УДК 681.3

С. В. Зимина
Нижегородский государственный университет им. Н. И. Лобачевского

Влияние флуктуаций весового вектора на характеристики нейронной сети с настройкой по критерию наименьшего среднего квадрата ошибки

Представлены результаты статистического анализа искусственных нейронных сетей (ИНС), настраивающихся по LMS-алгоритму с квадратичным ограничением, при учете флуктуаций весового вектора. Получены выражения для корреляционной функции выходного сигнала искусственных нейронов различных слоев сети. Показано, что флуктуации весового вектора приводят к искажениям статистических характеристик ИНС, причем искажения, вносимые флуктуациями, увеличиваются по мере увеличения номера слоя сети.

Искусственная нейронная сеть, флуктуации весовых коэффициентов, LMS-алгоритм с квадратичным ограничением

Искусственные нейронные сети позволяют решать широкий спектр задач, связанных с обработкой сигналов: выделение полезного сигнала на фоне помех, распознавание изображений, анализ неформализованных математически данных, в связи с чем постоянно привлекают к себе внимание исследователей [1], [2]. Скорость обработки данных и точность настройки ИНС во многом определяют флуктуации настраиваемых весовых коэффициентов передачи искусственных нейронов сети. В [3]–[5] проведен анализ статистических характеристик ИНС с различными алгоритмами настройки с учетом флуктуаций весового вектора и показано, что флуктуации приводят к искажениям выходного сигнала ИНС и к появлению в статистических характеристиках ИНС дополнительных членов, обусловленных флуктуациями.

Представляет интерес анализ характеристик ИНС, настраивающейся по LMS-алгоритму[1] с квадратичным ограничением на коэффициент усиления полезного сигнала [6], с учетом флуктуаций весовых коэффициентов. Указанный алгоритм является современной модификацией LMS-алгоритма, предложенного Уидроу [7]. Применение последнего алгоритма встретило трудности, выражающиеся в неконтролируемом подавлении полезного сигнала в случае, когда помеха коррелирована с сигналом. Модифицированный LMS-алгоритм этого недостатка не имеет. В настоящей статье приведены результаты статистического анализа корреляционных характеристик ИНС, настраивающейся по алгоритму [6], полученные при учете флуктуаций настраиваемых весовых коэффициентов.

Постановка задачи. Классические нейросетевые алгоритмы большей частью предназначены для решения других типов задач, например для задачи распознавания образов или задач предсказания, и в них чаще всего используется обучение с учителем, а также конкретный вид ИНС [1].

В задаче выделения полезного сигнала на фоне помех необходим алгоритм обучения ИНС без учителя, поэтому для решения задачи был выбран многослоевой персептрон.

Эти условия существенно ограничивают выбор алгоритма настройки сети, что и побудило обобщить LMS-алгоритм с квадратичным ограничением для работы в ИНС. Кроме того, использование одного и того же алгоритма для настройки адаптивных антенных решеток и ИНС позволяет сравнивать значения искажений, вносимых флуктуациями весового вектора в статистические характеристики обоих типов адаптивных систем.

Рассмотрим работу ИНС, имеющей M слоев и N нейронов в каждом слое, и решающей задачу выделения полезного сигнала на фоне помех. Такие задачи встречаются в радиолокации и в сотовой связи. Например, адаптивная система может обеспечить качественный прием сигнала в телефонии с гашением имеющихся помех и выделением голоса или в телевидении, обеспечивая в этом случае устойчивый прием телевизионного сигнала.

Настройка вектора весовых коэффициентов $\mathbf{W}_{pi}(k+1)$ i-го искусственного нейрона p-го слоя ИНС, настраивающейся по LMS-алгоритму с квад-

[1] LMS-алгоритм – алгоритм настройки по методу наименьшего квадрата (МНК – LMS) ошибки.

