ИСПОЛЬЗОВАНИЕ ВЕЙВНЕТОВ ПРОИЗВОЛЬНОЙ РАЗМЕРНОСТИ ДЛЯ МОДЕЛИРОВАНИЯ ОБЪЕКТОВ

Д.Е. Катаев, Я.М. Жаров (Москва)

Введение

В сферах имитационного моделирования и идентификации объектов управления существуют и используются подходы, основанные на искусственных нейронных сетях. Такой подход позволяет создавать вычислительно эффективные модели даже без полной информации об объекте, в том числе автоматически. Однако, нейросетевые модели имеют определенные недостатки. Главным из них можно считать общую невозможность интерпретировать ее параметры. Это ведет, с одной стороны, к тому, что валидация модели может быть выполнена только путем численных экспериментов, с другой стороны, к невозможности получить дополнительную информацию о моделируемом объекте. Еще одна проблема связана с качеством данных, используемых для обучения и валидации нейросетевых моделей [1, 3].

В рамках данного доклада мы рассмотрим использование вейвлетов в качестве активационной функции искусственной нейронной сети (ИНС) с одним скрытым слоем. Термин «вейвлон» будет обозначать нейрон, активационной функцией которого является вейвлет. Вейвлет-нейронная сеть - ИНС, скрытый слой которой состоит из вейвлонов, чьи сдвиги и сжатия задаются при инициализации и в процессе обучения не меняются. В свою очередь вейвнет определяется как вейвлет-нейронная сеть, сдвиги и сжатия вейвлонов которой так же обучаются.

Вейвлет-нейронные сети являются относительно изученными, в [1] показано, что такие сети способны быть устойчивы к шумам. В той же работе показано, что использование многомерных вейвлет-функций позволяет произвести более точную и тонкую настройку работы сети с отдельными входами или группами входов. Однако, использование обучаемых вейвлонов на данный момент слабо изучено. С другой стороны, такая архитектура открывает дорогу к аналитической интерпретации нейросетевой модели.

1.ВЕЙВЛЕТ-НЕЙРОННЫЕ СЕТИ

Вейвлет - функция с компактным носителем. Вейвлет преобразование было разработано как обобщение Фурье-преобразования, показавшего свою ограниченность при поиске закономерностей в некоторых сигналах. Вейвлет преобразование, в отличие от Фурье, переводит сигнал из временной области, в область вейвлет-коэффициентов - сдвиг · сжатие · уровень (или время · масштаб · уровень). При этом сдвиг s и сжатие τ - коэффициенты порождающие конкретный вейвлет $\psi_{s,\tau}$ из материнского ψ, по формуле (1). А уровень показывает насколько хорошо такой вейвлет аппроксимирует сигнал в момент времени t [2]. Таким образом мы получаем изображение, дающее нам понятие, как о частотных, так и о временных характеристиках сигнала.

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left((\frac{t-\tau}{s}\right) \quad (1)$$

Вейвлет-нейронная сеть представляет из себя ИНС, в качестве активационной функции нейронов скрытого слоя которой выступают вейвлеты порожденные от одной материнской функции. Так как вейвлеты обладают компактным носителем это позволяет конкретному нейрону реагировать не только на конкретный набор входов но и только в отдельном диапазоне. Это позволяет получить дополнительную помехоустойчивость без дополнительных мер, выходящих за рамки искусственной нейронной сети. Для обучения методом обратного распространения ошибки важным свойством является дифференцируемость активационной функции, поэтому некоторые вейвлеты не слишком хорошо подходят на роль активационной функции [1], [4].

2.ВЕЙВНЕТЫ

Вейвнет представляет из себя ИНС, в качестве активационной функции нейронов скрытого слоя которой выступают вейвлеты порожденные от одной материнской функции. Так как вейвлеты обладают компактным носителем это позволяет конкретному нейрону реагировать не только на конкретный набор входов но и только в отдельном диапазоне. В ходе обучения изменяются не только веса, но и коэффициенты конкретных вейвлетов.

