

УДК 519.6  
MSC 97R40

**CONSTRUCTION OF NEURAL ARCHITECTURES WITH  
DESIRED BEHAVIOUR UNDER GEOMETRIC  
TRANSFORMATIONS OF THE INPUT**

V. V. DUDAR, V. V. SEMENOV

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,  
E-mail: slavko123@ukr.net, semenov.volodya@gmail.com

**ПОБУДОВА АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ З  
БАЖАНОЮ ПОВЕДІНКОЮ ПІД ДІЄЮ ГЕОМЕТРИЧНИХ  
ПЕРЕТВОРЕНЬ ВХОДУ**

В. В. ДУДАР, В. В. СЕМЕНОВ

Київський національний університет імені Тараса Шевченка, Київ, Україна,  
E-mail: slavko123@ukr.net, semenov.volodya@gmail.com

**ABSTRACT.** We present a general method for analysis of convolutional layers under geometric transformations of the input that are linear with respect to pixel values. We also describe the algorithm for finding all possible types of behaviours of the output of convolutional layers under geometric transformations of the input. We also present a general method for construction of convolutional architectures with desired behaviour under geometric transformations of the input.

**KEYWORDS:** convolutional neural network, invariance of CNN, equi-invariance of CNN, steerable convolutional kernels, induced linear operators, invariant subspaces.

**АНОТАЦІЯ.** В статті запропоновано загальний метод для аналізу згорткових шарів під дією геометричних перетворень входу, які можна лінійно параметризувати відносно значень пікселів зображення. Запропоновано метод для знаходження всіх можливих типів поведінки виходу згорткових шарів під дією перетворень входу, та відповідних параметризацій ядер, а також загальний метод для побудови згорткових нейронних мереж з бажаною поведінкою виходу від дією даних геометричних перетворень входу. **КЛЮЧОВІ СЛОВА:** згорткова нейронна мережа, інваріантність згорткових нейронних мереж, еківаріантність згорткової мережі, керовані ядра згорток, звуження лінійних операторів, інваріантні підпростори.

## ВСТУП

Питання побудови згорткових нейронних мереж з заданою поведінкою при геометричних перетвореннях входу вважається одним з таких, що дозволить значно покращити ефективність моделей на тестових даних [3].

Один підходів — Spatial Transformer Networks [4]. Автори пропонують використати додаткову згорткову нейронну мережу, яка за заданим вхідним зображенням оцінить параметри афінного перетворення, яке потім застосовується до вхідного зображення, і подається на вхід другій згортковій нейронній мережі, яка проводить класифікацію. Цей підхід зміг показав кращі результати за звичайні CNN на декількох датасетах: Distorted Mnist, SVHN та CIFAR-10.

В роботі [5] автори пропонують повертати фільтри згортки для того щоб детектувати певні паттерни на зображенні під будь-яким кутом. Автори використали згорткову мережу з одним згортковим шаром, який складався з набору фільтрів, кожен з яких був циклічно повернутий. Автори показали покращення результатів класифікації текстур, особливо у випадку малої навчальної вибірки.

Архітектура згорткової мережі Scattering network [6] використовує вейвлет перетворення, нелінійності та усереднення по групі для отримання інваріантних ознак. Мережі такого типу були узагальнені для використання на групі зміщень, поворотів та масштабування, та застосовуються для розпізнавання об'єктів та текстур.

Багато робіт присвячено задачі обчислення еквіваріантних репрезентацій вхідних даних (без учителя), зокрема, можна виділити перетворюючі автоенкодера [7], еквіваріантні машини Больцмана [8, 9] та еквіваріантні фільтри на сферичних групах [10].

В [11] було показано, що у випадку звичайна згорткова мережа тренується на достатньо великій репрезентативній вибірці, такий як ImageNet, результуючі репрезентації є еквіваріантними до симетричних відображень, масштабування та поворотів.

В роботі [12] пропонується приблизно еквіваріантна згорткова архітектура, що використовує розріджені багатовимірні ознаки для обробки багатовимірних груп перетворень.

Робота [13] є першою, в якій авторам вдалося вбудувати додаткові інваріантності до геометричних перетворень в структуру згорткової нейронної мережі. Автори показали, що стандартна CNN може бути узагальнена щоб бути еквіваріантною (чи інваріантною) до наборів геометричних перетворень, що утворюють групи: зміщення, повороти, симетричні відображення, та їх комбінації.

В роботі [14] автори, дослідили групу перетворень  $r4m$ , щоб розробити мережі, інваріантні до поворотів та симетричних відображень, які були б ще більш ефективними. Автори показали, що кожне кероване представлення є композицією низьковимірних простих ознак (feature types), кожна з яких може бути керована незалежно від інших. Аналогічно до G-згортки, керовані згортки можуть бути використані замість звичайних згортки в згорткових нейронних мережах. Експерименти показують, що такий тип

мережі дозволяє зменшити кількість параметрів, та покращити генералізацію на тестовій вибірці.

Групові та керовані мережі також були узагальнені для 3D випадку по відношенню до групи обертань сфери [15].

Робота [16] розвиває загальну теорію еквіваріантних згорткових нейронних мереж по відношенню до групових перетворень, та покриває декілька попередньо розроблених методів.

Ще один підхід для досягнення інваріантності — це капсульні мережі [17, 18], що комбінують в собі декілька ідей. По-перше, кожному вектору ознак відповідає вектор орієнтації, довжина якого відповідає степені наявності ознаки в даній позиції. Вектор орієнтації може відображати напрямок, відносний зсув чи інші геометричні властивості певної ознаки на об'єкті. Прості об'єкти (такі як відрізки) комбінуються в більш складні (такі як кути) за допомогою процедури *dynamic routing*: вектори простих ознак "віддають голоси" за конфігурації складніших векторів ознак, а якщо їх "голос" узгоджується з загальною конфігурацією, їх внесок підсилюється. Приблизна еквіваріантність репрезентацій впливає з факту, що все залежить від відносних конфігурацій векторів ознак, наприклад, якщо два простіші ознаки будуть повернуті на певний кут, тоді орієнтація ознаки, що стоїть вище в ієрархії, також буде повернута таким же чином.

