МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт кибернетики

Молодёжь и современные информационные технологии

**Сборник трудов
XII Международной научно-практической
конференции студентов, аспирантов
и молодых учёных**

Том I

12–14 ноября 2014 г.

Томск 2014

УДК 378:004 (063)
ББК Ч481.23л0
М754

Молодежь и современные информационные технологии. Сборник трудов XII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии». Томск, 12-14 ноября 2014 г. – Томск: Изд-во ТПУ. – Т. 1 – 431 с.

Сборник содержит доклады, представленные на XII Международной научно-практическую конференцию студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии», прошедшей в Томском политехническом университете на базе института Кибернетики. Материалы сборника отражают доклады студентов, аспирантов и молодых ученых, принятые к обсуждению на секциях: «Микропроцессорные системы, компьютерные сети и телекоммуникации», «Математическое моделирование и компьютерный анализ данных, «Автоматизация и управление в технических системах», «Информационные и программные системы в производстве и управлении», «Компьютерная графика и дизайн», «Информационные технологии в гуманитраных и медицинских исследованиях».

Сборник предназначен для специалистов в области информационных технологий, студентов и аспирантов соответствующих специальностей.

УДК 378:004 (063)
ББК Ч481.23л0
М 754

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ОДНОЙ ОБРАТНОЙ ЗАДАЧИ

Е.В. Семёнов, А.В. Старченко
Национальный исследовательский Томский государственный университет
semenov.evgeny.92@gmail.com

Введение

Искусственный нейрон представляет собой математическую модель биологического нейрона (Рис. 1). Основными составляющими нейрона являются: тело клетки (сома); отростки клетки, по которым поступают сигналы от соседних нейронов (дендриты); хвостик, по которому клетка передаёт команды другим клеткам (аксон); ответвления аксона, цепляющиеся за дендриты близлежащих нейронов (синапсы).

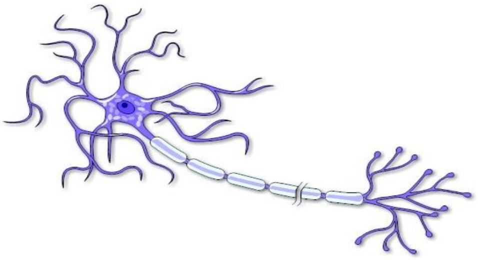

Рис. 1. Биологическая модель нейрона

В математической модели каждому входу искусственного нейрона в соответствие ставится некоторое число w_i, именуемое весом синаптической связи или просто весом, где физический смысл синапса, это его электропроводимость. Массив значений синаптических связей всех нейронов сети есть массив весовых коэффициентов (весов) и, обычно, обозначается буквой W. Выходное состояние каждого нейрона сети определяется функцией от скалярного произведения его весовых коэффициентов на входной вектор обучающей выборки $y = F(W\bar{x}^T)$. Функция на выходе нейрона носит название передаточная функция или функция активации[1].

Одним из типов обучения нейронных сетей является так называемое обучение с учителем, когда данные, предоставляемые нейронной сети, состоят не только из входного вектора $\bar{x}$, но и из соответствующего ему вектора $\bar{d}$, который мы ожидаем получить на выходе сети. Совокупность таких пар носит название обучающей выборки и обозначается, как $\{\bar{x}_1, \bar{d}_1\}, ..., \{\bar{x}_Q, \bar{d}_Q\}$. Ниже будет рассмотрен именно такой подход к обучению искусственной нейронной сети.

В 1960 году Бернард Уидроу и Тед Хофф произвели некоторые модернизации персептрона Френка Розеблатта и ввели в теорию нейронных сетей новое понятие ADALINE (Adaptive Linear Neuron - адаптивный линейный нейрон). Правило для нейронных сетей типа ADALINE (Рис. 2) они назвали “Дельта – правило”, оно основано на методе градиентного спуска, где в качестве направления градиента используется направление градиента ошибки. Алгоритм обучения нейронной сети типа ADALINE носит название LMS (Least Mean Square - наименьшее среднеквадратичное). Кардинальным отличием персептрона от сети ADALINE являлось не только правило обучения сети, сколько вид передаточной функции[2].

