УДК 004.89:534.1:531.1

ПРИМЕНЕНИЕ ФИЗИЧЕСКИ-ИНФОРМИРОВАННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРЯМОГО МОДЕЛИРОВАНИЯ БАЛЛИСТИЧЕСКОГО И ЗАТУХАЮЩЕГО КОЛЕБАТЕЛЬНОГО ДВИЖЕНИЙ

Козырев Д. В. (Университет «ИТМО»), Янышевская К. В. (-), Шаламов В. В. (-).

Научный руководитель — Демидова Г. Л. (Университет «ИТМО»)

Аннотация

В работе исследуется применение физически-информированных нейронных сетей (PINN) к задачам прямого моделирования динамических систем возрастающей сложности. На простейшей задаче баллистического движения (h''(t)-g=0) показано, что введение физического регуляризующего слагаемого в функцию потерь снижает ошибку инференса на 4 порядка при обучении всего на 10 зашумлённых точках. Для свободных затухающих колебаний с высокочастотными компонентами выявлена проблема спектрального смещения (spectral bias) PINN с гиперболическим тангенсом в качестве функции активации и предложено её решение на основе синусоидальных сетей представления (SIREN) с адаптивным взвешиванием потерь и warmup-фазой, обеспечивающее улучшение точности на 5 порядков.

Ключевые слова: *Физически-информированные нейронные сети, физический регуляризатор, спектральное смещение, баллистическое движение, колебательное движение, математическое моделирование*

Введение

Традиционные подходы к моделированию динамических систем основываются на численном решении дифференциальных уравнений, требующем точно известных параметров системы и использующем ресурсоемкие методы вычисления очередного шага моделирования. Аппроксимация с помощью суррогатных моделей позволяет существенно сократить время на вычисление сложный аппроксимируемых выражений, однако могут отличаться сложностью их разработки [4]. Классические нейросетевые методы обладают простотой реализации, но нуждаются в больших объёмах данных и не гарантирую физическую корректность предсказаний. Когда необходимо произвести аппроксимацию физико-математической модели возможно использовать специализированный нейросетевой подход, получивший развитие в последние 5 лет. Физически-информированные нейронные сети (PINN) [3] объединяют нейросетевой аппарат аппроксимации зависимостей и описанные физические законы для задания направления обучения.

PINN показали свои перспективы развития, однако стоит учитывать особенности решаемой задачи, чтобы корректно подобрать архитектуру PINN сети. Так, при наличии высокочастотных компонент в решаемой PINN задаче, монотонные функции активации могут демонстрировать эффект спектрального смещения (англ. spectral bias), преимущественно обучаясь низкочастотным компонентам задачи [2]. В настоящей работе проведено исследование особенностей применения PINN на задаче простейшей баллистической траектории и свободного затухающего колебания и предложены связанные с результатами рекомендации по их реализации.

1. Баллистическое движение

Траектория свободного падения материальной точки описывается простейшим дифференциальным уравнением второго порядка, в котором ускорение свободного падения наблюдаемой точки должно совпадать с ускорением свободного падения на Земле. Поставленный эксперимент заключался в обучении полносвязной нейронной сети с двумя скрытыми слоями по 32 нейрона в каждой, входным параметром которой является момент времени t, выходным параметром – положение точки по вертикали. Обучение производилось по 10 точкам на всей траектории полета, записанными с долей шума 40% от эталонного значения функции. Использовалась двухэтапная оптимизация: 1) метод адаптивной оценки моментов (Adam, 5000 эпох); 2) Алгоритм Бройдена – Флетчера – Гольдфарба – Шанно с ограниченным использованием памяти (L-BFGS, 1250 итераций).

В эксперименте по сравнительному анализу точности обучения в зависимости от используемого подхода к обучению сравнивались две одинаковые нейронные сети, обучаемые по двум различным функциям невязки: PINN модель, в отличии от классической, включала регуляризирующее слагаемое, основанное на дифференциальном уравнении свободного патения материальной точки.

Оценка среднеквадратической ошибки PINN-модели достигла примерно 10^{-6} против примерно 10^{-2} классической модели. Экспериментально полученное расхождение на 4 порядка подтвердило, что физический регуляризатор обеспечивает устойчивость к шуму и физически корректное предсказание движения даже при минимальном объёме данных.