ратичным ограничением на усиление полезного сигнала, в момент времени $(k+1)$ описывается уравнением

$$\mathbf{W}_{pi}(k+1)=\mathbf{W}_{pi}(k)-$$
$$-\mu\left\{\mathbf{Z}_{p-1}(k)F\left[\mathbf{Z}_{p-1}^{\text{H}}(k)\mathbf{W}_{pi}(k)\right]+\right.$$
$$\left.+\frac{\left|F\left[\mathbf{Z}_{p-1}^{\text{H}}(k)\mathbf{W}_{pi}(k)\right]\right|^{2}}{\left|F\left[\mathbf{S}^{\text{H}}\mathbf{W}_{pi}(k)\right]\right|^{2}}\mathbf{S}F\left[\mathbf{S}^{\text{H}}\mathbf{W}_{pi}(k)\right]\right\}, \quad (1)$$

где μ – коэффициент адаптации LMS-алгоритма с квадратичным ограничением; $\mathbf{Z}_{p-1}(k)$ – вектор выходного сигнала слоя $(p-1)$ ИНС в момент времени k; $F[\cdot]$ – нелинейная активационная функция рассматриваемого искусственного нейрона; $\mathbf{S}=[S_1, S_2, \ldots, S_N]^{\text{T}}$ – вектор полезного сигнала; $^{\text{"H"}}$, $^{\text{"T"}}$ – символы операций эрмитова сопряжения и транспонирования соответственно.

Выражение (1) – обобщение LMS-алгоритма с квадратичным ограничением для работы ИНС. Исходно данный алгоритм предназначался для настройки адаптивных антенных решеток в задачах выделения полезного сигнала на фоне помех [6]. В рассматриваемом примере ИНС выполняет аналогичную задачу.

Выходной векторный сигнал $(p-1)$-го слоя ИНС $\mathbf{Z}_{p-1}(k)$ является одновременно входным сигналом i-го нейрона p-го слоя. Выходной сигнал этого нейрона $Z_{pi}(k)$ может быть записан в виде

$$Z_{pi}(k)=F\left[\sum_{q=1}^{N}Z_{(p-1)q}(k)W_{piq}(k)\right]=$$
$$=F\left[y_{pi}(k)\right]=\sum_{l=1}^{N_1}a_l y_{pi}^{l}(k)=\mathbf{A}^{\text{T}}\mathbf{Y}_{pi}(k),$$

где N – количество нейронов в $(p-1)$-м слое; $y_{pi}(k)$ – выходной сигнал линейной части i-го нейрона p-го слоя;

$$\mathbf{A}=\left[a_1, a_2, \ldots, a_l, \ldots, a_{N_1}\right]^{\text{T}}$$

– вектор коэффициентов разложения функции F в ряд Вольтерра [8] размера N_1;

$$\mathbf{Y}_{pi}(k)=\left[y_{pi}(k), y_{pi}^{2}(k), \ldots, y_{pi}^{N_1}(k)\right]^{\text{T}}$$

– вектор степеней выходного сигнала линейной части i-го искусственного нейрона, принадлежащего слою p.

Рассмотрим ИНС, в которой все нейроны являются узкополосными. Тогда корреляционная матрица входных сигналов каждого из таких искусственных нейронов в p-м слое будет иметь вид

$$R_p(k,\ k+n)\equiv\left\langle\mathbf{Z}_p^{*}(k)\mathbf{Z}_p^{\text{T}}(k+n)\right\rangle=\mathcal{R}_p r^{|n|},$$

где n – временно́й сдвиг между сигналами; $\langle\cdot\rangle$ – символ усреднения; $\mathcal{R}_p$ – пространственная часть корреляционной матрицы входных сигналов; r – коэффициент корреляции между отсчетами входных сигналов.

Корреляционная матрица R_p имеет разный вид в зависимости от номера слоя ИНС. В рассмотренном в настоящей статье частном случае одинаковой помеховой обстановки, подаваемой на все нейроны входного слоя, для всех скрытых слоев ИНС корреляционная матрица будет иметь в качестве всех своих элементов единицы. Во входном слое эта матрица будет равна корреляционной матрице входных сигналов ИНС.

Коэффициент корреляции r между отсчетами принимаемого ИНС полезного сигнала определяет его вид. Если $r=0$, то ИНС принимает "белый" шум. Если же $r\approx1$, ИНС принимает близкий к детерминированному синусоидальный сигнал. Все промежуточные значения коэффициента корреляции обеспечивают плавный переход между этими двумя типами сигналов.

Значения коэффициента корреляции влияют на особенности выделения адаптивной системой полезного сигнала на фоне помех. В ранее опубликованных работах показано, что возможны два эффекта искажений полезного сигнала – эффект рассогласования [7], [9], [10] и эффект "перекомпенсации" [9], [11], [12]. Первый эффект заключается в том, что адаптивная система не способна в силу флуктуаций весовых коэффициентов полностью погасить помеховый сигнал. Это приводит к появлению остаточной мощности помехи на выходе системы. Второй эффект состоит в том, что адаптивная система из-за флуктуаций весовых коэффициентов, статистически связанных с вектором входных сигналов, подавляя помеху, также подавляет и полезный сигнал. В результате мощность сигнала на выходе адаптивной системы становится меньше мощности, найденной при постоянном стационарном весовом векторе.

