



АФАНАСЬСВ

Євгеній Володимирович — доктор філософії (математика), науковий співробітник відділу математичної фізики Фізико-технічного інституту низьких температур ім. Б.І. Веркіна Національної академії наук України

ШТУЧНИЙ ІНТЕЛЕКТ, НЕЙРОННІ МЕРЕЖІ ТА СПЕКТРАЛЬНІ ВЛАСТИВОСТІ ВИПАДКОВИХ МАТРИЦЬ

**За матеріалами наукового повідомлення
на засіданні Президії НАН України
28 травня 2025 р.**

У доповіді наведено результати досліджень штучного інтелекту на основі нейронних мереж, властивостей нейронних мереж та використання теорії випадкових матриць. Ці результати, отримані у відділі математичної фізики Фізико-технічного інституту низьких температур ім. Б.І. Веркіна НАН України, можуть бути застосовані до побудови глибоких нейронних мереж.

Останніми роками технологія штучних нейронних мереж набула дуже стрімкого розвитку. Наприклад, у 2014 р. було створено штучний інтелект (ШІ), який міг описувати зображення одним реченням. Для однієї з фотографій було отримано такі описи. Людина: «Група чоловіків, які грають у фрісбі в парку». Програма: «Група молодих людей, які грають у гру фрісбі». Комп'ютер описав зображення з дивовижною точністю! Звісно, це був один із найбільш вдалих прикладів роботи цієї програми. У деяких випадках комп'ютер видавав фразу, яка була мало пов'язана або зовсім не пов'язана із зображенням. Однак факт залишався фактом — було створено програму, яка може дати досить точний опис зображення. І стало це можливим саме завдяки розвитку нейронних мереж.

У наведеному вище прикладі було використано дві нейронні мережі: мережу для розпізнавання образів та частину мережі для машинного перекладу текстів. Проте застосування нейронних мереж аж ніяк не обмежується розпізнаванням образів та машинним перекладом. Штучний інтелект використовують для розпізнавання мови, облич, створення фейкових відео- та аудіозаписів, персоналізації реклами, пошуку в Інтернеті, автоматичної торгівлі акціями на фондових ринках тощо.

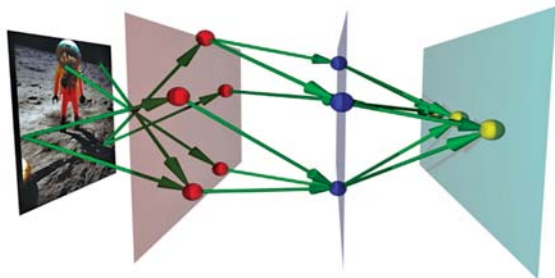


Рис. 1. Схема роботи штучних нейронних мереж. Нейронні мережі складаються з шарів, у кожному з яких є багато нейронів, тобто комірок пам'яті. Під час обробки інформації в нейрони пошарово записуються числа. Між сусідніми шарами нейронів є зв'язки, які визначають, як із чисел попереднього шару отримати числа наступного. Кожен зв'язок може бути більш або менш міцним, і ця міцність описується відповідним ваговим коефіцієнтом. Саме ці вагові коефіцієнти й визначають, як саме нейронна мережа обробляє вхідну інформацію

Зокрема, ШІ на основі нейронних мереж було використано в програмі AlphaGo, створеній компанією Google DeepMind. У 2016 р. AlphaGo перемогла майже всіх найбільш титулованих професійних гравців у давню китайську гру го. Слід зазначити, що гра в го набагато складніша за шахи — кількість можливих конфігурацій дошки в ній сягає 10^{170} . Це більше, ніж число атомів у відомому нам Всесвіті!

Одним із проривів у цій галузі стала поява великих мовних моделей, таких як GPT (Generative Pre-Trained Transformer). Наприкінці 2022 р. компанія OpenAI представила широкому загалу у вільному доступі демонстраційну версію ChatGPT — чат-бот на основі ШІ, що стимулювало «вибуховий» розвиток подібних програмних продуктів.

У 2024 р. досягнення в галузі штучного інтелекту було відзначено Нобелівською премією, причому одразу в двох номінаціях. Премію з фізики присуджено Джону Гопфілду (John J. Hopfield) і Джеффрі Гінтону (Geoffrey E. Hinton) за «фундаментальні відкриття та винаходи, що уможливили машинне навчання за допомогою штучних нейронних мереж», а з хімії — Девіду Бейкеру (David Baker) за «комп'ютерний дизайн білків», а також представникам компанії Google DeepMind Демісу Гассабісу (Demis Hassabis) і Джону Джамперу (John M. Jumper) за «прогнозування структури білка» з використанням програми AlphaFold.

