатися і для евклідового простору матриць.

Загалом, задача лінійної дискримінації зводиться до задачі мінімізації квадратичної формули $y^T Z(\wp_\alpha) y = 0$ у випадку існування мінімуму, значення якого дорівнює нулю, розв'язок існує, у іншому випадку – ні.

Як відомо, існують розвинені алгоритми квадратичної оптимізації невід'ємно визначених квадратичних норм. Хоча задача лінійної дискримінації не обмежена, але існує і в обмеженому варіанті. Тому для розв'язання задач лінійної дискримінації використовують обчислювальні методи оптимізації (мінімізації) симетричних невід'ємно визначених квадратичних форм в області, які мають ітераційний характер.

РОЗДІЛ 4

ФОРМУЛИ ГРЕВІЛЯ-КИРИЧЕНКА: ЕВКЛІДОВІ ПРОСТОРИ МАТРИЦЬ

У прикладних дослідженнях важливими є процедури рекурентного обчислення, які виникають при розширенні інформації. У задачах відновлення функції та групування таке розширення інформації втілюється в розширенні наявної вибірки. У задачах відновлення функції –додавання нового спостереження. У задачах класифікації, кластеризації, розпізнавання образів – це розширення навчальної вибірки відповідних класів. Зокрема, у задачах відновлення функції таке згадане розширення інформації про спостереження функції у лінійному варіанті регресії призводить на збільшення одинички рядків матриці класу і збільшення на одиничку компоненту правої частини СЛАР. У зв'язку з чим формується задача оцінювання квадратичної мінімізації. Осклькии всі розвязки СЛАР і псевдорозвязки описуються співвідношенням (гіперплощиною), складові описи якої використовують ПдО та згадана рекурентність зводиться до рекурентних обчислень за розширенням рядком чи стовпчиком.

Сказане вище повною мірою відноситься до відстаней відповідностей у випадку занурення у підпростір чи гіперплощину чи еліпсоїди. Повністю формули ПдО за розширенням матриці за рядком чи стовпчиком у евклідовому просторі числових векторів доведені М.Ф. Кириченком.

Як зазначалось вище, в задачах класифікації, кластеризації, розпізнавання образів така рекурентність зможе бути використана для опису засобів кластеризації для векторів числових ознак, які представляють класифіковані об'єкти. Повною мірою це відноситься до того випадку, коли навчальна вибірка кластерів складається із матриць.

Загалом, у матричному випадку, задача побудови визначення рекурентності, співвідношення ПдО зводиться до рекурентних формул ПдО для «кортежного оператора» чи його спряженого за розширенням кортежу додатковою матрицею. Як показано нижче, рекурентні формули ПдО для матриць переносяться на «кортежні оператори» із редукцією складності обчислень, як і у матричному випадку.

У цьому розділі розглядаються розширені рекурентні формули ПдО, так звані формули Гревіля – Кириченка за використанням «кортежних операторів». Загалом, на початку розділу наведено загальні формули Гревіля – Кириченка в прямому варіанті за розширенням матриці.

4.1. Залежність псевдообернених і проекційних матриць від приєднання довільних вектор-рядків до вихідної матриці

Якщо передбачити, що розширення матриці

$$A = (a(1) \dots a(n)) \equiv (a_{(1)} \dots a_{(m)})^T \in R^{m\times n}.$$

де $a(j) \in R^m$ - стовчики матриці A,

$a_{(i)} \in R^n$ – вектори рядків матриці A, $i=\overline{1,m}$, $j=\overline{1,n}$.

відбувається приєднанням нового рядка $a^T \in R^n$ після $(i-1)$ - ого рядка $(i=\overline{2,m+1})$, тобто утворюється матриця

$$A_{i,a} = (a_{(1)} \dots a_{(i-1)}\, a\, a_{(i)} \dots a_{(m)})^T \in R^{(m+1)\times n},$$

то при відомій псевдооберненій матриці $A^+ \in R^{n\times m}$ для рекурентного обчислення псевдооберненої матриці $A^+_{i,a} \in R^{n\times(m+1)}$ мають місце співвідношення – прямі формули Гревіля.

Прямі формули Гревіля для псевдообернених матриць. Якщо невідому псевдообернену матрицю $A^+_{i,a}$ представити в вигляді

$$A^+_{i,a} = (p(1) \dots p(i-1)\, p(i)\, p(i+1) \dots p(m+1)) \in R^{n\times(m+1)},$$

вважати, що матриця $A^+_{i,a}$ може бути отримана з матриці

$$P_i = \left(p(1) \ldots p(i-1)\; p(i+1) \ldots p(n+1)\right),$$

вставленням після $(i-1)$ - го стовпчика вектора $p(i)$, то для матриці P_i має місце формула:

$$P_i = (1 - p(i)a^T)A^+ .$$

При цьому невідомий вектор $p(i)$ визначається в залежності від властивості лінійної залежності вектора a від векторів підпростору вектор-рядків матриці A:

1) якщо вектор a лінійно незалежний від векторів підпростору вектор-рядків матриці A, тобто $a^T Z(A)a > 0$, то

$$p(i) = Z(A)a)\left\|Z(A)a\right\|^{-2},$$

2) якщо вектор a лінійно залежний від векторів підпростору вектор-рядків матриці A, тобто $a^T Z(A)a = 0$, то

$$p(i) = R(A)a(1 + a^T R(A)a)^{-1}$$

Наслідок. Операція псевдообернення комутативна з транспонуванням, тому прямі формули Гревіля аналогічно виписуються для варіанту розширення матриці стовпчиком.

Прямі формули Гревіля для проекційних матриць. Використовуючи прямі формули Гревіля для представлення псевдооберненої матриці A^+_{ia} можна отримати формули для обчислення проекційних матриць $Z(A_{i,a})$, $Z(A^T_{i,a})$, зважених проекційних матриць $R(A_{i,a})$, $R(A^T_{i,a})$, які визначені формулами:

$$Z(A) = I_n - A^+ A,\; Z(A^T) = I_m - AA^+ \qquad (1)$$

$$R(A) = A^+(A^+)^T,\; R(A^T) = (A^+)^T A^+ . \qquad (2)$$

а) при $a^T Z(A)a > 0$ (вектор a лінійно незалежний)

$$Z(A^T_{i,a}) = \begin{pmatrix} Z_{11}(A^T) & 0(i-1) & Z_{12}(A^T) \\ 0^T(i-1) & 0 & 0^T(m+1-i) \\ Z_{21}(A^T) & 0(m+1-i) & Z_{22}(A^T) \end{pmatrix} \in R^{(m+1)\times(m+1)}, \tag{4.1}$$

де матриці – блоки

$$Z_{11}(A^T) \in R^{(i-1)\times(i-1)},\ Z_{21}(A^T) \in R^{(m-i+1)\times(i-1)},$$
$$Z_{12}(A^T) \in R^{(i-1)\times(m-i+1)},\ Z_{22}(A^T) \in R^{(m-i+1)\times(m-i+1)}$$

визначаются згідно представлення незбуреної проекційної матриці $Z(A^T)$:

$$Z(A^T) = \begin{pmatrix} Z_{11}(A^T) & Z_{12}(A^T) \\ Z_{21}(A^T) & Z_{22}(A^T) \end{pmatrix} \in R^{m\times m}, \tag{4.2}$$

а $0(k) = (0,0,...,0)^T \in R^k$ - нульовий вектор;

для зважених проекційних матриць:

$$R(A_{i,a}) = R(A) - (p(i)a^T R(A) - R(A)a\,p^T(i))\|p(i)\|^{-2} +$$
$$+ p(i)\,p^T(i)\,(1 + a^T R(A)a)^{-1} \in R^{n\times n};$$

$$R(A^T_{i,a}) = \begin{pmatrix} R_{11}(A^T) & c(i-1) & R_{12}(A^T) \\ c^T(i-1) & 1 & c^T(m+1-i) \\ R_{21}(A^T) & c(m+1-i) & R_{22}(A^T) \end{pmatrix} \in R^{(m+1)\times(m+1)},$$

де

$$R_{11}(A^T) \in R^{(i-1)\times(i-1)}, R_{21}(A^T) \in R^{(m-i+1)\times(i-1)},$$
$$R_{12}(A^T) \in R^{(i-1)\times(m-i+1)},\ R_{22}(A^T) \in R^{(m-i+1)\times(m-i+1)}$$

- блоки незбуреної зваженої проекційної матриці $R(A^T)$:

$$R(A^T) = \begin{pmatrix} R_{11}(A^T) & R_{12}(A^T) \\ R_{21}(A^T) & R_{22}(A^T) \end{pmatrix} \in R^{m\times m}, \tag{4.3}$$

а вектори $c(i-1),\quad c(m+1-i)$ є вектор-блоки вектора c::

$$c(i-1) = (c_1, c_2, ...c_{i-i})^T, \qquad c(m+1-i) = (c_i, c_{i+1}, ...c_m)^T$$

$$(c(i-1) \vdots c(m+1-i))^T \equiv (c_1, c_2, ...c_{i-i}, c_i, c_{i+1}, ...c_m)^T \equiv c = -A^{+^T} a;$$

б) при $a^T Z(A)a = 0$ (вектор a лінійно залежний)

$$Z(A_{i,a}) = Z(A) \in R^{n\times n} \ ,$$

$$Z(A_{ia}^T) = \begin{pmatrix} Z_{11}(A_{i,a}^T) & d(i-1) & Z_{12}(A_{i,a}^T) \\ d^T(i-1) & -1 & d^T(m+1-i) \\ Z_{21}(A_{i,a}^T) & d(m+1-i) & Z_{22}(A_{i,a}^T) \end{pmatrix} \in R^{(m+1)\times(m+1)} ,$$

де матриці-блоки $Z_{kl}(A_{i,a}^T)$ визначаються матричною рівністю:

$$\begin{pmatrix} Z_{11}(A_{i,a}^T) & Z_{12}(A_{i,a}^T) \\ Z_{21}(A_{i,a}^T) & Z_{22}(A_{i,a}^T) \end{pmatrix} = Z(A^T) + dd^T(1 + a^T R(A)a) \in R^{m\times m} ,$$

проекційна матриця $Z(A^T)$ визначається формулою (4.2), а вектори $d(i-1),\quad d(m+1-i)$ – з векторної рівності:

$$(d(i-1) \vdots d(m+1-i))^T \equiv (d_1,...,d_{i-i}, d_i, d_{i+1},...d_m)^T \equiv$$
$$\equiv -A^{+^T} a(1 + a^T R(A)a)^{-1}$$