Эффект рассогласования наблюдается при сравнительно небольших значениях коэффициента корреляции. Наоборот, эффект "перекомпенсации" характерен для его высоких значений. Так, в случае адаптивной антенной решетки, настраивающейся по дискретному градиентному алгоритму с многократными линейными ограничениями, переход от одного эффекта к другому происходит при $r \approx 0.577$.

Алгоритм настройки (1) описывает скорее каскадное соединение искусственных нейронов, чем единую ИНС, поскольку весовые коэффициенты каждого отдельного нейрона зависят только от собственных значений в предыдущий момент времени и не связаны непосредственными математическими отношениями с весовыми коэффициентами других искусственных нейронов. Тем не менее, косвенное влияние весовых коэффициентов различных искусственных нейронов друг на друга существует и осуществляется через выходные сигналы элементов сети, что также следует из формулы (1).

Определим статистические характеристики описанной ИНС с учетом флуктуаций настраиваемых весовых коэффициентов. С этой целью на первом этапе анализа найдем статистические характеристики отдельного нейрона с учетом флуктуаций весового вектора, а затем обобщим полученные результаты на ИНС.

Методами теории возмущений в первом, так называемом борновском, приближении первоначально найдены статистические характеристики искусственного нейрона, настраивающегося по LMS-алгоритму с квадратичным ограничением на коэффициент усиления полезного сигнала с учетом флуктуаций весового вектора. Для упрощения обобщения полученных результатов на ИНС предположим, что на входные элементы каждого из нейронов входного слоя ИНС поступают одни и те же входные сигналы, т. е. существует одна и та же помеховая обстановка. В этом случае при старте с одних и тех же начальных условий выходные сигналы искусственных нейронов в каждом слое ИНС (и, следовательно, входные сигналы в каждом последующем слое сети) будут одинаковыми, что существенно упрощает вид корреляционной матрицы входных сигналов скрытых слоев сети.

Корреляционные характеристики искусственной нейронной сети. Расчет корреляционной функции широко используется в системах связи. В частности, на ее основе можно определить, имеется ли в принимаемом сигнале полезная составляющая.

Корреляционные характеристики выходного сигнала ИНС позволяют описать полезный сигнал, который приходит на вход ИНС в смеси с помехами. Форма корреляционной функции, быстрота ее спадания, время корреляции позволяют судить о том, какой полезный сигнал принят – "белый" шум, близкая к детерминированной синусоида или какой-то промежуточный вариант. Как указано ранее, тип полезного сигнала влияет также на эффекты (рассогласования или "перекомпенсации"), возникающие в адаптивной системе из-за флуктуаций весового вектора при выделении этого сигнала системой из смеси с помех.

Анализ влияния флуктуаций на статистические характеристики ИНС выполнен при условии, что векторы входных сигналов и весовых коэффициентов каждого искусственного нейрона ИНС связаны негауссовской статистической зависимостью. Это означает, что смешанная кумулянтная функция третьего порядка для этих величин не равна нулю: $\langle \tilde{\mathbf{\Phi}}_{XX} \tilde{\mathbf{W}} \rangle \neq 0$, где $\tilde{\mathbf{\Phi}}_{XX}$ – флуктуационная часть стохастической матрицы входных сигналов $M_{XX}(k,\, k+n) = \mathbf{X}^*(k)\mathbf{X}^{\mathrm{T}}(k+n)$; $\tilde{\mathbf{W}}$ – флуктуационная часть вектора весовых коэффициентов. Это наиболее общий случай зависимости между указанными величинами, который позволяет учесть все эффекты, вносимые флуктуациями в работу адаптивной системы (эффекты как рассогласования, так и перекомпенсации). Предположение же о гауссовской статистической зависимости между вектором входных сигналов и весовым вектором $\left(\langle \tilde{\mathbf{\Phi}}_{XX} \tilde{\mathbf{W}} \rangle = 0\right)$ не позволяет теоретически предсказать эффект перекомпенсации, который наряду с эффектом рассогласования наблюдается в экспериментах.

