

Зная, как вычисляются среднеквадратические отклонения погрешностей для всех элементов стандартного нейрона, можно вычислить среднеквадратические отклонения погрешностей сигналов для всей сети. Если заданы среднеквадратические отклонения погрешностей для выходных сигналов сети, можно вычислить среднеквадратические отклонения погрешностей для последнего слоя сети. Когда вычислены среднеквадратические отклонения погрешностей всех входных сигналов последнего слоя сети, переходим к вычислению среднеквадратических отклонений погрешностей предпоследнего слоя и так далее.

Рассмотрим пример на рис.10. Пусть дана сеть с тремя нейронами входного слоя, двумя нейронами скрытого слоя и одним выходным нейроном. На рисунке показаны сигналы, проходящие по сети при данном векторе входных сигналов, и веса связей. В данном примере элементы сети не имеют собственных погрешностей. Характеристическая функция нелинейных преобразователей имеет вид:  $\varphi(x) = x / (2 + |x|)$ , где  $x$  – входной сигнал нелинейного преобразователя. Среднеквадратическое отклонение вектора выходных сигналов сети  $\sigma_{out}$  равняется 0.01. Среднеквадратические отклонения погрешностей по входам сумматора вычисляются с использованием формулы для равномерного распределения среднеквадратических отклонений.

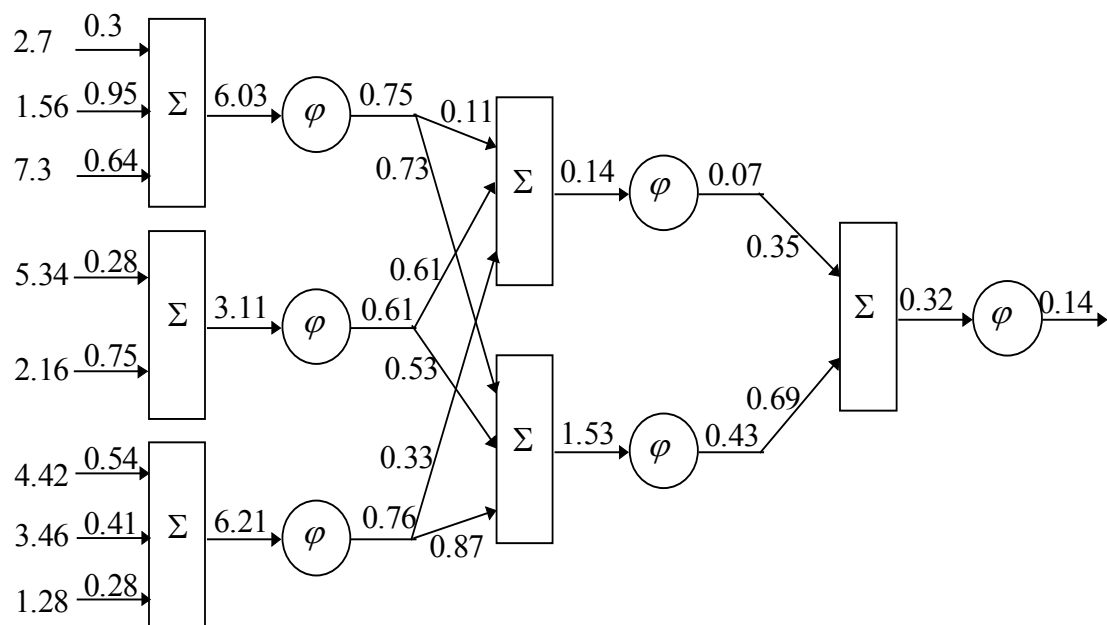


Рис.10

$\Sigma$  – адаптивные сумматоры,  $\varphi$  – нелинейные преобразователи.

Вычислим среднеквадратические отклонения для всех сигналов сети при данном векторе входных сигналов. Все вычисленные значения в этом примере округляются до двух знаков после запятой. На рис.11 показаны вычисленные среднеквадратические отклонения для данного примера.