для зважених проекційних матриць:

$$R(A_{i,a}) = R(A) + R(A)aa^T R(A)(1 + a^T R(A)a)^{-1} \in R^{n\times n} .$$

$$R(A_{i,a}^T) = \begin{pmatrix} R_{11}(A_{i,a}^T) & b(i-1) & R_{12}(A_{i,a}^T) \\ b^T(i-1) & 1 & b^T(m+1-i) \\ R_{21}(A_{i,a}^T) & b(m+1-i) & R_{22}(A_{i,a}^T) \end{pmatrix} \in R^{(m+1)\times(m+1)},$$

матриці – блоки $R_{kl}(A_{i,a}^T)$ визначається формулою:

$$\begin{pmatrix} R_{11}(A_{i,a}^T) & R_{12}(A_{i,a}^T) \\ R_{21}(A_{i,a}^T) & R_{22}(A_{i,a}^T) \end{pmatrix} = R(A^T) + A^{+^T} aa^T A^+ \|p(i)\|^2 -$$

$$- A^{+^T}(p(i)a^T + ap^T(i))A^+ \in R^{m\times m} ,$$

де зважена проекційна матриця $R(A^T)$ визначається формулою (4.3), а вектори $b(i-1),\quad b(m+1-i)$ - з векторної рівності:

$$(b(i-1) \vdots b(m+1-i))^T \equiv (b_1,...,b_{i-i}, b_i, b_{i+1},...b_m)^T \equiv b = A^{+^T} p(i) + \|p(i)\|^2 c ,$$

$$p(i) = R(A)a(1 + a^T R(A)a)^{-1},\ c = -A^{+^T} a .$$

4.2. Залежність псевдообернених і проекційних матриць від видалення довільних вектор-рядків з вихідної матриці

Якщо для матриці $A_{i,a} \in R^{(m+1)\times n}$ (після $(i-1)$ - го рядка $(i=\overline{2,m+1})$ матриці A стоїть вектор-рядок a^T), відома псевдообернена матриця $A^{+}_{i,a} \in R^{n\times(m+1)}$, то для визначення псевдооберненої матриці $A^{+} \in R^{n\times m}$ (рядок a^T з матриці $A_{i,a}$ видаляється) мають місце обернені формули Гревіля.

Обернені формули Гревіля для псевдообернених матриць.

1) У випадку лінійної залежності вектора a^T від векторів підпростору вектор-рядків матриці A, а це визначається умовою

$$a^T p(i) < 1$$

псевдообернена матриця $A^{+} \in R^{n\times m}$ має вид

$$A^{+} = \left(I_n + p(i)a^T(1-a^T p(i))^{-1}\right)P_i$$

При цьому ранг псевдооберненої матриці не змінюється, тобто $rank\ A = rank\left(A^T \vdots a\right)^T$.

2) У випадку лінійної незалежності вектора a^T від векторів підпростору вектор-рядків матриці A, а це визначається умовою

$$a^T p(i) = 1,$$

матриця $A^{+} \in R^{n\times m}$ визначається формулою

$$A^{+} = (I_n - p(i)p^T(i)\|p(i)\|^{-2})P_i.$$

При цьому ранг псевдооберненої матриці падає, тобто $rank\ A = rank\left(A^T \vdots a\right)^T - 1$.
Мають місце співвідношення

$$\|p(i)\|^2 = R_{ii}\left(A^T_{i,a}\right),\ 1 - p^T(i)a = Z_{ii}\left(A^T_{i,a}\right).$$

Обернені формули Гревіля для проекційних матриць.

Використовуючи обернені формули Гревіля для представлення псевдооберненої матриці $A^{+} \in R^{n \times m}$ при відомій псевдооберненій матриці $A^{+}_{i,a} \in R^{n \times (m+1)}$, можна отримати корисні для застосувань формули економного обчислення матриць $Z(A)$, $Z(A^T)$, $R(A)$, $R(A^T)$.

1) Якщо вектор-рядок a^T, який видаляється, від векторів підпростору вектор-рядків матриці A_{ia} лінійно залежний, тобто $a^T p(i) < 1$, то

$$Z(A) = Z(A_{i,a});$$

$$R(A) = R(A_{i,a}) + p(i) p^T(i) Z_{ii}^{-1}(A_{i,a}^T) \; ;$$

$$Z(A^T) = E_i Z(A_{i,a}^T) - E_i h_i h_i^T E_i^T Z_{ii}^{-1}(A_{i,a}^T) \; .$$

Тут розглядається матриця E_i – вона отримана вилученням i - ого $(i = \overline{2, m+1})$ вектор-рядка з одиничної матриці I_{m+1}:

$$E_i = \left(e_1(m+1), \ldots, e_{i-1}(m+1), e_{i+1}(m+1), \ldots, e_{m+1}(m+1) \right)^T \in R^{m \times (m+1)} .$$

Інші позначення визначаються формулами:

$$h_i = -A_{i,a}^{+\,T} a \, ;$$

$$R(A^T) = E_i R(A_{i,a}^T) E_i^T + E_i (h_i h_i^T (p^T(i) p(i)) + h_i t_i^T + t_i h_i^T) E_i^T ,$$

$$t_i = -A_{i,a}^{+\,T} p(i) Z_{ii}^{-1} (A_{i,a}^T) .$$

2) Якщо вектор-рядок a^T від векторів підпростору вектор-рядків матриці A_{ia} лінійно незалежний, тобто $a^T p(i) = 1$, то

$$Z(A) = Z(A_{i,a}) + p(i) p^T(i) \| p(i) \|^{-2} ,$$

$$R(A) = Z(p^T(i)) R(A_{i,a}) Z(p^T(i)) ,$$

$$Z(p^T(i)) = I_n - p(i) p^T(i) \| p(i) \|^{-2} .$$

З використанням позначення для матриці E_i

$$Z(A^T) = E_i Z(A_{i,a}^T) E_i^T ,$$

$$R(A^T) = E_i R(A_{i,a}^T) E_i^T .$$

4.3. Формули Гревіля- Кириченка: евклідові простори матриць

Задача поширення формул Гревіля-Кириченка розглядатиметься первісна для спряженого до «кортежного оператора» і полягатиме у рекурентному обчисленні через попередню розмірність, довжину кортежу, обчислення ПдО за розширенням кортежу додатковою матрицею. Зважаючи на те, що ПдО і спряженість комутують, та сама схема описує і аналогічну задачу для самого «кортежного оператора».

«Кортежний оператор» та спряжений $\wp_{\alpha,A}, \wp^*_{\alpha,A}$ визначаються співвідношеннями:

$$\wp_{\alpha,A}\begin{pmatrix} z \\ \beta \end{pmatrix} = \wp_\alpha \mathrm{M}A \begin{pmatrix} z \\ \beta \end{pmatrix} = \wp_\alpha z + \beta A = \sum_{k=1}^{K} z_k A_k + \beta A \quad : z \in R^K, \beta \in R^1,$$

$$z \in R^K, z_{K+1} \in R^1$$

$$\wp^*_{\alpha,A} X = \begin{pmatrix} (A_1, X) \\ \vdots \\ (A_K, X) \\ (A, X) \end{pmatrix} = \begin{pmatrix} l_{A_1} \\ \vdots \\ l_{A_m} \\ l_A \end{pmatrix} X = \begin{pmatrix} l_{A_1} X \\ \vdots \\ l_{A_m} X \\ l_A X \end{pmatrix} : R^{m\times n} \to R^{K+1}, \tag{4.4}$$

де $l_A, A \in R^{m\times n}$ позначає лінійний функціонал, що природним чином визначається скалярним добутком: $l_A X = (A,X), X, A \in R^{m\times n}$.

Зауваження 3. Очевидним чином, «кортежний оператор», що визначений у другому розділі $\wp^*_{\alpha,A}$ є спряженим до «кортежного оператора»:

$$\wp_{\alpha,A} : \ \wp_{\alpha,A}\ ^* = \wp^*_{\alpha,A}. \tag{4.5}$$

Обидва оператори визначаються за набором елементів $A_k \in R^{m\times n}, k = \overline{1,K}$, що очевидним чином визначають набори лінійних функціоналів $l_{A_k} X = (A_k, X), k = \overline{1,K}, X \in R^{m\times n}$ чи $l_A X = (A,X), X \in R^{m\times n}$ відповідно.

Зауваження 4. Спряжений до оператора $Ql_A : R^{m\times n} \to R^{m\times n}, Q, A \in R^{m\times n}$, дія якого визначається як $Ql_A X = Q\ A,X\ , X \in R^{m\times n}$, задається співвідношенням:

$Ql_A\ ^* = Al_Q$, тобто для довільного $Y \in R^{m\times n}$ $\ Ql_A\ ^* Y = A\ Q,Y\ = Al_Q$.

Загалом, в прямих формулах Гревіля - Кириченка, власне, мова йде про рекурентні формули, що визначають псевдообернення розширеного оператора через псевдообернення вихідного оператора для операторів-матриць між евклідовими просторами числових векторів. У цьому випадку теорема показує, як псевдообернення матриці, доповненої рядком, виражається через псевдообернену для вихідної матриці. Нижче цей результат переноситься на «кортежні оператори». Як і випадку матриць лінійних операторів вигляд рекурентної залежності визначається тим, чи є лінійний функціонал, яким розширюється оператор, лінійно залежним від решти функціоналів, що формують «кортежний оператор». Так само, умова лінійної залежності конструктивно описується в термінах слушних операторів ортогонального проектування, що будується за SVD - розкладом КорО .

Так само, як і для просторів числових векторів, з ПдО КорО асоціюються дві пари КорО: ортогональні проектори та групуючі оператори (зважені проекційні) із тими самим способом використання.

Крім того, у зв'язку із вихідною матрицею та ПдО визначається дві пари асоційованих з ПдО операторів:

- ортогональні проектори:

$Z(\wp_\alpha)=E_{m\times n}-\wp_\alpha^*\wp_\alpha$, $Z(\wp_\alpha^*)=E_m-\left(\wp_\alpha^*\right)^+\wp_\alpha^*\overset{\text{теорема}}{=}E_m-\wp_\alpha\wp_\alpha^+$, $E_{m\times n}$ - тотожний в матричному просторі, з необхідними і достатніми умовами лінійної залежності-незалежності доданої матриці від решти матриць-елементів кортежу: $A,Z(\wp_\alpha)A=0$ - лінійна залежність, $A,Z(\wp_\alpha)A>0$ - лінійна незалежність;

- групуючи оператори

$$R(\wp_\alpha)=\wp_\alpha^+\wp_\alpha^{+*}\overset{\text{теорема}}{=}\left(\wp_\alpha^*\wp_\alpha\right)^+,\quad R(\wp_\alpha^*)=\left(\wp_\alpha^*\right)^+\left(\left(\wp_\alpha^*\right)^+\right)^T\overset{\text{теорема}}{=}\left(\wp_\alpha\wp_\alpha^*\right)^+.$$

Загалом, формулювання результату має вигляд наступної теореми.