2. Свободные затухающие колебания

Если добавить к классической функции невязки нейронной сети прямого распространения регуляризационное слагаемое в виде левой части дифференциального уравнения свободного затухающего колебательного движения, требуя чтобы оно равнялось нулю, обучение нейронной сети будет происходить для предсказания требуемого колебательного движения. При этом важно установить коэффициент значимости λ, указывающий на относительный вклад регуляризационного слагаемого в функцию невязки. Он должен быть выбран таким, чтобы не позволить процессу обучения сойтись к тривиальному решению дифференциального уравнения.

Произведенное обучение нейронной сети прямого распространения с двумя скрытыми слоями по 128 нейронов, содержащих функцию активации вида гиперболический тангенс, для частоты свободных колебания порядка 10 рад/с продемонстрировало успешное схождение к ожидаемой траектории движения маятника. Однако, с ростом частоты наблюдались сложности в обучении нейронной сети, связанные с проявлением эффекта спектрального смещения (англ. spectral bias), т.е. склонности нейронных сетей обучаться медленным изменениям быстрее, чем быстрым. Найденным решением оказалось использовать периодические функции с промежутками противоположно направленной монотонности, в частности, синус. Используя синусоидальную сеть (SIREN) [1] представления, получилось пятикратно увеличить требуемую частоту колебаний, что было нетривиальной задачей в случае использования гиперболического тангенса, при этом среднеквадратическая ошибка сократилась на 5 порядков.

Заключение

Проведенные эксперименты по обучению физически-информированных нейронных сетей привело к демонстрации нескольких важных свойств нейронных сетей. Первое, это возможность нейронных сетей значительно улучшать точность предсказания траектории движения объекта при сохранении объема обучающей выборки, если снабдить функцию невязки регуляризационным членом, описывающим требуемую динамику выходных параметров. При этом важно учитывать, соотношение вкладов компонент невязок так, чтобы система не сходилась к тривиальному решению дифференциального уравнения, т.е. влияние обучающих данных было достаточно велико, но не на сколько, чтобы аппроксимация эталонной траектории основывалась только на обучающей выборке. Второе, следующее из предыдущего, это, регулируя соотношение влияния двух компонентов невязки, возможно снижать влияние шума, обнаруживаемого в обучающих данных, компенсируя его физико-математическим соотношением, описывающем эталонное поведение. Или напротив, с достаточной долей доверия к обучающим данным использовать их для поиска скрытых зависимостей в анализируемой системе, которая недостаточно описана дифференциальными уравнениями в функции невязки. Третье наблюдение касается разработки нейросетей для предсказания явно выраженного колебательного движения: использование тригонометрического синуса в качестве функции активации позволяет существенно сократить влияние эффекта спектрального смещения, т.е. значительно ускорить обучение.

Дальнейшие исследования будут касаться использования в различных архитектурах нейронных сетей разных функций активаций: часть монотонных функций активации, а часть – тригонометрический синус. Следует обратить внимание на соотношение слагаемых в функции невязки, и произвести исследование возможных комбинаций коэффициентов перед информационной и физической компонентами невязки с целью предсказания вывода о характере выходных данных после обучения. А также планируется произвести сравнение физически-информированных нейронных сетей, обучаемых с помощью кинематических и энергетических законов для предсказания движения материальной точки.

Список литературы

1. Sitzmann V. et al. Implicit neural representations with periodic activation functions // NeurIPS. 2020. Vol. 33. P. 7462–7473.
2. Nasim Rahaman, Aristide Baratin et. al. On the Spectral vias of Neural Networks // Proceedings of the 36th International Conference on Machine Learning, PMLR 97:5301-5310, 2019.
3. Raissi M., Perdikaris P., Karniadakis G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems // J. Comput. Phys. 2019. Vol. 378. P. 686–707.
4. Kudela, J., Matousek, R. Recent advances and applications of surrogate models for finite element method computations: a review. Soft Comput 26, 13709–13733 (2022). https://doi.org/10.1007/s00500-022-07362-8