

ИНФОРМАТИКА

УДК: 681.322:517.444

**ИССЛЕДОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ
МЕТОДАМИ ЛИНЕАРИЗАЦИИ И ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ**

В.А. ПТИЧКИН

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь*

Поступила в редакцию 19 декабря 2003

Дается обзор научных результатов, полученных автором за последние годы в области теории искусственных нейронных сетей, статических и динамических. Они относятся к анализу преобразований, осуществляемых многослойными нейронными сетями, процессов обучения статических сетей и показателей качества работы нейродинамических систем управления. Используется единая методика анализа нейронных сетей и систем, основанная на идее линеаризации и полиномиальной аппроксимации функций активации формальных нейронов и нейросетевых преобразований.

Ключевые слова: многослойные нейронные сети, нейродинамические системы, анализ нейросетевых преобразований, метод статистической линеаризации.

Введение

Объективной причиной, обуславливающей широкое распространение нейросетевой парадигмы в различных областях науки и техники, является (до конца еще не осознанная) роль основной математической модели искусственной нейронной сети в ряду основных структур математики. На наш взгляд, эта модель в своей чисто математической формулировке заслуживает пристального внимания и широкого применения даже без какого-либо биологического обоснования и вне связи с задачами искусственного интеллекта.

Основное достоинство нейросетевых моделей заключается в представлении функций многих переменных в виде композиции n -арных линейных и унарных нелинейных функций определенного вида (функций активации). Это позволяет посредством сравнительно небольшого числа параметров представлять функции многих переменных в широкой области определения аргументов. Необходимость такой унифицированной формы представления функций нескольких переменных уже давно остро ощущалась во многих прикладных науках, использующих математические методы и модели представления данных.

До появления нейросетевой модели казалось, что единственной универсальной формой представления нелинейных функций нескольких переменных являются полиномы. Однако число параметров, необходимых для достаточно точного представления функции полиномом n переменных в достаточно широкой области, так резко растет с увеличением числа переменных и степени полинома, что говорить о практическом использовании такой формы можно только в ограниченном числе случаев. Они касаются либо малого отклонения нелинейных характеристик от линейных, либо столь малых областей определения полиномов, что в них упомянутое отклонение тоже можно считать малым.

На этом основании можно предположить, что композиция n -арных линейных и унарных полиномиальных функций позволяет посредством сравнительно небольшого числа параметров представлять функции многих переменных в широкой области определения аргументов.

Несмотря на то что в данном случае речь идет о полиномах одной переменной, желательно использовать полиномы возможно меньшей степени, чтобы сократить число параметров при описании функции. В предельно простом случае может идти речь об использовании полиномов первой степени, т.е. о линеаризации функций активации и нейронной сети в целом. Линеаризационные представления не исчерпываются линеаризацией в бесконечно малой окрестности. Например, в методах статистической и гармонической линеаризации диапазон линеаризации может быть сколь угодно большим. Эти методы были ориентированы на решение задач анализа систем, состоящих из одномерных нелинейных статических и линейных динамических звеньев. Однако практическое применение данные методы находили в основном при единственном нелинейном звене и одном или небольшом числе линейных звеньев произвольного порядка.

Обобщение линеаризационных методов на случай нескольких нелинейных элементов связано со столь значительным увеличением трудоемкости расчетов, что казалось неоправданным из-за несущественного расширения класса исследуемых таким образом систем. Однако, как показывает опыт использования нейросетевых моделей, расширение класса нелинейных систем может иметь качественный скачок после достижения числа нелинейных элементов некоторого критического значения. Трудно заранее оценить его, но можно принять за рабочую гипотезу, что статические (безынерционные) нейронные сети уже при нескольких десятках формальных нейронов могут решать практически полезные задачи. Для динамических нейронных сетей это число может быть на порядок меньше. Таким образом, имеет смысл вернуться к идее исследования систем линеаризационными методами.

Возникающие при этом трудности могут быть преодолены. Наша уверенность базируется в первую очередь на возросшей значимости проблемы: если раньше речь шла о некотором расширении, казалось, узкого класса систем, то теперь она касается нейронных сетей и систем. Другим источником уверенности в возможности преодоления трудностей, связанных со значительным увеличением числа нелинейных элементов в системе, являются результаты, полученные нами ранее в рамках классической теории управления [1]. На многочисленных примерах показано, что общепринятые ограничения в форме предположения о нормальности распределения входных координат в методе статистической линеаризации или о гармоническом характере воздействий в методе гармонической линеаризации могут столь значительно нарушаться, что при определенном соотношении между параметрами одни коэффициенты могут заменять другие. Более того, имеется возможность использования гораздо более простых выражений коэффициентов линеаризации, чем известные, для типовых нелинейностей классических схем управления и регулирования [1]. Заметим, что многие функции активации формальных нейронов имеют в точности такое же аналитическое описание, что и упомянутые типовые нелинейности, или могли бы быть отнесены к ним.

К недостаткам линеаризационных методов в теории управления относится тот факт, что на завершающем этапе они требуют решения систем нелинейных уравнений, число которых может быть значительным в нейросетевых приложениях. Однако за последнее время отношение к необходимости решения систем нелинейных уравнений на завершающей стадии исследования существенно изменилось. Эта задача уже не выглядит столь сложной, как это было несколько десятилетий тому назад. Все это и обусловило нашу уверенность в полезности применения методов классической теории управления к решению задач теории нейронных сетей.

Некоторым препятствием на пути применения классических методов теории управления для решения задач теории нейронных сетей является принципиальное различие в приоритете задач, решаемой той и другой теориями. В классической теории управления основной, базовой задачей является задача анализа замкнутых динамических систем, а в теории нейронных сетей — алгоритмы обучения нейронных сетей. По классификации теории управления последние относятся к задачам параметрической оптимизации статических разомкнутых систем методом последовательных приближений. В теории управления такая задача даже не обсуждается.

В свою очередь, задача анализа преобразований, осуществляемых статическими нейронными сетями, в теории нейронных сетей не рассматривается.