Однак, незважаючи на стрімке розширення сфери практичного застосування ШІ, глибокого розуміння закономірностей того, як саме навчаються нейронні мережі, і досі немає. Принцип роботи вже навчених нейронних мереж показано на рис. 1. Нейронні мережі складаються з шарів, у кожному з яких є багато нейронів — комірок пам'яті. Під час обробки інформації в нейрони пошарово записуються числа. Між сусідніми шарами нейронів є зв'язки, які визначають, як саме з чисел попереднього шару отримати числа наступного. Кожен зв'язок, грубо кажучи, може бути більш або менш міцним. Ця міцність описується відповідним ваговим коефіцієнтом, тобто числом. Саме ці вагові коефіцієнти визначають, як саме нейронна мережа обробляє вхідну інформацію. По суті, процес навчання нейронної мережі — це поступовий підбір вагових коефіцієнтів. Станом на зараз цей процес являє собою «чорний ящик».

Одним із сучасних підходів до розв'язання цієї проблеми є застосування розвинених математичних теорій, зокрема теорії випадкових матриць. Який же тут зв'язок з випадковими матрицями? Річ у тім, що вагові коефіцієнти між сусідніми шарами нейронів можна впорядкувати у вигляді матриці великого розміру. Перед початком навчання вагові коефіцієнти обирають випадково, потім коефіцієнти покроково змінюються так, щоб нейронна мережа якомога краще класифікувала відомі об'єкти.

Отже, спочатку ми маємо випадкові вагові коефіцієнти, організовані у формі випадкових матриць. Тому для дослідження штучних нейронних мереж застосування теорії випадкових матриць є доцільним.

Предметом дослідження є спектральні властивості випадкових матриць, тобто, як і для звичайних матриць, — власні значення та власні вектори, а також усе, що з ними пов'язано, наприклад детермінанти. Ключова відмінність від звичайних матриць полягає в тому, що всі ці величини є випадковими.

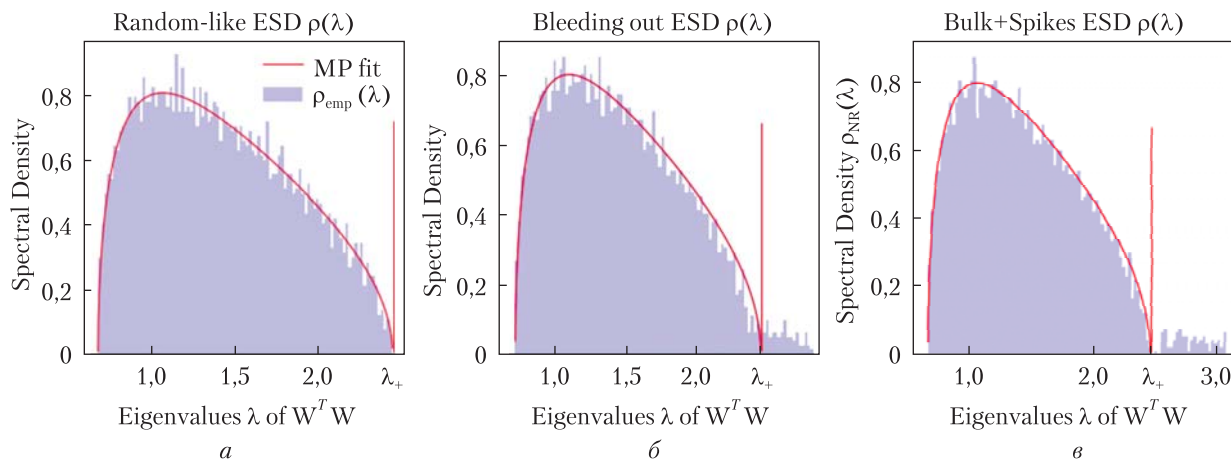


Рис. 2. Розподіл власних значень матриць вагових коефіцієнтів: *a* — до початку навчання; матриця вагових коефіцієнтів W — випадкова; *б* — після часткового навчання; *в* — після навчання; $W = R + S$, де матриця R — випадкова, а матриця S — не випадкова [8]

Теорія випадкових матриць виникла у фізиці в 1950-х роках у зв'язку з дослідженням рівнів збудження важких атомних ядер і з часом знайшла застосування в математиці, інших фізичних напрямках, економіці тощо. А в останні роки з'явилася нова сфера її використання — поліпшення алгоритмів ШІ. Сьогодні застосування випадкових матриць для поліпшення алгоритмів ШІ — це дуже перспективний напрям, яким активно займаються науковці у усьому світі (див., напр., [1–9]).

