УДК 519.62, 004.032.26

А.А. Шолохова, А.Н. Иванов

МОДЕЛИРОВАНИЕ ДИНАМИЧЕСКИХ СИСТЕМ НА ОСНОВЕ ПОЛИНОМИАЛЬНЫХ НЕЙРОННЫХ СЕТЕЙ

Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

В статье представлена полиномиальная архитектура нейронной сети, применяемая для идентификации динамических систем. Рассматриваемая модель является нейросетевым представлением матричного оператора Ли, который используется для решения систем нелинейных обыкновенных дифференциальных уравнений. Вычисление оператора Ли в матричной форме на основе полиномиальных степенных рядов позволяет строить отображение, описывающее динамику системы. Такое отображение может использоваться как эффективный численный метод исследования динамических систем, а нейросетевое представление оператора Ли позволяет объединить присущие искусственным нейронным сетям параллельную вычислительную архитектуру со строгой теорией динамических систем и теорией дифференциальных уравнений. Кратко представлена математическая формализация подхода и выведены основные формулы, связывающие предлагаемую нейросетевую архитектуру с системами дифференциальных уравнений. Рассматриваются численный метод решения систем нелинейных дифференциальных уравнений на основе численного матричного интегрирования, а также использование данного подхода для построения систем машинного обучение и идентификации систем. Приведены примеры применения алгоритма как на модельных задачах, так и в прикладной области идентификации движения морских судов. В заключении обсуждаются ограничения и перспективы развития рассмотренного метода.

Ключевые слова: полиномиальная нейронная сеть, идентификация систем, машинное обучение

Введение.

При моделировании и идентификации сложных динамических систем часто возникают задачи учета эффектов нелинейности и неопределённости. Кроме того, при практической реализации подходов требуется принимать во внимание вычислительную сложность алгоритмов, которая должна быть достаточно простой для обработки большого количества данных. Нейронные сети как инструмент моделирования систем могут выступать одним из вариантов решения указанных проблем. Применение нейронных систем в области исследования динамических систем отражено в различных приложениях и предметных областях, таких как решение обыкновенных дифференциальных уравнений [1,2], обработка сигналов [3], системы управления [4], теория моделирования и идентификации динамических систем [5] и др. В указанных работах подходы к применению нейронных сетей можно разделить на две группы. К одной группе следует отнести методы решения заранее известных систем дифференциальных уравнений при помощи нейронных сетей. В другой группе рассматривается использование аппарата искусственных нейронных сетей для идентификации динамических систем и решения обратных задач. В этом случае вид уравнений не задан, либо задан с точностью до параметров, а обученная нейронная сеть является моделью рассматриваемого процесса или динамического объекта.

В обоих случаях применимость нейронных сетей объясняется лишь универсальной теоремой аппроксимации [6] без строгих математических доказательств, почему именно предлагаемая нейронная сеть применима в рассматриваемой проблеме. Кроме того, такое моделирование на основе черного ящика не может гарантировать сохранение физических инвариантов, которые часто возникают в динамических системах. Такие инварианты играют ключевую роль в компьютерном моделировании и позволяют сохранять физические свойства реальных систем, однако практически не рассматриваются в теории искусственных нейронных сетей.

В статье рассматривается архитектура нейронной сети, которая решает указанные проблемы для достаточно широкого класса динамических систем, которые описываются системой нелинейных дифференциальных уравнений:

$$\frac{dX}{dt} = F(t, X), \tag{1}$$

где t является независимой переменной, $X \in \Re^n$ задает вектор состояния системы. На функцию F накладывается ограничение о разложимости в ряд Тейлора относительно вектора состояния X, переменная t может входить в уравнение как произвольная нелинейная функция [7]. Подобные системы описывают процессы в различных областях, таких как автоматическое управление, робототехника, механические и биологические системы, химические реакции, разработка лекарств, молекулярная динамика.

Рассматриваемая архитектура полиномиальной нейронной сети является матричным оператором Ли [7, 8], применяемым для описания динамических систем. Такое представление оператора Ли позволяет объединить вычислительную параллельную архитектуру и гибкую возможность учета неопределенностей, присущих нейронным сетям, со строгой теорией динамических систем на основе обыкновенных дифференциальных уравнений.