Вместе с тем она является не только начальной на пути проникновения методов теории управления в теорию нейронных сетей, но и основной, базовой в теории динамических нейронных систем. Основное содержание настоящей работы — выяснение взаимосвязи данной задачи с другими задачами теории нейронных сетей, необходимость решения которых не вызывает сомнений.

Задачи анализа и синтеза нейронных сетей и систем

Многослойные нейронные сети (МНС) и системы являются, хотя и новым, но все же частным случаем систем преобразования информации и управления. Поэтому все то, что можно сказать о классификации систем и задач теории управления имеет самое непосредственное отношение к нейронным сетям и системам. Системы (управления или преобразования информации) можно разделить на статические (безынерционные) и динамические. Методы исследования динамических систем базируются на методах анализа статических систем. Это не очевидно, точнее, завуалировано гораздо большей предысторией развития и применения математических моделей безынерционных преобразований во всех областях науки и техники, чем математические модели динамических систем, но это так. Системы управления относятся к динамическим системам, о статических преобразованиях говорят только при описании отдельных звеньев в структурном представлении динамических систем. В теории нейронных сетей статические режимы работы, как наиболее простые, долгое время были, если и не единственным, то основным предметом изучения.

С формальной точки зрения все задачи теории управления можно разделить на задачи анализа и синтеза. Анализ систем с содержательной точки зрения не является самостоятельной задачей. Она только выглядит таковой в достаточно устоявшихся разделах теории, а в действительности, явно или неявно, является составной частью процесса синтеза или оптимизации. В первом случае алгоритмы, методы и результаты анализа являются составной частью формальной процедуры синтеза, во втором – сама процедура синтеза может быть не до конца формализована, поэтому связь между задачами анализа и синтеза находится за пределами теории управления. Коротко говоря, методы синтеза систем базируются на методах анализа, а в своей простейшей постановке, как задачи параметрического синтеза, являются естественным обобщением задач анализа. В связи с этим применение математической модели нейронной сети для решения традиционных задач анализа статических и динамических систем управления является важной задачей на пути проникновения теории нейронных сетей в теорию управления. Не менее важно применение методов теории управления для решения традиционных задач теории и практики нейронных сетей.

Основное внимание в теории нейронных сетей и систем уделяется решению задач синтеза. Тем не менее, задачи анализа имеют право на существование. Говоря о нейронных системах, мы имеем в виду более общее понятие, чем нейронная сеть, даже в том случае, когда идет речь о статической системе. Нейронная сеть может быть центральным, но не единственным звеном нейронной системы. Например, нейронная сеть в процессе обучения может рассматриваться как часть системы, включающей в себя дополнительно блок вычисления показателя качества и источник сигнала "учителя". Эта система может рассматриваться как часть системы, включающей в себя блоки вычисления градиента показателя качества и механизмов реализации изменений параметров и т.д.

В современной теории нейронных сетей, как правило, исходят из того, что структура сети задана и целью обучения является нахождение значений параметров сети из условия минимума ошибки. На наш взгляд, такую постановку задачи можно отнести к задачам параметрической оптимизации. Последние имеют столь много общего с задачами анализа, что о параметрической оптимизации можно говорить как о естественном продолжении задачи анализа.

Более того, имеет смысл разделять задачи анализа систем на задачи анализа в узком и широком смысле этого слова. Под анализом в узком смысле следует понимать определение характеристик выходного сигнала по аналогичным характеристикам входного при известном

описании системы. Анализ системы в широком смысле этого слова включает в себя определение показателей качества работы системы и зависимостей показателя качества от параметров системы. Вопросы использования упомянутых зависимостей для обучения сетей или параметрической оптимизации их уже выходят за рамки задач анализа даже в широком смысле этого слова. Однако приходится различать не только задачи анализа в узком и широком смыслах, но и задачи анализа сетей и систем. Даже минимальное расширение сети до системы, вычисляющей квадратический показатель качества для каждого образца в отдельности, существенно усложняет решение задачи анализа.

В классической теории управления необходимость решения задач анализа даже не обсуждается. Может быть это и служило когда-то предметом обсуждения, но, вероятнее всего, они с самого начала были единственной формализованной составной частью процесса синтеза систем. В теории нейронных сетей эта задача (анализа в узком смысле) даже не ставилась, насколько нам известно.

Действительно, если под анализом нейронной сети понимать определение значений выходных координат по заданным значениям входных координат, то она имеет элементарное решение и не достойна такого названия. Однако условия функционирования МНС могут описываться не детерминированным значением входного вектора, а случайным. Тогда задача определения распределения выходных координат по распределению входных уже становится достаточно трудной даже для сравнительно простых законов распределения входного вектора.

В теории автоматического управления моментный подход к описанию случайных процессов изначально лежал в основе методов исследования систем. Может быть, причиной послужила главенствующая роль линейных моделей в классической теории управления. При исследовании таких систем можно ограничиться рамками корреляционной теории, т.е. описанием всех воздействий только моментами второго порядка и определением моментов того же порядка так называемых "выходных координат", в число которых может входить и сигнал ошибки. При исследовании нелинейных систем нередко тоже достаточно определения моментов второго порядка для суждения о точности работы системы. Однако для более точного определения моментов второго порядка какой-либо одной координаты (дисперсии ошибки, например), как правило, требуется оперировать моментами более высокого порядка всех координат. Обычно это требование формулируется не в терминах моментов, а в терминах функций распределения или плотностей вероятности.

В связи с этим в классический период развития теории автоматического управления большой популярностью пользовался так называемый метод статистической линеаризации, который предназначен для анализа точности нелинейных динамических систем в рамках корреляционной теории. В своей классической формулировке метод ориентирован на определение дисперсии ошибки замкнутой динамической системы, состоящей из нелинейного статического блока и линейного динамического звена, на вход которой поступает стационарный случайный процесс с известными математическим ожиданием и корреляционной функцией или спектральной плотностью.

МНС и сама по себе, т.е. вне связи с последующими динамическими преобразованиями, может решать достаточно важные и актуальные задачи преобразования информации. Для анализа работы даже простейшей нейросетевой системы необходимо располагать описанием МНС в более удобной для решения аналитических задач форме, чем исходное описание МНС. Такой может быть наилучшая линейная (желательно даже полиномиальная) аппроксимация математической модели МНС. Сказанное почти дословно повторяет постановку задачи анализа МНС как составной части нейродинамической системы. Можно продолжить эту аналогию, включив в нее ограничение самыми простыми законами распределения входного для МНС вектора.