В даній статті буде описаний загальний алгоритм для створення згорткових нейронних мереж, вихід яких змінюється певним наперед заданим чином під час перетворення вхідних зображень певними геометричними перетвореннями. Наприклад, виходом нейронної мережі може бути векторне поле, для якого ми можемо гарантувати, що вектори будуть повертатись при повороті зображення приблизно на такий же кут, як і саме зображення (і точно на такий же кут при повороті на кратні  $\pi/2$ ). Таким чином, описаний механізм, зокрема, дозволить будувати мережі, що апроксимують векторні поля, що еквіваріантні до зміщень та поворотів вхідного зображення, що може бути корисним, наприклад, при апроксимації векторного поля зміщень об'єктів на сусідніх кадрах відеопотоку. Інше застосування - мережа, виходом якої є єдиний вектор, еквіваріантний до поворотів, що дозволяє тренувати мережу, що знаходить кутову орієнтацію об'єкта. Якщо об'єкт на зображенні є симетричним відносно певної точки (наприклад, лінія під певним кутом), то для нього не можна визначити векторну орієнтацію (в межах  $[0, 2\pi)$ ), але можна знайти його кутову орієнтацію, яка змінюється в межах  $[0, \pi)$ . Буде наведений метод побудови мережі, вихід якої змінюється таким чином, що з нього можна отримати кутову орієнтацію об'єкта в межах  $[0, \pi)$ , що відповідним чином змінюється при повороті зображення.

Ми наведемо алгоритм, який дозволяє знаходити всі можливі типи поведінки виходів згорткової нейронної мережі для заданої групи перетворень вхідного зображення. Для кожного такого типу поведінки ми зможемо отримувати параметризацію ядер, згортки з якими поведуть себе таким наперед заданим чином, і спосіб знаходження похідних по відношенню до параметрів, що потрібно для навчання мережі. Даний метод є

достатньо загальним, і може бути застосований і у більш загальних випадках, якщо мережа тренується, наприклад, на зображеннях на сферах, на графах певної структури, та необхідно гарантувати певні властивості поведінки її виходу при зміні її входу певними перетвореннями.

### 1. ОПЕРАТОР СОБЕЛЯ

Для ілюстрації ефекту, якого ми намагаємось досягнути, розглянемо диференційний оператор Собеля [1, 2], що складається з двох згортки з ядрами  $3 \times 3$ , та є дискретною апроксимацією градієнта функції в точці.

Нехай  $I$  — вхідне зображення, а  $G_x, G_y$  — зображення, що містять похідні  $I$  по  $x$  та  $y$ . Тоді

$$\nabla_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad \nabla_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad (1)$$

$$G_x = \nabla_x * I, G_y = \nabla_y * I$$

(в класичному варіанті  $\nabla_y$  має протилежні знаки, ми змінили їх, щоб оперувати у правосторонній координатній системі).

Дослідимо, як поведуть себе величини  $G_x, G_y$ , якщо вхідне зображення повертається на величину, кратну  $\pi/2$  або симетрично відображається відносно вертикальної осі. Оскільки пара величин  $G_x, G_y$  є апроксимацією вектора градієнта, який еквіваріантний до поворотів та симетричних відображень функції, то ми очікуємо, що і його апроксимація також проявляє такі властивості. Щоб перевірити це, розглянемо 8-елементну групу перетворень зображення, що складається з комбінацій 4 поворотів на  $\pi/2$  та 2 симетричних відображень навколо вертикальної осі. Позначимо операцію повороту зображення на  $\pi/2$  за годинниковою стрілкою через  $r$ , операцію горизонтального симетричного відображення через  $f$ , операцію тотожного перетворення через  $e$ . Операції  $r$  та  $f$  задовольняють таким властивостям:

- $r^4 = e$ ,
- $f^2 = e$ ,
- $fr = r^3f$ ,
- $fr^2 = r^2f$ ,
- $fr^3 = rf$ .

Комбінуючи операції повороту  $r$  та симетричного відображення  $f$ , отримуємо 8-елементну групу перетворень, яка ізоморфна дієдральній групі  $D_4$  — групі симетрій квадрата.

Тепер, використовуючи перетворення цієї групи, побудуємо таблицю 1 як змінюються величини  $G_x, G_y$ , якщо вхідне зображення перетворюється відповідно до перетворень групи.

Таким чином, як і очікувалось, величини  $G_x, G_y$  поведуть себе як вектор під поворотами та симетричними відображеннями. Отже, згортки тензора з ядрами  $\nabla_x, \nabla_y$  можна використовувати, щоб отримувати вектор, що повертається та відображається разом з вхідним тензором. Для більш ефективного використання оператора Собеля в нейронній мережі, доцільним

ТАБЛ. 1. Поведінка оператора Собеля під дією поворотів та відображень.

$e$	$r$	$r^2$	$r^3$	$f$	$fr$	$fr^2$	$fr^3$
$G_x$	$G_y$	$-G_x$	$-G_y$	$-G_x$	$-G_y$	$G_x$	$G_y$
$G_y$	$-G_x$	$-G_y$	$G_x$	$G_y$	$-G_x$	$-G_y$	$G_x$

є його параметризація коефіцієнтами, які б навчались під час тренування, із збереженням властивостей, описаних в таблиці 1. Найбільш загальна форма пари згорток, яка поводить себе так само як оператор Собеля під перетвореннями вищеописаної групи, є такою:

$$g_x = \begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -a & 0 & a \end{bmatrix}, \quad g_y = \begin{bmatrix} a & b & a \\ 0 & 0 & -0 \\ -a & -b & -a \end{bmatrix}.$$

Бачимо, що структура ядер  $g_x, g_y$  аналогічна структурі диференційних операторів Собеля (1). Зокрема, при підстановці  $a = 1, b = 2$  отримуємо оператор Собеля.

Отже ми отримали параметризацію пари ядер згортки, яка поводить себе як вектор при поворотах та симетричних відображеннях зображення. Отримана параметризація є лінійною, а отже диференційовною, що дозволить шукати похідну функцію похибки по змінним параметризації, і проводити навчання мережі по відношенню до цих параметрів.

## 2. ЗНАХОДЖЕННЯ ВСІХ МОЖЛИВИХ ТИПІВ ПЕРЕТВОРЕНЬ ДЛЯ ЗАДАНОГО ОПЕРАТОРА

Раніше було розглянуто узагальнення оператора Собеля, яке поводить себе певним наперед заданим чином (як вектор) при поворотах і симетричних відображеннях вхідного зображення. Знайдемо всі можливі типи поведінки виходів шарів під дією лінійного перетворення вхідного зображення та параметризації операторів, що задають такі виходи.