В следующем параграфе представлена архитектура нейронной сети, ее математическое описание и эквивалентность преобразованию Ли для системы (1). В статье рассматривается применение предлагаемого подхода для идентификации модели Лотки – Вольтерра, а также приведено практическое применение нейронной сети для решения практической задачи идентификации модели движения морского судна.

Материалы и методы.

Динамика вектора X в уравнении (1) может быть описана при помощи хронологически упорядоченной экспоненты [7]

$$M(t|t_0) = T\exp(\int_{t_0}^{t} \Im_F(\tau)d\tau),$$

задаваемой оператором Ли $\Im_F(\tau)$. В предположении, что функция F в уравнении (1) может быть представлена в виде степенного ряда

$$F = \sum_{k=0}^{\infty} P_k(t) X^{[k]},$$

искомое решение внутри области своей сходимости также может быть представлено в виде полиномиального преобразования

$$X(t) = M(t) \circ X_0 = \sum_{k=0}^{\infty n} R_k(t) X_0^{[k]}. \qquad (2)$$

Здесь под $X^{[k]}$ понимается *k*-ая кронекеровская степень вектора X. В работе [8] показано, как можно вычислять матрицы $R_k(t)$ либо в символьном, либо в численном виде. Идея алгоритмов заключается в замене дифференциального уравнения (1) относительно вектора состояния X системой дифференциальных уравнений относительно коэффициентов матриц $R_k(t)$ [9]. Интегрируя уравнение (2) по времени

$$\frac{dX}{dt} = \sum_{k=0}^{n} \frac{d}{dt} R_k(t) X_0^{[k]},$$

и сопоставляя его с уравнением (1), в котором подставлены степенные разложения векторов $X^{[k]}$

$$\begin{aligned}\frac{dX}{dt} &= P_0(t) + P_1(t)X + P_2(t)X^{[2]} + \ldots + P_p(t)X^{[p]} = \\ &= P_0(t) + \\ &+ P_1(t)\left(R_0(t) + R_1(t)X_0 + \ldots + R_k(t){X_0}^{[k]}\right) + \\ &+ P_2(t)\left(R_0(t) + R_1(t)X_0 + \ldots + R_k(t){X_0}^{[k]}\right)^{[2]} + \\ &+ \ldots + \\ &+ P_p(t)\left(R_0(t) + R_1(t)X_0 + \ldots + R_k(t){X_0}^{[k]}\right)^{[p]},\end{aligned}$$

можно получить новую систему дифференциальных уравнений

$$\begin{aligned}\frac{dR_0(t)}{dt} &= \sum_{k=0}^{n} P_k(t)\left(R_k(t)\right)^{[k]}, \\ \frac{dR_j(t)}{dt} &= \sum_{k=0}^{p} P_k(t) \frac{\partial\left(X(t)\right)^{[k]}}{\partial\left(X_0^{[j]}\right)^T},\end{aligned} \qquad (3)$$

где вектор X задается равенством (2), а под $(\cdot)^T$ понимается операция транспонирования. Данную систему уравнений (3) можно решить подходящим численным методом интегрирования дифференциальных уравнений с нулевыми начальными условиями $R_1(t_0) = E, R_k(t_0) = \Theta$ при $k \neq 1$.

Предлагаемая архитектура нейронной сети реализует полиномиальное преобразование (2) до заданного порядка нелинейности k

$$Y = W_0 + W_1 X + W_2 X^{[2]} + \ldots + W_k X^{[k]}, \tag{4}$$

в котором $X \in \Re, Y \in \Re^m, W_i$ – матрицы весовых коэффициентов соответствующих размерностей. Для примера, на рис. 1 отображено графическое представление такой нейронной сети до третьего порядка нелинейности. На схеме входной вектор $X = (x_1, x_2)$ последовательно преобразуется в кронекеровские степени $X^{[2]} = (x_1{}^2, x_1 x_2, x_2{}^2)$ и $X^{[2]} = (x_1{}^3, x_1{}^2 x_2, x_1 x_2{}^2, x_2{}^3)$, которые затем умножаются на матрицы весовых коэффициентов.

