УДК 004.2-004.8

А.И. Андрюхин (канд.тех.наук, доц.)
Донецький национальный технический університет
alexandruckin@gmail.com

ОЦЕНКА ВАРИАЦИЙ ПАРАМЕТРОВ НЕЙРОМОРФНЫХ СЕТЕЙ

Выполнен анализ базового элемента операции распознавания в пороговой логике. Рассмотрены оценки вариаций для параметров базового элемента. Приведены численные результаты расчетов.

Ключевые слова: пороговая логика, XOR, вариация, параметр.

Введение

Основой будущей вычислительной техники и искусственного интеллекта будут нейрокомпьютерные системы. Идеальным решением реализации сенсорных систем восприятия на базе твердотельной электроники являются в настоящее время аналоговые СБИС. В случае успеха мы научимся создавать вживляемые кремниевые сетчатки для слепых и звуковые процессоры для глухих, а также дешевые и эффективные визуальные, звуковые и обонятельные чипы для роботов. Создание нейроморфных микрочипов представляет собой отображение (морфинг) нервных связей на кремниевые электронные цепи.

Обычно разделяют реализацию биологических и формальных нейронных систем.

Наиболее известный шаг при аппаратной реализации биологических нейронных систем связан с именами Миши Маховальд и известного специалиста по микроэлектронной технологии Карвера Мида (Carver Mead), которые впервые попытались изготовить глазную сетчатку из кремния в Калифорнийском технологическом институте. Они воспроизвели на основе нейрочипов три из пяти ее слоев. Другие исследователи моделировали (морфировали) остальные части зрительной и слуховой систем. В 2001 г. изготовили все пять слоев сетчатки и смоделировали визуальные сообщения, посылаемые мозгу ганглионарными клетками, т.е. выходными нейронами сетчатки. Кремниевый чип сетчатки, Visio1, воспроизводит реакцию основных четырех типов ганглионарных клеток сетчатки, которые вырабатывают сигналы и вместе составляют 90% оптического нерва.Чип имитирует процесс, в котором активированные напряжением ионные каналы заставляют ганглионарные клетки (и нейроны в остальной части мозга) вырабатывать нервные импульсы.

Для аналоговой реализации центральной нервной системы важен механизм генерации нервного импульса, т.е. функционирование клеточной мембраны, которая поддерживает постоянный состав цитоплазмы внутри клетки и обеспечивает проведение нервных импульсов. Импульсы по волокну передаются в виде скачков потенциала внутриклеточной среды по отношению к внешней среде, окружающей клетку.

Для описания этих процессов используют классическую модель электрогенеза нервной клетки Ходжкина-Хаксли или ее более простые модификации Морриса-Лекара, Фиц Хью - Нагумо, Хиндмарш-Розе и др.

Согласно современным представлениям, такие функции человеческого мозга, как сознание, распознавание и многие другие, объясняются синхронным поведением нейронов определенных структур. С другой стороны, движениями живого существа управляют специальные нервные подсистемы — центральные генераторы ритма(ЦГ). Хотя ЦГ может быть автономной подсистемой, его ритм и фаза должны быть синхронизированы с состоянием двигательной системы, что осуществляется на основе сенсорной обратной связи. Поэтому большое внимание уделяется синхронизации в нейронных структурах.

Регулярная и хаотическая синхронизация является одной из актуальных и наиболее активно исследуемых задач нелинейной динамики. В последние годы было показано, что посредством связи синхронизируются не только периодические системы, но и хаотические.

При реализации формальных нейронных систем широко используют тот факт, что многие булевы функции можно представить пороговыми функциями. Как известно, пороговая функция (TLF) –это булева функция $Y(\boldsymbol{x})=sgn(\sum_{i=1}^{n}\omega_i x_i - \epsilon$), где ω_i-веса аргументов x_i, ϵ-порог, $\boldsymbol{x}=(x_1, \ldots x_i,\ldots x_n)$. Более того, важность использования TL связана с доказательством того, что некоторые виды булевых функции могут быть реализованы с использованием TL сетей, которые требуют меньше вентилей при меньшем количестве ступеней схемной реализации в сравнении с реализациями, основанными на традиционных булевых вентилях. Это ясно при рассмотрении следующих простых примера $x_1\&(x_2\vee x_3)= sgn(2x_1+ x_2+ x_3-3)$ или $x_1\vee(x_2\wedge x_3) = sng(2x_1+x_2+x_3-1.5)$[1-2].