Однако если применительно к анализу нейродинамических систем предположение о нормальности распределения входного для МНС вектора кажется приемлемым во многих практических ситуациях, то при решении задач распознавания оно уже не кажется таковым. Более того, оно выглядит противоестественным или настолько стеснительным, что возникает опасение получить в результате анализа МНС в этих условиях бесполезные для практики результа-

ты. Однако это не так. Хотя данные ограничения весьма существенны, при их выполнении можно получить решение некоторых интересных задач анализа нейронных сетей и систем.

Надеяться на достаточно полное описание входных образов МНС столь бедными характеристиками как статистические моменты второго порядка можно только в исключительных случаях. Однако возможна аппроксимация распределения входных координат выпуклой линейной комбинацией нормальных распределений с различными параметрами. Это позволяет использовать разработанные для этого простейшего случая методы анализа при определении вероятностных характеристик выходных координат с учетом достаточно сложных законов распределения входных координат.

Другой областью применения предлагаемых здесь методов является линейная аппроксимация преобразования, осуществляемого МНС, в заданной области изменения входных координат. После замены исходного описания нейросетевого преобразования его линеаризованным представлением уже можно дать (в первом приближении) решение более сложных задач, чем формальная задача анализа. Примером могут служить градиентные методы оптимизации, где линейная аппроксимация используется для оптимизации нелинейных функций.

Как известно, наилучшей (в смысле минимума среднеквадратической ошибки) линейной аппроксимацией преобразования в условиях действия на входе сигнала с заданным законом распределения является аналогичная аппроксимация условного математического ожидания выходных координат относительно входных. Таким образом, определение коэффициентов линеаризации направлено и на оценку упомянутых условных математических ожиданий.

В теории нейронных сетей обычно рассматриваются преобразования детерминированных входов. Представление входного образа случайным вектором есть не что иное, как удобная форма описания множества образов. Например, на основании качественных рассуждений о характере представления образов разных классов в пространстве признаков, которое и есть пространство входных координат МНС, можно предположить, что образы каждого класса группируются около своего центра и чем дальше от него, тем реже встречаются точки (векторы) этого класса. Используя вероятностную терминологию, данную ситуацию можно представить как представление каждого класса векторной случайной величиной, а каждого образа этого класса — реализацией данной случайной величины. Простейшей характеристикой области группирования каждого класса могут служить координаты центра этой области и, если можно так выразиться, эллипсоида рассеяния. В вероятностных рассуждениях аналогичную роль играют первые два момента: математическое ожидание и ковариационная матрица. Первая характеризует центр, а вторая — эллипсоид рассеяния, т.е. дает представление о разбросе каждой координаты с помощью соответствующей дисперсии и об ориентации эллипсоида с помощью корреляционных моментов или коэффициентов корреляции.

Таким образом, расплывчатые термины, касающиеся расположения и формы областей сосредоточения образов одного класса, становятся строго определенными в вероятностных терминах и, что еще важнее, между ними существуют строго определенные взаимосвязи. От использования вероятностной терминологии детерминированная по своей сути задача не становится вероятностной, но позволяет использовать математический аппарат теории вероятностей и математической статистики.

Анализ нейронных сетей

Анализу нейронных сетей посвящено несколько наших работ. В самой первой из них [2] уже дается основное содержание всего сказанного выше о постановке задачи анализа статических нейронных сетей и о возможности использования для этой цели метода статистической линеаризации. Казалось, все самое существенное об анализе МНС в ней сказано. Осталось только привести хорошо известные соотношения между моментами при линейном многомерном и одномерном нелинейном преобразованиях, согласовав обозначения и терминологию. Работа [3] была посвящена решению этой задачи. В ней предложен приближенный метод определения моментов первого и второго порядков всех координат МНС и линейных аппроксимаций условных моментных функций выходных координат сети относительно других ее координат. Впоследствии к этой задаче пришлось вернуться, так как оказалось, что даже в рамках корре-

ляционной теории, т.е. используя только моменты первых двух порядков можно, существенно повысить точность предложенных в предыдущей работе методов. Работа [4] содержит изложение этого нового подхода. С предложенными методами можно познакомиться и по переводам двух последних работ [5, 6]. Как уже неоднократно отмечалось, анализ МНС не является самостоятельной задачей, она, как правило, выступает составной частью другой задачи, более близкой к практическому применению. В связи с этим вряд ли когда-нибудь задачу анализа можно считать окончательно решенной. Особенно, когда речь идет не о точных, а о приближенных методах решения. Повышение требований к точности анализа, как правило, связано с усложнением алгоритмов анализа и понижением их эффективности. Поэтому менее точные методы не всегда должны уступать место более точным методам решения той же задачи. При выборе между ними всегда желательно ограничиться достаточной точностью.

Исключение может представлять появление нового качественного результата с повышением точности анализа. Например, в рамках корреляционной теории даже точный результат анализа может претендовать только на определение линейной аппроксимации условных математических ожиданий выходной координаты в зависимости от параметров. Для определения, хотя бы и самых грубых, оценок квадратичных зависимостей необходимо выйти за рамки моментов второго порядка при описании распределений координат МНС и линеаризованных представлений функций активации. Использованию полиномиальной аппроксимации как естественного обобщения метода статистической линеаризации посвящена работа [7] которая тоже, вероятно, не будет последней, посвященной анализу МНС.

Анализ многослойной нейронной сети в узком смысле этого слова в детерминистическом случае не только не представляет труда, но даже сама постановка этой задачи кажется лишенной содержательного смысла. На первый взгляд кажется даже, что математическая модель такой сети и не преследует никакой иной цели кроме описания алгоритма вычисления выходных координат наиболее простым образом. Определение значения выходных координат сети с последовательными связями по значениям ее входных координат сводится к последовательному выполнению преобразований для каждого слоя нейронов в отдельности. В свою очередь, преобразование, выполняемое каждым слоем, сводится к линейному преобразованию входного вектора в вектор входных координат для нелинейных преобразований, заданных функциями активации нейронов этого слоя, вычитанию из каждой координаты полученного вектора величин смещений каждого нейрона и, наконец, нелинейному преобразованию каждой координаты в отдельности.