Теорема 4.1 Псевдообернення розширеного КорО $\wp_{\alpha,A}^*$ за лінійної незалежності чи залежності A від елементів $A_1,\ldots,A_K$ кортежу $\alpha=(A_1,\ldots,A_K)$ визначається співвідношенням

$$\begin{pmatrix} \wp_\alpha^* \\ l_A \end{pmatrix}^+ = \left(E_{m\times n} - Q l_A\right)\left(\wp_\alpha^{*+} \mid Q\right), \qquad Q = \begin{cases} \dfrac{Z(\wp_\alpha^*) l_A}{\left(A, Z(\wp_\alpha^*)A\right)} A & \text{лін. нез.} \Leftrightarrow \left(A, Z(\wp_\alpha^*)A\right) > 0 \\ \dfrac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)} & \text{лін.зал.} \Leftrightarrow \left(A, Z(\wp_\alpha^*)A\right) = 0 \end{cases}, \tag{4.6}$$

Доведення теореми доцільно розбити на дві теореми, що формулюються і доводяться нижче: для незалежного від інших елементів розширення кортежу, та - за його залежності.

Теорема 4.2 Псевдообернення розширеного КорО $\wp_{\alpha,A}$ за лінійної залежності A від елементів $A_1,\ldots,A_K$ кортежу $\alpha = (A_1,\ldots,A_K)$ визначається співвідношенням

$$(\wp_\alpha \mid A)^+ Y = \begin{pmatrix} \wp_\alpha^+ \left(E_{m\times n} - A\dfrac{\left(Y, R(\wp_\alpha^*)A\right)}{1+(A,R(\wp_\alpha^*)A)}\right) \\ \dfrac{\left(Y, R(\wp_\alpha^*)A\right)}{1+(A,R(\wp_\alpha^*)A)} \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ \left(E_{m\times n} - A l_{\frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}} Y\right) \\ l_{\frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}} Y \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ \left(E_{m\times n} - A l_Q\right) \\ l_Q \end{pmatrix} Y \quad \forall Y \in R^{m\times n} \tag{4.7}$$

де

$$Q = \frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}. \tag{4.8}$$

Еквівалентним чином:

$$(\wp_\alpha \mid A)^+ = \begin{pmatrix} \wp_\alpha^+ \left(E_{m\times n} - A l_Q\right) \\ l_Q \end{pmatrix}. \tag{4.9}$$

Доведення. За лінійної залежності останнього рядку розширеної матриці від решти рядків $\left(\left(A, Z(\wp_\alpha^*)A\right) = 0\right)$ справедливість теореми доводиться через побудову найменшого за нормою розв’язку СЛАР за кортежем $(\wp_\alpha \mid A)\begin{pmatrix} z \\ \beta \end{pmatrix} = Y, \ Y \in R^{m\times n}$. У цьому випадку розглядається СЛАР (розширена)

$$(\wp_\alpha \mid A)\begin{pmatrix} z \\ \beta \end{pmatrix} = Y : z \in R^m, \beta \in R^1, Y \in R^{m\times n} \Leftrightarrow \wp_\alpha z + A\beta = Y. \tag{4.10}$$

Розглядається за додаткової умови $A,Z(\wp_\alpha^*)A = 0$, яка еквівалентна умові, що $A \in L_{\wp_\alpha}$. Останнє означає, що існує лінійна комбінація елементів кортежу $\alpha = A_1,...,A_m$, що визначає $\wp_\alpha$ з коефіцієнтами, що утворюють вектор $z_0 : z_0 \in R^n$, - яка дає матрицю A. Це означає, що виконується рівність:

$$\wp_\alpha z_0 = A. \tag{4.11}$$

Будемо вважати, що z_0 вибране серед усіх можливих наборів коефіцієнтів, які представляють матрицю A, таким чином, що має мінімальну можливу норму. Це означатиме, що

$$z_0 = \wp_\alpha^+ A. \tag{4.12}$$

Повертаючись до розширеної СЛАР (4.10) і підставивши в неї представлення матриці A через z_0, отримуємо наступне представлення досліджуваної СЛАР: з

$$\wp_\alpha z + \beta A = Y$$

маємо

$$\wp_\alpha z + \beta \wp_\alpha z_0 = Y.$$

Остання СЛАР є еквівалентною наступній:

$$\wp_\alpha (z + \beta z_0) = Y.$$

Очевидним чином, вона має розв'язки щодо $z + \beta z_0$, серед яких мінімальний за нормою має вигляд:

$$z + z_0 \beta = \wp_\alpha^+ Y. \tag{4.13}$$

Підставляючи (4.13) вираз для z_0 з (4.12), отримуємо:

$$z + \beta \wp_\alpha^+ A = \wp_\alpha^+ Y. \tag{4.14}$$

Очевидним і єдиним розв'язком останньої СЛАР з мінімальною нормою, залежною від $\beta \in R^1$, є значення що визначається співвідношенням

$$z_\beta = \wp_\alpha^+ (Y - \beta A). \tag{4.15}$$

Вихідна задача побудови розв'язку $\begin{pmatrix} z \\ \beta \end{pmatrix}$ розширеної СЛАР з мінімальною нормою, співвідношенням (4.10), зводиться до умовної оптимізації з цим обмеженням за двома аргументами z, β. Вона може бути, зокрема, розв'язана через вираз одного із аргументів - z - через інший аргумент - β - з наступною мінімізацією $\left\| \begin{pmatrix} z \\ \beta \end{pmatrix} \right\|^2$ за β. Співвідношення (4.7) дає такий вираз. Отже, задача умовної оптимізації за $z \in R^m, \beta \in R^1$ зводиться до безумовної задачі за $\beta \in R^1$ для $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$.

Побудова для кожного β, z_β з мінімальною нормою за обмеження (4.10) та наступний вибір β так, щоб мінімізувати норму $\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}$, розв'яже задачу отримання розв'язку розширеної СЛАР з мінімальною нормою.

Лема. Значення $\hat{\beta}$, що мінімізує $\left\| \begin{pmatrix} z_\alpha \\ \beta \end{pmatrix} \right\|$ чи $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$ визначається співвідношенням:

$$\hat{\beta} = \frac{Y, R(\wp_\alpha^*) A}{1 + \ A, R(\wp_\alpha^*) A} = l_Q Y : \qquad Q = \frac{R(\wp_\alpha^*) A}{1 + \ A, R(\wp_\alpha^*) A}, \qquad R(\wp_\alpha^*) = \wp_\alpha^{*+} \wp_\alpha^+ .$$

Доведення. Оптимальне значення $\hat{\beta}$: може бути визначене з рівняння, що визначає необхідну умову екстремуму:

$$\frac{d}{d\beta} \left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2 = 0. \tag{4.16}$$

З урахуванням (4.7), для що визначає z_β, для $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$ маємо:

$$\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2 = || z_\beta ||^2 + \beta^2 = || \wp_\alpha^+ (Y - \beta A) ||^2 + \beta^2 =$$

$$= \ \wp_\alpha^+ (Y - \beta A), \wp_\alpha^+ (Y - \beta A) \ + \beta^2 = \ Y - \beta A, \wp_\alpha^{*+} \wp_\alpha^+ (Y - \beta A) \ + \beta^2 = \ Y - \beta A, R(\wp_\alpha^*)(Y - \beta A) \ + \beta^2 =$$

$$=\beta^2 \; A,R(\wp_\alpha^*)A \; -2\beta \; Y,R(\wp_\alpha^*)A \; + \; Y,R(\wp_\alpha^*)Y \; +\beta^2 =$$

$$=\beta^2\left[1+ \; A,R(\wp_\alpha^*)A \;\right]-2\beta \; Y,R(\wp_\alpha^*)A \; + \; Y,R(\wp_\alpha^*)Y \; .$$

Отже, остаточно

$$\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 =\beta^2\left[1+ \; A,R(\wp_\alpha^*)A \;\right]-2\beta \; Y,R(\wp_\alpha^*)A \; + \; Y,R(\wp_\alpha^*)Y \; .$$

Останній вираз є квадратним тричленом з додатнім старшим коефіцієнтом, тому точкою екстремуму є точка мінімуму, яка може бути визначена із необхідної умови екстремуму.

Оскільки градієнт має вигляд

$$grad_\beta\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 = grad_\beta \; \beta^2\left[1+ \; A,R(\wp_\alpha^*)A \;\right]-2\beta \; Y,R(\wp_\alpha^*)A \; + \; Y,R(\wp_\alpha^*)Y \; =$$

$$=2\beta\left[1+ \; A,R(\wp_\alpha^*)A \;\right]-2 \; Y,R(\wp_\alpha^*)A \; ,$$

рівняння – необхідна умова екстремуму має вигляд

$$2\beta\left[1+ \; A,R(\wp_\alpha^*)A \;\right]-2 \; Y,R(\wp_\alpha^*)A \; =0$$

чи еквівалентним чином –

$$\beta\left[1+ \; A,R(\wp_\alpha^*)A \;\right]= \; Y,R(\wp_\alpha^*)A$$

Розв’язок $\hat{\beta}$ останнього рівняння - точка мінімуму для $\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2$ - задається співвідношенням

$$\hat{\beta}=\frac{Y,R(\wp_\alpha^*)A}{A,R(\wp_\alpha^*)A \; +1}=\frac{Y,R(\wp_\alpha^*)A}{1+ \; A,R(\wp_\alpha^*)A}=l_Q Y, \qquad Q=\frac{R(\wp_\alpha^*)A}{1+ \; A,R(\wp_\alpha^*)A}.$$

що й доводить лему.

Наслідок. Розв'язком розширеної СЛАР з мінімальною нормою є вектор

$$\begin{pmatrix} z_{\hat{\beta}} \\ \hat{\beta} \end{pmatrix}=\begin{pmatrix} \wp_\alpha^+(Y-\hat{\beta}A) \\ \hat{\beta} \end{pmatrix}=\begin{pmatrix} \wp_\alpha^+ \; E-Al_Q Y \\ l_Q Y \end{pmatrix}==\begin{pmatrix} \wp_\alpha^+ \; E-Al_Q \\ l_Q \end{pmatrix}Y .$$

Доведення. Дійсно,

$$\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 == \begin{pmatrix} \wp_\alpha^+(Y-\hat\beta A) \\ \hat\beta \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+\left(Y - A\dfrac{Y,R(\wp_\alpha^*)A}{1+\ A,R(\wp_\alpha^*)A}\right) \\ \dfrac{Y,R(\wp_\alpha^*)A}{1+\ A,R(\wp_\alpha^*)A} \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+\left(Y - Al_{\frac{R(\wp_\alpha^*)}{1+\ A,R(\wp_\alpha^*)A}}Y\right) \\ l_{\frac{R(\wp_\alpha^*)}{1+\ A,R(\wp_\alpha^*)A}}Y \end{pmatrix} =$$

$$== \begin{pmatrix} \wp_\alpha^+\ E_{m\times n} - Al_Q \\ l_Q \end{pmatrix} Y\ ,\ l_Q Y = \frac{Y,R(\wp_\alpha^*)A}{1+\ A,R(\wp_\alpha^*)A}.$$

У відповідності до визначення за Пенроузом, матриця, що визначає найменший за нормою розв'язок як функцію від правої частини СЛАР і є псевдоберненою.