Позначимо через  $x \in \mathbb{R}^n$  патч зображення  $k \times k$ , розгорнутий у вектор довжиною  $n = k^2$ , через  $A \in \mathbb{R}^{n \times n}$  — матрицю лінійного оператора, що діє на  $x$  (це може бути оператор симетричного відображення, повороту на  $\pi/2$ , чи довільні оператори, що лінійно діють на значення пікселів). Тоді необхідно знайти такі вектори  $s_1 \dots s_m \in \mathbb{R}^n$  (ядра згорток), що згортки з лінійно перетвореним зображенням є певною лінійною комбінацією згорток початкового зображення. Тобто, існують такі коефіцієнти  $\lambda_{11}, \lambda_{12} \dots \lambda_{mm}$ , що

$$\begin{cases} s_1^T Ax = \lambda_{11} s_1^T x + \dots \lambda_{1m} s_m^T x, \\ \dots \\ s_m^T Ax = \lambda_{m1} s_1^T x + \dots \lambda_{mm} s_m^T x. \end{cases} \quad (2)$$

Введемо такі позначення:

$$\Lambda = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1m} \\ \dots & \dots & \dots \\ \lambda_{m1} & \dots & \lambda_{mm} \end{bmatrix}, \quad \Lambda \in \mathbb{R}^{m \times m},$$

$$S = \begin{bmatrix} s_1^T \\ \dots \\ s_m^T \end{bmatrix}, \quad S \in \mathbb{R}^{m \times n}.$$

Записуючи систему рівнянь (2) у матричній формі, отримаємо:

$$SAx = \Lambda Sx. \quad (3)$$

Оскільки рівність (3) має виконуватись для всіх  $x$ , отримаємо:

$$SA = \Lambda S. \quad (4)$$

Будемо також вимагати щоб рядки матриці  $S$  були лінійно незалежними. Дійсно, якщо деякий фільтр згортки можна виразити через інші, тоді і результат згортки з даним фільтром можна виразити через інші результати згортки таким же чином, незалежно від вхідного зображення. В згортковій нейронній мережі цього ефекту можна добитись введенням додаткового згорткового шару з ядром  $1 \times 1$ . Крім того, без введення такого обмеження множина матриць  $\Lambda$ , для яких існує нетривіальний розв'язок рівності (4) стає занадто широкою, і перестає відображати реальну структуру оператора  $A$ .

**Означення 1.** Нехай  $A \in \mathbb{R}^{n \times n}$  — лінійний оператор. Назвемо матрицю  $\Lambda \in \mathbb{R}^{m \times m}$  лівою приведеною матрицею оператора  $A$ , якщо існує  $S \in \mathbb{R}^{m \times n}$ ,  $\text{rank}(S) = m$ , така що виконується рівність (4).

**Означення 2.** Матрицю  $\Lambda \in \mathbb{R}^{m \times m}$  назвемо правою приведеною матрицею оператора  $A \in \mathbb{R}^{n \times n}$ , якщо існує  $S \in \mathbb{R}^{n \times m}$ ,  $\text{rank}(S) = m$  така що

$$AS = S\Lambda.$$

Встановимо зв'язок між правими і лівими приведеними матрицями оператора.

**Лема 1.**  $\Lambda$  — ліва (права) приведена матриця  $A$  тоді й лише тоді коли  $\Lambda^T$  — права (ліва) приведена  $A^T$ .

**Лема 2.** Якщо  $A$  — невироджена, тоді всі її приведені матриці невироджені.  $\Lambda$  — ліва (права) приведена матриця  $A$  тоді й лише тоді коли  $\Lambda^{-1}$  — ліва (права) приведена  $A^{-1}$ .

**Лема 3.** Якщо  $\Lambda$  — права (ліва) приведена матриця, тоді  $\Lambda^T$  — також права (ліва) приведені матриці.

**Лема 4.** Множина лівих (правих) приведених матриць дорівнює множині правих (лівих) приведених матриць.

ТАБЛ. 2. Зв'язок між лівими та правими приведеними матрицями та параметризаціями ядер.

	Ліві приведені	Праві приведені
$A$	$\{(\Lambda, S_l)\}$	$\{(\Lambda, S_r)\}$
$A^T$	$\{(\Lambda^T, S_r^T)\}$	$\{(\Lambda^T, S_l^T)\}$
$A^{-1}$	$\{(\Lambda^{-1}, S_l)\}$	$\{(\Lambda^{-1}, S_r)\}$
$(A^T)^{-1}$	$\left\{\left((\Lambda^T)^{-1}, S_r^T\right)\right\}$	$\left\{\left((\Lambda^T)^{-1}, S_l^T\right)\right\}$

Підведемо підсумок застосування лем 1-4 до довільного оператора у вигляді таблиці. В кожній клітині показана множина пар — приведена матриця та деяка матриця  $S$  ранку  $m$ , що задовольняє відповідній рівності.

Таблиця 2, наведена вище, може бути використана для того, щоб звести знаходження лівих приведених матриць та параметризацій просторів розв'язків  $S$  до знаходження правих приведених матриць та відповідних параметризацій через відповідне перетворення. У випадку, якщо оператор  $A$ , який досліджується, задовольняє одній з умов:  $A = A^T$  (симетричність),  $A = A^{-1}$ ,  $A^{-1} = A^T$  (ортогональність), тоді процес знаходження відповідних величин не вимагає введення в розгляд додаткового оператора. У випадках, що розглядатимуться далі в розділі, це може бути використано, оскільки оператор симетричного відображення - симетричний, а оператор повороту на  $\pi/2$  — ортогональний.

Рівність (4) вимагає знаходження лівих приведених матриць та їх параметризацій, в той же час поняття правої приведеної матриці нам видається більш відповідним наступним крокам та загальноприйнятим позначенням в лінійній алгебрі, тому надалі будуть розглядатись властивості та алгоритм знаходження саме правих приведених матриць. Надалі під терміном приведена матриця розумітимемо саме праву приведену матрицю.

Нехай  $\Lambda_A = \{\Lambda \in \mathbb{R}^{m \times m} | \exists S \in \mathbb{R}^{m \times n}, rank(S) = m : AS = S\Lambda\}$  — множина правих приведених матриць для оператора  $A$ ,  $S_{A,\Lambda} = \{S | AS = S\Lambda\}$  — множина розв'язків (4) для фіксованої приведеної матриці  $\Lambda$  та оператора  $A$ .