Формальное обобщение данной задачи на случай, когда на входе действует случайный вектор, с теоретической точки зрения не представляет принципиальных трудностей. Проблема может заключаться только в выборе наиболее эффективных процедур и аналитических выражений для анализа преобразований двух типов: многомерного линейного и покоординатного нелинейного преобразования случайного вектора. Последовательное выполнение (суперпозиция) этих преобразований описывает преобразование, осуществляемое отдельным слоем. В свою очередь, суперпозиция послойных преобразований полностью описывает МНС с последовательными связями. Таким образом, достаточно рассмотреть только два указанных выше типа преобразований случайных векторов в такой форме, чтобы выполнение суперпозиции между ними не вызывало затруднений.

Определение характеристик линейного преобразования случайного вектора представляется столь простой задачей, что любая формализация этой задачи кажется излишней. Чаще всего в рассматриваемой ситуации ограничиваются определением только моментов второго порядка. В этом случае проще всего по отдельности определять моменты первого и второго порядков в виде взаимосвязи между математическими ожиданиями и корреляционными (ковариационными) матрицами на входе и выходе линейного преобразователя. Такие выражения явно содержат матрицу линейного преобразования и потому широко используются в практике вычислений.

В цитированных работах предлагается другой подход к анализу линейных преобразований: определение столбца начальных моментов второго порядка выходного вектора по столбцу моментов того же порядка входного вектора. Причиной такого нетрадиционного подхода даже в рамках корреляционной теории явилось желание представить результат анализа в той же

форме, что и результат нелинейного по координатам преобразования для простоты представления результата суперпозиции этих преобразований.

Действительно, при линеаризации нелинейных преобразований (функций активации формальных нейронов) классическими для теории управления методами можно получить совокупность коэффициентов линеаризации (по математическому ожиданию и дисперсии, к тому же по первому и второму способам) для каждой функции активации в отдельности. Выполнение композиции многомерного линейного и по координатам нелинейного в этих условиях, конечно, возможно, но полученный результат будет лишен наглядности и труден для использования при композиции с преобразованием следующего слоя. Если же оба преобразования (линейное и нелинейное) представить в матричной форме как соответствие между столбцами моментов определенного порядка, то выполнение суперпозиции сводится к перемножению матриц.

Для получения соответствия между столбцами моментов входного и выходного векторов в матричной форме можно даже не прибегать к вероятностным рассуждениям, если известна матрица связи между этими векторами. Достаточно выразить столбец всех одночленов определенного порядка от всех составляющих выходного вектора через столбец всех одночленов того же порядка от составляющих входного вектора и представить результат в матричной форме. Полученная матрица будет и матрицей связи между столбцами моментов. Более того, эта матрица описывает и выражение условных моментных функций выходных координат в зависимости от "точечных" значений входных координат.

Если при ограничении моментами второго порядка использование подобной матрицы, как уже отмечалось, представляет собой не более чем удобную форму представления известных результатов, получаемых иным путем, то при использовании полиномиальной аппроксимации общепринятой формы представления соотношения между моментами выше второго порядка, насколько нам известно, не существует. Предложенный способ определения этого соотношения в таком случае не имеет альтернативы и должен использоваться не только для представления данного результата, но и для фактического вычисления столбца моментов выходного вектора по заданному значению столбца моментов входного вектора [7].

Из сказанного следует, что для завершения анализа преобразования, осуществляемого одним слоем сети, описанным выше способом необходимо представить нелинейное по координатам преобразование в матричном виде. Очевидно, что это описание может быть только приближенным. Так как исходное нелинейное описание представляет собой совокупность функций активации, то приближенного описания каждой из них в полиномиальном виде не достаточно для представления преобразования столбца моментов входного вектора в столбец моментов выходного вектора. Для этого необходимо располагать также приближенным полиномиальным представлением произведения k унарных функций активации, если в число моментов выходного вектора, которые желательно определить, входят смешанные моменты k -го порядка.

Совершенно ясно, что речь может идти только о моментах и, следовательно, полиномах весьма низкого порядка. Не будь этого обстоятельства, не было бы проблемы в полиномиальном представлении произведения двух функций активации после такого же представления каждой функции в отдельности. Для этого бы достаточно было перемножить полиномиальные представления каждой из них. Однако порядки полиномов при перемножении складываются. Такой подход приемлем только при линеаризационных представлениях всех функций активации [3]. В дальнейших работах предложены иные подходы к ограничению полиномами заданной степени, которые позволяют отразить нелинейные свойства как отдельных функций активации, так и их произведений. Одни из них существенным образом используют свойства функций активации [4], а другие — аппроксимацию полиномов нескольких переменных полиномами тех же переменных меньшей степени [7].

В теории управления при анализе замкнутых динамических систем используется предположение о нормальности законов распределения координат, подвергающихся нелинейному преобразованию. Основанием для этого является выполнение гипотезы фильтра. В теории нейронных сетей ее роль может играть центральная предельная теорема, согласно которой можно считать, что сумма достаточно большого числа случайных величин с произвольными законами распределения подчиняется нормальному закону распределения.

На наш взгляд, даже значительное нарушение предположения о характере закона распределения координат, подвергающихся нелинейному преобразованию активационной функцией, не ведет к катастрофическим последствиям. Самое большое влияние на коэффициенты линеаризации функций активации оказывают значения первых двух моментов ее аргумента. В обоснование этой точки зрения можно привести немало как рассуждений качественного порядка, так и примеров сравнения коэффициентов линеаризации некоторых типовых нелинейностей, полученных при разных законах распределения [1]. Коэффициенты линеаризации функций активации, полученные в предположении о нормальности законов распределения их аргументов, применялись при анализе тестовой задачи, в которой закон распределения принципиально иной [3, 4, 7]. В этих примерах аргументы пороговых функций активации могли принимать свои значения из двухэлементного множества. Тем не менее, результаты анализа имели допустимую точность.