Наступна теорема є очевидним наслідком попередньої .

Теорема 4.3 Псевдообернення розширеного КорО $\wp_{\alpha,A}$ за лінійної залежності A від елементів $A_1,...,A_K$ кортежу $\alpha=(A_1,...,A_K)$ визначається співвідношенням

$$\wp_\alpha \mathrm{M}A^+ Y = \begin{pmatrix} \wp_\alpha^+\left(E_{m\times n} - A\dfrac{Y,R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}\right) \\ \dfrac{Y,R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)} \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+\left(E_{m\times n} - Al_{\frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}}Y\right) \\ l_{\frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}}Y \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+\ E_{m\times n} - Al_Q \\ l_Q \end{pmatrix} Y\ , \quad (4.17)$$

$Y \in R^{m\times n}$,

де
$$Q = \frac{R(\wp_\alpha^*)A}{1+(A,R(\wp_\alpha^*)A)}\ . \quad (4.18)$$

Еквівалентним чином:

$$\wp_\alpha \mathrm{M}A^+ = \begin{pmatrix} \wp_\alpha^+\ E_{m\times n} - Al_Q \\ l_Q \end{pmatrix}. \quad (4.19)$$

Доведення. За лінійної залежності останнього рядку розширеної матриці від решти рядків ($A,Z(\wp_\alpha^*)A\ = 0$) справедливість теореми доводиться через побудову найменшого за нормою розв'язку СЛАР за кортежем $(\wp_\alpha MA)\begin{pmatrix} z \\ \beta \end{pmatrix} = Y,\ \ Y \in R^{m\times n}$. У цьому випадку розглядається СЛАР (розширена)

$$\wp_\alpha MA \begin{pmatrix} z \\ \beta \end{pmatrix} = Y : z \in R^m, \beta \in R^1, Y \in R^{m\times n} \Leftrightarrow \wp_\alpha z + A\beta = Y. \tag{4.20}$$

Розглядається за додаткової умови $A, Z(\wp_\alpha^*)A = 0$, яка еквівалентна умові, що $A \in L_{\wp_\alpha}$. Останнє означає, що існує лінійна комбінація елементів кортежу $\alpha = A_1, \ldots, A_m$, що визначає $\wp_\alpha$ з коефіцієнтами, що утворюють вектор $z_0 : z_0 \in R^n$, - яка дає матрицю A. Це означає, що виконується рівність:

$$\wp_\alpha z_0 = A. \tag{4.21}$$

Будемо вважати, що z_0 вибране серед всіх можливих наборів коефіцієнтів, що представляють матрицю A, таким чином, що має мінімальну можливу норму. Це означатиме, що

$$z_0 = \wp_\alpha^+ A. \tag{4.22}$$

Повертаючись до розширеної СЛАР і підставивши в неї представлення матриці A через z_0, отримуємо наступне представлення досліджуваної СЛАР: з

$$\wp_\alpha z + \beta A = Y$$

маємо

$$\wp_\alpha z + \beta \wp_\alpha z_0 = Y.$$

Остання СЛАР є еквівалентною наступній:

$$\wp_\alpha (z + \beta z_0) = Y.$$

Очевидним чином, вона має розв’язки щодо $z + \beta z_0$, серед яких мінімальний за нормою має вигляд:

$$z + z_0 \beta = \wp_\alpha^+ Y. \tag{4.23}$$

Підставляючи (4.23) вираз для z_0 з (4.22), отримуємо:

$$z + \beta \wp_\alpha^+ A = \wp_\alpha^+ Y. \tag{4.24}$$

Очевидним і єдиним розв’язком останньої СЛАР з мінімальною нормою, залежною від $\beta \in R^1$, є значення що визначається співвідношенням

$$z_\beta = \wp_\alpha^+ (Y - \beta A). \tag{4.25}$$

Вихідна задача побудови розв'язку $\begin{pmatrix} z \\ \beta \end{pmatrix}$ розширеної СЛАР з мінімальною нормою, співвідношенням (4.17), зводиться до умовної оптимізації з цим обмеженням за двома аргументами z, β. Вона може бути, зокрема, розв’язана через вираз одного із аргументів - z - через інший аргумент - β - з наступною мінімізацією $\left\| \begin{pmatrix} z \\ \beta \end{pmatrix} \right\|^2$ за β. Співвідношення (4.17) дає такий вираз. Отже, задача умовної оптимізації за $z \in R^m, \beta \in R^1$ зводиться до безумовної задачі за $\beta \in R^1$ для $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$.

Побудова для кожного β, z_β з мінімальною нормою за обмеженням (4.21) та наступний вибір β так, щоб мінімізувати норму $\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}$, розв'яже задачу отримання розв'язку розширеної СЛАР з мінімальною нормою.

Лема. Значення $\hat{\beta}$, що мінімізує $\left\| \begin{pmatrix} z_\alpha \\ \beta \end{pmatrix} \right\|$ чи $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$ визначається співвідношенням:

$$\hat{\beta} = \frac{Y, R(\wp_\alpha^*) A}{1 + A, R(\wp_\alpha^*) A} = l_Q Y: \qquad Q = \frac{R(\wp_\alpha^*) A}{1 + A, R(\wp_\alpha^*) A}, \qquad R(\wp_\alpha^*) = \wp_\alpha^{*+} \wp_\alpha^+.$$

Доведення. Оптимальне значення $\hat{\beta}$: може бути визначене з рівняння, що визначає необхідну умову екстремуму:

$$\frac{d}{d\beta} \left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2 = 0. \tag{4.26}$$

З урахуванням (4.17), для що визначає z_β, для $\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2$ маємо:

$$\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 = || z_\beta ||^2 + \beta^2 = || \wp_\alpha^{+}(Y - \beta A) ||^2 + \beta^2 =$$

$$= \ \wp_\alpha^{+}(Y-\beta A), \wp_\alpha^{+}(Y-\beta A) \ + \beta^2 = \ Y - \beta A, \wp_\alpha^{*+} \wp_\alpha^{+}(Y-\beta A) \ + \beta^2 = \ Y-\beta A, R(\wp_\alpha^{*})(Y-\beta A) \ + \beta^2 =$$

$$= \beta^2 \ A, R(\wp_\alpha^{*})A \ - 2\beta \ Y, R(\wp_\alpha^{*})A \ + \ Y, R(\wp_\alpha^{*})Y \ + \beta^2 =$$

$$= \beta^2 \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] - 2\beta \ Y, R(\wp_\alpha^{*})A \ + \ Y, R(\wp_\alpha^{*})Y \ .$$

Отже, остаточно

$$\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 = \beta^2 \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] - 2\beta \ Y, R(\wp_\alpha^{*})A \ + \ Y, R(\wp_\alpha^{*})Y \ .$$

Останній вираз є квадратним тричленом з додатнім старшим коефіцієнтом, тому точкою екстремуму є точка мінімуму, яка може бути визначена із необхідної умови екстремуму.

Оскільки градієнт має вигляд

$$grad_\beta \left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2 = grad_\beta \ \beta^2 \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] - 2\beta \ Y, R(\wp_\alpha^{*})A \ + \ Y, R(\wp_\alpha^{*})Y \ =$$

$$= 2\beta \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] - 2 \ Y, R(\wp_\alpha^{*})A \ ,$$

рівняння – необхідна умова екстремуму має вигляд

$$2\beta \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] - 2 \ Y, R(\wp_\alpha^{*})A \ = 0$$

чи еквівалентним чином –

$$\beta \left[1 + \ A, R(\wp_\alpha^{*})A \ \right] = \ Y, R(\wp_\alpha^{*})A$$

Розв’язок $\hat{\beta}$ останнього рівняння - точка мінімуму для $\left\|\begin{pmatrix} z_\beta \\ \beta \end{pmatrix}\right\|^2$ - задається співвідношенням

$$\hat{\beta} = \frac{Y, R(\wp_\alpha^{*})A}{A, R(\wp_\alpha^{*})A \ + 1} = \frac{Y, R(\wp_\alpha^{*})A}{1 + \ A, R(\wp_\alpha^{*})A} = l_Q Y, \qquad Q = \frac{R(\wp_\alpha^{*})A}{1 + \ A, R(\wp_\alpha^{*})A} .$$

що й доводить лему.

Наслідок. Розв'язком розширеної СЛАР з мінімальною нормою є вектор

$$\begin{pmatrix} z_{\hat\beta} \\ \hat\beta \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ (Y - \hat\beta A) \\ \hat\beta \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ \quad E - Al_Q Y \\ l_Q Y \end{pmatrix} == \begin{pmatrix} \wp_\alpha^+ \quad E - Al_Q \\ l_Q \end{pmatrix} Y .$$

Доведення. Дійсно,

$$\left\| \begin{pmatrix} z_\beta \\ \beta \end{pmatrix} \right\|^2 == \begin{pmatrix} \wp_\alpha^+ (Y - \hat\beta A) \\ \hat\beta \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ \left(Y - A \dfrac{Y, R(\wp_\alpha^*) A}{1 + A, R(\wp_\alpha^*) A} \right) \\ \dfrac{Y, R(\wp_\alpha^*) A}{1 + A, R(\wp_\alpha^*) A} \end{pmatrix} = \begin{pmatrix} \wp_\alpha^+ \left(Y - Al_{\frac{R(\wp_\alpha^*)}{1 + A, R(\wp_\alpha^*) A}} Y \right) \\ l_{\frac{R(\wp_\alpha^*)}{1 + A, R(\wp_\alpha^*) A}} Y \end{pmatrix} =$$

$$== \begin{pmatrix} \wp_\alpha^+ \quad E_{m\times n} - Al_Q \\ l_Q \end{pmatrix} Y , \quad l_Q Y = \frac{Y, R(\wp_\alpha^*) A}{1 + A, R(\wp_\alpha^*) A} .$$

У відповідності до визначення за Пенроузом, матриця, що визначає найменший за нормою розв'язок як функцію від правої частини СЛАР і є псевдооберненою.

Наступна теорема є очевидним наслідком попередньої .

