

ВІДГУК

офіційного опонента на дисертаційну роботу ТКАЧЕНКА Олександра Миколайовича «**ПОЛІНОМІАЛЬНІ МЕТОДИ ТА ЗАСОБИ ОЦІНЮВАННЯ ПАРАМЕТРІВ РЕГРЕСІЇ З ВИКОРИСТАННЯМ МОДЕЛЕЙ НЕГАУСОВИХ ПОМИЛОК**», яку подано на здобуття наукового ступеня доктора філософії зі спеціальності 122 — Комп'ютерні науки.

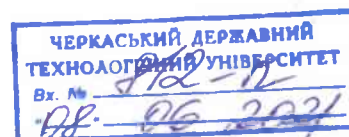
На експертизу представлено дисертаційну роботу обсягом 138 стор. і опубліковані за темою дисертації 8 публікації.

Предметом дисертаційного дослідження є *математичні моделі адитивної взаємодії детермінованої складової регресивних залежностей і негаусових випадкових регресійних помилок, засновані на застосуванні статистик вищих порядків (моментів та кумулянтів), а також методи і засоби моделювання процесів оцінювання інформативних параметрів регресії, орієнтовані на створення засобів їх комп'ютерної реалізації.*

Дисертаційна робота ТКАЧЕНКА О. М. виконувалась в Черкаському державному технологічному університеті (ЧДТУ) відповідно до основних наукових напрямів та найважливіших проблем фундаментальних досліджень у галузі природничих, технічних, суспільних і гуманітарних наук НАНУ на 2019–2023 роки, зокрема: «Дослідження математичних моделей, проблем комп'ютерної математики, оптимізації, оцінювання, ідентифікації», в рамках держбюджетної науково-дослідної роботи (НДР) «Розробка технології та пристроїв адитивного виробництва індивідуальних хірургічних імплантатів та протезів з біосумісних полімерних матеріалів» (державна реєстрація № 0117U000937).

1. АКТУАЛЬНІСТЬ ТЕМИ

На сучасному етапі розвитку науково-технічної революції регресійний аналіз розглядається як одна із найбільш поширених груп задач, що вирішуються в межах сучасного прикладного напрямку розвитку комп'ютерних наук та інформаційних технологій відомого як «Наука про Дані» (Data Science). В цьому сенсі опрацювання значних обсягів даних, що генерується в усіх галузях діяльності людини (наука, виробництво, культура, суспільне життя, тощо), не можливо представити без використання обчислювальної техніки і відповідних програмних засобів, орієнтованих на комп'ютерне моделювання і статистичний аналіз. Серед найбільш поширених програмних засобів слід, насамперед, виділити, наприклад, спеціалізовані комерційні продукти типу MATLAB, SAS, STATISTICA та високорівневі мови програмування, зокрема, R та Python, орієнтовані на задачі аналізу даних. Формалізовану основу зазначених програмних інструментальних засобів є різноманітні *моделі та обчислювальні методи*. Основним фактором подальшого розвитку існуючих та розробки нових моделей та обчислювальних методів при створенні і застосуванні інформаційних технологій слід розглядати відповідні критерії, пов'язані з



підвищенням точності опрацювання реальних даних з урахуванням специфіки останніх.

В ретроспективному плані оцінювання інформативних параметрів регресійних залежностей є однією із самих ранніх задач статистичного аналізу, яка поширюється на прикладні задачі технічного, геофізичного, біомедичного, економічного та інших аспектів людської діяльності. У більшості з них для віднаходження оцінок параметрів застосовуються різні варіації методу найменших квадратів (МНК), теоретичні основи якого започатковано німецьким математиком та механіком Гаусом. Хоча МНК дозволяє отримувати розв'язки у замкнутій формі, не потребує додаткової апріорної інформації про ймовірнісні властивості моделі похибок, однак в реальних ситуаціях застосування МНК може давати оцінки з невисокою точністю, що зумовлено суттєвим відхиленням розподілу даних від нормального (гаусового) закону.

Дослідження останніх років показують, що намагання підвищення точності методу МНК стикається зі значним ускладненням алгоритмів статистичного аналізу в умовах досягнення бажаної повноти ймовірнісного опису експериментальних даних.

Враховуючи вказане вище, *актуальність* чинної дисертаційної роботи зумовлено необхідністю модифікації існуючих регресійних моделей, які містять адитивну суміш детермінованої та негаусової випадкової складової у вимірювальній інформації. Крім того гостро постає задача забезпечення адаптивності алгоритмам обробки інформації в умовах апріорної невизначеності щодо властивостей випадкової складової регресійних моделей, яка є характерним чинником для реальних ситуацій використання таких моделей. Останнє також посилює *актуальність досліджень*, виконаних у дисертаційній роботі, що представляється до захисту.