В настоящее время чаще всего используются однородные нейронные сети с непрерывными функциями активации сигмоидного типа. В связи с этим вполне естественным было бы в первую очередь рассмотреть вопросы полиномиальной аппроксимации совокупности таких функций активации. Однако в цитированных выше работах использовались только сигнум - функции активации, униполярные и биполярные.

Для такого предпочтения есть несколько оснований. Первоначально рассматривались только пороговые функции активации. Утрата лидирующих позиций пороговой функцией активации, на наш взгляд, связана с появлением алгоритма обратного распространения, в основе которого лежит определенный способ вычисления производных выходной координаты от других координат сети.

Пороговые функции активации могут вновь играть ведущую роль в архитектуре нейронных сетей, как только отпадет необходимость использования производных функций активации. Например, решение задачи распознавания или классификации, являющейся центральной задачей теории статических нейронных сетей, практически всегда можно представить себе как задачу разбиения пространства признаков (входных координат МНС) на связные области значений признаков одного класса. Упомянутое разбиение с достаточной степенью точности всегда можно провести с помощью набора гиперплоскостей в пространстве признаков и последующей логической обработки информации о принадлежности образа той или иной области. Кстати, логическую обработку в нейросетевом базисе удобнее всего проводить на сетях с пороговыми функциями активации.

И все же за долгие годы использования непрерывных функций активации сложился определенный круг задач, где способность функций активации принимать свои значения из континуального множества является достоинством. В первую очередь это относится к задачам аппроксимации функций многих переменных в нейросетевом базисе. В этом случае ограниченная линейная функция активации может заменить большинство из известных в настоящее время непрерывных функций активации. Сравнительный анализ функций активации может быть предметом самостоятельного исследования, ориентированного на упрощение структурной модели сети [8].

Таким образом, в рамках предлагаемых методов анализа можно ограничиться всего двумя функциями активации: пороговой и ограниченно линейной. При этом речь может идти именно о двух функциях, а не о двух типах функций, так как за счет структурных преобразований произвольные параметры этих функций, которые по содержательному смыслу желательно использовать при проектировании сети, можно отнести к параметрам линейных преобразований между слоями сети [8].

Это обстоятельство весьма существенно для обоснования возможности практического использования предлагаемого подхода к анализу МНС. Действительно, трудоемкость матричного описания слоя из нелинейных по координатам преобразований столь велика при выходе за рамки линеаризации, что, будучи составной частью предлагаемого метода анализа, она могла бы служить труднопреодолимым препятствием на пути практического использования предлагаемого метода, если бы это было необходимо делать при анализе каждой конкретной системы. Однако, к счастью, это не так. Достаточно один раз проделать работу по матричному описанию совокупности из k скалярных функций активации одного типа в предположении, что их аргу-

менты подчиняются k -мерному закону распределения, например, нормальному, чтобы потом воспользоваться этими результатами при анализе однородных нейронных сетей с данным типом функций активации и любым числом нейронов в каждом слое. Во многих случаях можно ограничиться малой степенью полиномов ($k \leq 4$), но даже квадратичные аппроксимации всех зависимостей ($k=2$) представляют интерес и целый ряд еще неиспользованных по сравнению с методом статистической линейаризации ($k=1$) возможностей. Вообще, моментный анализ покоординатного нелинейного преобразования вектора с заданным, например, нормальным распределением представляет собой наиболее трудоемкую часть предлагаемого метода.

Несмотря на то что при выходе за рамки корреляционной теории порядок используемых моментов и полиномов теоретически может быть любым, практически можно рассчитывать только на использование моментов и полиномов не выше третьего-четвертого порядка. Причиной тому является резкий рост числа моментов и одночленов с ростом размерности n вектора координат и порядка полинома k . Проблема повышения порядка полиномов n переменных, которые можно использовать при полиномиальной аппроксимации нейросетевых преобразований, лежит за рамками теории управления и теории нейронных сетей. Вызывает некоторое недоумение отсутствие в современных системах программирования какой-либо программной поддержки для выполнения алгебраических операций над полиномами многих переменных.

В этом направлении нами сделаны определенные шаги [9], которых в настоящее время достаточно только для "внутреннего использования" при расчетах коэффициентов полиномиальной аппроксимации покоординатных нелинейных преобразований. Продолжение работ по программному обеспечению операций над полиномами многих переменных позволит повысить порядок используемых полиномов при анализе систем методом полиномиальной аппроксимации.

Линеаризационные методы в алгоритме обратного распространения

Алгоритм обратного распространения сыграл самую значительную роль в теории и практике нейронных сетей. По существу, он представляет собой нейросетевую реализацию одного из градиентных методов поиска минимума квадратического показателя качества в пространстве параметров нейронной сети. После установления этого факта оказалось возможным применение практически всех других методов оптимизации функций для обучения нейронных сетей. Таким образом, многообразие методов обучения не отменяет центрального положения алгоритма обратного распространения.

Использование давно устоявшейся терминологии методов оптимизации скрывает линеаризационную суть градиентных методов вообще и алгоритма обратного распространения в частности. При традиционном изложении данного алгоритма на первый план выступает техника определения производных. Цепное правило вычисления производных может быть и полезно (при очень большом числе слоев МНС), но для понимания существа дела технику определения производных лучше не принимать во внимание или использовать менее эффективные, но прозрачные для теоретических рассуждений приемы. Те же самые выражения значений производных, что и в алгоритме обратного распространения могут быть получены на основании представления каждой нелинейной функции, использованной при описании МНС, рядом Тейлора и ограничения линейными членами. Формальное доказательство этой возможности приведено в работе [10]. Приведем содержательную сторону дела, чтобы затем, оперируя другими методами линеаризации, предложить определенную модификацию алгоритма обратного распространения, которая не требует вычисления производных. Это позволяет применять данный алгоритм к сетям с не дифференцируемыми, разрывными функциями активации.

