АНАЛИЗ ВЛИЯНИЯ ФЛУКТУАЦИЙ ВЕСОВОГО ВЕКТОРА НА СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ИСКУССТВЕННОГО НЕЙРОНА С ОГРАНИЧЕНИЯМИ, НАСТРАИВАЮЩЕГОСЯ ПО ДИСКРЕТНОМУ ГРАДИЕНТНОМУ АЛГОРИТМУ

Зимина С.В.

Нижегородский государственный университет им. Н.И. Лобачевского

ВВЕДЕНИЕ.

Известно, что флуктуации регулируемых весовых коэффициентов определяют качество настройки адаптивных систем и, в частности, адаптивных антенных решёток (ААР). Изучение влияния флуктуаций весового вектора на статистические характеристики ААР с различными алгоритмами настройки проводилось в работах [1-4]. В приведённых исследованиях было показано, что алгоритм настройки определяет качественную и количественную специфику влияния флуктуаций на точность настройки адаптивных антенных решёток. В последнее время активно развивается научное направление, связанное с изучением искусственных нейронных сетей (ИНС). Это связано с тем, что ИНС способны решать задачи большей степени трудности, чем линейные адаптивные системы.

Представляет интерес изучение влияния флуктуаций весового вектора на статистические характеристики искусственного нейрона (ИН), поскольку он является простейшим обобщением адаптивной антенной решётки, у которой на выходе присутствует нелинейная активационная функция. Можно предположить, что точность настройки ИН так же, как и в случае линейной адаптивной системы, будет определяться флуктуациями весовых коэффициентов.

В данной работе представлены результаты анализа влияния флуктуаций весовых коэффициентов на статистические характеристики искусственного нейрона с ограничениями, настраивающегося по дискретному градиентному алгоритму. Решение данной задачи осуществлялось методами теории возмущений. В первом (борновском) приближении были найдены выражения для мощности выходного сигнала ИН и корреляционной матрицы флуктуаций весового вектора. При статистическом анализе предполагалось, что искусственный нейрон содержит N_1-раз дифференцируемую функцию, без уточнения её конкретного вида.

ПОСТАНОВКА ЗАДАЧИ И ЕЁ РЕШЕНИЕ.

Рассмотрим работу искусственного нейрона, схема которого представлена на рис.1. Настройка вектора весовых коэффициентов $\vec{W}$ ИН с ограничениями, работающего по дискретному градиентному алгоритму, описывается N -мерным векторным уравнением:

$$\vec{W}(k+1) = \mathbf{P}\{\vec{W}(k) - \mu \cdot \vec{X}^*(k) \cdot Z(k)\} + \vec{W}_q . \qquad (1)$$

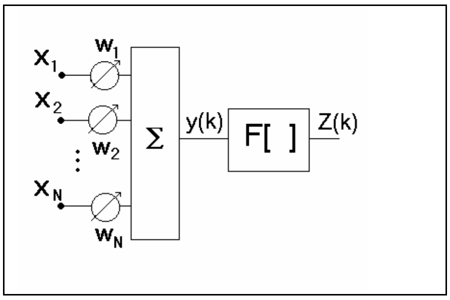

Рис.1.

Здесь $\vec{X}(k)$ - вектор входных сигналов, $\mathbf{P}$ - проекционная матрица, обеспечивающая введение многократных линейных ограничений на диаграмму направленности ИН; μ - коэффициент адаптации; $\vec{W}_q$ - вектор комплексных весовых коэффициентов, соответствующих диаграмме направленности покоя (при отсутствии внешних помех); H, *, + соответственно знаки эрмитовского и комплексного сопряжения, а также псевдообращения.

В формуле (1) $Z(k)$ - выходной сигнал искусственного нейрона, который может быть записан в виде:

$$Z(k) = F[y(k)] = \sum_{j=1}^{N_1} a_j \cdot y^j(k) = \vec{A}^T \vec{Y}(k), \qquad (2)$$

где $F[\ ]$ - нелинейная активационная функция на выходе ИН, a_j - коэффициенты разложения нелинейности F в ряд Вольтерра [5], $y(k)$ - выходной сигнал линейной части искусственного нейрона. $\vec{A} = [a_1\ a_2 \cdots a_{N_1}]^T$ - вектор коэффициентов разложения активационной функции в ряд Вольтерра, $\vec{Y}(k) = \left[y(k)\ y^2(k) \cdots y^{N_1}(k)\right]^T$ - вектор степеней выходного сигнала линейной части искусственного нейрона.

Будем рассматривать узкополосный искусственный нейрон с корреляционной матрицей входных сигналов следующего вида:

$$\mathbf{R}_{XX}(k, k+n) \equiv \langle \vec{X}^*(k) \vec{X}^T(k+n) \rangle = \mathbf{R}_{XX} r^{|n|},$$

где r - коэффициент корреляции между отсчётами входных сигналов,

$\mathbf{R}_{XX}$ - пространственная часть корреляционной матрицы входных сигналов.