2. ОСНОВНІ НАУКОВІ РЕЗУЛЬТАТИ

У дисертаційній роботі ТКАЧЕНКА О. М. вирішено науково-технічну задачу, мета якої полягає у створенні методів і засобів математичного і комп'ютерного моделювання процесів оцінювання параметрів регресійних залежностей шляхом адаптації ймовірнісних моделей досліджуваних даних для покращення точності отримуваних оцінок на основі урахування відмінностей статистичних властивостей регресійних помилок від гаусової ідеалізації.

Зокрема автором отримано такі наукові результати.

Вперше розроблено:

- методи адаптивного поліноміального оцінювання параметрів регресії, які за рахунок використання удосконалених моделей регресійних залежностей в умовах апріорної невизначеності щодо ймовірнісного характеру негаусових регресійних помилок дозволяють підвищити точність оцінок параметрів регресії.

Удосконалено:

- моделі регресійних залежностей на основі використання статистик вищих порядків для опису їх випадкової складової, що дозволяє реалізаційно просто враховувати негаусовість статистичних даних.

Отримали подальший розвиток:

- елементи теорії оцінювання параметрів регресії з використанням методу максимізації полінома, що дозволяє отримувати субефективні, адаптивні і компромісні з точки зору практичної реалізації рішення з оцінювання інформативних параметрів регресії за негаусового характеру випадкових помилок.

Із зазначеного вище можна зробити висновок, що наукова новизна дисертаційної роботи ТКАЧЕНКА О. М. відповідає вимогам п. 13 діючого «Порядку присудження наукових ступенів».

Автором отримано та опубліковано нові науково-обґрунтовані результати, які, у сукупності, розв'язують важливу наукову-прикладну задачу. Робота містить нові, раніше не захищені наукові положення.

Оформлення дисертації відповідає вимогам ДСТУ. Мова і стиль викладення дисертації коректно висвітлюють одержані науково-практичні результати, які відповідають меті досліджень.

3. ПРАКТИЧНА ЦІННІСТЬ РЕЗУЛЬТАТІВ РОБОТИ

Практична цінність виконаних досліджень підтверджується наведеними в роботі результатами прикладного впровадження (у навчальний процес ЧДТУ) розроблених в дисертаційній роботі моделей, методів та програмних засобів.

При цьому слід зазначити, що практична корисність роботи полягає у доведенні отриманих наукових результатів до конкретних інженерних методик, алгоритмів і комп'ютерних програм, які може бути використано для адаптивного оцінювання параметрів однофакторних і багатфакторних лінійних, поліноміальних та нелінійних регресійних залежностей

Результати дисертаційного дослідження ТКАЧЕНКА О.М. використано в учбовому процесі на кафедрі радіотехніки, телекомунікаційних і робототехнічних систем ЧДТУ при викладанні спецкурсу «Теорія нелінійної статистичної радіотехніки».

4. ДОСТОВІРНІСТЬ ОТРИМАНИХ РЕЗУЛЬТАТІВ

Достовірність наукових результатів роботи підтверджується обґрунтованим коректним використанням наведеного у розділах 2 та 3 математичного апарату теорії матриць та регресійного аналізу, а також наведеними у розділі 4 результатів порівняльного аналізу теоретичних значень параметрів лінійної регресії з результати статистичного моделювання, зокрема, при асиметрії регресійних помилок.

Основні положення та результати дисертаційного дослідження пройшли апробацію на Міжнародних та Всеукраїнських науково-практичних

конференціях і семінарах та отримали підтримку науковців із даного напрямку досліджень.

5. СТРУКТУРА РОБОТИ

Дисертаційна робота структурно складається із вступу, чотирьох розділів з висновками, загальних висновків, списку використаних джерел і додатків. Загальний обсяг дисертаційної роботи становить 138 сторінок, у тому числі 123 сторінок основного тексту, ілюстрованого 17 рисунками і 8 таблицями.

Вступ містить обґрунтування актуальності роботи; об'єкт, предмет, мету та задачі дисертаційного дослідження; наукову новизну та практичне значення отриманих результатів, а також і інші відомості щодо дисертаційної роботи.

Перший розділ присвячено формулюванню проблеми досліджень та аналітичному огляду підходів до знаходження оцінок параметрів регресії при негаусовому розподілі помилок. При цьому, зокрема, розглянуто задачі та моделі регресійного аналізу; програмні засоби для регресійного аналізу; класичні методи оцінювання параметрів регресії; способи оцінювання параметрів за умов негаусового розподілу регресійних помилок; особливості застосування стохастичних поліномів Кунченка для задач регресійного аналізу.