У Фізико-технічному інституті низьких температур ім. Б.І. Веркіна НАН України функціонує всесвітньо відома наукова школа з теорії випадкових матриць, до якої я належу і в якій уже понад 70 років проводять фундаментальні дослідження спектральних властивостей різноманітних видів випадкових матриць. Яскравим прикладом є фундаментальні роботи академіка Л.А. Пастура і В.В. Славина [2, 3], в яких вони за найбільш загальних на сьогодні припущень знайшли розподіл власних значень нейронних мереж як функції від випадкових матриць.

У наших дослідженнях спектральних властивостей випадкових матриць, які можна застосовувати до вивчення глибоких нейронних мереж, було отримано такі результати [10].

Теорема. Нехай вагова матриця має вигляд $W = R + S$, де R — випадкова матриця, а S —

невипадкова. Тоді при $a > 1$ кількість власних значень $W^T W$ на відрізку $[(1 + a^2)/a; (1 + b^2)/b] \approx$ кількість власних значень $S^T S$ на відрізку $[a; b]$.

Знання про спектр таких матриць буде корисним для дослідження попередньо навчених штучних нейронних мереж.

На рис. 2 наведено гістограми розподілу власних значень матриць вагових коефіцієнтів нейронної мережі на різних стадіях навчання, отримані американськими фахівцями з комп'ютерних наук Ч. Мартіном і М. Махонеем [8]. На рис. 2а зображено розподіл власних значень при ініціалізації, тобто до початку навчання. Цей розподіл збігається з асимптотичним розподілом власних значень випадкових матриць ансамблю Марченка—Пастура. У процесі навчання (рис. 2б) частина власних значень усе ще розподілена, так само як для ансамблю Марченка—Пастура, а частина власних значень виходить за праву межу цього розподілу. Після навчання (рис. 2в) ті власні значення, які відійшли від основної групи, відділяються ще більше.

Моделлю вагових матриць навчених штучних нейронних мереж є ті випадкові матриці, які досліджувалися в роботі [10], — матриці виду $W = R + S$, де R — випадкова матриця, а S — не випадкова. Можна припустити, що саме великі власні значення матриці, які знаходять-

ся поза розподілом Марченка—Пастура, відображають те, чого навчилася нейронна мережа. Тоді всі інші власні значення можна замінити нулями без істотних втрат якості роботи нейронної мережі — зробити так звану «обрізку».

Є різні способи робити «обрізку», але найбільш значущою є саме «обрізка» власних значень, оскільки цей метод потребує менше обчислювальних потужностей при зіставних результатах. Більше того, чисельні експерименти професора Берлянда зі співавторами показують [4], що в деяких випадках точність штучних нейронних мереж після «обрізки» може навіть зростати. Для теоретичного обґрунтування можна використати отриманий результат про розподіл власних значень матриць W [10]. Це ми плануємо зробити в подальшому.

Крім уже зазначеної роботи [10], ми дослідили також спектральні властивості випадкових матриць у локальному режимі, тобто в малому околі однієї точки (розмір околу прямує до нуля). Було отримано асимптотичну поведінку детермінантів для різних ансамблів випадкових матриць [11—15]. Знаходження асимптотичної поведінки детермінантів є першим кроком у дослідженні локального розподілу

власних значень. А знання про локальний розподіл у перспективі також можуть виявитися корисними для розвитку штучних нейронних мереж або для застосувань в інших галузях.

Підсумовуючи свій виступ, зазначу, що за останні шість років мною (зі співавторами) було отримано такі результати [10—15]:

- одержано асимптотичний розподіл великих власних значень випадкових матриць деформованого ансамблю Марченка—Пастура виду $(R + S)^T(R + S)$;
- отримано асимптотичну поведінку детермінантів кількох ансамблів випадкових матриць.

Ці результати мають значний потенціал застосування для теоретичного обґрунтування методу «обрізки» власних значень матриць вагових коефіцієнтів штучних нейронних мереж.

Надалі ми плануємо розвивати такі напрями досліджень:

- вивчення асимптотичної поведінки власних векторів випадкових матриць деформованого ансамблю Марченка—Пастура;
- теоретичне обґрунтування методу «обрізки» власних значень матриць вагових коефіцієнтів штучних нейронних мереж.