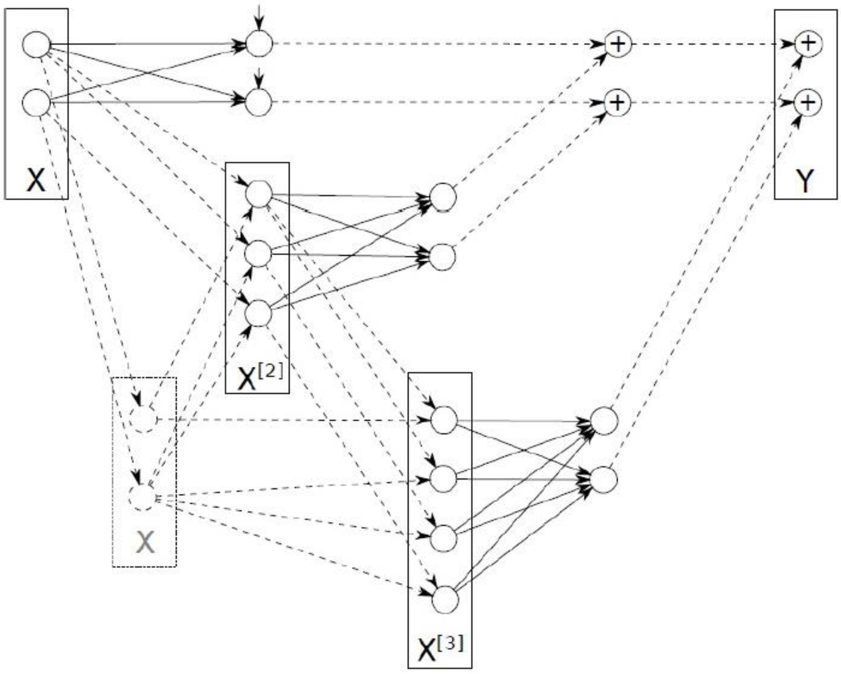

Рисунок 1 – Нейросетевое представление матричного преобразования Ли

Такая архитектура нейронной сети описывает динамическую систему (1) и может быть использована для решения задачи идентификации. В зависимости от сложности искомой системы можно выбрать отображение требуемого порядка. При этом задача идентификации сводится к поиску коэффициентов матриц W_i. Отметим также, что наложение физических свойств и инвариантов на рассматриваемую модель приводит лишь к дополнительным ограничениям, накладываемым на искомые коэффициенты. Например, в работе [8] представлен алгоритм симплектификации отображения (2), который приводит к тому, что некоторые из весовых коэффициентов нейронной сети будут связаны между собой.

Результаты.

В этом параграфе рассмотрено применение предлагаемой архитектуры нейронных сетей для идентификации модели Лотки – Вольтерра. Полиномиальная нейронная сеть на основе матричного представления оператора Ли сравнивается с классическим перцептроном (полносвязная нейронная сеть с сигмоидальными функциями активации). Идентификация отображения производится путем настройки коэффициентов нейронной сети на одном частном решении. Затем проверяется способность нейронной сети оценивать другие частные решения системы, соответствующие различным начальным условиям.

Уравнение Лотки – Вольтерра может служить простейшим примером нелинейных систем

$$\dot{x} = \alpha x - \beta xy,$$
$$\dot{y} = \delta xy - \gamma y.$$

Система имеет стационарную точку, а частные решения описывают осцилляции. Например, на рис. 2 представлен пример частных решений для обучающего (*training data*) и тестового (*test data*) набора данных.