Теорема 4.4 Псевдообернення розширеного КорО $\wp_{\alpha,A}^*$ за лінійної незалежності A від елементів $A_1,...,A_K$ кортежу $\alpha = (A_1,...,A_K)$ визначається співвідношенням
$$\begin{pmatrix} \wp_\alpha^* \\ l_A \end{pmatrix}^+ = E_{m\times n} - Ql_A \quad \wp_\alpha^{*+} \vdots Q \quad , Q = \frac{Z(\wp_\alpha^*) A}{A, Z(\wp_\alpha^*) A} , \tag{4.27}$$

де
$$Q = \frac{Z(\wp_\alpha^*) A}{A, Z(\wp_\alpha^*) A} . \tag{4.28}$$

Таким чином,для $\begin{pmatrix} z \\ \beta \end{pmatrix} \in R^{K+1} : z \in R^K , \beta \in R^1$

$$\begin{pmatrix} \wp_\alpha^* \\ l_A \end{pmatrix}^+ \begin{pmatrix} z \\ \beta \end{pmatrix} = E_{m\times n} - Ql_A \quad \wp_\alpha^{*+} \vdots Q \begin{pmatrix} z \\ \beta \end{pmatrix} = E_{m\times n} - Ql_A \quad \wp_\alpha^{*+} z + Q\beta .$$

Доведення. Дійсно, для

$$\begin{pmatrix} \wp_{\alpha}^{*} \\ l_A \end{pmatrix} X = \begin{pmatrix} y \\ y' \end{pmatrix}, X \in R^{m\times n}, y \in R^{m}, y' \in R^{1},$$

чи

$$\begin{pmatrix} \wp_{\alpha}^{*} \\ l_A \end{pmatrix} X = \begin{pmatrix} y \\ y' \end{pmatrix} \Leftrightarrow \begin{cases} \wp_{\alpha}^{*} X = y \\ (A, X) = y' \end{cases}.$$

позначимо множину розв’язків останньої СЛАР через $\Omega_{\wp_{\alpha}^{*},A}$.

Тоді

$$\Omega_{\wp_{\alpha}^{*},A} = \{X : \wp_{\alpha}^{*} X = y\} \cap \{X : (A, X) = y'\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, V \in R^{m\times n}\} \cap \{X : (A, X) = y'\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, (A, \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V) = y', V \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, (A, \wp_{\alpha}^{*+} y) + (A, Z(\wp_{\alpha}^{*})V) = y', V \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, (A, Z(\wp_{\alpha}^{*})V) = y' - (A, \wp_{\alpha}^{*+} y), V \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, (Z(\wp_{\alpha}^{*})A, V) = y' - (A, \wp_{\alpha}^{*+} y), V \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, l_{Z(\wp_{\alpha}^{*})A} V = y' - (A, \wp_{\alpha}^{*+} y), V \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, V = l^{+}_{Z(\wp_{\alpha}^{*})A} \left(y' - (A, \wp_{\alpha}^{*+} y) \right) + Z\left(l_{Z(\wp_{\alpha}^{*})A} \right) W, W \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, V = l^{+}_{Z(\wp_{\alpha}^{*})A} \left(y' - (A, \wp_{\alpha}^{*+} y) \right) +$$

$$+ \left(W - l^{+}_{Z(\wp_{\alpha}^{*})A} l_{Z(\wp_{\alpha}^{*})A} W \right), W \in R^{m\times n}\} =$$

$$= \{X : X = \wp_{\alpha}^{*+} y + Z(\wp_{\alpha}^{*})V, V = \frac{Z(\wp_{\alpha}^{*})A}{\| Z(\wp_{\alpha}^{*})A \|^{2}} \left(y' - (A, \wp_{\alpha}^{*+} y) \right) +$$

$$+ \left(E_{m\times n} - \frac{Z(\wp_{\alpha}^{*})A}{\| Z(\wp_{\alpha}^{*})A \|^{2}} l_{Z(\wp_{\alpha}^{*})A} \right) W, W \in R^{m\times n}\} =$$

$$=\left\{X: X=\wp_\alpha^{*+}y+Z(\wp_\alpha^*)V, V=\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}\ \ y'-(A,\wp_\alpha^{*+}y)\ \ +\right.$$

$$\left.+\left(W-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W)\right), W\in R^{m\times n}\right\}=$$

$$=\left\{X: X=\wp_\alpha^{*+}y+Z(\wp_\alpha^*)\left\{\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}\ \ y'-(A,\wp_\alpha^{*+}y)\ \ +\right.\right.$$

$$\left.\left.+\left(W-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W)\right)\right\}, W\in R^{m\times n}\right\}=$$

$$=\left\{X: X=\wp_\alpha^{*+}y+\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}\ \ y'-(A,\wp_\alpha^{*+}y)\ \ +\right.$$

$$\left.+Z(\wp_\alpha^*)\left(W-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W)\right), W\in R^{m\times n}\right\}=$$

$$=\left\{X: X=\wp_\alpha^{*+}y+\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}\ \ y'-(A,\wp_\alpha^{*+}y)\ \ +\right.$$

$$\left.+\left(Z(\wp_\alpha^*)W-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W)\right), W\in R^{m\times n}\right\}=$$

$$=\left\{X: X=\wp_\alpha^{*+}y+\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}y'-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(A,\wp_\alpha^{*+}y)+\right.$$

$$\left.+\left(Z(\wp_\alpha^*)W-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W)\right), W\in R^{m\times n}\right\}=$$

$$=\left\{X: X=\wp_\alpha^{*+}y-\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(A,\wp_\alpha^{*+}y)+\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}y'+\right.$$

$$+\left(Z(\wp_\alpha^*)W - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W) \right), W \in R^{m\times n} \right\} =$$

$$= \left\{ X : X = \wp_\alpha^{*+} y - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} l_A \wp_\alpha^{*+} y + \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} y' + \right.$$

$$+\left(Z(\wp_\alpha^*)W - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W) \right), W \in R^{m\times n} \right\} =$$

$$= \left\{ X : X = \left(E_{m\times n} - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} l_A \right) \wp_\alpha^{*+} y + \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} y' + \right.$$

$$+\left(Z(\wp_\alpha^*)W - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W) \right), W \in R^{m\times n} \right\} =$$

$$= \left\{ X : X = \left(\underbrace{\left(E_{m\times n} - \underbrace{\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} l_A}_{Ql_A} \right) \wp_\alpha^{*+}}_{\Pi} \vdots \underbrace{\frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}}_{Q} \right) \begin{pmatrix} y \\ y' \end{pmatrix} + \right.$$

$$+\left(Z(\wp_\alpha^*)W - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W) \right), W \in R^{m\times n} \right\} =$$

$$= \left\{ X : X = \left(\Pi, Q \right) \begin{pmatrix} y \\ y' \end{pmatrix} + \right.$$

$$+\left(Z(\wp_\alpha^*)W - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2}(Z(\wp_\alpha^*)A,W) \right), W \in R^{m\times n} \right\},$$

$$\Pi = \left(E_{m\times n} - \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2} l_A \right) \wp_\alpha^{*+} = \left(E_{m\times n} - Ql_A \right) \wp_\alpha^{*+}, \qquad Q = \frac{Z(\wp_\alpha^*)A}{\|Z(\wp_\alpha^*)A\|^2},$$

що й доводить теорему.

Висновки до розділу 4.

Загалом, у прямих формулах Гревіля-Кириченка, власне, мова йде про рекурентні формули, що визначають псевдообернення розширеного оператора через псевдообернення вихідного оператора для операторів-матриць між евклідовими просторами числових векторів. У цьому випадку теорема вказує, як псевдообернення матриці, доповненої рядком, виражається через псевдообернену для вихідної матриці. Нижче цей результат переноситься на «кортежні оператори». Як і у випадку матриць лінійних операторів вигляд рекурентної залежності визначається тим, чи є лінійний функціонал, яким розширюється оператор, лінійно залежним від решти функціоналів, що формують «кортежний оператор». Так само, умова лінійної залежності конструктивно описується в термінах слушних операторів ортогонального проектування, що будується за SVD-розкладом КорО .

Так само, як і для просторів числових векторів, з ПдО КорО асоціюються дві пари КорО: ортогональні проектори та групуючі оператори (зважені проекційні) із тими самим способом використання.

ВИСНОВКИ

Дисертація є завершеним науковим дослідженням, в якому розв`язується важлива наукова проблема – побудова математичних засобів групування інформації для евклідових просторів матриць фіксованої розмірності. У роботі запропонований розвиток математичного апарату псевдообернення за Муром-Пенроузом на основі розвитку теорії псевдообернення для евклідових просторів матриць фіксованої розмірності. Запропоновано реалізацію математичного апарату дослідження базових структур в евклідовому просторі матриць фіксованої розмірності через реалізацію спеціального класу операторів названих в роботі кортежними. Отримані результати дозволяють реалізувати методи, моделі та алгоритми кластеризації, загалом групування інформації для важливих класів задач з матричними представниками. При цьому отримані наступні нові наукові та практичні результати:

- реалізовано концепцію «кортежних операторів» для матричних кортежів, що дало можливість побудувати конструктивну теорію SVD (сингулярний розклад) та псевдообернення для таких операторів;
- вичерпно та конструктивно досліджено систему лінійних алгебраїчних рівнянь для кортежних операторів та спряжених до них;
- сформульовано та розв’язано задачу лінійної робастної дискримінації в матричних евклідових просторах;
- сформульовано та розв'язано задачу рекурентного обчислення псевдобернення для «кортежних операторів» та спряжених до них за розширення кортежів новими матричними елементами для матриць: побудовано аналоги прямих формул Ґревіля-Кириченка;
- удосконалено засоби дослідження лінійних та нелінійних базових структур в матричних евклідових просторах;
- побудовано низку методів розв'язання задач групування інформації для представників з матричних евклідових просторів.

Результати наукового дослідження методів та алгоритмів псевдообернення в евклідовому просторі матриць для розв’язання прикладних задач у сучасних інформаційних системах, отримані в рамках підготовки кандидатської дисертації, впроваджені у 2013-2014 н.р. у навчальний процес кафедри системного аналізу та теорії прийняття рішень факультету комп’ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка при викладанні курсу “Обробка інформації в умовах невизначеності” для студентів 1 курсу магістратури за спеціальністю “Системи і методи прийняття рішень”, а також застосовано при виконанні науково-дослідної теми №11БФ015-06 “Проблеми теорії прийняття рішень та системного аналізу стохастичних мереж”.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових фахових виданнях

1. Donchenko V. S. Development of pseudo inverse techniques: cortege operators and applications in grouping information problem / V. S. Donchenko, F. M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2014. – No. 1. – С. 85–90.