Будем считать также, что при разложении активационной функции искусственного нейрона в ряд Вольтерра (2) первый член ряда оказывается значительно больше, чем последующие слагаемые, и в силу этого ими можно пренебречь.

Методами теории возмущений по малому параметру μ при использовании указанных предположений в первом приближении было получено выражение для выходной мощности искусственного нейрона при учёте флуктуаций весового вектора:

$$< \left|Z\right|^2 >_{CT} = < \left|Z\right|^2 >_0 \times$$
$$\times \left\{ 1 + \mu^2 a_1^2 Sp(\mathbf{PR}_{XX}\mathbf{PR}_{XX}) \cdot \left(\frac{1+r}{1-r} - \frac{r}{(1-r)^2}\right) + \mu^2 a_1^2 Sp^2(\mathbf{PR}_{XX}) \frac{1}{(1-r)^2} \right\}, \quad (3)$$

где $< \left|Z\right|^2 >_0 \equiv y_{CT}^H a_1^2 y_{CT} = a_1^2 \vec{W}_{CT}^H \mathbf{R}_{XX} \vec{W}_{CT}$ - выходная мощность без учёта флуктуаций весовых коэффициентов.

Из данной формулы видно, что второе и третье слагаемое в выражении мощности на выходе ИН определяются флуктуациями весового вектора. Необходимо подчеркнуть, что вклад, вносимый флуктуациями в величину выходной мощности, может быть по знаку как положительным, так и отрицательным. По этой причине мощность выходного сигнала искусственного нейрона также может быть как больше, так и меньше мощности, найденной при постоянном стационарном весовом векторе, т.е. в ИН так же, как и в адаптивной антенной решётке, может иметь место как эффект рассогласования [6], так и эффект «перекомпенсации» [1-2].

Матрица ковариации вектора весовых коэффициентов в стационарном режиме работы в совпадающие моменты времени в борновском приближении имеет следующий вид:

$$\mathbf{K}_{\tilde{W}} = \mu^2 \frac{1-r-r^2}{(1-r)^2} Sp(\vec{Y}_{CT}^* \vec{Y}_{CT}^T \mathbf{R}_{AA}) \cdot (\mathbf{PR}_{XX}\mathbf{P})^* , \quad (4)$$

где $\vec{Y}_{CT} = \left[y_{CT} \; y_{CT}^2 \cdots y_{CT}^{N_1} \right]^T = \left[\vec{X}^T \vec{W}_{CT} \; (\vec{X}^T \vec{W}_{CT})^2 \cdots (\vec{X}^T \vec{W}_{CT})^{N_1} \right]^T$ - стационарный выходной сигнал линейной части искусственного нейрона.

Из приведённой формулы видно, что поскольку собственные числа корреляционной матрицы входных сигналов искусственного нейрона $\mathbf{R}_{XX}$ при воздействии внешних помех различны, то в общем случае собственные числа корреляционной матрицы флуктуаций весового вектора $\mathbf{K}_{\tilde{W}}$ также различны. Поэтому флуктуации весового вектора в ИН, настраивающемся по дискретному градиентному алгоритму, неизотропны в пространстве весовых коэффициентов. Они равны нулю в направлении ограничений, максимальны в направлении помех, а во всех остальных направлениях имеют минимальную ненулевую величину, определяемую мощностью собственного шума искусственного нейрона. Заметим, что для градиентного алгоритма в отсутствии нелинейной активационной функции флуктуации весов изотропны, т.е. являются одинаковыми по величине во всех направлениях подпространства ограничений [1-2].

Таким образом, проведённое исследование показало, что флуктуации весового вектора приводят к искажениям выходного сигнала искусственного нейрона, и флуктуации неизотропны в пространстве весовых коэффициентов.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (РФФИ гранты N 03-02-17141, N НШ-1729.2003.2, 03-02-06378).

ЛИТЕРАТУРА.

1. Игнатенко С. В., Мальцев А. А. Статистические характеристики адаптивных антенных решёток при обработке дискретных сигналов с коррелированными отсчётами // Изв. ВУЗов. Радиофизика, 1994. т. 37. N12. С. 1532 - 1545.
2. А. А. Мальцев, С. В. Зимина. Спектрально-корреляционные характеристики выходного сигнала адаптивных антенных решёток с учётом флуктуаций весового вектора // Радиотехника и электроника, 2001. т.46. N11. с.1350– 1355.
3. Krolik J. L., Swingler D. N. On the mean-square error performance of adaptive minimum variance beamformers based on the sample covariance matrix // IEEE Trans. Signal Processing. 1994. v. 42. N2. P. 445 - 448.
4. Мальцев А.А., Зимина С.В. Влияние флуктуаций весового вектора на статистические характеристики адаптивной антенной решётки с быстрым рекуррентным алгоритмом настройки // Изв. ВУЗов. Радиофизика, 2002. т. 45. N8. с. 708 - 721.
5. Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем. М.: Наука, 1976.