

М.А. СТОЛЯРЕНКО, канд. техн. наук

ПРОБЛЕМЫ СХОДИМОСТИ ОБУЧЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

У статті розглянуті проблеми навчання штучних нейронних мереж методом зворотного розповсюдження помилки. Проведений аналіз та порівняння різновидів методу та виявлені основні проблеми збіжності. Запропонована методика визначення початкових значень вагових коефіцієнтів, що поліпшує швидкість збігу процедури навчання.

Постановка проблеми. В современной практике искусственных нейронных сетей большую популярность завоевали сети прямого распространения с одним скрытым слоем, обучаемые методом обратного распространения ошибки. Этот метод был предложен в 1986 г. Руммельхартом, Хинтоном и Вильямсом [1], и на сегодняшний день является одним из самых популярных для обучения сетей данного типа.

Рассматриваемый метод имеет два существенных недостатка. Первый заключается в низкой скорости сходимости метода и соответственно, большие вычислительные и временные затраты. Вторым недостатком является возможность возникновения «паралича сети» - невозможности обучения сети при определенных начальных условиях [2].

Анализ последних достижений. Классически метод обратного распространения ошибки реализован с помощью градиентного алгоритма наискорейшего спуска. Данная методика имеет линейную скорость сходимости, а также резкое замедление процесса в окрестности точки оптимального решения, что делает алгоритм малоэффективным. Более эффективными методами настройки весов многослойной сети являются квазиньютоновские алгоритмы, такие как Бройдлени-Флетчера-Гольдфарба-Шено, Девидона-Флетчера-Пауэла и др. При большом числе настраиваемых весов (десятки тысяч и более) хорошо зарекомендовал себя алгоритм сопряженных градиентов, хотя при меньшей размерности он уступает квазиньютоновским методам [3]. Однако нет строгого доказательства сходимости всех этих методов и для увеличения вероятности нахождения глобального экстремума необходимо проводить обучение несколько раз с разными начальными весами нейронов.

Цель статьи. Произвести анализ метода обратного распространения ошибки и предложить методы улучшения сходимости.

Основной материал. На основе анализа метода было установлено, что обучение производится с помощью градиентного алгоритма поиска минимума функционала ошибки нейронной сети. Т.к. в прикладных задачах функционал ошибки является многоэкстремальной функцией, то применяющийся в большинстве случаев случайный выбор начальных значений весовых коэффициентов нейросети приводит к тому, что метод обратного распространения ошибки может остановиться

в локальном минимуме, не обеспечив необходимой точности прогноза. Таким образом, актуальной задачей является начальный выбор весовых коэффициентов нейронной сети.

Как известно, взвешенная сумма классического нейрона представляет собой скалярное произведение вектора входов нейрона и вектора весовых коэффициентов:

$$d = w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n = (\vec{w}, \vec{x}^*), \quad (1)$$

где \vec{w} - вектор весовых коэффициентов \vec{x}^* - вектор входов нейрона, в котором $x_0=1$:

$$\vec{x}^* = (1, x_1, \dots, x_n). \quad (2)$$

Т.к. скалярное произведение принимает максимальное значение при условии коллинеарности векторов, то максимальное значение взвешенной суммы (1) будет достигаться при условии (3):

$$\vec{w} = k \cdot \vec{x}^*, k > 0, \quad (3)$$

а минимальное при:

$$\vec{w} = -k \cdot \vec{x}^*, k > 0. \quad (4)$$

В силу того, что сигмоидальная активационная функция нейрона является строго монотонно возрастающей, то максимальное (минимальное) значение на выходе нейрона возникнет в случаях достижения соответствующего экстремума взвешенной суммы.