Для узкополосного искусственного нейрона p-го слоя, настраивающегося по LMS-алгоритму с квадратичным ограничением, в случае негауссовской статистической зависимости между векторами входных сигналов и весовым вектором корреляционная функция выходного сигнала с учетом флуктуаций весовых коэффициентов может быть представлена в виде суммы:

$$K_{Z_p}(k,\, k+n) = A_1 + B_1 + C_1 + D_1. \qquad (2)$$

Слагаемые этой функции определяются следующими выражениями:

$$A_1 = a_1^2 \mathbf{W}_{\text{ст}}^{\mathrm{H}} R_{XX} \mathbf{W}_{\text{ст}} r^{|n|}, \qquad (3)$$

где R_{XX} – корреляционная матрица входного сигнала; $\mathbf{W}_{\text{ст}}$ – постоянный стационарный весовой вектор произвольного искусственного нейрона первого слоя ИНС;

$$B_1 = \mu^2 a_1^4 \frac{M_{\text{уср}}^{-1}}{1-r^2} r^{|n|} \times$$
$$\times \left[\mathrm{Sp}(R_{XX}) \mathbf{W}_{\text{ст}}^{\text{н}} R_{SS}^{\text{н}} R_{XX} \mathbf{W}_{\text{ст}} + \right.$$
$$\left. + \mathrm{Sp}(R_{XX}) \mathbf{W}_{\text{ст}}^{\text{н}} R_{XX} R_{SS} \mathbf{W}_{\text{ст}} \right], \quad (4)$$

где $M_{\text{уср}}$ – постоянный коэффициент, возникающий при усреднении в LMS-алгоритме с квадратичным ограничением; $\mathrm{Sp}(\cdot)$ – след матрицы; R_{SS} – корреляционная матрица полезного сигнала;

$$C_1 = \mu^2 a_1^4 r \frac{1+r-r^n}{(1-r)^2} r^{|n|} \mathbf{W}_{\text{ст}}^{\text{н}} R_{XX}^3 \mathbf{W}_{\text{ст}} +$$
$$+ \mu^2 a_1^4 \frac{r}{(1-r)^2 (1+r)} r^{|n|} \mathrm{Sp}(R_{XX}) \mathbf{W}_{\text{ст}}^{\text{н}} R_{XX}^2 \mathbf{W}_{\text{ст}} +$$
$$+ \mu^2 a_1^4 r^{|n|} \left[\frac{2}{\left(1-r^2\right)^2} + \frac{1-r^{2n+2}}{1-r^2} \right] \times$$
$$\times \mathrm{Sp}\left(R_{XX}^2\right) \mathbf{W}_{\text{ст}}^{\text{н}} R_{XX} \mathbf{W}_{\text{ст}}; \quad (5)$$

$$D_1 = \mu^2 a_1^4 \frac{r^{|n|+3}}{1-r^2} \left(\frac{1-r^n}{1-r^2} + \frac{1}{1-r} \right) \times$$
$$\times \mathrm{Sp}\left(R_{XX}^2\right) \mathbf{W}_{\text{ст}}^{\text{н}} R_{XX}^2 \mathbf{W}_{\text{ст}}. \quad (6)$$

Из (3)–(6) видно, что корреляционная функция выходного сигнала произвольного искусственного нейрона первого слоя ИНС зависит только от сдвига во времени n между рассматриваемыми сигналами и не зависит от абсолютного времени k. Кроме того, учтем, что все нейроны входного слоя сети будут иметь на выходе один и тот же сигнал (и, следовательно, одну и ту же корреляционную функцию). Тогда можно упростить обозначение корреляционной функции в (2), заменив $K_{Z_p}(k, k+n)$ на $K_1(n)$.