Наиболее известные представления булевых функций пороговыми функциями

$$OR(x_1,x_2,..,x_n)=sgn(-1+ x_1+ x_2+\ldots+ x_n),$$
$$AND(x_1,x_2,..,x_n)=sgn(-n+ x_1+ x_2+\ldots+ x_n),$$
$$NAND(x_1,x_2,..,x_n)=sgn(n-1- x_1- x_2-\ldots- x_n),$$
$$MAJ(x_1,x_2,..,x_n)=sgn(-\lceil n/2\rceil+ x_1+ x_2+\ldots+ x_n).$$

Однако легко показывается, что функция $XOR(x_1, x_2)$ не представима функцией вида $sgn(w_0+w_1\ x_1+w_2\ x_2)$, т.е. не является пороговой (ЛРБФ-линейно разделимая булева функция) функцией.

Целью исследования является оценка возможности определения вариаций параметров искусственных нейронных сетей, при которых возможно правильное функционирование последних. Это позволит более точно и эффективно выполнять проектирование и диагностирование ИНС.

Задачей исследования является построение математической модели влияния вариаций параметров ИНС на ее функционирование по проекту. Это позволит определять граничные значения параметров ИНС.

Общая постановка задачи

Согласно [3], степень соответствия характеристик исследуемого объекта его проектным нормам определяются целевой функцией, т.е. некоторой численной метрикой отклонений от проектных решений. Обозначим через $P=(p_1,p_2,.,p_{n-1},p_n)$ ($P+\Delta P=(p_1+\Delta\ p_1,p_2+\Delta\ p_2,.,p_{n-1}+\Delta\ p_{n-1},p_n+\Delta\ p_n)$) вектор параметров ИНС и вектор параметров с вариациям соответственно. Известно[4], что наиболее широко используется для описания систем автоматная модель $Y_t=Y(X_t,S_t,P)$, $S_{t+\Delta t}=S(X_t,S_t,P)$, (X, Y,S –входы, выходы, состояния ИНС).

Определим общую целевую функцию выхода ИНС выражением

$$F(\Delta P)=\sum_{i=1}^{N}\sum_{j=1}^{M}\mathrm{sign}(Y(X_i,S_j,P+\Delta P)-Y(X_i,S_j,P))^2$$

Здесь N(M)-число различных значений входов X и состояний S автоматной модели ИНС. Ясно, что в общем случае $N=2^K$, где K равно числу внешних входных полюсов ИНС.

Определяя граничные значения ΔP, мы выполняем анализ чувствительности проектного решения, что в принципе представляет задачу параметрической оптимизации с ограничением $F(\Delta P)=0$. Последнее уравнение определяет, что нет различия в выходных реакциях проектного решения и проекта с определенными вариациями параметров.

Анализ целевой функции для произвольного проекта ИНС весьма сложен, но если мы ограничимся проектами ИНС для реализаций только булевых выражений, все значительно упрощается, так как:

а) имеем высокую степень регулярности ИНС и следовательно N будет равно $C*2^L$, где L равно максимальному числу входов нейронов в ИНС, а C равно общему числу нейронов с L входами;

б) так как мы используем пороговую логику, можно не анализировать состояния ИНС (нет обратных связей при реализации булевых выражений), и следовательно целевая функция упрощается и будет иметь вид

$$F(\Delta P) = \sum_{i=1}^{N} \operatorname{sign}((Y(X_i, P + \Delta P) - Y(X_i, P))^2$$

в) функция f активации нейрона является пороговой функцией и следовательно монотонно возрастающей. Поэтому справедливо, что если мы меняем какой -либо из параметров его функционирования $P_i = z$ на интервале (a,b) и f(a)=f(b), то для любого $z \in (a,b)$ справедливо f(z)= f(a)=f(b).