2. Donchenko V.S. Cortege operators in grouping information problem: linear discrimination of matrix ‘Feature Vectors’ / V. S. Donchenko, F. M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2014. – No. 3. – С. 101–104.

3. Donchenko V. S. Direct formulas Greevil – Kirichenko for matrix Euclidean spaces / V. S. Donchenko, F. M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2015. – No. 1. – С. 119–123.

4. Скотаренко Ф. Алгоритм лінійної дискримінації для матричних просторів та його основи / Ф.М Скотаренко // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2015. – No 2. – С. 121–127.

Статті [1-4] опубліковані у виданні України, яке включене до міжнародної наукометричної бази ВІНІТІ РАН, української реферативної бази даних «Україніка наукова» та українського реферативного журналу «Джерело».

5. Donchenko, V. S. "'Feature vectors' in grouping information problem in applied mathematics: vectors and matrices" / V. S. Donchenko, T. P. Zinko, F. M. Skotarenko // Problems of Computer Intellectualization, Kyiv, Ukraine, ITHEA.– 2012. – P. 111–124.

6. Donchenko V. S. Spreading the Moore-Penrose pseudo inverse on matrixes Euclidean spaces: theory and applications / V. S. Donchenko, T. P. Zinko, F. M. Skotarenko // ITHEA International Journal Information content and processing (IJI

CP), Varna, Bulgaria – 2014. Vol. 1, No. 2. –P.115–123.

7. Donchenko V. S. Matrix ‘Feature Vectors’ in Grouping Information Problem: linear discrimination / V. S. Donchenko, F. M. Skotarenko // ITHEA International Journal Information Technologies and Knowledge" (IJ ITK) Varna, Bulgaria– 2014. – Vol. 21, No. 1. –P. 40–47.

Статті [5-7] опубліковані у міжнародному журналі, який включений до міжнародних наукометричних баз Worldcat, ROAD Directory of Open Access scholarly Resources, Google Scholar, CiteseerX, ITHEA.

8. Донченко В.С Концепція кортежності для лінійних операторів та її реалізація для матричних кортежів / В. С. Донченко, Т. П. Зінько, Ф. М. Скотаренко // Журнал обчислювальної та прикладної математики. –2015, –№3. –С. 127–140. ISSN – 0868–6912.

Стаття 8 опублікована у виданні України, яке включене до міжнародних наукометричних баз DOAJ (Directory of Open Access Journals) , Index Copernicus International, Google Scholar, eLIBRARY, BIHITI PAH, української реферативної бази даних «Україніка наукова» та українського реферативного журналу «Джерело».

Тези наукових конференцій

9. Скотаренко Ф. Лінійна Дискримінація в задачах розпізнавання з матричними представниками / Ф. Скотаренко // 2-га Міжнародна Науково-Практична Конференція 'Інформаційні технології та взаємодія' (IT&I) 3-5 листопада 2015. –С. 273–274.

10. Скотаренко Ф. Формули Гревіля-Кириченка для матричних евклідових просторів / Ф. Скотаренко // 5-та Міжнародна Науково-Практична Конференція «Проблеми інформатики та комп’ютерної техніки» ПІКТ–2016. –С. 123–124.

11. Скотаренко Ф. Базові структури евклідових просторів в прикладних задачах: кортежні оператори для матричних кортежів / В.С. Донченко, Т.П. Зінько, Ф.М. Скотаренко // 18-th International Conference SAIT 2016 (System

Analysis and Information Technology) –2016–C. 78–79.

12. Skotarenko F. Grouping Information Problem: Linear Discrimination Matrix ‘Feature Vectors’ / F.M. Skotarenko // XXVII International Conference Problems Of Decision Making Under Uncertainties –2016–C. 215–216.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ballard D.H. Generalizing the Hough transform to detect arbitrary shapes. - Pattern recognition, v.13, N 2, 1981.–pp. 111-122. Hough P.V.C. Method and Means for Recognizing Complex Patterns. - U.S. Patent 3069354, 1962.

2. Bengio Y. Global Optimization of a Neural Network-Hidden Markov Model Hybrid / Yoshua Bengio, Renato De Mori, Giovanni Flammia and Ralf Kompe // IEEE Transactions on Neural Networks, vol. 3, №.2, 1992. –pp. 252–259

3. Ben-Israel A. and Charnes A. Contributions to the Theory of Generalized Inverses. // Journal of the Society for Industrial and Applied Mathematics Published by the Society Philadelphia, Pennsylvania.-September 1963.- Vol. 11, №3 – 667-699 p.

4. Bishop C.M. Neural Networks for Pattern Recognition / Bishop C.M. // Oxford University Press, –NY, 1995. – 488 p.

5. Christopher M. Bishop. Pattern Recognition and Machine Learning / Christopher M. Bishop // Series: Information Science and Statistics, 1st ed. 2006. Corr. 2nd printing 2011, – 740 p.

6. Donchenko V. “Feature vectors” in grouping information problem in appliedmathematics: vectors and matrixes / V. Donchenko, T. Zinko, F. Skotarenko // Problems of computer Intellectualization. ITHEA, Kiyv, Ukraine-Sofia, Bulgaria, 2012. –pp. 111-124.

7. Donchenko V. General scheme of the Hough Transform and properties of the Hough estimates in the special case of Discrete Spaces. //Statistical Research Report, 1994-6, University of Umea, S-901 87, Umea, Sweden, 1994.- 10 pp.

8. Donchenko V. S. Matrix ‘Feature Vectors’ in Grouping Information Problem: linear discrimination / V. S. Donchenko, F. M. Skotarenko // ITHEA International Journal Information Technologies and Knowledge" (IJ ITK) Varna, Bulgaria– 2014. – Vol. 21, No. 1. –P. 40–47.

9. Donchenko V. S. Development of pseudo inverse techniques: cortege operators and applications in grouping information problem / V. S. Donchenko, F.

M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2014. – No. 1. – С. 85–90.

10. Donchenko V. S. Direct formulas Greevil – Kirichenko for matrix Euclidean spaces / V. S. Donchenko, F. M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2015. – No. 1. – С. 119–123.

11. Donchenko V. S. Spreading the Moore-Penrose pseudo inverse on matrixes Euclidean spaces: theory and applications / V. S. Donchenko, T. P. Zinko, F. M. Skotarenko // ITHEA International Journal Information content and processing (IJI CP), Varna, Bulgaria – 2014. Vol. 1, No. 2. –P.115–123.

12. Donchenko V.S. Cortege operators in grouping information problem: linear discrimination of matrix 'Feature Vectors' / V. S. Donchenko, F. M. Skotarenko // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2014. – No. 3. – С. 101–104.

13. Dongarra J. J., Bunch J. R., Moler C. B., Stewart G. W. LINPACK User's Guide. Philadelphia, 1979.

14. Elden L. A weighted pseudoinverse, generalized singular values, and constrained least squares problems // BIT. 1982. - V. 22. - P. 487-502.

15. Groetsch C.W. Regularization with linear equality constraints // Lect. Notes Math. 1986. - № 1225. - P. 168-181.

16. Hart P. E. How the Hough Transform was Invented / Hart P. E. // IEEE Signal Processing Magazine, V. 26, Issue 6, November, 2009. – pp. 18-22.

17. Holmes R.B. Course on optimization and best approximation. Berlin: Springer-Verlag, 1972. - 233 p.

18. Hough P.V.C. Method and Means for Recognizing Complex Patterns. - U.S. Patent 3069354, 1962.

19. Huang X. Spoken language processing: a guide to theory, algorithm, and system development / X. Huang, A. Acero, H. Hon. // Prentice Hall PTR, 2001. – 936p.

20. Illingworth J. and Kittler J.V. The Adaptive Hough Transform //IEEE Transactions on Pattern Analysis and Machine Intelligence. –v.9, №5.–1987. –pp. 690-698.

21. Illingworth J., Kittler J. A Survey of the Hough Transform // Computer Vision, Graphics, and Image Processing, v. 44, 1988.–pp. 87-116.

22. Intelligence and Decision Making.– International book series "INFORMATION

23. Juang B. H. Hidden Markov Models for Speech Recognition / B. H. Juang, L. R. Rabiner // Technometrics, Vol. 33, №. 3. Aug. 1991. –pp. 251-272.

24. Kirichenko N.F. Analytical representation of perturbation of pseudoinverse matrices / Kirichenko N.F. // Cybernetics and system analysis. –V.33, №2. –March-April, 1997. –pp. 230-239.

25. Kirichenko N.F. Nonlinear recursive nonlinear transformation: dynamic systems and optimization / N.F. Kirichenko, V.S. Donchenko, D.P. Serbaev // Cybernetics and system analysis. –V.41, №3. –May 2005. –pp. 364-373.

26. Kohonen T. Self-Organizing Maps, -3-d ed. - Tokyo: Springer, 2001.-501 p.

27. KurzhanskiiA.B. On evolution equations in estimation problems for systems with uncertainty.// Laxenburg: International institute for Applied Systems Analysis. WP-82-84, 1982.

28. Lawson C. L., Hanson R. J. Solving Least Squares Problems. Prentice-Hall, 1974.

29. Lucchese L., Mitra S. Colour image segmentation: A state-of-the-art survey // Proc. of the Indian National Science Academy (INSA-A), 67(2):207–221, 2001.

30. M. Zuhair Nashed. Generalized Inverses and Applications, 1st Edition Proceedings of an Advanced Seminar Sponsored by the Mathematics Research Center, the University of Wisconsin—Madison, October 8 - 10, 1973

31. Minarnide N. Nakarnura K. A restricted pseudoinverse and its application to cotrained minima // SIAM J. Appl. Math. 1970. - V. 19. -P. 167-177.

32. Moler C. B., Stewart G. W. An Algorithm for Generalized Matrix Eigenvalue Problems//SIAM J. Numer. Anal. 1973. N 2. Vol. 10.

33. Moore E.H. On the reciprocal of the general algebraic matrix / Moore E.H. // Bulletin of the American Mathematical Society, 1920. –pp. 394–395.

34. Nashed M. Zuhair ,Votruba G.F. A Unified Operator Theory of Generalized Inverse.//Proceedings of an Advanced Seminar Sponsored by the Mathematical Research Center, The University of Wisconsin, Madison, October 8-10, 1973. – New York, Academic Press, 1976.

35. Pardalos M. Optimization and Control of Bilinear Systems: Theory, Algorithms, and Applications / Panos M. Pardalos, Vitaliy A. Yatsenko // Springer

36. Penrose R. A generalized inverse for matrices / Penrose R. // Proceedings of the Cambridge Philosophical Society, 1955. – pp. 406–413.