Найдем корреляционную матрицу входных сигналов второго слоя сети с учетом флуктуаций весовых коэффициентов. Согласно определению, корреляционная функция сигнала $Z(n)$ имеет вид

$$K(n) = \left\langle Z^{\text{н}}(k) Z(k+n) \right\rangle.$$

Тогда из описанных свойств внутренних сигналов ИНС следует, что любые взаимные корреляционные функции между выходными сигналами любых нейронов i и j первого слоя будут равны между собой:

$$\left\langle Z_{i1}^{\text{н}}(k) Z_{j1}(k+n) \right\rangle = K_1(n), \quad i, j = \overline{1, N_1}. \quad (7)$$

Учитывая (7), запишем выражение для корреляционной матрицы выходных сигналов первого слоя ИНС (и, соответственно, входных сигналов второго слоя ИНС):

$$\mathcal{K}_1(n) =$$
$$= \begin{bmatrix} \left\langle Z_{11}^{\text{н}}(k) Z_{11}(k+n) \right\rangle & \cdots & \left\langle Z_{11}^{\text{н}}(k) Z_{N_1 1}(k+n) \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle Z_{N_1 1}^{\text{н}}(k) Z_{11}(k+n) \right\rangle & \cdots & \left\langle Z_{N_1 1}^{\text{н}}(k) Z_{N_1 1}(k+n) \right\rangle \end{bmatrix} =$$
$$= K_1(n) E, \quad (8)$$

где $E = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$ – матрица, все элементы которой являются единицами.

Запишем далее корреляционную функцию произвольного искусственного нейрона i, принадлежащего второму слою нейронной сети. Для этого учтем, что требуемое выражение получается заменой корреляционной матрицы входных сигналов $R_{XX}(k, k+n)$ на корреляционную матрицу $\mathcal{K}_1(n)$ (8), описывающую корреляционную связь входных сигналов данного слоя. Тогда на основании (2) искомая корреляционная функция может быть представлена полиномом по степеням корреляционной матрицы $\mathcal{K}_1(n)$, а с учетом (8) – по степеням корреляционной функции $K_1(n)$. Проведя описанные преобразования, получим:

$$K_{Z2}(k, k+n) = K_2(n) =$$
$$= A_2 K_1(n) + B_2 K_1^2(n) + C_2 K_1^3(n) + D_2 K_1^4(n), \quad (9)$$

где

$$A_2 = a_1^2 \mathbf{W}_{\text{ст}2i}^{\text{н}} E \mathbf{W}_{\text{ст}2i} r^{|n|}; \quad (10)$$

$$B_2 = \mu^2 a_1^4 \frac{M_{\text{уср}}^{-1}}{1-r^2} r^{|n|} \times$$
$$\times \left(\mathrm{Sp}(E) \mathbf{W}_{\text{ст}2i}^{\text{н}} R_{SS}^{\text{н}} E \mathbf{W}_{\text{ст}2i} + \right.$$
$$\left. + \mathrm{Sp}(E) \mathbf{W}_{\text{ст}2i}^{\text{н}} E R_{SS} \mathbf{W}_{\text{ст}2i} \right); \quad (11)$$

$$C_2 = \mu^2 a_1^4 r \frac{1+r-r^n}{(1-r)^2} r^{|n|} \mathbf{W}_{\text{ст}2i}^{\text{н}} E^3 \mathbf{W}_{\text{ст}2i} +$$
$$+ \mu^2 a_1^4 \frac{r}{(1-r)^2 (1+r)} r^{|n|} \mathrm{Sp}(E) \mathbf{W}_{\text{ст}2i}^{\text{н}} E^2 \mathbf{W}_{\text{ст}2i} +$$
$$+ \mu^2 a_1^4 r^{|n|} \left[\frac{2}{\left(1-r^2\right)^2} + \frac{1-r^{2n+2}}{1-r^2} \right] \times$$
$$\times \mathrm{Sp}\left(E^2\right) \mathbf{W}_{\text{ст}2i}^{\text{н}} E \mathbf{W}_{\text{ст}2i}; \quad (12)$$

$$D_2 = \mu^2 a_1^4 \frac{r^{|n|+3}}{1-r^2}\left(\frac{1-r^n}{1-r^2}+\frac{1}{1-r}\right)\times$$
$$\times \mathrm{Sp}\left(E^2\right)\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}2i}E^2\mathbf{W}_{\mathrm{ст}2i}. \qquad (13)$$

По аналогии запишем выражение для корреляционной функции искусственного нейрона, принадлежащего третьему слою искусственной нейронной сети:

$$K_{Z3}(k,\ k+n) = K_3(n) =$$
$$= A_3K_2(n) + B_3K_2^2(n) + C_3K_2^3(n) + D_3K_2^4(n).$$