В задачах классификации и распознавания, где достаточно широко применяется данный метод, выход сети может иметь несколько фиксированных значений (чаще всего два), соответствующих распознаваемым классам. Как правило, значение выхода нейронной сети, соответствующего одному из них, принимается близким к 1 (максимально возможному значению в случае сигмоидальной функции активации), а второму – к 0 (минимально возможному). В этом случае выходы нейронов обученной нейронной сети принимают свои экстремальные значения при условии подачи на их входы векторов, соответствующих объектам рассматриваемых классов.

В предположении разделимости исследуемых классов максимальные и минимальные значения выходов нейронов должны приниматься в центрах соответствующих областей. В общем случае центры (масс) объектов класса могут находиться за его пределами, поэтому необходимо рассматривать только линейно разделимые области.

Таким образом, целесообразно начальные значения весовых коэффициентов задавать близкими к координатам экстремума выхода нейрона. Наибольшую эффективность показал прием выбора в качестве начальных значений координат центров масс линейно разделимых областей нейронов, соответствующих уровню логической единицы на выходе нейросети:

$$w_i = \frac{1}{m} \sum_{j=1}^m x_{ij}, \quad (5)$$

где x_{ij} – значение i -го входа j -го элемента выборки, m – количество выборок обучающей последовательности.

Проведенное численное исследование подтвердило предлагаемую методику. Была произведена серия обучений нейронной сети с одним скрытым слоем методом обратного распространения ошибки (градиентным методом с постоянным шагом). В первом случае начальные значения весовых коэффициентов были заданы случайным образом. В этом случае среднее количество итераций в серии из 100 обучений составило 11979. Во втором случае при прочих равных условиях начальные значения весовых коэффициентов скрытого слоя были приняты как координаты центров масс линейно разделимых областей нейронов. Количество итераций при этом составило 7223, что на 35% лучше первого случая.

Выводы. Предлагаемая методика выбора начальных значений весовых коэффициентов позволяет улучшить сходимость метода обратного распространения ошибки и уменьшить вероятность возникновения «паралича» сети.

Список литературы: 1. Rumelhart D.E., G.E. Hinton and R.J. Williams. "Learning internal representations by error propagation", in D.E. Rumelhart and J.L. McClelland, eds., vol 1, Chapter 8, Cambridge, MA: MIT Press, 1986. 2. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр.: Пер. с англ. – М: Вильямс, 2006. – 1104 с. 3. Усаков А.А., Кузьмин А.В. Интеллектуальные технологии управления. Искусственные нейронные сети и нечеткая логика. – М.: Горячая Линия – Телеком, 2004. – 143 с.

Поступила в редколлегию 07.04.2008

УДК 621.396.98

В.И. БАРЫШЕВ, А.В. МАЗУРЕНКО, В.В. ПЕЧЕНИН

МЕТОДИКА И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ ПОМЕХОУСТОЙЧИВОСТИ ОДНОКАНАЛЬНОГО ФИЛЬТРА ДОПЛЕРОВСКОЙ ЧАСТОТЫ

В статті запропонована і досліджується методика оцінки завадостійкості одноканального фільтра доплерівської частоти, реалізованого на електромагнітній лінії затримки. Основою методики є використання цифрового еквівалента аналогової структури фільтра, що дозволяє підбирати параметри фільтра за заданою завадостійкістю та величиною розлагодження. Наведено приклад оцінки завадостійкості доплерівського фільтра одержаної моделюванням на ЕОМ.

Введение

Вопросам разработки резонансных фильтров доплеровской частоты (ФДЧ), повышению их помехоустойчивости, методам расчета, способам уменьшения полосы пропускания, улучшению стабильности работы фильтров при расширении полосы обработки сигнала посвящено достаточное количество работ [1-3]. Особое внимание уделяется разработке многоканальных ФДЧ, в которых применяются резонансные индуктивно-емкостные элементы. При этом возникает ряд труднопреодолимых ограничений, а именно, использование неоправданно большого количества каналов, усложнение их практической реализации, ухудшение стабильности положения и формы частотной характеристики отдельного фильтрующего канала и т.д.