Заметим, что функция XOR, являющаяся элементарным распознавателем самых простых сигналов или объектов, реализуется достаточно сложной нейронной сетью. Известно, что операция распознавания является базовой операцией для интеллектуальных систем. Формальный нейрон, который описывается пороговой функцией, является базовым элементом ИНС. Используя ИНС, мы пытаемся построить системы искусственного разума, и естественно предположить, что базовый элемент может выполнять операцию распознавания. Но распознавание является эмерджентным свойством систем, которые могут быть построены на пороговых элементах, как и память для базового элемента памяти. Последний элемент легко реализуется двумя нейронами с обратными связями. Второе объяснение связано с неэквивалентностью биологического и формального нейронов.

Пример расчета

Конкретизируем основные этапы предлагаемой методики на примере классической проблемы реализации логического элемента XOR двухслойной ИНС. Эта функция является единственной (и ее отрицание) булевой функцией от двух переменных из 16, которая не может быть реализована однослойной ИНС и это обстоятельство было одной из причин угасания внимания к нейронным сетям в 60-е годы.

Заметим, что в настоящее время нет полной ясности о числе линейно разделимых (или неразделимых) булевых функций от n переменных при n>8 [2].

Таблица 1.Таблица истинности функции XOR.

	X1	X2	Y=XOR(X1,X2)
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0

Согласно [5], общее решение проблемы реализации XOR неоднослойной ИНС можно представить рисунок 1, который соответствует рисунку 3.4 из[5].

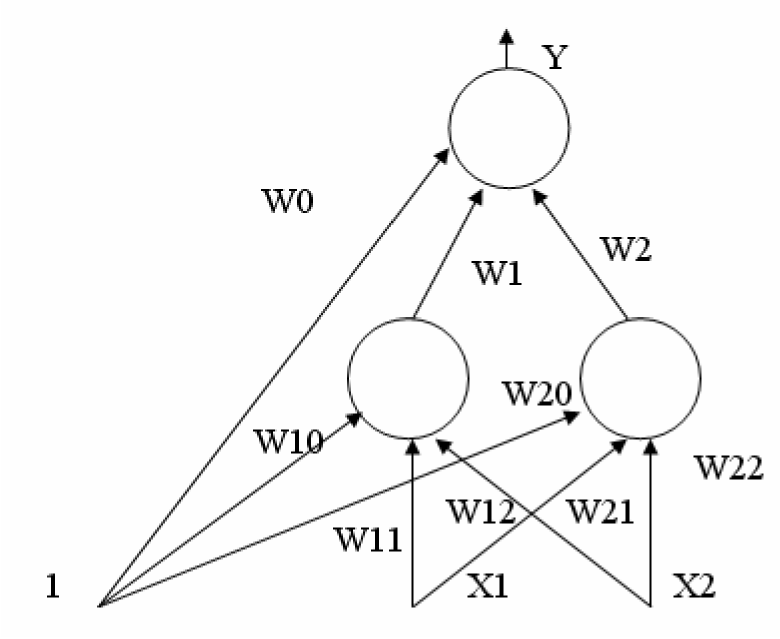


Рисунок 1 – Структура ИНС для функции XOR.

Учитывая, что первый элемент нижнего слоя нейронов, обозначенный символом 1 на рис.1, обеспечивает вариацию порогов остальных элементов, можем представить решение на рис.2 в виде 3 двухвходовых нейронов

$$Y3=f(w31*Y1+w32*Y2-TH3),$$

где выходы нейронов нижнего слоя определяются выражениями

$$Y1= f(w11*x1+w12*x2-TH1)$$

$$Y2= f(w21*x1+w22*x2-TH2)$$

Для дальнейшего более краткого изложения обозначим через P параметры (w11,w12,TH1,w21,w22,TH2,w31,w32,TH3).

Следовательно, выход ИНС Y3 есть функция F(x1,x2,P) для каждой пары булевых значений x1,x2.