Переобозначив $A_{31} = A_3$; $A_{32} = B_3$; $A_{33} = C_3$; $A_{34} = D_3$ и учтя, что $K_2(n)$ описывается выражением (9), получим:

$$K_3(n) = \sum_{i=1}^{4} A_{3i}K_2^i(n) = \sum_{i=1}^{4} A_{3i}\left(\sum_{j=1}^{4} A_{2j}K_1^j(n)\right)^i.$$

Обобщив это выражение по принципу математической индукции, запишем итеративную формулу для корреляционной функции выходного сигнала искусственного нейрона, принадлежащего произвольному слою p ИНС:

$$K_p(n) = \sum_{i_1=1}^{4} A_{pi_1}\left[K_{p-1}(n)\right]^{i_1} =$$
$$= \sum_{i_1=1}^{4} A_{pi_1}\left\{\sum_{i_2=1}^{4} A_{(p-1)i_2}\left[K_{p-2}(n)\right]^{i_2}\right\}^{i_1} =$$
$$= \sum_{i_1=1}^{4} A_{pi_1}\left(\sum_{i_2=1}^{4} A_{(p-1)i_2}\cdots\right.$$
$$\left.\cdots\left\{\sum_{i_{p-2}=1}^{4} A_{2i_{p-2}}\left[K_1(n)\right]^{i_p-2}\cdots\right\}^{i_2}\right)^{i_1}, \qquad (14)$$

где

$$A_{p1} = a_1^2\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\ \mathcal{R}_p\mathbf{W}_{\mathrm{ст}pi}r^{|n|}; \qquad (15)$$

$$A_{p2} = \mu^2 a_1^4\,\mathrm{Sp}\left(\mathcal{R}_p\right)\frac{M^{-1}_{\mathrm{уср}}}{1-r^2}\times$$
$$\times\left(\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\mathcal{R}_pR^{\mathrm{H}}_{SS}\mathbf{W}_{\mathrm{ст}\,pi}+\right.$$
$$\left.+\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}r^2\mathcal{R}_pR_{SS}\mathbf{W}_{\mathrm{ст}\,pi}\right)r^{|n|}; \qquad (16)$$

$$A_{p3} = \mu^2 a_1^4 r\frac{1+r-r^n}{(1-r)^2}\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\mathcal{R}_p^3\mathbf{W}_{\mathrm{ст}\,pi}r^{|n|}+$$
$$+\mu^2 a_1^4\frac{r}{(1-r)^2(1+r)}\times$$
$$\times\mathrm{Sp}\left(\mathcal{R}_p\right)\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\mathcal{R}_p^2\mathbf{W}_{\mathrm{ст}\,pi}r^{|n|}+$$
$$+\mu^2 a_1^4\left[\frac{2}{\left(1-r^2\right)^2}+\frac{1-r^{2n+2}}{1-r^2}\right]\times$$
$$\times\mathrm{Sp}\left(\mathcal{R}_p^2\right)\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\mathcal{R}_p\mathbf{W}_{\mathrm{ст}\,pi}r^{|n|}; \qquad (17)$$

$$A_{p4} = \mu^2 a_1^4\frac{r^3}{1-r^2}\left(\frac{1-r^n}{1-r^2}+\frac{1}{1-r}\right)\times$$
$$\times\mathrm{Sp}\left(\mathcal{R}_p^2\right)\mathbf{W}^{\mathrm{H}}_{\mathrm{ст}\,pi}\mathcal{R}_p^2\mathbf{W}_{\mathrm{ст}\,pi}r^{|n|}. \qquad (18)$$

Сравнение выражений (15)–(18) с формулами (3)–(6) и (10)–(13) показывает, что (15)–(18) обобщают выражения для коэффициентов из (3)–(6) и (10)–(13). Результат зависит от номера слоя ИНС и вида матрицы $\mathcal{R}_p$:

$$\mathcal{R}_p = \begin{cases} R_{XX},\ p=1; \\ E,\ p=2,\ 3,\ \ldots . \end{cases} \qquad (19)$$

Из (19) следует, что для входного слоя ИНС матрицей $\mathcal{R}_p$ является корреляционная матрица входных сигналов. При этом формулы (15)–(18) переходят в (3)–(6). Для скрытых слоев и выходного слоя ИНС матрицей $\mathcal{R}_p$ является матрица E. Тогда для второго слоя формулы (15)–(18) перейдут в выражения (10)–(13). Для последующих слоев ИНС коэффициенты разложения корреляционной функции выходного сигнала произвольного искусственного нейрона i данного слоя также определяются формулами (15)–(18) при соответствующем значении номера слоя p.