Наведемо деякі властивості приведених матриць.

**Лема 5.** Для фіксованої приведеної матриці  $\Lambda$  множина  $S_{A,\Lambda}$ , утворює лінійний підпростір, а отже, може бути лінійно параметризована.

**Лема 6.** Для кожної матриці  $S \in S_{A,\Lambda}$ , лінійна оболонка її стовпців є лінійно інваріантним підпростором відносно оператора  $A$ .

**Лема 7.** Якщо  $\Lambda$  — це приведена матриця, тоді для довільної невиродженої матриці  $G \in \mathbb{R}^{m \times m}$  матриця  $G^{-1}\Lambda G$  також є приведеною.

**Лема 8.** Нехай  $A \in \mathbb{R}^{n \times n}$  — лінійний оператор. Тоді для довільної невиродженої матриці  $G \in \mathbb{R}^{n \times n}$  множини приведених матриць  $A$  та  $GAG^{-1}$  співпадають.

З леми 7 та факту що відношення подібності між операторами є відношенням еквівалентності, випливає, що множину лівих приведених матриць можна розбити на класи еквівалентності.

**Теорема 1.** *Множина лівих приведених матриць співпадає з множиною усіх можливих звужень оператора  $A$  на лінійно інваріантні підпростори.*

Зауважимо, що хоча множини звужень оператора та лівих приведених матриць співпадають, взаємно однозначну відповідність між приведеними матрицями та інваріантними підпросторами побудувати не можна: кожній такій матриці може відповідати декілька різних інваріантних підпросторів. Проте, виконується така лема.

**Лема 9.** *Кожному лінійно інваріантному підпростору відповідає рівно один клас еквівалентності приведених матриць.*

Таким чином, щоб отримати керований набір згорток необхідно знайти лінійну параметризацію множини розв'язків (4) для фіксованої матриці  $\Lambda$ . Після того як відповідну параметризацію знайдено, оптимальні параметри можуть бути знайдені під час тренування нейронної мережі.

Доведемо деякі додаткові властивості множини лівих приведених матриць.

Оскільки накладена умова  $\text{rank}(S) = m$  (лінійна незалежність стовпців  $S$ ), отримуємо обмеження на можливі розмірності приведених матриць:  $1 \leq m \leq n$ .

**Лема 10.** *Множина лівих приведених матриць розміром  $m = 1$  співпадає з множиною власних чисел оператора  $A$ , а множини відповідних розв'язків  $S_{A,\Lambda}$  співпадають з відповідними власними підпросторами оператора.*

Таким чином, поняття лівої приведеної матриці є узагальненням поняття власного числа, а відповідні їм матриці  $S_{A,\Lambda}$  — узагальненнями власних векторів.

**Лема 11.** *Множина лівих приведених матриць для  $m = n$  співпадає з множиною матриць, подібних до  $A$ .*

Розглянемо деякі методи для знаходження всіх приведених матриць оператора  $A$ . Знаходження таких матриць дозволить проаналізувати поведінку оператора в різних його інваріантних підпросторах, а також знаходити параметризовані згортки, які виділяють ознаки, що поводять себе деяким наперед заданим чином під дією оператора на зображення. Буде показано, що деякі з цих ознак мають геометричну інтерпретацію в термінах елементів зображення, які вони можуть детектувати.

Для побудови загальної множини приведених матриць для заданого оператора доведемо ще декілька допоміжних лем.

**Лема 12.**  $\Lambda$  — приведена матриця оператора  $A$  тоді і лише тоді, коли існує невірроджена матриця  $G$ , така що

$$G^{-1}AG = \begin{pmatrix} \Lambda & P \\ 0 & Q \end{pmatrix}.$$

**Лема 13.** Нехай  $\Lambda$  — приведена матриця оператора  $A$ . Тоді її характеристичний поліном є дільником характеристичного полінома  $A$ , а її мінімальний поліном є дільником мінімального полінома  $A$ .

**Лема 14.** Нехай  $\Lambda \in \mathbb{R}^{m \times m}$  — приведена матриця оператора  $A \in \mathbb{R}^{n \times n}$ . Тоді для довільного полінома  $p$  виконується:  $\text{Null}(p(\Lambda)) \leq \text{Null}(p(A))$ ,  $\text{rank}(p(\Lambda)) \leq \text{rank}(p(A))$ .

**Лема 15.** Нехай  $\Lambda \in \mathbb{R}^{m \times m}$  — приведена матриця оператора  $A \in \mathbb{R}^{n \times n}$ . Тоді для довільного полінома  $p$  та довільного натурального числа  $i \in \mathbb{N}$  виконується:

$$\text{Null}(p^{i+1}(\Lambda)) - \text{Null}(p^i(\Lambda)) \leq \text{Null}(p^{i+1}(A)) - \text{Null}(p^i(A))$$

**Лема 16.** Нехай  $M_A(\lambda) = p_1^{n_1}(\lambda)p_2^{n_2}(\lambda)\dots p_k^{n_k}(\lambda)$  — мінімальний поліном оператора  $A$  та його розклад на степені лінійних та квадратичних множників. Нехай  $p_i(\lambda) = \lambda - \lambda_i$  — деякий лінійний дільник мінімального полінома, що входить зі степенню  $n_i$ . Тоді для кожного  $k = 1 \dots n_i$  жорданова клітина розміром  $k \times k$

$$J_k(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix} \quad (5)$$

є приведеною матрицею.

Нехай  $p_i(\lambda) = \lambda^2 - 2a_i\lambda + (a_i^2 + b_i^2)$  — незвідний квадратичний поліном з комплексними коренями  $a_i \pm b_i i$ , що входить в розклад зі степенню  $n_i$ . Тоді для кожного  $k = 1 \dots n_i$  матриця розміром  $2k \times 2k$

$$J_k(a_i, b_i) = \begin{pmatrix} a_i & b_i & 1 & 0 & 0 & 0 & \dots & 0 \\ -b_i & a_i & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & a_i & b_i & 1 & 0 & \dots & 0 \\ 0 & 0 & -b_i & a_i & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & 0 & a_i & b_i \\ 0 & 0 & \dots & \dots & 0 & 0 & -b_i & a_i \end{pmatrix} \quad (6)$$

є приведеною.