В самом кратком изложении алгоритм обратного распространения сводится к определению производных квадратического показателя качества в пространстве параметров нейронной сети. Их можно представить в виде произведения двух сомножителей. Одним из сомножителей выражения производных является значение ошибки после сравнения ответа "учителя" и ответа сети. Вторым сомножителем являются производные выходной координаты сети по параметрам. В свою очередь эти производные выражаются через производные выходной координаты по аргументам функций активации отдельных нейрона. Действительно, аргументом

функций активации нейрона является сумма порогового значения этого нейрона и линейной комбинации его синаптических весов — параметров МНС. Таким образом, вычисление производных выходной координаты МНС по параметрам сводится к вычислению производных выходной координаты по координатам, которые являются входными для нелинейных элементов сети, т.е. аргументами функций активации.

Заметим, что производные используются только для линейной аппроксимации нелинейных зависимостей, но после упомянутой аппроксимации сеть представляет собой последовательное соединение линейных преобразователей. Пользуясь терминологией теории управления, можно сказать, что нейронная сеть представляет собой последовательно-параллельное соединение статических элементов с заданными коэффициентами усиления. Вычисление коэффициентов усиления между любыми двумя точками такой сети не представляет никакого труда, особенно если использовать введенные в работе [10] матричные обозначения и соотношения.

Предпринятые в данной работе усилия были направлены не столько на иной способ получения известных результатов, сколько на то, чтобы исключить само упоминание о производных в случае, когда получены линейные аппроксимации всех нелинейных функций каким-либо иным путем, например методом статистической линеаризации [11].

Итак, если вместо производных функций активации использовать коэффициенты их линеаризации, то по правилам анализа линейных систем можно определить коэффициенты линеаризации нелинейных зависимостей выходной координаты МНС от координат на входах в нелинейные элементы. Это по существу и составляет задачу анализа МНС в широком смысле этого слова. Все сказанное об анализе МНС не трудно перенести и на анализ нейронной системы, включающей в себя, кроме МНС, источник сигнала учителя и блок вычисления квадратического показателя качества. Такие системы используются для решения задачи распознавания и аппроксимации неизвестной функции нейросетевой моделью. Зависимость выходной координаты такой системы от аргументов функций активации или ее линейная аппроксимация содержит информацию о направлении изменения упомянутых аргументов для уменьшения показателя качества. Другими словами, таким образом можно определить градиент показателя качества в пространстве аргументов функций активации. Последние линейно выражаются через синаптические коэффициенты и пороговые значения формальных нейронов. Коэффициентами упомянутых линейных зависимостей в данном случае являются значения математических ожиданий выходных координат формальных нейронов. Так как перед определением коэффициентов статистической линеаризации (или параллельно с ним) проводится определение моментов до второго порядка включительно всех координат МНС, то проблему вычисления градиента показателя качества в пространстве параметров можно считать решенной.

Непосредственному применению описанной последовательности действий для решения задач распознавания или аппроксимации мешает только то обстоятельство, что описание условий работы (совместного распределения входных координат МНС и сигнала учителя) ограничивается моментами второго порядка. Они могут отразить особенности множества входных образов только тогда, когда совокупность моментных характеристик одного класса существенно отличается от аналогичных характеристик другого класса. По-видимому, это не так часто встречается в практике распознавания образов, хотя в качестве тестовых примеров для разработки и проверки различных методов обучения такое описание условий функционирования распознающей системы можно использовать. Распространение же предложенных методов на случай использования полиномов и моментов выше второго порядка еще не завершено.

Анализ и синтез нейродинамических систем

В теории автоматического управления описание условий функционирования моментами второго порядка часто считается достаточным. Точнее, этот вопрос не обсуждается в пределах теории управления, а только перед практическим применении результатов теории управления.

По своему назначению нейродинамические системы имеют столь много общего с системами управления, что возникает желание воспользоваться методами классической теории регулирования и управления для их исследования. Это тем более актуально что, по мнению

многих авторитетных специалистов, недалеко то время, когда все системы регулирования и управления станут нейродинамическими [12]. Причиной такого развития событий станут, по-видимому, не только и не столько соображения, связанные с интеллектуализацией систем управления, сколько использование технических возможностей сравнительно легко реализуемых нейроконтроллеров в традиционных схемах управления, например, при использовании нейроконтроллера в замкнутом контуре управления или модели объекта управления в нейросетевом базисе в самонастраивающейся системе.

Говоря о методах классической теории управления, мы не имеем в виду исключительно спектральные методы, а только многообразие исходных форм описания систем, связанных со структурными представлениями. Другими словами, описание систем в пространстве состояний нисколько не препятствует применению классических методов исследования систем управления, особенно когда речь идет о стационарных системах.

Точно так же, говоря о классических методах, мы не имеем в виду их непосредственное применение, а только использование их методологических особенностей, которые вытекают из представления системы в виде соединения подсистем и звеньев, каждое из которых имеет отдельное, как правило, приближенное описание своей собственной математической моделью. Другими словами, никак не формализованная, но, тем не менее, часто используемая в приложениях замена аппроксимации композиции преобразований композицией аппроксимаций тех же преобразований лежит в основе предлагаемых методов исследования нейронных систем и сетей.

Такая замена может быть строго обоснована только при аппроксимации нелинейных функций рядами Тейлора в окрестности "положения равновесия". При расширении области аппроксимации до конечной с той или иной весовой функцией точность композиции аппроксимаций функций, как правило, может быть только ниже аппроксимации их композиций. Тем не менее (за неимением лучшего), предлагается осуществлять наилучшую аппроксимацию каждой нелинейной функции, входящей в композицию, с использованием техники разложения по ортогональным полиномам, а затем композицию разложений. После этого, не претендуя на наилучшую аппроксимацию композиции преобразований, предлагается оперировать ею как достаточно точной в том классе функций или преобразований, к которому относится окончательный результат.

МНС представляется суперпозицией (последовательным соединением) многомерных линейных и покоординатных (унарных) нелинейных преобразований. Анализ и тех и других по отдельности не представляет принципиальных трудностей. Однако возникают чисто технологические трудности при создании эффективных алгоритмов совместного исследования этих преобразований, особенно при выходе за рамки линеаризационных представлений и корреляционной теории. Преодолению их должно способствовать то обстоятельство, что единственным типом нелинейных функций в описании МНС являются функции активации формальных нейронов. Поэтому можно пойти даже на значительные трудозатраты по приближенному описанию нелинейного покоординатного преобразования системой уравнений моментов, условных или безусловных, в матричной форме, когда каждая координата преобразуется в соответствии с функцией активации. Действительно, полученные таким образом результаты могут использоваться при исследовании любой другой однородной МНС с той же функцией активации формальных нейронов.