Определим целевую функцию F отклонения значений выхода исследуемой ИНС от значений функции XOR. Для всех значений аргументов и значений самой функции представленной в таблице 1, можем записать

$$F(P)= (f(0,0,P)-0)^2+(f(0,1,P)-1)^2+(f(1,0,P)-1)^2+(f(1,1,P)-0)^2$$

Параметры w11,w12,TH1,w21,w22,TH2,w31,w32,TH3 в процессе изготовления и эксплуатации могут иметь вариации Δw11, Δw12, ΔTH1, Δw21, Δw22, ΔTH2, Δw31, Δw32, ΔTH3 соответственно от проектных значений. Ясно, что ИНС выполняет правильно свою функцию при различных вариациях параметров, если F(P+ΔP)=0.

Мы можем зафиксировать проектные значения всех параметров кроме рассматриваемого P_i и определить его интервал изменения ($P_i-\Delta P_i$, $P_i+\Delta P_i$) для которого F(P+ΔP)=0 и F(P-ΔP)=0 , где естественно $\Delta P=(0,0,..\Delta P_i,\ldots,0)$.

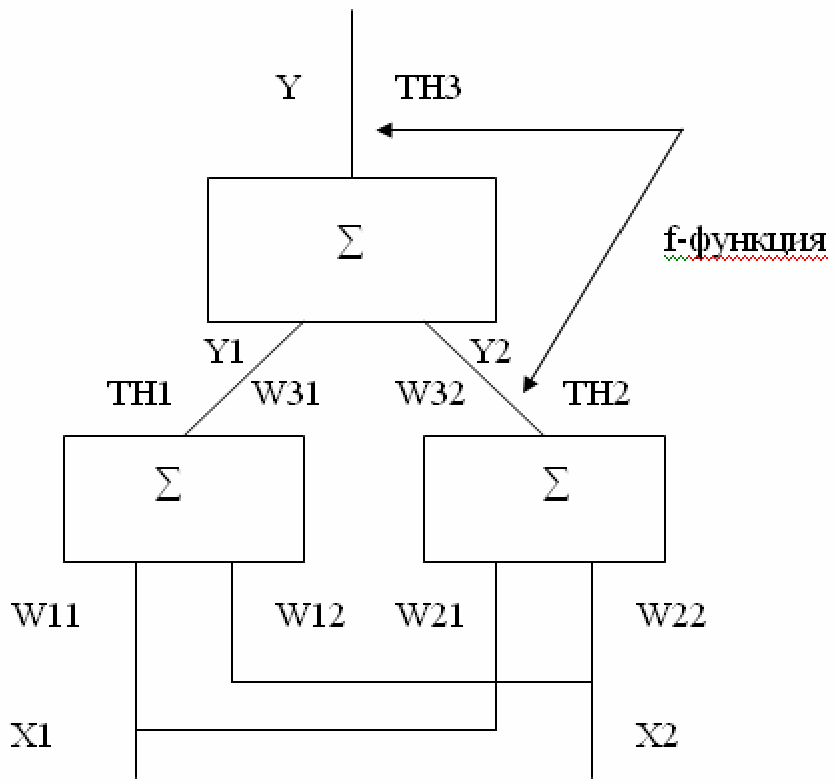


Рисунок 2 – Классическая реализация функции XOR трехнейронной сетью.

Обозначим Δ P_i через z и конкретно, пусть z есть вариация порогов нейронов TH1, TH2,TH3 при определенных значениях их весов.

Возьмем случай w11 =w22=1, w12=w21=-1, w31=w32=1, TH=TH2=0.41,TH3=0.7. При таких значениях параметров

$$F(P)= (F(0,0,P)-0)^2+(F(0,1,P)-1)^2+(F(1,0,P)-1)^2+(F(1,1,P)-1)^2=0$$

Для того чтобы F(P)= 0 достаточно, чтобы значения выходов в нейронов не изменялись при изменении их пороговых значений в интервалах (0.41-z,0.41+z), (0.41-z,0.41+z) ,(0.7-z,0.7+z) для TH1,TH2,TH3 соответственно.

Это означает определение max z для всех возможных значений x1,x2 при следующих условиях:

1) для Y1 f(x1-x2-TH1-z)= f(x1-x2-TH1+z)= f(x1-x2-TH1);
2) для Y2 f(-x1+x2-TH2-z)= f(-x1+x2-TH2+z)= f(-x1+x2-TH2);
3) для Y3 f(Y1+Y2-TH3-z)=f(Y1+Y2-TH3+z)=f(Y1+Y2-TH3).