Многомерные вейвлеты могут как создаваться в виде отдельных специальных функций, так и собираться из более простых [2]. Для сохранения компактности носителя функции целесообразно перемножать. Таким образом формулой многомерного вейвлета в общем виде будет (2).

$$\Psi(t1,...,tn) = \prod_{i=1}^{n}[\frac{1}{\sqrt{s_i}}\psi\left(\frac{t_i-\tau_i}{s_i}\right)] \quad (2)$$

Как видно из формулы (2), для каждой из переменных отдельно задается смещение и сжатие. Что позволяет точно и просто взаимодействовать с формой и положением вейвлета. Многомерный вейвнет отличается от вейвнета с одномерными вейвлетами тем, что при подаче в скрытый слой не суммирует аргументы, а подает каждый из аргументов на отдельный вход функции. Благодаря этому появляется возможность тоньше настраивать реакцию сети на конкретные входы.

3.РАБОТА И ОБУЧЕНИЕ ВЕЙВНЕТОВ

В случае вейвлет-нейронной сети и вейвнета вектор выходных значений формируется из элементов, рассчитываемых по формуле (3)

$$y_k = \sum_{j=1}^{hid}[\mu_{jk}\prod_{i=1}^{inp}\{\psi(\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}})\}] + \chi_k \quad (3)$$

где ψ - материнская вейвлет-функция, u_i - i-й входной сигнал, ω_{ij} - вес связи между i-м входом и j-м вейвлоном, t_{ij}, λ_{ij} - коэффициенты переноса и сжатия i-того вейвлона по j-тому входу, μ_{jk} - вес связи между j-м вейвлоном и k-м выходом, χ_k - добавочный вес для k-го выхода.

Адаптируем алгоритм обратного распространения ошибки [4] для случая обучаемых вейвлетов произвольной размерности в качестве активационных функций. Для удобства работы с производными введем обозначение (4)

$$\Xi_{qj} = \prod_{i=1}^{q-1}\{\psi(\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}})\}\dot{\psi}(\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}})\prod_{i=q+1}^{inp}\{\psi(\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}})\} \quad (4)$$

Исходя из этого, производные для параметров должны рассчитываться по формулам:

$$\frac{\partial y_k}{\partial \chi_k} = 1$$
$$\frac{\partial y_k}{\partial \mu_{jk}} = \prod_{i=1}^{inp}\{\psi(\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}})\}$$

$$\frac{\partial y_k}{\partial \omega_{ij}} = \Xi_{ij}\mu_{jk}\frac{u_i}{\lambda_{ij}}$$
$$\frac{\partial y_k}{\partial t_{ij}} = -\Xi_{ij}\frac{\mu_{jk}}{\lambda_{ij}}$$
$$\frac{\partial y_k}{\partial \lambda_j} = -\Xi_{ij}\mu_{jk}\frac{u_i\omega_{ij}-t_{ij}}{\lambda_{ij}^2}$$

Для удобства описания алгоритма перепишем эти функции в матричном виде. Для простоты записи матричных выражений введем несколько используемых промежуточных матриц, поэлементно определяемых как:

$$\rho_{ij} = [\prod_{k=1}^{i-1} z_{kj}]z'_{kj}[\prod_{k=i+1}^{inp} z_{kj}]$$
$$\eta_{ij} = \prod_{k=1}^{inp} z_{kj}$$
$$\nu_{ij} = \sum_{k=1}^{hid} \mu_{ik}$$
$$\phi_{ij} = u_j$$

Теперь, используя эти обозначения, можно достаточно просто записать единичные изменения в коэффициентах и весах сети на каждом шаге обучения по алгоритму обратного распространения ошибки:

$$\delta X = Err$$
$$\delta M = H.Err$$
$$\delta\Omega = \frac{PN\Phi}{\Lambda}$$
$$\delta T = \frac{PNT}{\Lambda}$$
$$\delta\Lambda = \frac{PN(d(U).\Omega - T)}{\Lambda}$$

где *Err* - вектор ошибки, а *Z*, *Z'* - вектора результатов вычисления активационной функции скрытого слоя и их производных соответственно. Операция «.» - матричное умножение, а *d()* - операция, которая обращает вектор в диагональную матрицу соответствующей размерности. Остальные действия производятся поэлементно.