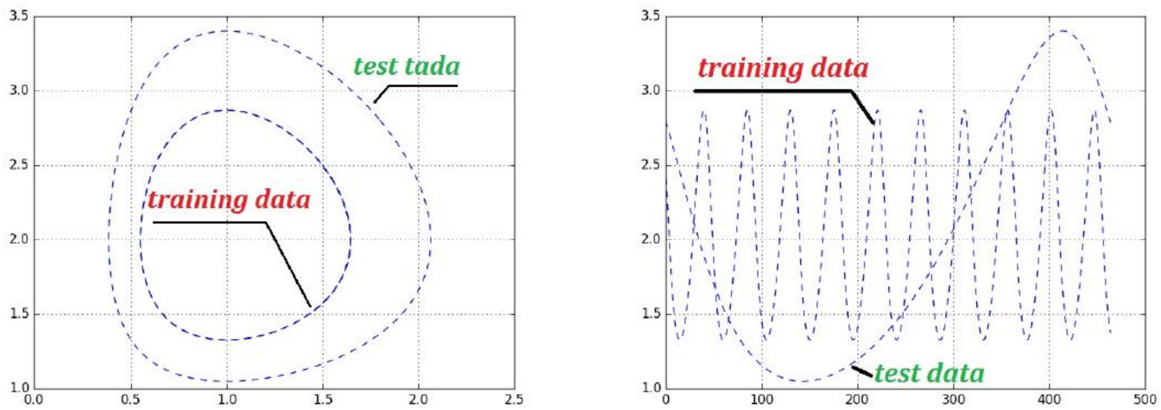

Рисунок 2 – Визуализация решений модели Лотки – Вольтерра в фазовом (слева) и временном (справа) пространствах

Коэффициенты нейронных сетей подбираются при помощи методов численной оптимизации только по набору данных *training data*. На рис. 3 представлены результаты работы обученных нейронных сетей. Классическая нейронная сеть оказывается неспособной оценивать траектории решений системы, отличные от той, на которой обучалась. При попытке построить частные решения, соответствующие другим начальным условиям, такая нейронная сеть стремится сойтись к траектории, предоставленной ей на этапе обучения (рис. 3, слева). Полиномиальная нейронная сеть успешно распознает характер осцилляции решений (рис. 3, справа) и корректно оценивает динамику системы при других начальных условиях. Отметим, что оценка точности и сходимости предлагаемого подхода выходит за рамки данной работы и требует отдельного исследования.

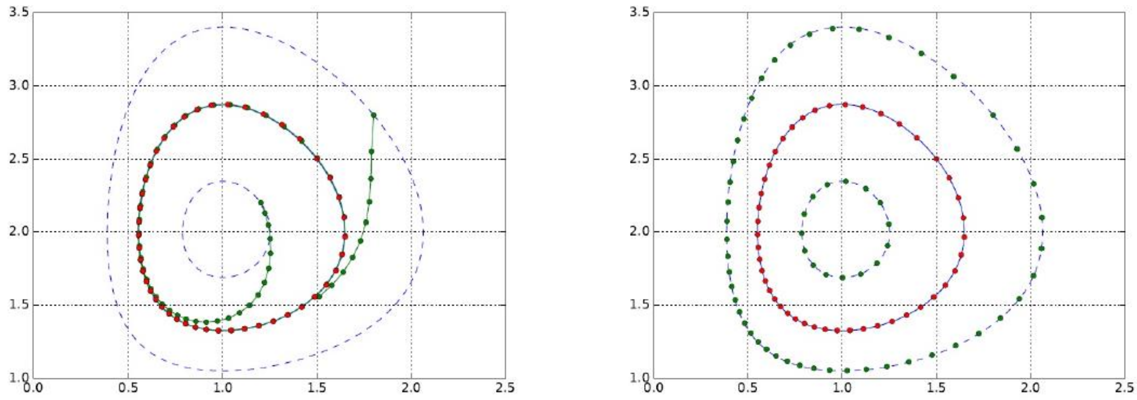

Рисунок 3 – Оценка динамики системы при помощи обученного многослойного перцептрона (слева) и полиномиальной нейронной сети (справа)

В качестве практического применения предлагаемой архитектуры нейронной сети рассмотрена задача идентификации модели движения судна. В работе [10] представлено несколько систем дифференциальных уравнений, описывающих динамику движения судна. Отметим, что уравнения могут иметь различные вектора состояния, т.к. на разных судах могут присутствовать разные датчики (направление и скорость ветра, параметры качки, характеристики работы двигателей и др.). Кроме того, коэффициенты этих уравнений отличаются от судна к судну и зависят от их физических параметров. Однако в общем виде все модели представимы в виде уравнения (1) внутри некоторого дискретного шага по времени.