Значения этих выражений по проекту нам известны, мы просто можем вычислить последние правые части этих уравнений.

Как было указано ранее, вид целевой функции можно значительно упростить путем подстановки вместо используемых выражений для выходов нейронов (в нашем случае Y1,Y2 для Y3) комбинаций x1,x2. Действительно выходы нейронов принимают булевы значения, но на практике не все возможные комбинации. Так при наших параметрах невозможен случай Y1=Y2=1.

Тем самым, мы увеличиваем число ограничений на вариацию порога, но значительно упрощаем ее вид. Ясно, что такие рассуждения справедливы, тогда когда число внешних входов ИНС не меньше, чем число входов для нейрона, выходное значение которого мы рассматриваем. Однако легко показать, что это ограничение не является существенным путем добавления фиктивных первичных входов ИНС, которые не влияют на функционирование нейронов.

С учетом вышесказанного получаем в итоге целевую функцию отклонения

$$F(z)=F_1(z)+ F_2(z)+ F_3(z),$$

где части функции отклонения в выходах Y1,Y2,Y3 соответственно равны

$F_1(z)=(f(-0.41-z)-0)^2+ f(-0.41+z)-0)^2+ f(-1-0.41-z)-0)^2+ f(-1-0.41+z)-0)^2+(1-0.41-z)-1)^2+ f(1-0.41+z)-1)^2+(-0.41-z)-0)^2+ f(-0.41+z)-0)^2$,

$F_2(z)=(f(-0.41-z)-0)^2+ f(-0.41+z)-0)^2+ f(1-0.41-z)-1)^2+ f(1-0.41+z)-1)^2+(-1-0.41-z)-0)^2+ f(-1-0.41+z)-0)^2+(-0.41-z)-0)^2+ f(-0.41+z)-0)^2$,

$F_3(z)=(f(-0.7-z)-0)^2+ f(-0.7+z)-0)^2+ f(1-0.7-z)-1)^2+ f(1-0.7+z)-1)^2+(1-0.7-z)-1)^2+ f(1-0.7+z)-1)^2+(2-0.7-z)-1)^2+ f(2-0.7+z)-1)^2$.

На основании вышесказанного можем сформулировать следующую задачу условного экстремума:

определить max z при F(z)=0.

Используя метод множителей Лагранжа, определим функцию Лагранжа Ф(z)=z+L1*F(z), получаем решение z=0.2500029 из системы уравнений F(z)=0, $\Phi'(z)=0$. Заметим, что решение для неупрощенной целевой функции, которая эквивалентна F(z) без двух последних слагаемых в $F_3(z)$, является таким же. Основные результаты расчетов наглядно представлены на рисунках 3-8.

Анализ полученных результатов

Как было отмечено ранее, приведенные результаты расчетов были получены с использованием идеальной разрывной функции активации нейронов f равной 1, если взвешенная сумма превышает порог и 0 в противном случае (функция Хевисайда при нулевом значении в точке 0). Естественно на практике переключение происходит за ненулевое время и