Остановимся более подробно на выражении (14). Из него следует, что по мере увеличения номера слоя ИНС коэффициенты при корреляционных функциях выходных сигналов слоя возводятся во все более высокую степень и перемножаются между собой. В результате в корреляционной функции выходного сигнала каждого последующего слоя сети возрастает вклад быстро спадающих компонентов, пропорционально величине $r^{\vartheta|n|}$ (ϑ – коэффициент, увеличивающийся по мере продвижения от входа ИНС к выходу). В результате в структуре выходных сигналов искусственных нейронов по мере возрастания номера слоя сети возрастает влияние компонентов с меньшим временем автокорреляции, чем у входного полезного сигнала, который подавался на входной слой сети одновременно с помехой.

Для иллюстрации данного утверждения приведем формулу для корреляционной функции выходного сигнала произвольного нейрона, принадлежащего слою p ИНС, полученную без учета флуктуаций весового вектора:

$$K_p(n) = a_1^{2p} r^{p|n|} \prod_{i=1}^{p} \mathbf{W}_{\text{ст}i}^{\text{H}} \mathcal{R}_i \mathbf{W}_{\text{ст}i}, \qquad (20)$$

где $\mathcal{R}_i$ определено по (19), а весовой вектор произвольного нейрона, принадлежащего слою i $\mathbf{W}_{\text{ст}i}$, является постоянным и стационарным.

Сравнив (14) и (20), приходим к выводу, что во втором случае в корреляционной функции отсутствуют слагаемые, пропорциональные коэффициенту адаптации μ. Это свидетельствует о том, что без учета флуктуаций ИНС выделяет только полезный сигнал на фоне помех. Вместе с тем как при учете флуктуаций, так и без их учета с увеличением числа слоев ИНС временна́я часть корреляционной функции содержит все большее число быстро спадающих компонентов.

Из (20) следует, что выходной сигнал искажается даже без учета флуктуаций. По мере увеличения номера слоя ИНС меняется все сильнее вид временной части корреляционной функции выходного сигнала, пропорциональной $r^{p|n|}$. Искажений временной части нет только в однослойной ИНС: в этом случае временная часть корреляционной функции выходного сигнала такая же, как и функции входного полезного сигнала: $r^{|n|}$. Учет флуктуаций показывает, что выходной сигнал искажается еще больше уже за счет флуктуационных слагаемых, начиная с первого слоя ИНС. Однако искажения за счет флуктуаций весовых коэффициентов имеют второй порядок малости по коэффициенту адаптации μ, т. е. достаточно малы.

Из представленного в настоящей статье математического анализа следует, что слагаемые корреляционной функции выходного сигнала зависят также от первого коэффициента разложения в ряд Вольтерра нелинейной функции F a_1, которая присутствует на выходе искусственного нейрона. В частном случае наиболее распространенного вида логистической нелинейной функции или сигмоида $f(x) = 1/\left(1+e^{-\rho x}\right)$ [13] первый коэффициент разложения составляет $a_1 = \rho/4$, причем ρ определяет крутизну сигмоида. При $\rho = 0$ сигмоид вырождается в горизонтальную линию на уровне 0.5, при увеличении ρ сигмоид приближается к функции единичного скачка [13]. Для практических задач достаточно брать значение $\rho = 0.5...1$. Тогда a_1^4 не превосходит значения $1/4^4 = 1/256 = 0.0039$. В результате флуктуационные слагаемые будут малы не только в силу малости коэффициента адаптации, но и по причине малости коэффициента a_1^4.

Если сравнить эффекты искажений из-за флуктуаций весового вектора выходного сигнала в адаптивной антенной решетке без нелинейной функции в цепи обратной связи, настраивающейся по LMS-алгоритму с квадратичным ограничением и содержащей в первом слое ИНС указанный алгоритм, то для первой сети эффект будет пропорционален коэффициенту адаптации μ, а во втором – величине $\mu^2 a_1^4$. Если принять $\mu = 0.1$, $a_1 = 1/4$, получим, что искажения, вносимые флуктуациями, в линейной адаптивной антенне будут выше в $\frac{\mu}{\mu^2 a_1^4} = \frac{0.1}{0.01 \cdot 0.0039} \approx 2564$ раза. Следовательно, нелинейная обработка существенно уменьшает вклад флуктуаций в выделяемый адаптивной системой сигнал.