4.ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА РАБОТОСПОСОБНОСТИ

Нахождение оценки функции или аппроксимация сигнала - одна из задач, решаемых вейвнетом. В качестве простого примера приведем задачу аппроксимации сигнала, образованного суммой двух синусоид разных частот и амплитуд с наложенными на него аддитивным и мультипликативным шумами. Графически сигнал представлен на рисунке 1. Вейвлет-нейронная сеть при случайной инициализации сдвигов и сжатий в заранее заданном диапазоне показала среднюю погрешность в 0,15%, аналогичный показатель для вейвнета составил 0,08%. В обоих случаях в качестве активационной функции использовался вейвлет Морле. На рисунке 2 представлен выход вейвнета, разложенный на выходы отдельных вейвлонов. Видно, что, фактически, каждый вейвлон отвечает за тренды с различными характеристиками в зависимости от сдвига и сжатия своего активационного вейвлета.

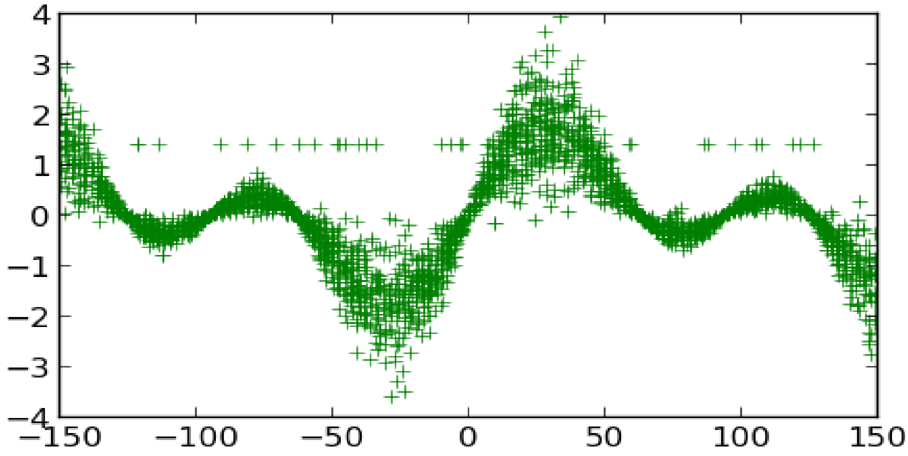

Рисунок 1: Зашумленный аппроксимируемый сигнал

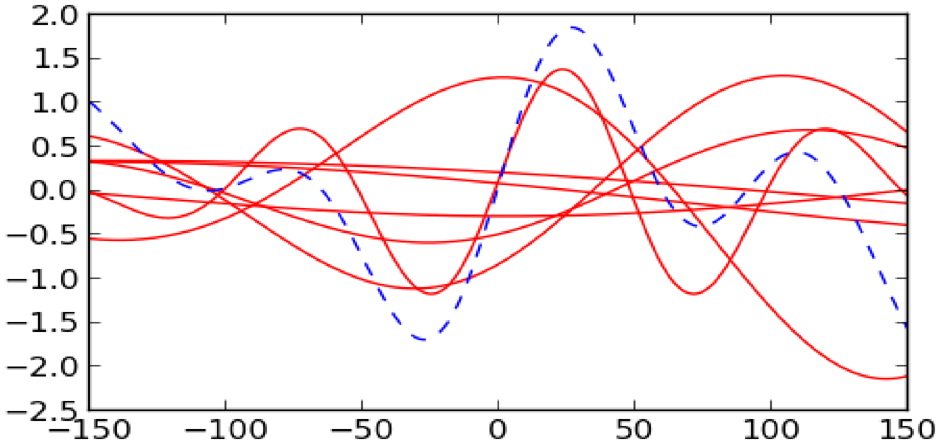

Рисунок 2: Аппроксимируемый вейвнетом сигнал (пунктирная линия) и выходы отдельных вейвлонов