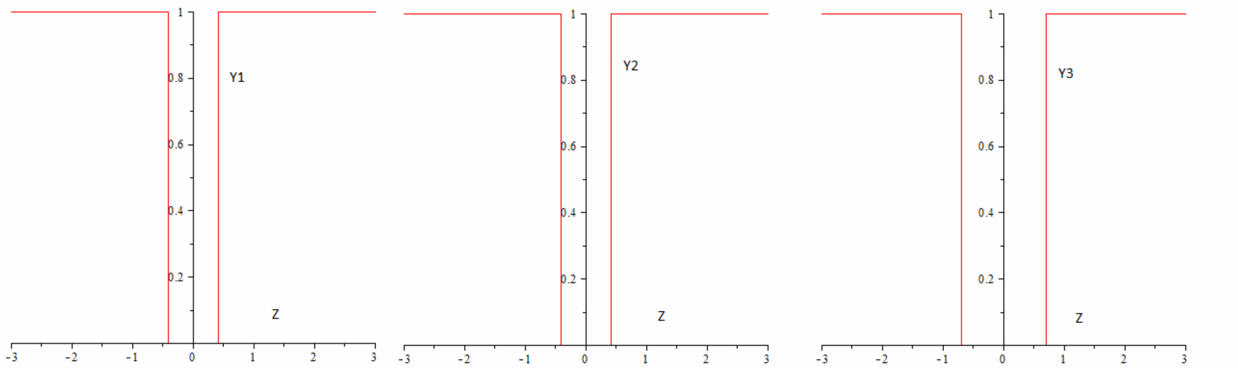


Рисунок 3 –Значения выходов нейронов при X1=X2=0 при изменении z.

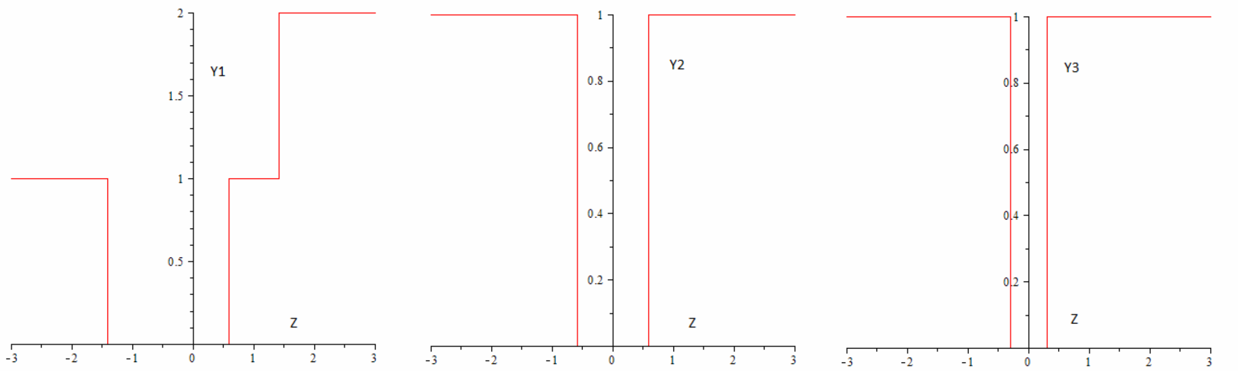


Рисунок 4 – Значения выходов нейронов при X1=0,X2=1 при изменении z.

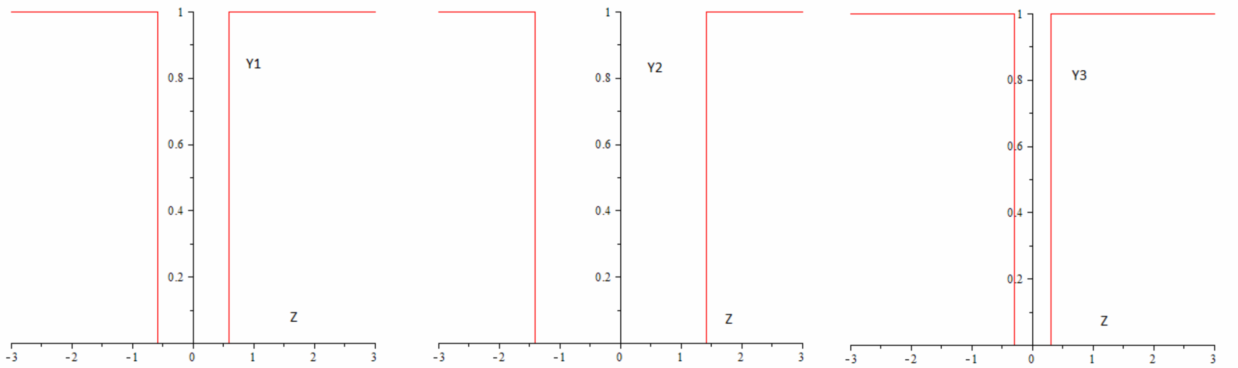


Рисунок 5– Значения выходов нейронов при X1=1,X2=0 при изменении z.