Таким чином, якщо жорданова форма матриці містить жордановий блок розміру  $k$ , тоді йому може відповідати блок розміром  $0, 1, \dots, k$  у жордановій формі приведеної матриці. Розглядаючи прями суми таких блоків, отримуватимемо різні жорданові форми приведених матриць. Покажемо, що всі отримані таким чином жорданові форми приведених матриць дійсно покривають всі можливі жорданові форми приведених матриць.

**Теорема 2.** Нехай  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  — заданий лінійний оператор,  $\Lambda \in \mathbb{R}^{m \times m}$  — деяка матриця. Нехай для кожного дійсного власного числа  $\lambda$  та розміру  $k$  оператор  $A$  має в жордановій формі  $n_k(\lambda)$  жорданових блоків  $J_k(\lambda)$

(розміром  $k \times k$ ), а оператор  $\Lambda$  —  $c_k(\lambda)$  жорданових блоків. Нехай  $n_k(a, b)$  — це кількість жорданових блоків  $J_k(a_i, b_i)$  оператора  $A$  для кожної пари комплексних власних чисел  $a \pm bi$ ,  $c_k(a, b)$  — кількість таких блоків у жордановій формі  $\Lambda$ .

Нехай  $K_A(\lambda)$  — максимальний розмір жорданового блоку оператора  $A$  з дійсним власним значенням  $\lambda$ ,  $K_A(a, b)$  — половина максимального розміру жорданового блоку оператора  $A$  для пари комплексних власних чисел  $a \pm bi$ ,  $K_\Lambda(\lambda)$ ,  $K_\Lambda(a, b)$  — відповідні значення для  $\Lambda$ . Якщо  $\lambda$  чи  $a \pm bi$  не є власними числами оператора, тоді відповідні значення вважаємо нулями.

$\Lambda$  — є правою приведеною матрицею  $A$  тоді й лише тоді, якщо виконуються умови:

1) Множина власних чисел (дійсних і пар комплексних)  $\Lambda$  є підмножиною власних чисел  $A$ .

2) Для довільного  $\lambda$  — дійсного власного числа  $A$ :

$$K_\Lambda(\lambda) \leq K_A(\lambda),$$

$$\begin{cases} c_1(\lambda) + \dots + c_{K_A(\lambda)}(\lambda) \leq n_1(\lambda) + \dots + n_{K_A(\lambda)}(\lambda), \\ c_2(\lambda) + \dots + c_{K_A(\lambda)}(\lambda) \leq n_2(\lambda) + \dots + n_{K_A(\lambda)}(\lambda), \\ \dots \\ c_{K_A(\lambda)}(\lambda) \leq n_{K_A(\lambda)}(\lambda). \end{cases}$$

3) Для довільної пари комплексних власних чисел  $a \pm bi$  оператора  $A$ :

$$K_\Lambda(a, b) \leq K_A(a, b),$$

$$\begin{cases} c_1(a, b) + \dots + c_{K_A(a, b)}(a, b) \leq n_1(a, b) + \dots + n_{K_A(a, b)}(a, b), \\ c_2(a, b) + \dots + c_{K_A(a, b)}(a, b) \leq n_2(a, b) + \dots + n_{K_A(a, b)}(a, b), \\ \dots \\ c_{K_A(a, b)}(a, b) \leq n_{K_A(a, b)}(a, b). \end{cases}$$

Отже, теорема 2 для заданого оператора описує повну множину його приведених матриць за допомогою опису множини можливих жорданових форм приведених матриць та леми 7. Таким чином, отримано опис усіх класів еквівалентності приведених матриць для заданого оператора. Проте, кількість таких класів еквівалентності є достатньо великою і швидко зростає з розмірністю оператора. Проте, наступна лема показує, що немає необхідності розглядати всі можливі класи еквівалентності — достатньо дослідити певний набір базових матриць.

**Лема 17.** Нехай  $\Lambda = \begin{pmatrix} \Lambda_1 & 0 & \dots & 0 \\ 0 & \Lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \Lambda_k \end{pmatrix} \in \mathbb{R}^{m \times m}$  — приведена матриця

оператора  $A$ , що є прямою сумою матриць  $\Lambda_1 \in \mathbb{R}^{m_1 \times m_1} \dots \Lambda_k \in \mathbb{R}^{m_k \times m_k}$ ,  $m_1 + \dots + m_k = m$ . Тоді  $\Lambda_j, j = 1 \dots k$  — приведені матриці оператора  $A$ , а множина розв'язків  $S_{A, \Lambda}$  є прямою сумою множин розв'язків  $S_{A, \Lambda_1}, S_{A, \Lambda_k}$ .

Таким чином, лема 17 зводить знаходження параметризацій  $S$  для приведеної матриці, що є прямою сумою інших матриць, до знаходження відповідних параметризацій для цієї множини матриць. Оскільки за теоремою 2 кожна приведена матриця має жорданову форму, яка складається з прямої суми певних жорданових блоків, з певними обмеженнями на їх кількість, тому є сенс досліджувати лише певний набір базових приведених матриць  $\Lambda$  та параметризації розв'язків  $S_{A,\Lambda}$ , оскільки для всіх інших приведених матриць відповідні параметризації можна знайти за допомогою операцій прямої суми та зміни базису.

Для зручності введення поняття базової матриці від жорданової нормальної форми перейдемо до фробеніусової нормальної форми. Це дозволить уніфікувати два випадки: жорданові клітини, що відповідають лінійним та квадратичним множникам, будуть розглядатись в рамках однієї системи. Крім того, це дозволить виписати явний вигляд параметризації розв'язків  $S_{A,\Lambda}$ .

Нагадаємо, що супутньою матрицею монома  $p(\lambda) = c_0 + c_1\lambda + \dots + c_{n-1}\lambda^{n-1} + \lambda^n$  називають матрицю

$$C(p) = \begin{bmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{bmatrix}. \quad (7)$$

Використаємо такі факти:

- Жорданова клітина вигляду (5)  $J_k(\lambda_i)$  подібна супутній матриці полінома  $C(\lambda) = (\lambda - \lambda_i)^k$ .

- Жорданова клітина вигляду (6)  $J_k(a_i, b_i)$  подібна супутній матриці полінома  $C(\lambda) = (\lambda^2 - 2a_i\lambda + (a_i^2 + b_i^2))^k$ .