Нейродинамическую систему в простейшем случае можно представить как соединение МНС линейной динамической системы по схеме обратной связи [13, 14]. Первая из них может описывать нейроконтроллер, а вторая — объект управления. Если последний описывается нелинейным дифференциальным или разностным уравнением, которое, как это часто бывает на практике, структурно можно представить в виде соединений линейных динамических и одномерных нелинейных статических звеньев, то и в этом случае о нейродинамической системе можно говорить как о композиции линейных динамических и нелинейных статических. Последние описываются уже не только функциями активации формальных нейронов, но и так называемыми типовыми нелинейными характеристиками. Среди них наиболее употребительными являются пороговые (релейные) и ограниченно линейные (линейные с насыщением), с зоной нечувствительности или без нее. Не трудно показать, что любую из этих характеристик с зоной

нечувствительности можно получить в виде линейной комбинации характеристик без зон нечувствительности. Другими словами, наличие типовых характеристик в описании системы, как правило, не привносит ничего нового в математическую модель динамической системы, содержащую функции активации.

Ограничение линейными динамическими и нелинейными унарными функциями в областях пространства состояний практически не ограничивает класс математических моделей динамических систем [15]. Практически все известные модели динамических нейронных систем представимы как нейросетевые модификации общей модели нелинейной динамической системы в пространстве состояний. Это открывает возможности применения линеаризационных методов для исследования более широкого класса систем, чем классические системы, состоящие из одного линейного динамического звена и одного нелинейного, заданного МНС.

Однако чисто технические сложности исследования динамических систем таковы, что представляет интерес преодоление их для классической схемы систем управления в рамках корреляционной теории, т.е. с учетом моментов не выше второго порядка. С принципиальной точки зрения анализ таких систем почти не отличается от анализа классической системы управления с одной нелинейностью. Единственное отличие рассматриваемого здесь случая от классического – существенное увеличение числа нелинейных уравнений, которые необходимо решать на завершающей стадии исследования [16]. Решение таких систем можно получить только методом последовательных приближений. Решающим обстоятельством, позволяющим надеяться на то, что подобные системы могут быть решены за небольшое число итераций, является предлагаемое разделение неизвестных (моментов) на два класса. К первому из них следует отнести собственно неизвестные переменные (моменты) и моменты параметры, определяющие метрическое пространство, в котором происходит линеаризация. Несмотря на то что они обозначаются одинаково, в итерационной процедуре к выбору следующего приближенного значения моментов из каждой группы можно относиться по-разному. На качественном уровне рассуждений показано и на тестовом примере продемонстрировано, что значительные ошибки в выборе значений моментов параметров метрического пространства не ведут к существенному изменению моментов, получаемых в результате решения системы нелинейных уравнений [17].

При параметрической оптимизации итерационный процесс решения системы уравнений относительно моментов процесса, протекающего в системе, включается в итерационную процедуру поиска наилучших значений параметров. На первый взгляд, это делает решение задачи оптимизации на порядок сложнее, чем решение задачи анализа при заданных значениях параметров. Однако столь существенного увеличения сложности можно избежать, если после завершения циклического процесса анализа при фиксированных значениях параметров придавать им малые приращения и во внутреннем цикле ограничиться одним-двумя циклами, не требуя завершения процесса анализа при новых значениях параметров [18]. Разумеется, подобные рекомендации носят эвристический характер, для повышения доверия к ним необходимо привести как можно больше примеров успешного их применения.

Все сказанное о возможности применения классических методов анализа линейных динамических систем относится только к анализу в рамках корреляционной теории, т.е. при использовании только линеаризационных представлений функций активации и моментов не выше второго порядка. Использование полиномов второго порядка и выше при полиномиальной аппроксимации МНС как отдельного блока в структурном представлении нейродинамической системы влечет за собой необходимость определения моментов выше второго порядка (спектральных плотностей и интегралов от них) определенных координат линеаризованной системы. Конечно, в теории линейных систем нет таких проблем, которые не могли бы быть решены. Однако определение таких характеристик, насколько нам известно, не служило предметом исследования в классической теории управления, воспользоваться известными результатами, таким образом, не представляется возможным.

В так называемой современной теории управления, использующей методы пространства состояний, насколько нам известно, определение моментных характеристик высших порядков также не рассматривалось. Как и в классической теории управления, причиной такого положения вещей является не принципиальная сложность проблемы, а ее невостребованность. Для практического решения поставленной задачи, т.е. для определения моментов высших по-

рядков различных координат динамической системы, представленной уравнениями состояний, можно воспользоваться приемом, описанным для линейных статических преобразований. Сравнительно просто это можно сделать для дискретных моделей в пространстве состояний. Действительно, такая модель в явной форме представляет выражение вектора координат в следующий момент дискретного времени в зависимости от того же вектора координат в текущий момент времени и случайных воздействий. Для теоретических исследований последние можно представить как процессы на входах формирующих фильтров, которые описываются линейными уравнениями в пространстве состояний. Объединение состояний системы и формирующих фильтров позволяет описать нейродинамическую систему в пространстве состояний традиционным образом с единственным отличием от общей модели нелинейной системы, которое заключается в использовании только унарных нелинейных функций в правой части уравнений состояния. Это отличие носит принципиальный характер, поскольку позволяет применить все сказанное об анализе МНС методами линеаризации и полиномиальной аппроксимации к исследованию правой части уравнения состояний нейродинамической системы.

