УДК 004.032.22, 519.24, 330.1, 316.4

Соловьёв А.С.
пенсионер
Россия, г. Ростов-на-Дону

К СЕТЕВОМУ ПРЕДСТАВЛЕНИЮ. ОБУЧЕНИЕ.

Аннотация: В работе предлагается метод обучения нейронной сети с учётом её количественных и качественных особенностей.

Ключевые слова: нейронная сеть, качество, количество, мера, сравнение, обучение, цель, кватернионы.

Soloviev A.S.
retiree
Russia, Rostov-on-don

TO THE NETWORK VIEW. TRAINING.

Annotation: The paper proposes a method of training a neural network with regard to its quantitative and qualitative features.

Key words: neural network, quality, quantity, measure, comparison, training, goal, quaternions.

Пусть в некоторой нормальной системе выделен тернарный узел

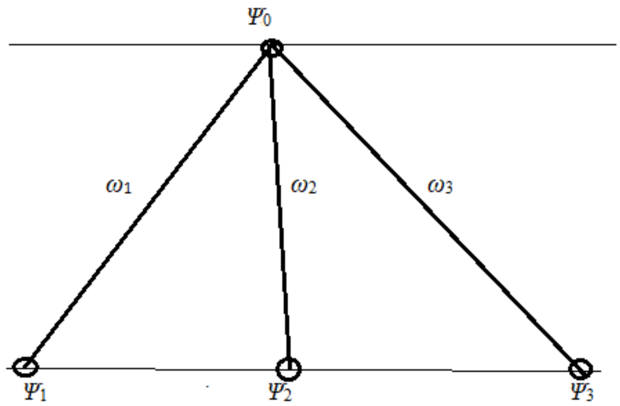

Рис. 1. Узел нормальной сети прямого распространения со смежными узлами нижестоящего среза.

(1) Для нормальной системы имеем соотношение

(2) $$\Psi_0 = \omega_1\Psi_1 + \omega_2\Psi_2 + \omega_3\Psi_3 = \omega_i\Psi_i, i \in N = \{1, 2,3\},$$

при этом для коэффициентов выполняется условие

(3) $\omega_1 + \omega_2 + \omega_3 = \omega^i{}_i = 1,$

а узлы системы определяются собственными функциями Ψ_i с собственными значениями равными единице

(4) $\mathbf{x}_i = x^i\Psi_i, \quad x^i = 1.$

В общем случае для стоящего в иерархии узла на более высокой ступени в сети прямого распространения имеем зависимость от смежных узлов нижележащего среза

(5) $\boldsymbol{x}_0 = \omega_i x^i \Psi_i = \omega_1 x^1 \Psi_1 + \omega_2 x^2 \Psi_2 + \omega_3 x^3 \Psi_3, \quad i \in N.$

Соотношение, аналогичное (1), справедливое для зависимости собственных функций, будет справедливо и для собственных значений

(6) $x^0 = \omega_i x^i, \quad i \in N.$

Элемент сети будем рассматривать как вектор

(7) $\boldsymbol{x}_i = x^i \Psi_i$

и знаком (*) обозначать переход к сопряжённому вектору, опуская вниз у сопряжённого вектора индекс. Тогда, если ввести векторы

(8) $\boldsymbol{x} = (x^i | \, i \in N), \quad \boldsymbol{\omega} = (\omega^i | \, i \in N),$

равенство (4) можно записать в виде внутреннего произведения векторов

(9) $\boldsymbol{x}_0 = \boldsymbol{\omega}^* \boldsymbol{x}.$

Рассмотрим две различные сети с одинаковой архитектурой. Выделим в них узел с собственной функцией Ψ_0 и рассмотрим его действие со смежными узлами. Поскольку архитектура сетей одинакова, то если узел одной сети имеет представление, изображённое на рис. 1, то аналогичное представление имеет и выделенный фрагмент второй сети, но описание их действия будет различными уравнениями – уравнением (8) и равенством

(10) $\boldsymbol{y}_0 = \boldsymbol{\varphi}^* \boldsymbol{y}.$

С учётом нормальности систем будем иметь

(11) $\boldsymbol{y} = \boldsymbol{x}.$

Таким образом, данные фрагменты будут отличаться только пропускной способностью связей. Будем полагать, что (8) и (9) состояния узла Ψ_0 из пространства $\boldsymbol{X}$ возможны состояний системы. Более того, считать состояние (9) целевым (плановым), а состояние (8) – фактическим. И полагать, что требуется последовательно сближать фактическое состояние с эталоном, естественно, с учётом ограничений на показатели функционирования.