У **другому розділі** виконано синтез моделей та методів поліноміального оцінювання параметрів регресії. Основну увагу приділено застосуванню методу максимізації полінома (ММПл) при знаходженні оцінок векторного параметра при неоднаково розподілених даних. Показано, що при використанні поліномів степеня $S \geq 2$ віднаходження ММПл-оцінок векторного параметру у переважній більшості випадків (аналогічно до ситуації із використанням методу максимальної правдоподібності (ММП)) потребує застосування числових методів розв'язку систем нелінійних рівнянь. Зокрема, часто використовується підхід, який базується на ітераційній числовій процедурі Ньютона-Рафсона, в основі якого лежить принцип лінеаризації шляхом розкладання лівої частини кожного нелінійного рівняння системи, що описує шукані оцінки параметра (який визначається), у ряд Тейлора в околі істинного значення вектору параметрів інформаційного сигналу.

Виконано постановку задачі оцінювання параметрів лінійної регресії та показано застосування методу ММПл для оцінювання параметрів лінійної регресії при різних степенях поліному.

Теоретично підтверджено, що При симетричному характері розподілу регресійних помилок для оцінювання параметрів доцільним є застосування стохастичних поліномів лише при степені $S=3$. За відсутності апріорної інформації про властивості випадкової складової регресійної моделей, що є характерним для реальних ситуацій, може бути застосований адаптивний

підхід, заснований на використанні апостеріорних оцінок статистик вищих порядків регресійних МНК-залишків

У **третьому розділі** проведено аналіз теоретичної ефективності поліноміальних оцінок параметрів регресії. Отримано асимптотичні дисперсії ММПл-оцінок векторного параметра при неоднаково розподілених даних та здійснено порівняльний аналіз теоретичної ефективності ММПл-оцінок із МНК-оцінками.

Важливим теоретичним результатом досліджень стало твердження того, що ступінь ефективності ММПл-оцінок (відносно МНК) не залежить від типу регресійних моделей та є однаковою для всіх складових векторного параметру, який оцінюється. При цьому теоретична величина коефіцієнтів зменшення дисперсії ММПл-оцінок залежить від ступеня негаусовості випадкової складової регресійної моделі, а зростання відносної точності ММПл-оцінок може бути достатньо суттєвим із наближенням величини кумулянтних коефіцієнтів до меж своїх областей допустимих значень.

Четвертий розділ присвячено створенню програмних засобів статистичного моделювання поліноміального оцінювання параметрів регресії та практиці їх (засобів) застосування у прикладних дослідженнях.

Детально розроблено структуру та функціонал програмних засобів поліноміального регресійного аналізу, та наведено результати розв'язування типових задач статистичного моделювання при оцінюванні параметрів лінійної регресії.

Сукупність отриманих результатів статистичного моделювання на основі багаторазових випробувань методом Монте-Карло, у цілому, підтвердило теоретично доведену у розділі 3 ефективність поліноміальних оцінок відносно лінійних МНК-оцінок для всіх проаналізованих типів регресійних моделей.

Список використаних джерел складається із 117 наукових праць.

У **додатку А** дисертаційної роботи наведено результати застосування поліноміальних моделей, методів та програмних засобів при регресійному аналізі реальних даних.

У **додатку Б** представлено список публікацій, в яких відображено основні наукові результати дисертації.

Додаток В містить Документи про впровадження дисертаційної роботи.

6. ПОВНОТА ВИКЛАДЕННЯ НАУКОВИХ ПОЛОЖЕНЬ, ВИСНОВКІВ І РЕКОМЕНДАЦІЙ, СФОРМУЛЬОВАНИХ У ДИСЕРТАЦІЇ, ТА ОПУБЛІКОВАНИХ В РОБОТАХ АВТОРА

Основні результати дисертаційного дослідження опубліковано у 8 наукових роботах. Серед них: 1 стаття у фаховому виданні України (*категорія А, індексовано у наукометричній базі Scopus*), 2 публікації — у зарубіжних наукових періодичних виданнях (*індексовано у наукометричній базі Scopus*), 5 тез доповідей у матеріалах наукових конференцій.

Наукові положення, висновки і рекомендації дисертаційної роботи відображено в публікаціях рівномірно по розділах. Одночасно, вони відбивають пріоритети автора дисертації у розв'язуванні наукових і прикладних задач обраного ним наукового напрямку досліджень.