При решении задач поиска аномалий в движении судна или прогнозирования потребления топлива [11] возникает необходимость наличия общей модели движения судна и ее идентификации по измеренным данным. В качестве примера используются данные о движении судна в течение месяца с дискретностью в одну минуту. Всего собрана информация о 16ти маршрутах. На рисунке 4 представлены нормализованные траектории, отмеченные разными цветами. Полный вектор состояния судна описывается набором из 8ми параметров $X_i = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)_i$, где i обозначает дискретный интервал времени, а переменные x_j – географические широту и долготу, скорость судна относительно воды и земли, глубину, загрузку двигателя, направление и скорость ветра.

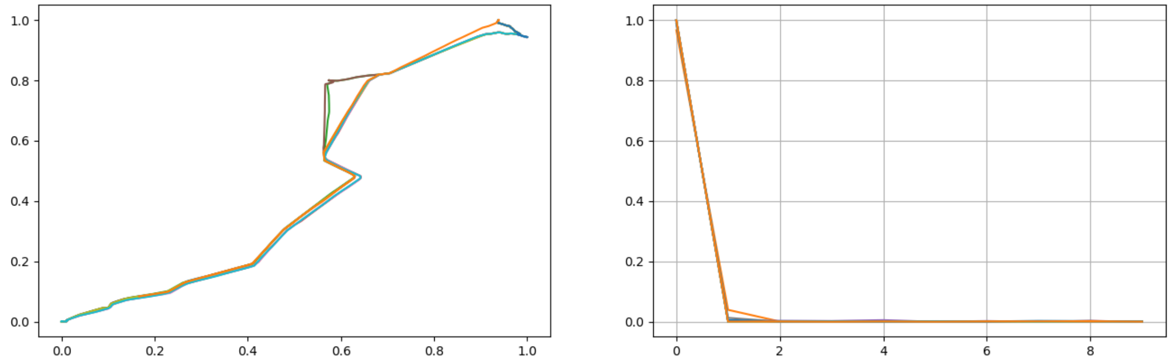

Рисунок 4 – Нормализованные координаты траекторий движения судна (слева) и коэффициенты полиномиальных нейронных сетей (справа), описывающих каждую траекторию

Задачей идентификации модели движения судна является построение такой нейросетевой модели, оптимальным образом отражающей динамику векторов состояния X_i. В такой постановке задача сводится к восстановлению регрессии (4)

$$Y \equiv (x_1, x_2)_{i+1} = W_0 + W_1 X_i + W_2 X_i^{[2]} + \ldots + W_k X_i^{[k]},$$

когда географические координаты предсказываются по предыдущему вектору состояния. Для решения этой задач были использованы полиномиальные нейронные сети второго порядка нелинейности, которые обучались для каждого маршрута по отдельности.

Отметим, что наиболее значимы вклад в модель вносят первые коэффициенты, соответствующие преобразованиям координат и скорости в линейном приближении (см. рис. 5 справа). Однако если проанализировать первые десять коэффициентов в увеличенном масштабе, можно заметить, что одна из построенных моделей заметно выбивается от усредненных значений. На рис. 4 такой маршрут отмечен красным цветом. На графике с траекториями видно, что во время движения осуществлялись резкие маневры, отсутствующие для других маршрутов, что и отражено в коэффициентах построенного отображения.

Обсуждение.

В статье представлен подход построения полиномиальной нейронной сети, которая может выступать в роли модели динамической системы, описываемой системой нелинейных обыкновенных дифференциальных уравнений. Основным преимуществом предлагаемой архитектуры является тот факт, что веса нейронной сети связаны с динамикой системы в целом и интерпретируются как соответствующие нелинейные члены в уравнениях. Следует также отметить, что если система дифференциальных уравнений полностью определена, веса можно вычислить с заданной точностью вплоть до требуемого порядка нелинейности. Однако если исходное уравнение неизвестно или задано с некоторой точностью, указанный подход можно применять для решения задачи идентификации системы по измеренным векторам состояния.