Сравним данные фрагменты системы опираясь на основное метрическое тождество и используя свойства скобки процедуры Кэли-Диксона в алгебре Клиффорда в пространстве $\boldsymbol{X} \otimes \boldsymbol{X}$ в качестве элементов тензорного произведения $\boldsymbol{c} = \boldsymbol{x}_0 \boldsymbol{y}_0$. Для этого, на их бинарном соответствии $\boldsymbol{z} = (\boldsymbol{x}_0, \boldsymbol{y}_0)$, введём функции для внутреннего $a = \boldsymbol{x}_0^* \boldsymbol{y}_0$ и внешнего $\boldsymbol{b} = \boldsymbol{x}_0 \wedge \boldsymbol{y}_0 = \mathrm{i}\boldsymbol{n}b$ произведений, где $\boldsymbol{n}$ единичный вектор в представлении $\boldsymbol{x}_0 \times \boldsymbol{y}_0 = b\boldsymbol{n}$ и гомоморфизм в пространстве определим функционалом $D = c^2 = D(\boldsymbol{z})$. Получаем, что на данном бинарном соответствии справедливо тождество Пифагора

(12) $a^2 + b^2 = c^2,$

а для его тензорной кватерниональной формы имеет место волновое представление

(13) $\boldsymbol{c} = c\Psi, \quad \Psi = exp(\mathrm{i}\boldsymbol{n}\theta),$

Будем полагать, что фактическое состояние системы есть образ (модель) оригинала, т.е. его расчётного состояния – эталон, в практической деятельности естественно последовательное стремление фактического состояния к эталону, и считать, что, не нарушая общности, ограничения (накладываемые на производственные факторы) для подобного совершенствования отсутствуют. Как скульптор из глыбы мрамора высекает Аполлона, так и мы будем последовательно совершенствовать образ, последовательно отсекая мешающие обстоятельства в нашем совершенствовании. Для производственной системы это модернизация производства, стимулирование и обучение персонала и т. п.

Как правило, если рассматривать производственные сети, то это будут многоуровневые иерархические структуры и алгоритм их совершенствования, с учётом сети прямого распространения, можно строить поэлементно, либо послойно, например, снизу вверх, либо сверху вниз. Из статистического сходства соответствующих слоёв вытекает основа алгоритма – последовательное улучшение корреляционного сходства.

Коэффициент корреляции определяется формулой

(14) $r = a/c$.

Применяя неравенства Коши-Буняковского

(15) $a <= c$,

в котором равенство достигается при полном соответствии сравниваемых слоёв, и полагая $r>0$,находим

(16) $a = rc <= c$.

При монотонном росте коэффициента корреляции образ будет последовательно стремиться к эталону. Предположим, что при темпах прироста $T(x)$ некоторого фактора сети темпы прироста коэффициента корреляции равны $T(r)$. При условии, что эластичность (чувствительность) коэффициента корреляции относительно данного изменения равна s получаем

(17) $r(t + 1) = (1 + sT(x))r(t)$.

Заключаем, что корреляция будет возрастать, если показатели s и $T(x)$ будут одного знака.

Пусть $\boldsymbol{x}$ и $\boldsymbol{y}$ произвольные фрагменты исходных систем, находящихся в структурном бинарном соответствии, а x и y – некоторые соответствующие их узлы. Для эластичности коэффициента корреляции данных фрагментов по показателю x получаем выражение

(18) $s= xy/a(\boldsymbol{x},\boldsymbol{y}) – x^2/\boldsymbol{c}^2(\boldsymbol{x},\boldsymbol{y})$,

т.е. сумма $S = \sum s$ эластичностей коэффициентов корреляции (17) по всем узлам данных фрагментов равна нулю. Следовательно, среди показателей эластичности узлов есть как положительные, так и отрицательные. Сумму положительных слагаемых обозначим S^+, а сумму абсолютных величин отрицательных слагаемых обозначим S^-. Приходим к выражению

(18) $S = S^+ - S^- = 0$.

Представление (18) такого расслоения показывает, что качественное сближение модели с эталоном можно проводить двумя методами улучшая качество модели за счёт совершенствования узлов:

1) путём последовательного роста корреляции узлов с положительными эластичностями действия, и

2) качественного сближения путём улучшения корреляции за счёт совершенствования элементов с отрицательным показателем эластичности.

Волновое представление (12) сходства фрагментов сети свидетельствует, что качественное сближение модели с эталоном можно проводить:

a) обращая внимание только на показатели с положительными коэффициентами эластичности, т.е. случай 1);

b) обращая внимание только на элементы сети с отрицательными коэффициентами эластичности, случай 2);

c) применяя в сближении одновременно как случай 1), так и случай 2).

Это свидетельствует, что "скульптор" из одной и той же " мраморной глыбы" может получить множество количественно различных, но качественно подобных моделей, т.е. в представлении (12) у таких моделей могут быть близкими по значению аргументы θ, но сильно варьироваться количественные показатели – амплитуды σ, т.е. $\boldsymbol{x} = k\boldsymbol{y}$, $\sigma(\boldsymbol{x}) = k\sigma(\boldsymbol{y})$.

На примере ненормированного фрагмента сети рис.1 определим состояния $\boldsymbol{x} = (1; 2; 3)$ и $\boldsymbol{y} = (1; 1; 2)$. При одинаковом качестве состояний их компоненты пропорциональны. Предъявляя $\boldsymbol{y}$ как эталон, состояние $\boldsymbol{x}$ можно привести к качественно подобному $\boldsymbol{y}$ состоянию двумя методами: 1) уменьшением компонент x^2 и x^3, и 2)увеличением в два раза компоненты x^1 и увеличением в полтора раза компоненты x^3. Если в первом случае получаем $\boldsymbol{x} = \boldsymbol{y}$, то в случае 2) имеем $\boldsymbol{x} = 2\boldsymbol{y}$. Это объясняют и значения коэффициентов эластичности: $s_1 = 5/126$, $s_2 = - 8/126$, $s_3 = 3/126$.

Отметим, что волновое представление (12) даёт возможность наблюдать процесс обучение в цветовой гамме.