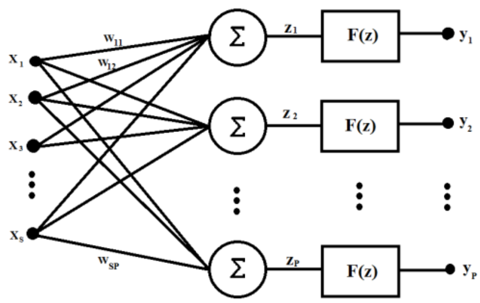

Рис. 2. ADALINE

В персептроне обычно использовали пороговую функцию $F(x) = \begin{cases} 0, & x < b \\ 1, & x \geq b \end{cases}$, где значение b выступало в качестве порогового значения возбуждения нейрона, или позже сигмоидальную функцию $F(x) = \dfrac{1}{1+e^{-\alpha x}}$, где параметр α отвечал за интенсивность изгиба функции, и при увеличении приближал сигмоидальную функцию к пороговой. В сети же типа ADALINE передаточная функция F является линейной[2]: $F(x) = x$.

$$\bar{x}_i \in \mathbb{R}^S; \bar{d}_i, \bar{y}_i \in \mathbb{R}^P; W \in \mathbb{R}^{P\times S}$$

$$\bar{y}_i = F(W\bar{x}_i^T) = W\bar{x}_i^T; i = \overline{1,Q}$$

Искусственные нейронные сети позволяют решать задачи, где обычные алгоритмы и методы применять затруднительно. Формула $\bar{y} = W\bar{x}^T$ описывает систему линейных алгебраических уравнений относительно вектора переменных $\bar{x}$: $A\bar{x} = \bar{d}$. Для случая, когда известны параметры A и $\bar{d}$, существует множество методов, позволяющих отыскать значение вектора $\bar{x}$ (прямая задача). Но когда известны $\bar{x}$ и $\bar{d}$, и требуется найти коэффициенты матрицы A (обратная задача), появляются трудности. В свою очередь, для искусственных нейронных сетей решение обратной задачи не составляет труда. Нейронные сети обучаются путем изменения коэффициентов весов, тем самым любая искусственная нейронная сеть вида $\bar{y} = W\bar{x}^T$ способна решить обратную задачу по нахождению коэффициентов матрицы СЛАУ.

Правило, по которому сеть обучается, носит название наименьшее среднеквадратичное. В свою очередь, функция ошибки представляла среднеквадратичное отклонение $\bar{y}$ от $\bar{d}$.

$$E(\mathrm{W}) = \sum_{i=1}^{Q}\left(\bar{d}_i - \bar{y}_i\right)^2 \xrightarrow[W]{} \min;$$

$$E(W) = \sum_{i=1}^{Q}\left(\bar{d}_i - \bar{y}_i\right)^2 = \sum_{i=1}^{Q}\left(\bar{d}_i - W\bar{x}_i^T\right)^2 =$$

$$= \sum_{i=1}^{Q}\bar{d}^2 - \sum_{i=1}^{Q}2\bar{d}W\bar{x}_i^T + \sum_{i=1}^{Q}W\bar{x}_i^T W\bar{x}_i^T$$

При помощи аппроксимации градиента функции ошибки мы получаем правило, по которому должны изменяться веса.

Рассмотрим градиент целевой функции:

$$\nabla E(W) = \frac{\partial\left(d - W\bar{x}^T\right)^2}{\partial W} =$$

$$= 2\left(d - W\bar{x}^T\right)\frac{\partial\left(d - W\bar{x}^T\right)}{\partial W} = -2\left(d - W\bar{x}^T\right)\bar{x}$$

Тогда в соответствии с методом наискорейшего спуска построим итерационный алгоритм:

$$W_{k+1} = W_k - \alpha\nabla E(W_k);$$

$$W_{k+1} = W_k + 2\alpha\left(d - W_k\bar{x}^T\right)\bar{x}, \quad k \ge 0$$

Стоит отметить, что нахождение параметра a не нуждается в одномерной минимизации, этот параметр (параметр скорости обучения) находится из условия сходимости правила обучения сети [2].

