

Teoria macierzy losowych to dział matematyki, głównym obiektem badań którego jest widmo dużych macierzy o losowych współczynnikach, wybranych z różnych rozkładów prawdopodobieństwa oraz posiadających różne typy symetrii. Oprócz teorii prawdopodobieństwa, algebry i analizy łączy się z innymi dziedzinami matematyki, w tym z teorią liczb, kombinatoryką, asymptotyczną analizą geometryczną i teorią grafów losowych. Macierze losowe mają zastosowania w fizyce, statystyce, naukach technicznych, informatyce teoretycznej, stając się efektywnym narzędziem do modelowania systemów złożonych. Zagadnienia, w których odgrywają one kluczową rolę to m. in. analiza widma jąder atomowych, chaos kwantowy, komunikacja bezprzewodowa, analiza danych. Jednym z powodów sukcesu teorii macierzy losowych jest zjawisko uniwersalności – statystyki spektralne macierzy okazują się być często niezależne od szczegółów modeli z rozpatrywanej klasy, co pozwala uchwycić istotne czynniki odpowiadające za ich zachowanie.

W ostatnich latach teoria macierzy losowych znalazła zastosowanie w szybko rozwijającej się dziedzinie uczenia maszynowego, m. in. w teorii głębokich sieci neuronowych, szeroko używanych w analizie danych. Są to sztuczne sieci neuronowe zbudowane w oparciu o wiele ukrytych warstw neuronów pomiędzy warstwami wejścia i wyjścia. Uczone w oparciu o optymalizacyjne algorytmy gradientu prostego, rozpoznają złożone wzorce w dużych zbiorach danych, często istotnie przewyższając wcześniejsze metody w zagadnieniach takich jak widzenie komputerowe, przetwarzanie języka naturalnego, rozpoznawanie mowy, tłumaczenie automatyczne. Pomimo tego, uczenie głębokich sieci neuronowych i zrozumienie ich działania pozostaje znaczącym wyzwaniem, gdyż wymaga dużych zbiorów danych, potężnej mocy obliczeniowej, precyzyjnego dopasowania parametrów i efektywnych technik regularizacyjnych. Ponadto są one nadal słabo poznane z teoretycznego punktu widzenia, co zazwyczaj sprawia, że wdrażanie sieci neuronowych jest bardziej sztuką niż nauką. Istnieje zatem pilna potrzeba dogłębnego matematycznego zrozumienia zachowania sieci neuronowych, niezbędnego do ulepszenia istniejących algorytmów i opracowania nowych.

Analiza sieci neuronowych zależy od ogromnej liczby parametrów zawartych w dużych macierzach, takich jak macierz Jacobiego wejścia-wyjścia, macierz kowariancji danych wejściowych oraz hesjan i macierz informacji Fishera. To sprawia, że macierze losowe dobrze nadają się do opracowywania nowych podstaw teoretycznych w celu poprawy wydajności sieci neuronowych. Z drugiej strony problemy pojawiające się w uczeniu głębokim stwarzają nowe, fundamentalne wyzwania dla teorii macierzy losowych, prowadzące do owocnej i obiecującej synergii między tymi dwoma dziedzinami.

Dzięki podejściu opartemu o macierze losowe uzyskano wiele ważnych wyników dotyczących wpływu informacji na wejściu na wektor wyjścia głębokich sieci neuronowych. Wykazano, że inicjalizacja sieci przy pomocy losowych wag o rozkładzie gaussowskim pozwala na wyeliminowanie wykładniczego wzrostu lub zaniku gradientów, co w wielu przypadkach pozwala istotnie zwiększyć tempo uczenia. Kolejnym zastosowaniem jest dobór kryteriów uczenia. Wiele wskazuje na to, że kryteria oparte o analizę wartości szczególnych związanej z siecią neuronową macierzy Jacobiego pozwolą uniknąć niekorzystnych efektów przetrenowania lub niedotrenowania sieci. Ponadto analiza oparta na teorii macierzy losowych może poprawić zbieżność i szybkość uczenia sieci neuronowych i zmniejszyć złożoność obliczeniową algorytmów uczących.

Zastosowanie macierzy losowych do badania sieci neuronowych prowadzi do wielu problemów otwartych. Istniejące techniki często bazują na uproszczonych założeniach niezależności lub gaussowskości danych wejściowych oraz wag sieci, które mogą nie być spełnione w praktyce. Rozszerzenie analizy do bardziej realistycznych sytuacji jest dużym wyzwaniem. Dotyczy to m. in. problemów nieliniowych aktywacji, danych zależnych, stochastycznych metod gradientu prostego, sieci splotowych i rekurencyjnych, wymagających uogólnienia znanych metod teorii macierzy losowych i rozszerzenia ich stosowalności.

Celem projektu jest zbadanie własności spektralnych macierzy losowych związanych z sieciami neuronowymi za pomocą metod analitycznych i numerycznych oraz rozwinięcie narzędzi pozwalających na zwiększenie dokładności i redukcję złożoności algorytmów uczących. Zajmiemy się takimi kwestiami, jak zwiększenie efektywności treningu poprzez przycinanie parametrów, znalezienie optymalnych wag inicjalizujących sieci neuronowych, wyznaczenie warunków stopu, a także uzasadnienie poprawności odpowiednich algorytmów numerycznych. Oczekujemy, że wyniki badań zapewnią nowy wgląd w teoretyczne podstawy głębokich sieci neuronowych, a także umożliwią opracowanie nowych narzędzi i technik teorii macierzy losowych, ukierunkowanych na zastosowania w sieciach neuronowych.