37. Penrose R. On a generalized inverse for matrices // Proc. Cambridge Philos. Soc. 1955. - V. 51, № 3. - P. 406-413.

38. Petryshyn W.V. On generalized inverses and on the Uniform convergence of (/ — /3K)n with application to iterative methods //J. Math. Anal. Appl. 1967. - V. 18. - P. 417-439.

39. Rabiner L.R. A tutorial on Hidden Markov models and selected applications in Speech Recognition / Rabiner L.R. // Proceeding of the IEEE. vol. 77, №2, February 1989. –pp. 257-286.

40. Richard O. Duda. Pattern Classification / Richard O. Duda, Peter E. Hart and David G. Stork // A Wiley-Interscience Publication. –2 ed, 2001. – 654 p.

41. Schweppe F.C. Recursive state estimation; unknown but bounded errors and system input.// IEEE Trans. Automat. Control.- v.AC-13,№1, 1968, p.22-28.

42. Schweppe F.C. Uncertain dynamic systems.- Englewood Cliffs: Prentce Hall, 1973.

43. Science+Business Media, LLC, – NY, 2008. – 370 p.

44. Shlaefer F.M., Schweppe F.C. Continious-time state estimation under disturbances bounded by convex sets.// IEEE Trans. Automat. Control.- v.AC-17,№2, 1972, p197-205.

45. Simon Haykin. Neural Networks, A Comprehensive Foundation.-Prentice Hall, Upper Saddle River, New Jersey 07458. – 1999.– p. 842.

46. Soffer M., Kiryati N. Guaranteed Convergence of the Hough Transform.// Computer Vision and Image Understanding.–v.69, №2.– February 1998. –pp.119-134.

47. Strang G. Introduction to Applied Mathematics. Wellesley-Cambridge, 1986.

48. Vapnik V.N. Statistical Learning Theory / Vapnik V.N. // Wiley-Interscience. NY: John Wiley, 1998. – 736 p.

49. Vapnik, V.N. Statistical Learning Theory. New York: Wiley, 1998

50. Wedin P. Perturbation theory for pseudo-inverses // BIT(SVER). -1973. V. 13, № 2. - P. 217-232.

51. Wikipedia. Сигнал // Материал из Википедии – свободной энциклопедии [Электронный ресурс]. – Режим доступа: http://ru.wikipedia.org/wiki/Сигнал.

52. Xu L. Further Developments on RHT: Basic Mechanisms, Algorithms, and Computational Complexities // Xu L. Oja E. ICPR(92): Proceedings, International Conference on Pattern Recognition.– Vol.1, 1992. –pp.125-132.

53. Алберт А. Регресія, псевдоінверсія, рекурентне оцінювання;; [пер. З англ.] / Алберт А. – М.: Наука, 1977. –305 с.

54. Бакан Г.М. Алгоритмы построения гарантированных размытых оценок в линейных системах на основе метода наименьших квадратов//Проблемы управления и информатики.-№3, 1995, с.117-129.

55. Бакан Г.М.Фильтрация в условиях нестатистически заданной неопределённости//Автоматика.-№2, 1980, с.13-32.

56. Беклемишев Д.В. Дополнительные главы линейной алгебры. – М.: Наука, 1983.–336 с.

57. Бублик Б.Н., Гаращенко Ф.Г, Кириченко Н.Ф. Структурно-параметрическая оптимизация и устойчивость динамических пучков.- Київ: Наукова думка, 1985.- 304 с.

58. Вапник В.Н. Теория распознавания образов (статистические проблемы обучения) / В.Н. Вапник, А.Я. Червоненкис // Главная редакция физ.-мат. литературы. –М.: Наука, 1974. – 416 с.

59. Верченко А.П., Кириченко Н.Ф., Лепеха Н.П. Рекуррентные средства кластеризации в применении к задачам сегментации изображений. Проблемы управления и информатики, 2005, №2 , с. 62–71

60. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов / Винцюк Т.К // Киев: Наукова думка, 1987. – 264 с.

61. Галунов В.И. Акустическая теория речеобразования и система фонетических признаков / В.И. Галунов, В.И. Гарбарук // Материалы международного конгресса 100 лет экспериментальной фонетике в России, 2001. –с. 58-62.

62. Гантмахер Ф.Р. Теория матриц. - М.:Наука, 1967.-287 с.

63. Голік А.О., Донченко В.С. Застосування еліпсоїдальної та ортогональної відстаней відповідності у задачах розпізнавання мови та жестів. // Вісник Київського національного університету імені Тараса Шевченка. Серія: фіз.-мат. науки, вип.1, 2013. – с. 146-155.

64. Гонсалес Р. Цифровая обработка изображений: пер. с англ. П. Чочиа. / Р. Гонсалес, Р. Вудс // –М.: Техносфера, 2005. – 1072 с.

65. Грузман И.С. Цифровая обработка изображений в информационных системах: Учебное пособие / И.С. Грузман, В.С. Киричук, В.П. Косых, Г.И. Перетягин, А.А. Спектор // Новосибирск: Изд-во НГТУ, 2002. – 352 с.

66. Дегтяров Н.П. Параметрические и информационное описание речевых сигналов / Дегтяров Н.П. – Минск, 2003. – 216 с.

67. Деркач М.Ф. Динамические спектры речевых сигналов / М.Ф. Деркач, Р.Я. Гумецкий, Б.М. Гура, М.Е. Чабан // –Л.: Вища школа. Из-во при Львов. ун-те, 1983. – 168 с.

68. Донченко В. Базовые структуры евклидовых пространств: конструктвные методы описания и использования / В. Донченко, Ю.

Кривонос, В. Омардибирова // New Trends in Classification and Data Mining. – ITHEA, Sofia, Bulgaria. - 2010. – pp. 155-170.

69. Донченко В.С Концепція кортежності для лінійних операторів та її реалізація для матричних кортежів / В. С. Донченко, Т. П. Зінько, Ф. М. Скотаренко // Журнал Обчислювальної та прикладної математики. –2015, – №3. –С. 127–140. ISSN – 0868–6912.

70. Донченко В.С. Евклидовы пространства: конструктивные методы описания базовых структур и их использование / Донченко В.С. // Information Models of Knowledge. –Editors: Krassimir Markov, Vitalii Velichko, Oleksy Voloshin – ITHEA. –Kiev, Ukraine – Sofia, Bulgaria. Number 19, –2010. –P. 362-376.

71. Донченко В.С., Кириченко Н.Ф. Псевдообращение в задачах управления с ограничениями//Кибернетика и системный анализ.- №6. - 2003.

72. Донченко В.С., Кириченко Н.Ф. Быстрое преобразование Хока и псевдообращение.//Проблемы управления и информатики.- №2. 2002, с.115-125.

73. Донченко В.С. Множинний підхід до опису невизначеності в математичному моделюванні: дис. на здобуття наук. ступеня докт. фіз.-мат.наук:спец. 01.05.04 «Системний аналіз і теорія оптимальних рішень» / Донченко В.С.; Київський нац. ун-тет ім. Т. Шевченка. –Київ, 2007. –349 с.

74. Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах / Зайченко Ю.П. // Учебное пособие для ВУЗов. Информатика. Прикладное обеспечение. –К.:Слово, 2008. –344 с.

75. Зайченко Ю.П. Основи проектування інтелектуальних систем / Зайченко Ю.П. // Навчальний посібник. –К.: Видавничий Дім «Слово», 2004. –352 с.

76. Ивахненко А.Г. Самоорганизующиеся системы распознавания и автоматического управления. – К.: Техника, 1969.–395 с.

77. Карпов О. Н. Технология построения устройств распознавания речи. Монография / Карпов О. Н. // – Д.: ДНУ, 2001. – 184 с.

78. Киедзи Асаи. Распознавание речи / Киедзи Асаи, Дзюндзо Ватада, Сокуке Иваи и др. // Прикладные нечёткие системы. Под редакцией Тэрано Т., Асаи К., Сугено М. – М.: Мир, 1993. – с. 157-170.

79. Кириченко Н. Ф. Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей / Н.Ф. Кириченко, Ю.В. Крак, А.А. Полищук // Кибернетика и системный анализ, 2004. –№3. –с. 116–129.

80. Кириченко Н.Ф., Лепеха Н.П. Псевдообратные и проекционные матрицы в применении к исследованию задач управления, наблюдения и идентификации.//Кибернетика и системный анализ.–2002.№4.–С.107-123.

81. Кириченко Н.Ф. Алгебраический Jack Knife: кластеризация по гиперплоскостям / Н.Ф. Кириченко, В.С. Донченко // Proceedings: XIII-th International Conference "Knowledge –Dialog – Solution".–June 18-24, 2007, Varna (Bulgaria). – 2007. – V.1.–P.89-95.

82. Кириченко Н.Ф. Аналитическое конструирование минимаксных результатов в линейных системах.//ДАН УССР. -№7, 1977, с591-594.

83. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц. // Кибернетика и системный анализ. – 1997.–№2.–С.98-107.

84. Кириченко Н.Ф. Гиперплоскости в «множествах и расстояниях соответствия»: кластеризация / Н.Ф. Кириченко, В.С. Донченко // Artificial Intelligence and Decision Making.– International book series "INFORMATION SCIENCE&COMPUTING", Number 7.– Sofia 2008.– pp. 25-36..

85. Кириченко Н.Ф. Нелинейные рекурсивные регрессионные преобразователи: динамические системы и оптимизация / Н.Ф. Кириченко, В.С. Донченко, Д.П. Сербаев // Кибернетика и системный анализ, 2005. –№3. –с. 58–68.

86. Кириченко Н.Ф. Псевдообращение матриц и их рекуррентность в задачах моделирования и управления. // Проблемы управления и информатики. –1995.-№1. – С.114-127.

87. Кириченко Н.Ф. Синтез систем нейрофункциональных преобразователей в решении задач классификации / Н.Ф. Кириченко, Ю.Г. Кривонос, Н.П. Лепеха // Кибернетика и системный анализ, 2007. – №3. –с. 47–57.

88. Кириченко Н.Ф., Донченко В.С. Псевдообращение в задачах кластеризации. // Кибернетика и системный анализ. – 2007. – №4. – С. 73-92

89. Кириченко Н.Ф., Донченко В.С., Сербаев Д.П. Нелинейные рекурсивные регрессионные преобразователи: динамические системы и оптимизация. // Кибернетика и системный анализ. – 2005. – №3. – С. 58-68

90. Кириченко Н.Ф., Крак Ю.В., Полищук А.А. Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей. // Кибернетика и системный анализ. – 2004. – №3. – С. 116-129

91. Кириченко Н.Ф., Кривонос Ю.Г., Лепеха Н.П. Оптимизация синтеза гиперплоскостных кластеров и нейрофункциональных преобразований в системах классификации сигналов // Кибернетика и системный анализ. – 2008.–№6.– С. 50-58

92. Кириченко Н.Ф., Кривонос Ю.Г., Лепеха Н.П. Синтез систем нейрофункциональных преобразователей в решении задач классификации. // Кибернетика и системный анализ. – 2007. – №3. С. 47-57

93. Кириченко Н.Ф., Лепеха Н.П. Применение псевдообратных и проекционных матриц к исследованию задач управления, наблюдения и идентификации. // Кибернетика и системный анализ. – 2002. – №4. – С. 107-124

94. Кириченко Н.Ф., Лепеха Н.П. Псевдообращение в задачах управления и наблюдения. // Автоматика. –1993.–№5. – С.69 –81.