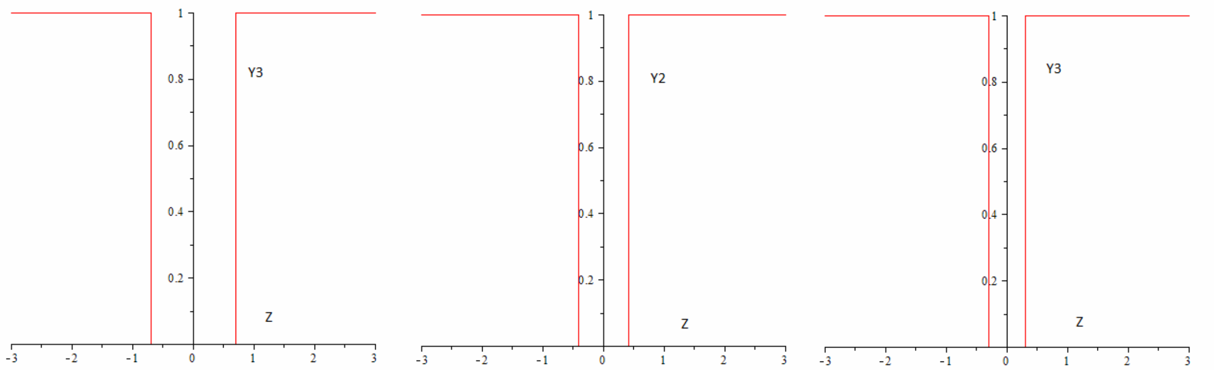


Рисунок 6 – Значения выходов нейронов при X1=X2=1 при изменении z.

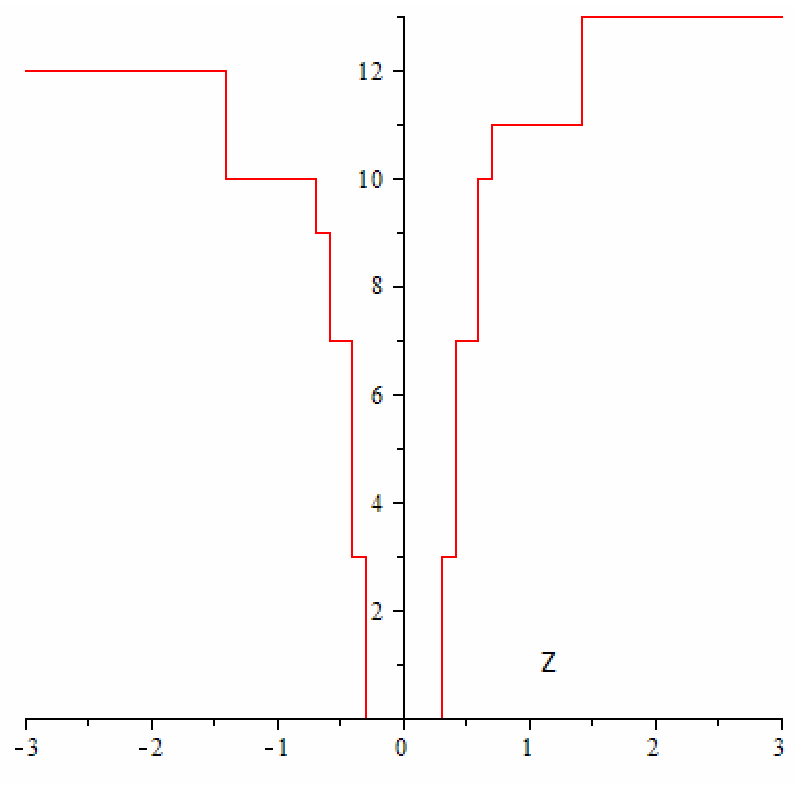
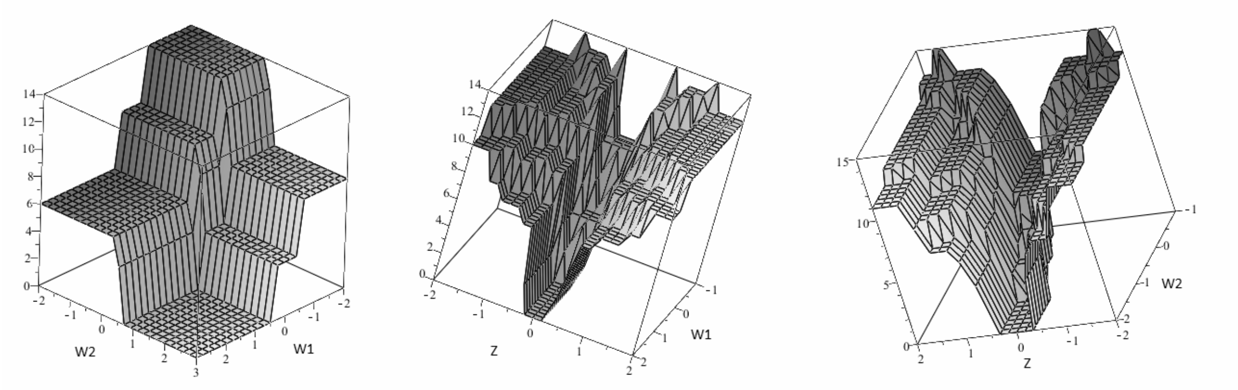