Публікації ТКАЧЕНКА О. М. задовольняють всім відповідним вимогам чинного «Порядку присудження наукових ступенів». Повнота відображення результатів дисертаційного дослідження і вимоги щодо кількості публікацій відповідає вимогам, які пред'являються до дисертацій на здобуття ступеня доктора філософії за обраною спеціальністю.

7. Рекомендації щодо використання результатів

Результати дисертаційного дослідження доцільно спрямувати в організації, які займаються проектуванням і виготовленням сучасних інформаційно-вимірвальних систем, систем діагностики, моніторингу, контролю, розвиток яких характеризується підвищеними вимогами до точності та якості обробки інформації, що дозволить підвищити ефективність їх функціонування.

8. Відсутність порушення академічної доброчесності

Дисертація містить результати власних досліджень здобувача. Порушень академічної доброчесності в дисертації та наукових публікаціях не виявлено. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело. Інші запозичення, виявлені в роботі є загальноновживаними і не є плагіатом.

9. Зауваження до дисертаційної роботи

1. У розділі 3 представлено результати порівняння дисперсій оцінок параметрів, отриманих різними методами, але не наведено аналітичні вирази дисперсії отриманих оцінок.
2. В роботі бажано було б розглянути вплив потенційної гомоскедактичності (нерівномірності дисперсії помилок) на ефективність запропонованих поліноміальних методів.
3. В роботі здійснено порівняння ефективності оцінок методу максимізації полінома з оцінками максимальної правдоподібності лише для симетричного випадку з використанням експоненціального степеневого розподілу. При цьому відсутнє подібне порівняння для асиметричних розподілів.
4. В тексті дисертації наявні скорочення, пояснення яких здійснюється значно пізніше по тексту, що викликає незручності при ознайомленні з дисертаційною роботою.
5. В тексті дисертації наявні стилістичні та технічні помилки, проте їх кількість цілком природна для друкованих робіт такого обсягу.

Однак, слід вказати, що наведені зауваження не знижують позитивного враження від дисертації і, частково, є побажаннями, які варто враховувати у подальшій роботі.

ВИСНОВОК

Зміст дисертаційної роботи ТКАЧЕНКА Олександра Миколайовича відповідає спеціальності 122 — Комп'ютерні науки.

Публікації автора у повному обсязі відображають результати виконаних досліджень. Крім того, повнота відображення результатів дисертаційного дослідження і кількість публікацій відповідають діючим чинним вимогам. Отримані результати наукових досліджень достовірні, достатньо обґрунтовані, мають наукову та прикладну значимість. На основі аналізу дисертаційної роботи та опублікованих робіт ТКАЧЕНКА О.М. вважаю, що здобувачем вирішено важливу науково-практичну задачу, яка полягає у створенні методів і засобів математичного і комп'ютерного моделювання процесів оцінювання параметрів регресійних залежностей шляхом адаптації ймовірнісних моделей досліджуваних даних для покращення точності отримуваних оцінок на основі урахування відмінностей статистичних властивостей регресійних помилок від гаусової ідеалізації.

Дисертаційну роботу оформлено із дотриманням необхідних вимог, прийнятих правил та норм. Дисертаційна робота ТКАЧЕНКА Олександра Миколайовича **«ПОЛІНОМІАЛЬНІ МЕТОДИ ТА ЗАСОБИ ОЦІНЮВАННЯ ПАРАМЕТРІВ РЕГРЕСІЇ З ВИКОРИСТАННЯМ МОДЕЛЕЙ НЕГАУСОВИХ ПОМИЛОК»** є завершеним науковим дослідженням, у якому поставлено і вирішено важливу науково-технічну задачу. Одержані результати мають наукову новизну і практичне значення.

Отже, дисертаційна робота за ступенем актуальності обраної теми, обґрунтованості основних наукових положень, висновків і рекомендацій, що були сформовані в роботі, їх новизни, повноти викладу в наукових публікаціях, відсутності порушення академічної доброчесності цілком відповідає вимогам «Порядку проведення експерименту з присудження ступеня доктора філософії», затвердженому Кабінету Міністрів України від 06.03.2019 р. № 167, а її автор — ТКАЧЕНКО Олександр Миколайович, заслуговує на присудження йому наукового ступеня доктора філософії за спеціальністю 122 Комп'ютерні науки.

Офіційний опонент:

Завідувач кафедри «Комп'ютеризовані системи управління» Державного університету «Одеська політехніка», доктор технічних наук, професор

 С. А. ПОЛОЖАЄНКО

Підпис професора ПОЛОЖАЄНКА С. А. засвідчую.
Вчений секретар Ради Державного університету «Одеська політехніка»
к.т.н., доцент

 В. І. ШЕВЧУК