Заключение

Основные задачи теории искусственных нейронных сетей после надлежащей формализации аналогичны задачам технической кибернетики и теории управления. Методы решения этих задач в упомянутых теориях также должны совпадать или развиваться в направлении сближения. Поэтому представляет интерес применение методов классической теории управления для решения задач теории нейронных сетей. Анализ статических МНС должен занимать центральное место в методах исследования нейронных систем, статических и динамических. Основной задачей анализа МНС является определение аппроксимаций условных моментов выходных координат сети от других ее координат. Коэффициенты аппроксимации условных моментов в классе линейных функций называются здесь коэффициентами линеаризации. Коэффициенты линеаризации зависимости выходных координат МНС от ее параметров предназначены для той же цели, что и соответствующие производные в алгоритме обратного распространения при обучении статической сети.

С использованием этих коэффициентов можно осуществить параметрическую оптимизацию нейродинамической системы, включающей в себя МНС в качестве нейроконтроллера. Принципиальным для практического применения моментом в этой процедуре является совмещение двух вложенных циклов: итерационной процедуры решения системы уравнений моментов и поиска наилучших значений параметров.

Вообще, исследование нейронных систем сводится к своего рода композиции методов анализа совокупности унарных нелинейных статических и многомерных линейных статических и динамических преобразований. Это является следствием того, что математические модели нейронных систем представляют собой композицию преобразований такого вида.

INVESTIGATION OF ARTIFICIAL NEURAL NETWORKS BY METHODS OF LINEARIZATION AND POLYNOMIAL APPROXIMATION

V.A. PTITCHKIN

Abstract

Scientific results obtained by the author during last years in the sphere of theory of artificial neural networks, both static and dynamic, are under review. These results are related to the analysis of transformations implemented by multi-layer neural networks, processes of learning of static networks, characteristics of quality of operation of neural dynamic control systems. Unified procedure is applied for the analysis of neural networks and systems; this procedure is based on the idea of linearization

and polynomial approximation of activation functions of formal neurons and neural-network transformations.

Литература

1. *Птичкин В.А.* Анализ нелинейных стохастических систем методами уравнений моментов. Мн.: Наука и техника, 1980.
2. *Птичкин В.А.* Анализ нейронных сетей методом статистической линеаризации // Сб. докл. V Всерос. конф. "Нейрокомпьютеры и их применение". М., 1999. С.408-410.
3. *Птичкин В.А.* Анализ нейросетевых преобразований линеаризационными методами. // Нейрокомпьютеры: разработка, применение. М., 1999. №1. С. 25-37.
4. *Птичкин В.А.* Анализ нейронных сетей с последовательными связями методом статистической линеаризации. // Нейрокомпьютеры: разработка, применение. М., 2001. №6. С. 3-13.
5. *Ptitchkin V.A.* Analysis of Neuronet Transformation by Linearization Techniques // Neurocomputers: Design and Applications. New York, Begell House Inc. Publishers. 2000. Vol. 1. Issue 1. P. 30-44.
6. *Ptitchkin V.A.* The Analysis of Neural Networks with Sequential Connections by the Statistical Linearization Method // Neurocomputers: Design and Applications. New York, Begell House Inc. Publishers 2001. Vol. 2, No 6. P. 2-17.
7. *Птичкин В.А.* Анализ многослойных нейронных сетей методами статистической линеаризации и полиномиальной аппроксимации // Нейрокомпьютеры: разработка, применение. М., 2004. №1.
8. *Ptitchkin V.A.* Comparative Analysis of Main Types of Formal Neuron Activation Functions for Structural Transformations of Multi-layer Neural Networks // Proceedings of Third International Conference on Neural Networks and Artificial Intelligence (ICNNAI'2003). November 12-14. Minsk, Belarus. 2003. P. 166-172.
9. *Птичкин В.А.* Класс полиномов многих переменных // Тез. докл. Всерос. научн. конф. "Проектирование научных и инженерных приложений в среде MATLAB". 28-29 мая 2002 г. М.: ИПУ РАН, 2002. С. 42-43.
10. *Ptitchkin V.* Determination of Features of the Dependence of Neuron Net Output Coordinates on Other Coordinates and Parameters of Neuron Net // Proceedings International Conference on Neural Networks and Artificial Intelligence (ICNNAI'99). 12-15 October 1999. Brest, Belarus. 1999. P. 40-45.
11. *Птичкин В.А.* Использование статистической линеаризации в процедуре обучения нейронных сетей // Сб. докл. VI Всерос. конф. "Нейрокомпьютеры и их применение". НКП-2000. М., 16-18 февраля 2000 г. С. 539-542.
12. *Сигеру Омату, Марзуки Халид, Рубия Юсоф.* Нейроуправление и его приложения / Пер. с англ. Ю.В. Батина; Под ред. А.И. Галушкина и В.А.Птичкина. М.: ИПРЖ Радиотехника, 2000.
13. *Птичкин В.А.* Анализ стационарных процессов в динамических нейронных сетях методом статистической линеаризации // Тр. VII Всерос. конф. "Нейрокомпьютеры и их применение". НКП-2001. М., 14-16 февраля 2001 г. С. 176-179.
14. *Ptitchkin V.A.* Analysis of Stationary Random Processes in Dynamic Neural Networks // Proceedings of Sixth International Conference Pattern Recognition and Information Processing. PRIP'2001. Minsk, Belarus – Szczecin, Poland, 18-20 May 2001. Vol. 1. P. 165–169.
15. *Ptitchkin V.A.* Models of Dynamic Neural Networks and Automatic Control Systems // Proceedings of Second International Conference on Neural Networks and Artificial Intelligence. ICNNAI'2001. Minsk, Belarus, October 2-5 2001. P. 17-23.
16. *Птичкин В.А.* Анализ динамических нейронных сетей как систем автоматического управления // Тр. VIII Всерос. конф. "Нейрокомпьютеры и их применение". НКП-2002. Москва, 21-22 марта 2002 г. С. 350-354.
17. *Ptitchkin V.A.* Analysis of Stationary Dynamic Neural Networks on The Basis of Statistical Linearization Method // Proceedings of 8th International Conference on Neural Information Processing. ICONIP'2001. Shanghai, China, November 14-18, 2001. P. 1544-1549.
18. *Ptitchkin V.A.* Probability-Based Analyses and Parametric Synthesis of Dynamic Neural Systems // Proceedings of 9th International Conference on Neural Information Processing. ICONIP'02. Orchid Country Club, Singapore, 18-22 November 2002. P. 1708-1712.