95. Кириченко Н.Ф., Матвиенко В.Т. Оптимальный синтез структур для линейных систем. // Проблемы управления и информатики. –1996.-№1-2. – С.162-171.

96. Кириченко Н.Ф., Н.П. Лепеха. Возмущение псевдообратных и проекционных матриц и их применение к идентификации линейных и нелинейных зависимостей. //Проблемы управления и информатики.–2001.-№1.-С.6-22.

97. Кириченко Н.Ф., Наконечный А.Г. Минимаксный подход к рекуррентному оцениванию состояний линейных динамических систем. // Кибернетика. №4, 1977, с 44-55.

98. Кириченко Н.Ф., Резник А.М., Щетенюк С.П. Псевдообращение матриц в проблеме проектирования ассоциативной памяти. // Кибернетика и системный анализ. – 2001. – №3. С. 18-28

99. Крак Ю. В. Відновлення параметрів мовного тракту з акустичного рівняння Клейна-Гордона / Ю. В. Крак, І. О. Стеля // Искусственный интеллект. 2007. –№ 3. –с. 421-427.

100. Крак Ю. В. Комп'ютерна модель голосових зв'язок та мовного тракту людини/ Ю. В. Крак, І. О. Стеля // Искусственный интеллект, 2008. №4. –с. 758-762.

101. Крак Ю. В. Синтез звуків голосу людини на основі фізичних моделей голосових зв'язок та мовного тракту / Ю. В. Крак, І. О. Стеля // Искусственный интеллект, 2009. № 4. – с.74-79.

102. Крак Ю. В. Чисельне моделювання голосових зв'язок за двомассовою моделлю/ Ю. В. Крак, І. О. Стеля // Журнал обчислювальної та прикладної математики, 2007. № 94, –с. 55-60.

103. Красовский Н.Н. Управление динамической системой. Задача о минимуме гарантированного результата.-М.: Наука, 1985.

104. Кривонос Ю. Г. Структура, свойства, характеристики объектов и элементов синтеза речи / Ю. Г. Кривонос, Ю. В. Крак, Н. Н. Шатковский // Компьютерная математика, 2006. № 1. – с. 61-69.

105. Кунцевич В.М. Лычак М.М. Синтез оптимального управления динамическими объектами с неизвестными параметрами.//Автоматика. -№2, 1980, с.22-23.

106. Куржанский А.Б. Управление и наблюдение в условиях неопределённости. -М.: Наука, 1977.

107. Куссуль Н.М., Шелестов А.Ю., Лавринюк А.М. Інтелектуальні обчислення: Навчальний посібник. – Київ: Наукова думка, 2006.–186 с.

108. Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов. /Пер. с англ. – М.: Наука, 1986.–232 с.

109. Наконечный А.Г. К минимаксной теории оценивания функционалов от решений операторных уравнений.// ДАН УССР. - Сер.А, №5, 1982.

110. Наконечный А.Г. Минимаксное оценивание решений квазилинейных стохастических уравнений.// Проблемы управления и информатики.-№1,2, 1996.

111. Наконечный А.Г. Минимаксное оценивание функционалов от решений вариационных уравнений в гильбертовых пространствах.-Киев: КГУ, 1983.–82 с.

112. Нечёткие множества и теория возможностей. Последние достижения. // Под ред. Р.Р.Ягера.–М.:Радио и связь, 1986.– 408 с.

113. Путятин Е.П. Обработка изображений в робототехнике / Е.П. Путятин, С.И. Аверин // –М: Машиностроение, 1990. – 320 с.

114. Путятін Є.П. Методи та алгоритми комп’ютерного зору: навчальний посібник / Є.П. Путятін, В.О. Гороховатський, О.О. Матат. – Х.: ТОВ «Компанія СМІТ», 2006. – 236 с.

115. Рабинер Л.Р. Цифровая обработка речевых сигналов / Л.Р. Рабинер, Р.В. Шафер. –М.: Радио и связь, 1981. – 395 с.

116. Савченко В.В. Различение случайных сигналов в частотной области / Савченко В.В. // Радиотехника и электроника, 1997, Т.42, №4, –с.426-429.

117. Скопецький В.В., Стоян В.А., Кривонос Ю.Г. Математичне моделювання прямих та обернених задач динаміки систем з розподіленими параметрами.- Київ, Наукова думка, 2002.- 361 с.

118. Скотаренко Ф. Алгоритм лінійної дискримінації для матричних просторів та його основи / Ф.М Скотаренко // Вісник Київського

національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2015. – No 2. – С. 121–127.

119. Ту Дж. Принципы распознавания образов / Дж. Ту, Р. Гонсалес // Пер. С англ. –М.: Мир, 1978. – 414 с.

120. Форсайт, Понс. Компьютерное зрение. Современный подход. Пер. с англ.- М., Изд-во «Вильямс», 2004.-928 с.

121. Хіміч О.М. Гібридний алгоритм для лінійної задачі найменших квадратів з напіввизначеною розрідженою матрицею // Хіміч О.М., Сидорук В.А. / «Теорія оптимальних рішень». − 2014, м. Київ. - С. 106-113.

122. Хіміч О.М. Гібридний алгоритм розв'язування систем лінійних рівнянь з розрідженими матрицями на основі блочного LLT методу // Хіміч О.М., Си-дорук В.А. / «Компьютерная математика». – 2015, Вип. 1. - С. 59-62.

123. Хіміч О.М. Паралельні однокрокові ітераційні методи розв'язання алгебра-їчної проблеми власних значень для розріджених матриць // Хіміч О.М., Чис-тяков О.В. / «Компьютерная математика» – 2014, Вип. 2. – С. 81-88.

124. Черников С.Н. Линейные неравенства.–М.: Наука, 1968.–488с.

125. Черноусько Ф.Л. Оптимальные гарантированные оценки неопределённостеё при помощи еллипсоидов.//Изв. АН СССР. Техническая кибернетика. №3, 1980, с.3-11; №4 с.3-11; №5 с.5-11.

126. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи.– М.: Радио и связь, 2000.– 454 с.

127. Шеннон К. Работы по теории информации и кибернетике / Шеннон К. – М.: Изд. иностр. лит., 1963. –830 с.

128. Шлезингер М.И. Математические средства обработки изображений / Шлезингер М.И. // – К.: Наукова думка, 1989. – 196 с.

129. Шлезингер М.И. Алгоритм выделения отрезков осевой линии на графических изображениях / В. М. Кийко, М. И. Шлезингер. Киев: ИК, 1983. – 22 с.

130. Шлезингер М.И., Главач В. Десять лекций по статистическому и структурному распознаванию. – Киев: Наук. думка, 2004. – 546 с.

ДОДАТОК А
АКТИ ВПРОВАДЖЕННЯ ДИСЕРТАЦІЙНОЇ РОБОТИ

"ЗАТВЕРДЖУЮ"
Виконуючий обов'язки проректора з
наукової роботи
Київського національного університету
імені Тараса Шевченка
____________ **Я.В. Лавренюк**
____02. 11.________2015 р.

Довідка
про впровадження у навчальний процес
результатів дисертаційної роботи Скотаренка Федора Миколайовича
"Розробка математичних методів групування інформації з матричними
ознаками", що представлена на здобуття наукового ступеня кандидата
фізико-математичних наук за спеціальністю 01.05.04 – системний аналіз і
теорія оптимальних рішень.

Результати наукового дослідження методів та алгоритмів псевдо обернення в евклідовому просторі матриць для розв'язування прикладних задач в сучасних інформаційних системах, отримані в рамках підготовки кандидатської дисертації Скотаренка Федора Миколайовича "Розробка математичних методів групування інформації з матричними ознаками", впроваджені у 2013-2014 н.р, у навчальний процес кафедри системного аналізу та теорії прийняття рішень факультету кібернетики Київського національного університету імені Тараса Шевченка при викладанні курсу "Обробка інформації в умовах невизначеності" для студентів 1 курсу магістратури за спеціальністю "Системи і методи прийняття рішень".

Декан факультету кібернетики
доктор фізико-математичних наук, професор А.В. Анісімов

Завідувач кафедри системного аналізу та
теорії прийняття рішень
доктор фізико-математичних наук, професор О.Г. Наконечний

Професор кафедри системного аналізу та
теорії прийняття рішень
доктор фізико-математичних наук, професор В.С. Донченко

«З А Т В Е Р Д Ж У Ю»
Перший проректор
Київського національного університету
імені Тараса Шевченка
______________________Закусило О.К.
«______»____________________2016 р.

Україна * Київський національний університет імені Тараса Шевченка * Ідентифікаційний код 02070944

АКТ
про впровадження результатів дисертаційного роботи
Скотаренка Федора Миколайовича на тему «Розробка математичних методів групування інформації з матричними ознаками» поданої на здобуття наукового ступеня кандидата фізико-математичних наук за спеціальністю 01.05.04 – системний аналіз і теорія оптимальних рішень

Комісія у складі:

Голова комісії: декан факультету кібернетики, член-кореспондент НАН України, доктор фізико-математичних наук, професор Анісімов А.В.
Члени комісії: доктор фізико-математичних наук, професор, завідувач кафедри "Системного аналізу та теорії прийняття рішень" Наконечний О.Г., доктор фізико-математичних наук, професор Мащенко С.О.,

склали цей акт про те, що результати дисертаційної роботи Скотаренка Федора Миколайовича на тему «Розробка математичних методів групування інформації з матричними ознаками», поданої на здобуття наукового ступеня кандидата фізико-математичних наук, застосовано при виконанні науково-дослідної теми №11БФ015-06 "Проблеми теорії прийняття рішень та системного аналізу стохастичних мереж" (державний номер реєстрації 0111U006680, термін виконання 2011-2015р.р. за програмою "Інформатизація суспільства") Київського національного університету імені Тараса Шевченка. При проведенні науково-дослідницьких робіт була використана концепція кортежних операторів, яка була розроблена у дисертаційній роботі.

Голова комісії: ____________Анісімов А.В.

Члени комісії: ____________ Наконечний О.Г.

____________ Мащенко С.О.