REFERENCES

1. Pennington J., Schoenholz S., Ganguli S. The emergence of spectral universality in deep networks. Proc. of the 21st Int. Conf. on Artificial Intelligence and Statistics (AISTATS). 2018. P. 1924. arXiv:1802.09979. <https://doi.org/10.48550/arXiv.1802.09979>
2. Pastur L. On random matrices arising in deep neural networks: Gaussian case. *Pure Appl. Funct. Anal.* 2020. 5(6): 1395—1424.
3. Pastur L., Slavin V. On random matrices arising in deep neural networks: general I.I.D. case. *Random Matrices Theory Appl.* 2023. 12(1): 2250046. <https://doi.org/10.1142/S2010326322500460>
4. Berlyand L., Sandier E., Shmalo Y., Zhang L. Enhancing Accuracy in Deep Learning Using Random Matrix Theory. *Journal of Machine Learning Research.* 2024. 3: 347—412. <https://doi.org/10.4208/jml.231220>
5. Price I., Tanner J. Dense for the Price of Sparse: Improved Performance of Sparsely Initialized Networks via a Subspace Offset. Proc. of the 38th Int. Conf. on Machine Learning. 2021. P. 8620—8629. <https://doi.org/10.48550/arXiv.2102.07655>
6. Saada T.N., Tanner J. On the Initialization of Wide Low-Rank Feedforward Neural Networks. arXiv:2301.13710. 2023. <https://doi.org/10.48550/arXiv.2301.13710>
7. Couillet R., Liao Z. *Random Matrix Methods for Machine Learning*. Cambridge University Press, 2022. <https://doi.org/10.1017/9781009128490>
8. Martin C.H., Mahoney M.W. Implicit self-regularization in deep neural networks: evidence from random matrix theory and implications for learning. *Journal of Machine Learning Research.* 2021. 22(165): 1—73. <http://jmlr.org/papers/v22/20-410.html>
9. Martin C.H., Mahoney M.W. Heavy-Tailed Universality Predicts Trends in Test Accuracies for Very Large Pre-Trained Deep Neural Networks. Proc. of the SIAM Int. Conf. on Data Mining. 2020. P. 505—513. <https://doi.org/10.1137/1.9781611976236.57>

10. Afanasiev I., Berlyand L., Kiyashko M. Asymptotic behavior of eigenvalues of large rank perturbations of large random matrices. arXiv:2507.12182. 2025. <https://doi.org/10.48550/arXiv.2507.12182>
11. Afanasiev I., Shcherbina T. Characteristic Polynomials of Sparse Non-Hermitian Random Matrices. *J. Stat. Phys.* 2025. **192**: 12. <https://doi.org/10.1007/s10955-024-03379-5>
12. Afanasiev I. On the Correlation Functions of the Characteristic Polynomials of Non-Hermitian Random Matrices with Independent Entries. *J. Stat. Phys.* 2019. **176**(6): 1561—1582. <https://doi.org/10.1007/s10955-019-02353-w>
13. Afanasiev I. On the Correlation Functions of the Characteristic Polynomials of the Sparse Hermitian Random Matrices. *J. Stat. Phys.* 2016. **163**(2): 324—356. <https://doi.org/10.1007/s10955-016-1486-z>
14. Afanasiev I. On the Correlation Functions of the Characteristic Polynomials of Random Matrices with Independent Entries: Interpolation Between Complex and Real Cases. *J. Math. Phys. Anal. Geom.* 2022. **18**(2): 159—181. <https://doi.org/10.15407/mag18.02.159>
15. Afanasiev I. On the Correlation Functions of the Characteristic Polynomials of Real Random Matrices with Independent Entries. *J. Math. Phys. Anal. Geom.* 2020. **16**(2): 91—118. <https://doi.org/10.15407/mag16.02.091>

Ievgenii V. Afanasiev

*B. Verkin Institute for Low Temperature Physics and Engineering
of the National Academy of Sciences of Ukraine, Kharkiv, Ukraine*
ORCID: <https://orcid.org/0000-0002-9238-7331>

ARTIFICIAL INTELLIGENCE, NEURAL NETWORKS AND SPECTRAL PROPERTIES OF RANDOM MATRICES

According to the materials of report at the meeting of the Presidium of the NAS of Ukraine, May 28, 2025

The report is devoted to recent results in the field of Artificial Intelligence and Neural Networks (NNs), both based on the spectral properties of random matrices. Nowadays, the development of Artificial Intelligence is skyrocketing, but we still lack an understanding of how it works. One of the modern approaches to solving this problem is the application of advanced mathematical theories, in particular, Random Matrix Theory. We present our results on the spectral properties of random matrices, obtained at the Department of Mathematical Physics at B. Verkin Institute for Low Temperature Physics and Engineering of the National Academy of Sciences of Ukraine. These matrices are closely connected with NNs: weight matrices of trained NNs could be represented in the form $W = R + S$, where R is random and S is deterministic. The spectrum of such matrices plays a key role in the rigorous underpinning of the novel pruning technique based on Random Matrix Theory.

Cite this article: Afanasiev I.V. Artificial intelligence, neural networks and spectral properties of random matrices (according to the materials of report at the meeting of the Presidium of the NAS of Ukraine, May 28, 2025). *Visn. Nac. Akad. Nauk Ukr.* 2025. (8): 69—73. <https://doi.org/10.15407/visn2025.08.069>