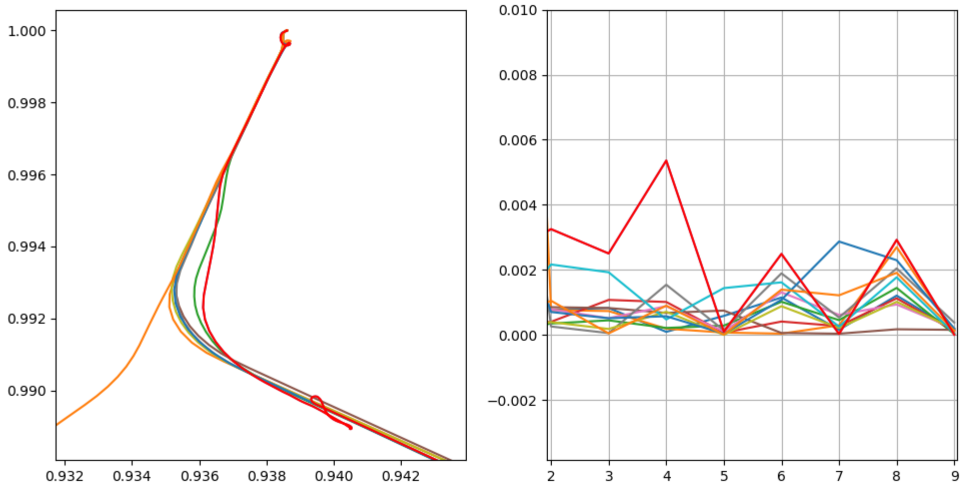

Рисунок 5 – Определение аномального движения судна по коэффициентам нейросетевых моделей

При применении классических архитектур нейронных сетей нет математических гарантий, что обученная сеть способна корректно отображать поведение системы. Обучение в этом случае происходит по прецедентам, когда нейронная сеть способна предсказывать только те данные, которые ей подавались на вход. Данное ограничение продемонстрировано на модельном примере. Применение полиномиальной архитектуры нейронных сетей позволяет избежать указанного недостатка. Кроме того, как подтверждено математическими выкладками и продемонстрировано на практические задачи по идентификации модели движения судна, коэффициенты полиномиальной нейронной сети описывают динамику системы и могут использоваться не просто как черный ящик, а как объяснимая модель рассматриваемой системы.

Заключение.

Приведенные математические выкладки предоставляют строгий аппарат для построения и исследования полиномиальных нейронных сетей, задаваемых матричным преобразованием Ли. Дальнейшим развитием рассмотренного подхода являются его обобщение на теорию дифференциальных уравнений в частных производных и уравнения с запаздыванием, а также развитие специализированных эффективных методов обучения полиномиальных нейронных сетей. Также важной является необходимость теоретических оценок точности и сходимости рассмотренного подхода.

Полученные в статье результаты демонстрируют возможность применения рассмотренной нейронной сети в качестве модели машинного обучения для решения задач восстановления регрессии и оценки общего решения систем обыкновенных дифференциальных уравнений.

Авторы выражают благодарность проф. Андрианову С.Н. за пояснения теоретических аспектов в матричном представлении оператора Ли и его применении для решения дифференциальных уравнений.