Виходячи з цього, введемо означення базової приведеної матриці.

**Означення 3.** Нехай  $A \in \mathbb{R}^{n \times n}$  - лінійний оператор. Нехай  $M_A(\lambda) = p_1^{n_1}(\lambda)p_2^{n_2}(\lambda) \dots p_k^{n_k}(\lambda)$  - мінімальний поліном оператора  $A$  та його розклад на степені лінійних та квадратичних множників. Для кожного  $p_i$  та кожного  $1 \leq k \leq n_k$  матрицю  $C(p_i^k)$  назвемо базовою приведеною матрицею.

**Лема 18.** Кожна базова приведена матриця є приведеною матрицею.

**Лема 19.** Кожна приведена матриця подібна прямій сумі базових приведених матриць.

**Теорема 3.** Нехай  $A \in \mathbb{R}^{n \times n}$  - лінійний оператор,  $p_i$  - лінійний або незвідний квадратичний дільник мінімального полінома, що входить до його розкладу зі степенню  $n_i$ ,  $k$  - натуральне число, що не перевищує  $n_i$ ,  $m = \deg(p_i^k)$  (тобто  $m$  може дорівнювати  $k$  або  $2k$ ),  $p_i^k(\lambda) = \lambda^m + a_1\lambda^{m-1} + \dots + \lambda_m$ . Нехай  $V = \text{Ker}(p_i^k(A))$ . Тоді  $\Lambda = C(p_i^k) \in \mathbb{R}^{m \times m}$  - базова незвідна матриця, а параметризація розв'язків рівності має вигляд

$$S_{A,\Lambda} = \{ [v|Av| \dots |A^{m-1}v] \mid v \in V \}.$$

Алгоритм знаходження базових приведених матриць оператора  $A$  та відповідних параметризацій  $S_{A,\Lambda}$ .

1) Знайти факторизацію мінімального полінома оператора  $A$ :  $M_A(\lambda) = p_1^{n_1}(\lambda)p_2^{n_2}(\lambda)\dots p_k^{n_k}(\lambda)$  на степені лінійних та незвідних квадратичних множників.

2) Для кожного множника  $p_i$  (що має степінь  $n_i$ ) та кожної степені  $1 \leq j \leq n_i$  базовою приведеною матрицею є  $C(p_i^j)$  (рівність (7)).

3) Для кожного полінома  $p_i^j$  та відповідної базової приведеної матриці  $\Lambda$  знайти базис  $V \in \mathbb{R}^{n \times k}$  (базисні вектори розташовані по стовпцям матриці  $V$ ) інваріантного підпростору  $\text{Ker}(p_i^j(A)) = \{V\alpha \mid \alpha \in \mathbb{R}^k\}$ .

4) Тоді параметризація розв'язків  $S_{A,\Lambda}$  така:

$$S_{A,\Lambda} = \left\{ [V\alpha \mid AV\alpha \mid \dots \mid A^{j_{n_i}-1}V\alpha] \mid \alpha \in \mathbb{R}^k \right\}.$$

Розглянемо природне узагальнення міркувань, наведених вище, для випадку одночасного застосування декількох операторів.

**Означення 4.** Нехай  $A_1 \dots A_l \in \mathbb{R}^{n \times n}$  — набір операторів. Набір матриць  $\Lambda_1 \dots \Lambda_l \in \mathbb{R}^{m \times m}$  назовемо приведеними, якщо існує  $S \in \mathbb{R}^{n \times m}$ , така що  $\text{rank}(S) = m$  та виконуються рівності:

$$\begin{cases} A_1 S = S \Lambda_1, \\ \dots \\ A_l S = S \Lambda_l. \end{cases} \quad (8)$$

**Лема 20.** Нехай  $(\Lambda_1 \dots \Lambda_l)$  — набір приведених матриць деяких операторів. Тоді для довільної невинродженої матриці  $G \in \mathbb{R}^{m \times m}$  набір матриць  $(G^{-1}\Lambda_1 G \dots G^{-1}\Lambda_l G)$  є приведеним.

Задача пошуку всіх наборів приведених матриць для довільного набору операторів наразі не є вирішеною. Проте, можна знайти деяку підмножину приведених матриць та відповідних параметризацій використовуючи базові приведені матриці для кожного оператора  $A_1 \dots A_l \in \mathbb{R}^{n \times n}$  окремо. Для цього необхідно розглянути всі можливі комбінації приведених матриць окремих операторів (використовуючи операцію прямої суми між базовими матрицями одного оператора для узгодження розмірів матриць між різними операторами), та для кожної комбінації знайти розв'язок системи (8) для знаходження параметризацій ядер.

### 3. Деякі часткові випадки

Проаналізуємо деякі лінійні оператори на зображеннях з точки зору описаного вище підходу.

Припустимо ми використовуємо згорткову нейронну мережу з ядрами згортки розміром  $3 \times 3$  (найпоширеніший випадок на практиці). Розглянемо оператор симетричного відображення зображення навколо вертикальної осі  $A_s$ . Розглянемо дію цього оператора на патч зображення розміром  $3 \times 3$ .

$$I = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix}, \quad A_s I = \begin{bmatrix} I_{13} & I_{12} & I_{11} \\ I_{23} & I_{22} & I_{21} \\ I_{33} & I_{32} & I_{31} \end{bmatrix}.$$

Якщо розглядати патч зображення  $I$  як вектор довжини 9, тоді на  $A_s$  можна дивитись як на лінійний оператор  $A_s : \mathbb{R}^9 \rightarrow \mathbb{R}^9$ .

Таким чином матриця лінійного оператора  $A_s$  має вигляд:

$$A_s = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Матриця  $A_s$  симетрична:  $A_s^T = A_s$ , отже, можемо шукати праві придені матриці оператора.

Характеристичний поліном матриці  $A_s$ :

$$\chi_{A_s}(\lambda) = -(\lambda - 1)^6(\lambda + 1)^3.$$

Оскільки  $A_s^2 = I_9$  (оскільки подвійне застосування оператора симетричного відображення - це одиничний оператор), то мінімальний поліном оператора такий:

$$M_{A_s}(\lambda) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Оскільки мінімальний поліном містить лише лінійні множники піднесені до першого степеня, тому оператор діагоналізований.