В отличие от условия дельта-правила, в ходе практического исследования нейронной сети было установлено новое правило для выбора параметра скорости обучения. Так называемая нормировка входных данных:

$$\alpha = \frac{1}{\|\bar{x}\|},\ где\ \|\bar{x}\| = \frac{1}{Q}\sum_{i=1}^{Q}\sum_{j=1}^{S}x_{i,j}$$

Постановка задачи: Для системы $A\bar{x} = \bar{d},\ A \in \mathbb{R}^{S\times S}$ дан набор пар $\{\bar{x}_1, \bar{d}_1\}, ..., \{\bar{x}_Q, \bar{d}_Q\}$, сама матрица неизвестна. Необходимо решить обратную задачу по нахождению коэффициентов исходной матрицы по известному решению и правой части.

Задача решалась для размера матрицы $S = 240$. В работе были использованы два размера обучающей выборки: $Q = 24$ и $Q = 120$.

При подготовке обучающей выборки (прямая задача) использовались два подхода:

1. Задавался вектор $\bar{x}$, после чего рассчитывался вектор $\bar{d} = A\bar{x}$.
2. Задавалась правая часть системы $\bar{d}$, после чего с помощью LU-факторизации находился вектор $\bar{x}$.

При использовании первого подхода нейронная сеть находила коэффициенты исходной матрицы с точностью 10^{-12} за $0,3-0,4$ секунды. Причём метод работал для обоих размеров обучающей выборки. В свою очередь, при использовании второго подхода, при размере выборки равном 24 нейронная сеть исчерпала отведённое ей количество итераций, не предоставив искомого результата. При использовании второй выборки сеть решала задачу с той же точностью, но решение оказывалось неверным. Происходило это потому, что сеть достигала локального минимума, не доходя до глобального.

Также, чем ближе сеть приближалась к глобальному минимуму, тем больше становилась осцилляция параметра ошибки. Решить эту проблему удалось изменением параметра скорости обучения в зависимости от номера итерации и используемых компонент входного вектора $\bar{x}$. Такой же подход может быть применён для выхода из локального минимума целевой функции.

Если входные данные содержали ошибку, то она суммировалась на выходе сети:

$$\tilde{y}_i = \sum_{j=1}^{S}w_{i,j}(\tilde{x}_l)_j = \sum_{j=1}^{S}w_{i,j}\left(x_{l,j} + \delta_{l,j}\right) =$$

$$= \sum_{j=1}^{S}w_{i,j}x_{l,j} + \sum_{j=1}^{S}w_{i,j}\delta_{l,j} = y_i + \Delta_l;\ i = \overline{1,P};\ l = \overline{1,Q}$$

Также искусственные нейронные сети поддаются распараллеливанию. Это связано с тем, что каждый нейрон каждого слоя рассчитывается отдельно от других, у него есть свой набор весов, характеризующих соединение с нейронами предыдущего слоя. Для подсчёта значения нейрона на любом слое, необходимо лишь предоставить вектор выходов нейронов предыдущего слоя каждому процессору. В таком случае мы получим ускорение работы программы на любом слое примерно во столько раз, сколько нейронов на этом слое расположено.

Заключение

В результате проведённой работы были изучены искусственные нейронные сети, принцип их работы и применение. Для решения поставленной задачи была разработана параллельная программа на языке Fortran с использованием технологий MPI и OpenMP. Тестирование сети проводилось на суперкомпьютере ТГУ СКИФ Cyberia.

Работа выполнена по Государственному Заданию Министерства образования и науки РФ, №5.628.2014/К.

Литература

1. Тархов Д. А., Нейронные сети, модели и алгоритмы. Кн. 18. – М.: Радиотехника, 2005.
2. Widrow B. and Stearns S.D., Adaptive Signal Processing, Prentice Hall, Englewood Cliffs, NJ, 1985.