ЛИТЕРАТУРА

1. Mall S., Chakraverty S. Comparison of artificial neural network architecture in solving ordinary differential equations // Advances in Artificial Neural Systems, 2013.
2. Lagaris I. E., Likas A., Fotiadis D. I. Artificial neural networks for solving ordinary and partial differential equations // Technical report, 1997. URL https: //arxiv.org/pdf/physics/9705023.pdf.
3. Lapedes A., Farber R. Nonlinear signal processing using neural networks: Prediction and system modelling // IEEE International Conference on Neural Networks, San Diego, CA, USA, 1987.
4. Lewis F. L., Ge S. S. Neural Networks in Feedback Control Systems, in Mechanical Engineers // Handbook: Instrumentation, Systems, Controls, and MEMS, volume 2. John Wiley and Sons, Hoboken, NJ, USA, third edition (ed m. kutz) edition, 2005.
5. Chen S., Billings S. A. Neural networks for nonlinear dynamic system modelling and identification // International Journal of Control, 56(2):319–346, 1992.
6. Cybenko, G. V. Approximation by Superpositions of a Sigmoidal function // Mathematics of Control Signals and Systems,V. 2, P. 303-314, 1989.
7. Dragt A. Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics. 2011. URL http://inspirehep.net/record/955313/files/TOC28Nov2011.pdf.
8. Андрианов С. Н. Динамическое моделирование систем управления пучками частиц. Изд-во СПбГУ, СПб. 2004. 368 с.
9. Иванов А. Н., Кузнецов П.М. Идентификация динамических систем на основе нелинейного матричного преобразования Ли // Вестник УГАТУ. — 2014. — Т. 18. — No 2(63). — С. 251–256.
10. E. Tu, G. Zhang, L. Rachmawati, E. Rajabally and G. B. Huang, Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey From Data to Methodology // *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1-24.
11. Шолохова А.А. Поиск аномалий в сенсорных данных на примере анализа движения морского судна // Моделирование, оптимизация и информационные технологии. — 2017. — № 3(18). — 19 С.

A.A. Sholokhova, A.N. Ivanov

DYNAMICAL SYSTEMS MODELING BASED ON POLYNOMIAL NEURAL NETWORKS

Saint Petersburg State University,
Saint Petersburg, Russia

In the article, a polynomial neural network architecture is presented. This architecture is utilized for dynamical systems identification. The given approach is based on matrix representation of Lie transform, that is useful for investigation of nonlinear systems of ordinary differential equations. The polynomial neural network, in this case, can play a role of an effective and efficient method of investigation of dynamical systems. Moreover, it joints advantages of parallel computing architecture with the strong mathematical theory of differential equations. The key concepts and formulations are briefly described. The numerical matrix integration of the systems of differential equations is also presented. As an example, the identification of the simple model problem is considered as well as an application of the technique for modeling of vessel motion is presented. In the conclusion the limitations and further development of the method is indicated.

Keywords: polynomial neural networks, system identification, machine learning

REFERENCES

1. Mall S., Chakraverty S. Comparison of artificial neural network architecture in solving ordinary differential equations // Advances in Artificial Neural Systems, 2013.
2. Lagaris I. E., Likas A., Fotiadis D. I. Artificial neural networks for solving ordinary and partial differential equations // Technical report, 1997. URL https: //arxiv.org/pdf/physics/9705023.pdf.
3. Lapedes A., Farber R. Nonlinear signal processing using neural networks: Prediction and system modelling // IEEE International Conference on Neural Networks, San Diego, CA, USA, 1987.
4. Lewis F. L., Ge S. S. Neural Networks in Feedback Control Systems, in Mechanical Engineers // Handbook: Instrumentation, Systems, Controls, and MEMS, volume 2. John Wiley and Sons, Hoboken, NJ, USA, third edition (ed m. kutz) edition, 2005.
5. Chen S., Billings S. A. Neural networks for nonlinear dynamic system modelling and identification // International Journal of Control, 56(2):319–346, 1992.
6. Cybenko, G. V. Approximation by Superpositions of a Sigmoidal function // Mathematics of Control Signals and Systems,V. 2, P. 303-314, 1989.
7. Dragt A. Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics. 2011. URL http://inspirehep.net/record/955313/files/TOC28Nov2011.pdf.
8. Andrianov S. N. Modeling of Beam Dynamics Control Systems. Saint-Petersburg, 2004. 368 c.

9. Ivanov A. N., Kuznetcov P.M. Identification of dynamical systems based on nonlinear matrix Lie transform // Vestnik UGATU. — 2014. — T. 18. — No 2(63). — C. 251–256.
10. E. Tu, G. Zhang, L. Rachmawati, E. Rajabally and G. B. Huang, Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey From Data to Methodology // *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1-24.
11. Sholokhova A.A. Anomaly detection in Sensor Data in Application to the Analysis of Maritime Vessels Motion // Journ. of Modelling, Optimization and Information Technology. — 2017. — № 3(18). — 19 C.