Власні вектори для  $\lambda_1 = -1$ :

$$\left\{ v_1^{-1} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_2^{-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_3^{-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Тоді множина розв'язків системи  $AS = S\Lambda$  для  $\Lambda = [-1]$  задається як  $\text{span}(v_1^{-1}, v_2^{-1}, v_3^{-1})$ . Таким чином, кожен розв'язок подається у вигляді

$$av_1^{-1} + bv_2^{-1} + cv_3^{-1} = \begin{bmatrix} -a \\ 0 \\ a \\ -b \\ 0 \\ b \\ -c \\ 0 \\ c \end{bmatrix}.$$

Записуючи це знову як матрицю  $3 \times 3$ , отримаємо параметризацію ядра згортки:

$$\begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -c & 0 & c \end{bmatrix}. \quad (9)$$

Для власного числа  $\lambda = 1$  власні вектори такі:

$$\left\{ v_1^1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_2^1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_3^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_4^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; v_5^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; v_6^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Аналогічним чином отримуємо загальну параметризацію ядра згортки для  $\lambda = 1$  у вигляді матриці:

$$\begin{bmatrix} d & e & d \\ f & g & f \\ h & i & h \end{bmatrix}. \quad (10)$$

Таким чином, для  $\lambda = 1$  маємо 6-вимірний, а для  $\lambda = -1$  3-вимірний підпростір ядер згортки.

Таким чином отримуємо дві базових приведених матриці для оператора симетричного відображення:  $\Lambda = [-1]$  та  $\Lambda = [1]$  яким відповідають відповідно антисиметричне та симетричне ядра згорток.

Проаналізуємо базові приведені матриці оператора. Якщо  $\Lambda = [1]$  отримаємо, що величини, отримані при застосуванні відповідного параметризованого ядра (10), яке є симетричним відносно вертикальної осі, не змінюються при симетричному відображенні зображення. Таким чином, ядра з такою параметризацією можна застосовувати для знаходження величин на зображенні, які повинні бути інваріантними при симетричних перетвореннях. Використовуючи параметризоване ядро (9) для  $\lambda = [-1]$  отримаємо

величину, яка змінює знак при симетричному відображенні зображення. Таким чином, такі величини можна використовувати для знаходження орієнтації об'єктів на зображенні (визначення напрямку об'єкта — «вліво» чи «вправо»).

Розглянемо аналогічним чином оператор повороту  $A_R$  на  $\pi/2$  (за годинниковою стрілкою), означеному на патчах зображень  $3 \times 3$ .

$$I = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix}, \quad A_R(I) = \begin{bmatrix} I_{31} & I_{21} & I_{11} \\ I_{32} & I_{22} & I_{12} \\ I_{33} & I_{23} & I_{13} \end{bmatrix}.$$

Знову представляючи зображення у вигляді вектора-стовпця, отримаємо, що матриця оператора  $A_R$  має вигляд:

$$A_R = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Оператор  $A_R$  — ортогональний, тому шукаємо праві приведені матриці та відповідні параметризації, і потім перетворимо їх за допомогою таблиці 2.

Характеристичний поліном:

$$\chi_{A_R}(\lambda) = -(\lambda + 1)^2(\lambda - 1)^3(\lambda^2 + 1)^2.$$

Мінімальний поліном:

$$M_{A_R}(\lambda) = (\lambda - 1)(\lambda + 1)(\lambda^2 + 1) = \lambda^4 - 1.$$

Це дійсно мінімальний поліном, оскільки, з одного боку, до нього входять всі незвідні множники характеристичного поліному в першій степені (отже, не існує полінома меншого степеня, який занулює цю матрицю), а з іншого боку  $A_R^4 = I_9$ , оскільки операція повороту на  $\pi/2$  циклічна з періодом 4.

Власні вектори для  $\lambda = -1$ :

$$\left\{ v_1^{-1} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}; v_2^{-1} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Параметризуючи власний підпростір для  $\lambda = -1$  (беручи лінійну комбінацію  $av_1^{-1} + bv_2^{-1}$ ) та записуючи результуючий вектор у вигляді матриці, отримаємо параметризоване ядро згортки:

$$\begin{bmatrix} b & a & -b \\ -a & 0 & a \\ -b & a & b \end{bmatrix}.$$

Власні вектори для  $\lambda = 1$ :

$$\left\{ v_1^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; v_2^1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}; v_3^1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Параметризуючи власні вектори з коефіцієнтами  $c, d, e$  отримуємо параметризоване ядро:

$$\begin{bmatrix} e & d & e \\ d & c & d \\ e & d & e \end{bmatrix}.$$

Розглянемо множник  $\lambda^2 + 1$  мінімального полінома. Він відповідає інваріантному підпростору, базис якого може бути знайдений як  $Ker(A^2 + I)$ .

Лінійний підпростір  $Ker(A^2 + I)$  є чотирьохвимірним, і може бути параметризованим як:

$$\begin{bmatrix} f & g & h \\ -i & 0 & i \\ -h & -g & -f \end{bmatrix}. \quad (11)$$

Тоді, застосовуючи оператор  $A$  до (11), отримаємо параметризацію пари ядер згорток:

$$\begin{bmatrix} f & g & h \\ -i & 0 & i \\ -h & -g & -f \end{bmatrix}, \begin{bmatrix} -h & i & f \\ -g & 0 & g \\ -f & -i & h \end{bmatrix}.$$

що відповідають приведеній матриці  $\Lambda = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ .

Використовуючи таблицю 2 отримаємо, що для знаходження лівої приведенної матриці, необхідно транспонувати і обернути  $\Lambda$ , і транспонувати параметризацію. Але  $\Lambda = (\Lambda^T)^{-1}$ , а транспонування  $S$  не змінює параметризації — лише представлення елементів ядра згортки зі стовпців на рядки.

Знайдемо ядра згорток, які є керованими одночасно для оператора повороту на  $\pi/2$  та симетричного відображення. Зауважимо, що комбінація цих операторів генерує групу перетворень  $D_4$ . Щоб знайти базові матриці для пари операторів, необхідно розглядати комбінації базових матриць  $(\Lambda_1, \Lambda_2)$

для кожного з операторів, і для кожної комбінації шукати параметризацію ядер, які поводять себе відповідно до  $\Lambda_1$  під дією першого оператора та відповідно до  $\Lambda_2$  під дією другого оператора. В таблиці 3 для кожної пари базових матриць наведені відповідні параметризації ядер. Зауважимо, що оскільки обидва оператори є ортогональними, ми знову шукаємо праві приведені матриці для обох операторів.