Таким образом, суммарное воздействие флуктуаций весовых коэффициентов и числа слоев ИНС искажает выделяемый сетью полезный сигнал. Однако необходимо отметить, что выражения получены для частного и наиболее простого случая, когда на все искусственные нейроны входного слоя ИНС подаются одни и те же входные сигналы. При различной помеховой обстановке на входах сети результаты будут иными и, возможно, более сложными.

Полученные результаты позволяют утверждать, что флуктуации весовых коэффициентов вносят искажения в корреляционные характеристики ИНС, которые заключаются в появлении в выражении корреляционной функции выходного сигнала искусственного нейрона произвольного слоя ИНС дополнительных обусловленных флуктуациями слагаемых. Однако эти искажения имеют второй порядок малости влияния на усиление полезного сигнала по коэффициенту адаптации LMS-алгоритма с квадратичным ограничением. С возрастанием номера слоя влияние флуктуаций увеличивается, что приводит к снижению качества выделения сигнала на фоне помех с помощью ИНС.

СПИСОК ЛИТЕРАТУРЫ

1. Haykin S. Neural networks: A comprehensive foundation. 2nd ed. / пер. с англ. под ред. В. В. Шахгильдяна. Upper saddle river: Prentice-Hall, 1999. 874 c.

2. Галушкин А. И. Теория нейронных сетей. М.: ИПРЖР, 2000. 416 с.

3. Зимина С. В. Статистические характеристики искусственной нейронной сети с дискретным градиентным алгоритмом настройки с учетом флуктуаций весовых коэффициентов // Нейрокомпьютеры: разработка и применение. 2006. № 10. С. 9–15.

4. Зимина С. В. Влияние флуктуаций весовых коэффициентов на статистические характеристики искусственной нейронной сети с алгоритмом рекуррентного обращения выборочной оценки корреляционной матрицы входных сигналов // Нейрокомпьютеры: разработка и применение. 2007. № 5. С. 3–7.

5. Литвинов О. С., Зимина С. В. Статистический анализ флуктуаций весовых коэффициентов искусственной нейронной сети, настраивающейся по алгоритму Хэбба // Нейрокомпьютеры: разработка и применение. 2009. № 3. С. 33–43.

6. Орешкин Б. Н., Бакулев П. А. Алгоритм LMS с квадратичным ограничением // Антенны. 2007. № 9. С. 29–34.

7. Уидроу Б., Стирнз С. Адаптивная обработка сигналов. М.: Радио и связь, 1989. 440 с.

8. Пупков К. А., Капалин В. И., Ющенко А. С. Функциональные ряды в теории нелинейных систем. М.: Наука, 1976. 448 с.

9. Зимина С. В. Флуктуации весового вектора в адаптивных антенных решетках // Антенны. 2004. № 6. С. 27–35.

10. Зимина С. В. Флуктуации весового вектора в адаптивных антенных решетках с нелинейной функцией в цепи обратной связи, настраивающихся по алгоритму рекуррентного обращения выборочной оценки корреляционной матрицы входных сигналов // Изв. вузов. Радиофизика. 2006. Т. 49, № 2. С. 164–173.

11. Игнатенко С. В., Мальцев А. А. Статистические характеристики адаптивных антенных решеток при обработке дискретных сигналов с коррелированными отсчетами // Изв. вузов. Радиофизика. 1994. Т. 37, № 12. С. 1532–1545.

12. Мальцев А. А., Зимина С. В. Влияние флуктуаций весового вектора на статистические характеристики адаптивной антенной решетки с быстрым рекуррентным алгоритмом настройки // Изв. вузов. Радиофизика. 2002. Т. 45, № 8. С. 708–721.

13. Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия–Телеком, 2002. 382 с.

S. V. Zimina

Lobachevsky state university of Nizhny Novgorod

Correlative characteristics of the neural network with LMS-algorithm tuning with square constraint with weight vector jitter

The results of statistical analysis of neural network tuning by LMS-algorithm with square constraint for weight vector jitter are represented. The expressions for correlation function of output signal of neuron of the neural of different network layers are obtained. It is shown, that the weight vector jitter leads to distortions of statistical characteristics of neural network and these distortions increase with increasing of number of neural network layer.

Neural network, weight vector jitter, LMS algorithm tuning with square constraint

Статья поступила в редакцию 9 апреля 2015 г.