Рисунок7– Значения целевой функции отклонения F при изменении z .



а) б) в)

Рисунок 8-Значения целевой функции отклонения F при вариации двух параметров: а)при z=0 ; б) при w2=1; в) при w1=1.

поэтому выполнялись расчеты для непрерывных функций активации, которые являлись известными аппроксимациями идеальной функции.

Рассматривались функции $f_1(x)=(th(a*x)+1)/2$ и $f(x)=(1/(1+exp(-b*x))$ $(a>0,b>0)$.Эти функции эквивалентны при b=2a.

Для первой функции получены следующие результаты: при a=50,60,75,100,125,233,234,620,621 имеем z=0.0468..,0.0833...,0.12500.., 0.1699.., 0.1699..,0.1875..,0.1875..,0.2500.. соответственно. Результаты расчетов с этой функцией совпадают с результатом расчета по функции Хевисайда начиная с a=621.Наблюдаем эффект переполнения разрядной сетки компьютера, обычный при вычислении трансцендентной экспоненты[3] .

Заключение и перспектива дальнейших исследований

Основное направление дальнейших исследований связано с изучением связи характеристик и параметров различных динамических алгоритмов изменения весов, что является основным процессом при обучении ИНС с устойчивыми состояниями функционирования ИНС и переходами между ними.

Список литературы

1.Valeriu Beiu, *Senior Member, IEEE*, José M. Quintana, and María J. Avedillo. VLSI Implementations of Threshold Logic—A Comprehensive Survey// IEEE Transactions on Neural Networks, Vol. 14, No. 5, September 2003,pp.1217-1243.

2.Андрюхин А.И.Моделирование и диагностирование дискретных устройств на переключательном уровне: монография .Донецк: ГВУЗ «ДонНТУ», 2012. – 257 с.:

3.Петренко А.И.Основы автоматизации проектирования.-К.:Техніка,1982.-295 с.

4. Пешель М. Моделирование сигналов и систем. М.: Мир.-1981. 300 с.

5.Осовский С. Нейронные сети для обработки информации/Пер. с польскогоИ.Д.Рудинского.-М.:Финансы и статистика,2004.-344 с.

Надійшла до редколегії 10.09.2012 р.
Рецензент:канд.тех.наук, доц. Григор'єв О.В.

О.І.Андрюхін
Донецький національній техничеcкий університет

Оцінка варіації параметрів нейроморфніх мереж. Виконано аналіз базового елементу операції розпізнавання в порогової логіці. Розглянуто оцінки варіацій для параметрів базового елементу. Наведено чисельні результати розрахунків.
Ключові слова: порогова логіка, XOR, варіація, параметр.

A.I.Andruckin
Donetsk National Technical University

Estimation of neural networks parameters variation. Analysis of the basic elements of the operation of recognition in threshold logic is carried out. Estimates of the variations for the basic element parameters are considered. Numerical results of calculations are presented.
Keywords: threshold logic, HOR, variation, parameter.