ТАБЛ. 3. Базові приведені матриці операторів повороту та симетричного відображення.

$\Lambda_{sym}$	$\Lambda_{rot}$	Параметризації ядер	
[1]	[1]	$\begin{bmatrix} a & b & a \\ b & c & b \\ a & b & a \end{bmatrix}$	
[-1]	[1]	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
[1]	[-1]	$\begin{bmatrix} 0 & a & 0 \\ -a & 0 & -a \\ 0 & a & 0 \end{bmatrix}$	
[-1]	[-1]	$\begin{bmatrix} a & 0 & -a \\ 0 & 0 & 0 \\ -a & 0 & a \end{bmatrix}$	
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -a & 0 & a \end{bmatrix}$	, $\begin{bmatrix} -a & -b & -a \\ 0 & 0 & 0 \\ a & b & a \end{bmatrix}$
$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} a & b & a \\ 0 & 0 & 0 \\ -a & -b & -a \end{bmatrix}$	, $\begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -a & -0 & a \end{bmatrix}$
$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

#### 4. Висновки: ПОБУДОВА ЗГОРТКОВИХ АРХІТЕКТУР З БАЖАНОЮ ПОВЕДІНКОЮ ВИХОДУ

Вищеописані методи дозволяють будувати архітектури згорткових нейронних мереж, що змінюються відповідно до заданих перетворень під дією геометричних перетворень входу. Для цього необхідно побудувати мережу як суперпозицію двох мереж. Перша мережа — це group convolutional

neural network відповідно до необхідної гупи перетворень (наприклад, 4-елементної групи поворотів на  $\pi/2$  чи 8-елементної групи поворотів та симетричних відображень), виходом якої є тензор, що еквіваріантно змінюється відносно перетворень входу. Друга частина мережі - це набір згорток, параметризованих відповідно до бажаної поведінки виходу (прикладі таких параметризацій наведені в Таблиці 3). Виходом суперпозиції підмережі та відповідним чином параметризованого шару згортки є тензор, кожен елемент якого змінюється відповідно до бажаної поведінки.

Вищеописаний метод є достатньо загальним, та може бути застосованим для геометричних перетворень, які є лінійними відносно значень пікселів, а бажана поведінка виходу нейронної мережі описується приведеною матрицею оператора геометричного перетворення входу.

#### ЛІТЕРАТУРА

1. Sobel I., Feldman G. A 3x3 Isotropic Gradient Operator for Image Processing. *Pattern Classification and Scene Analysis*. 1973. P. 271–272.
2. Kanopoulos N., Vasanthavada N., Baker R. L. Design of an image edge detection filter using the Sobel operator. *IEEE Journal of Solid-State Circuits*. 1988. 23 (2). P. 358–367.
3. Goodfellow I., Bengio Y., Courville A. Deep learning. MIT press. 2016.
4. Jaderberg M., Simonyan K., Zisserman A., Kavukcuoglu K. Spatial Transformer Networks. arXiv: 1506.02025, 2015
5. Diego Marcos, Michele Volpi and Devis Tuia: Learning rotation invariant convolutional filters for texture classification. *arXiv: 1604.06720*. 2016.
6. Bruna J., Mallat S. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. 2013. 35 (8). P. 1872–1886.
7. Hinton G. E., Krizhevsky A., Wang S. D. Transforming autoencoders. *ICANN-11: International Conference on Artificial Neural Networks*. Springer. 2011. P. 44–51.
8. Kivinen J. J., Williams Ch. K. I. Transformation equivariant Boltzmann machines. In: *21st International Conference on Artificial Neural Networks*. 2011.
9. Sohn K., Lee H. Learning Invariant Representations with Local Transformations. *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*. 2012.
10. Skibbe H. Spherical Tensor Algebra for Biomedical Image Analysis. PhD thesis. Albert-Ludwigs-Universität Freiburg im Breisgau. 2013.
11. Lenc K., Vedaldi A. Understanding image representations by measuring their equivariance and equivalence. In: *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2015.
12. Gens, R. and Domingos, P. Deep Symmetry Networks. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
13. Cohen T. S., Welling M. Group Equivariant Convolutional Networks. *arXiv: 1602.07576*. 2016.
14. Cohen T. S., Welling M. Steerable CNNs. *arXiv:1612.08498*. 2016.
15. Weiler M., Geiger M., Welling M., Boomsma W. Cohen T. 3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data. *arXiv:1807.02547*. 2018.

16. Cohen T. S., Geiger M., Weiler M. A general theory of equivariant cnns on homogeneous spaces. In: *Advances in Neural Information Processing Systems*. 2019. P. 9142–9153.
17. Sabour S., Frosst N., Hinton G. E. Dynamic Routing Between Capsules. *arXiv:1710.09829*. 2017.
18. Hinton G. E., Sabour S., Frosst N. Matrix capsules with EM routing. *ICLR*. 2018.

Надійшла: 14.05.2018 / Прийнята: 02.06.2020

## ПОСТРОЕНИЕ АРХИТЕКТУР НЕЙРОННЫХ СЕТЕЙ С ЖЕЛАЕМОМ ПОВЕДЕНИЕМ ПОД ДЕЙСТВИЕМ ГЕОМЕТРИЧЕСКИХ ПРЕОБРАЗОВАНИЙ ВХОДА

В. В. ДУДАРЬ, В. В. СЕМЕНОВ

Киевский национальный университет имени Тараса Шевченка, Киев, Украина,  
E-mail: slavko123@ukr.net, semenov.volodya@gmail.com

**АННОТАЦИЯ.** В работе предложен общий метод для анализа сверточных слоев под действием геометрических преобразований входа, которые можно линейно параметризовать относительно значений пикселей изображения. Предложен метод для нахождения всех возможных типов поведения выхода сверточных слоев под действием преобразований входа, и соответственных параметризаций ядер, а также общий метод для построения сверточных нейронных сетей с желаемым поведением выхода под действием геометрических преобразований входа.

**КЛЮЧЕВЫЕ СЛОВА:** сверточная нейронная сеть, инвариантность сверточных нейронных сетей, эквивариантность сверточной сети, управляемые ядра сверток, сужения линейных операторов, инвариантные подпространства.