Для исследования влияния размерности активационной функции на работу вейвнета проведен схожий эксперимент с нелинейной функцией двух переменных, графически представленной на рисунке 3. Эту задачу можно интерпретировать как простейший случай идентификации нелинейного объекта управления. Сравнивалась работа вейвнета с одномерными вейвлетами Морле и вейвнета с многомерными вейвлетами Морле при равно количестве эпох обучения. Погрешности составили, соответственно, 8,3% и 7,4%. Однако, более подробный анализ выявил, что одномерный вейвнет минимизировал ошибку лишь для нескольких локальных экстремумов функции, игнорируя при этом остальные. Это связано с частичной потерей компакности. Многомерный вейвнет, в свою очередь, проявил схожее качество работы во всей области определения аппроксимируемой функции.

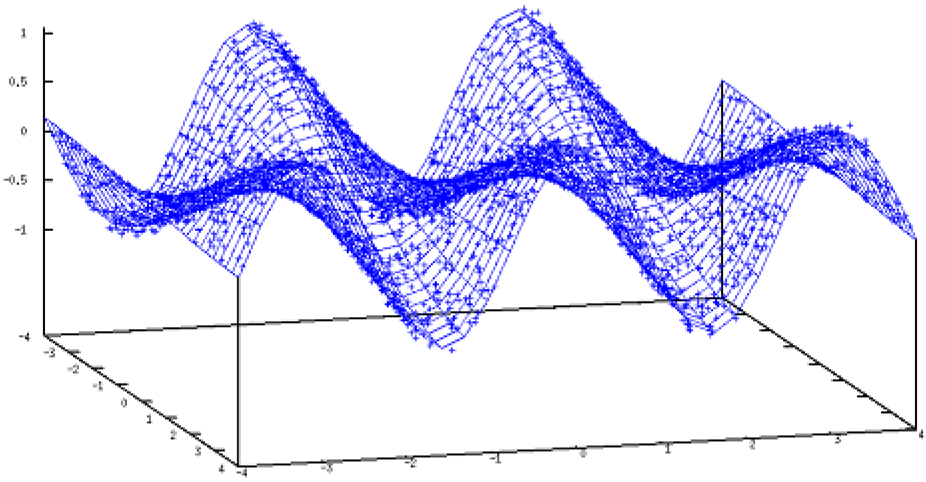

Рисунок 3: Аппроксимируемая функция двух переменных

Выводы

В рамках работы, представленной в данном докладе вейвлет-нейронные сети обобщены до вейвнетов путем адаптации алгоритма обратного распространения ошибки к обучению не только весов сети, но и параметров активационных функций произвольной размерности. Полученная новая архитектура нейронных сетей унаследовала от вейвлет-нейронных сетей повышенную устойчивость к шумам в данных, сохранив при этом преимущества нейросетевых моделей, однако, повысив при этом свою вычислительную сложность. Первичные эксперименты показывают, что при правильном выборе материнской вейвлет-функции такой подход может как улучшить качество работы, так и предоставить возможность интерпретации отдельных элементов нейросетевой модели. В свою очередь использование многомерных вейвнетов представляет перспективным с точки зрения имитационного моделирования и идентификации нелинейных объектов управления.

Литература

1. Veitch D. Wavelet Neural Networks and their application in the study of dynamical system. Departament of Mathematics University of York. 2005
2. Нагорнов О.В., Никитаев В.Г., Простокишин В.М., Тюфлин С.А., Проничев А.Н., Бухарова С.А., Чистов К.С., Кашафутдинов Р.З., Хоркин В.А. Вейвлет анализ в примерах. М.: НИЯУ МИФИ, 2010. – 120 с. ISBN 978-5-7262-1387-3
3. Комашинский В.И., Смирнов Д.А. К63 Нейронные сети и их применение в системах управления и связи. М.: Горячая линия-Телеком, 2003. 94 с. ISBN 5-93517-094-9
4. Николенко С.И., Тулупьев А.Л. Н63 Самообучающиеся системы. М.: МЦН-МО, 2009. 288 с. ISBN 978-5-940570-506-1