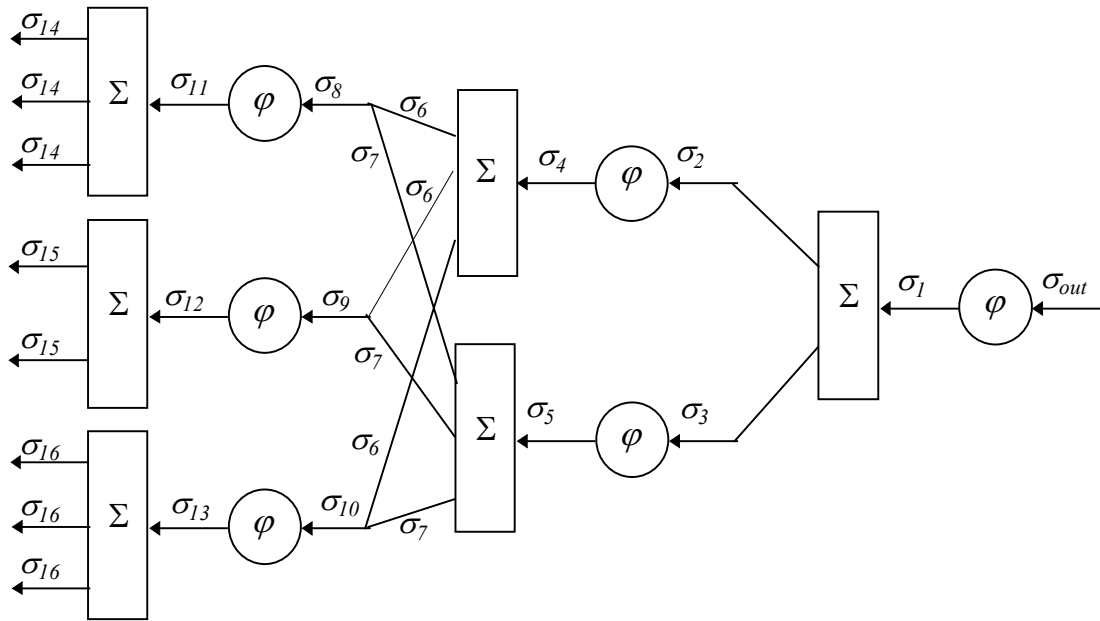


Рис. 11

$$\sigma_1 = \frac{\sigma_{out}}{|\varphi'(A)|} = \frac{0.01}{2/(2+0.32)^2} = 0.03; \sigma_2 = \sigma_3 = \frac{0.03}{0.35^2 + 0.69^2} = \frac{0.03}{0.77} = 0.04;$$

$$\sigma_4 = \frac{0.04}{2/(2+0.14)^2} = \frac{0.04}{0.44} = 0.08; \sigma_5 = \frac{0.04}{2/(2+1.53)^2} = \frac{0.04}{0.16} = 0.23;$$

$$\sigma_6 = \frac{0.08}{0.11^2 + 0.61^2 + 0.38^2} = \frac{0.08}{0.7} = 0.11;$$

$$\sigma_7 = \frac{0.23}{0.73^2 + 0.53^2 + 0.87^2} = \frac{0.23}{1.25} = 0.18;$$

$$\sigma_8 = \sigma_9 = \sigma_{10} = \min\{\sigma_6, \sigma_7\} = 0.11; \sigma_{11} = \frac{0.11}{2/(2+6.03)^2} = \frac{0.11}{0.03} = 3.55;$$

$$\sigma_{12} = \frac{0.11}{2/(2+3.11)^2} = \frac{0.11}{0.08} = 1.43; \sigma_{13} = \frac{0.11}{2/(2+6.21)^2} = \frac{0.11}{0.03} = 3.67;$$

$$\sigma_{14} = \frac{3.55}{0.3^2 + 0.95^2 + 0.64^2} = \frac{3.55}{1.49} = 2.38; \sigma_{15} = \frac{1.43}{0.28^2 + 0.75^2} = \frac{1.43}{0.8} = 1.79;$$

$$\sigma_{16} = \frac{3.67}{0.54^2 + 0.41^2 + 0.28^2} = \frac{3.67}{0.73} = 5.03;$$

Таким образом, получены формулы для вычисления среднеквадратических отклонений погрешностей сигналов сети, в предположении, что погрешности являются независимыми случайными величинами.

## 7. Анализ реализуемости сетей с собственными погрешностями элементов методом обратного распространения точности для оценки среднеквадратических отклонений

Все изложенные выше соображения о выполнимости метода обратного распространения точности справедливы и для метода обратного распространения точности для среднеквадратических отклонений погрешностей с учетом собственных погрешностей элементов. Отличие состоит в способе вычисления промежуточных среднеквадратических отклонений погрешностей.

Как и выше, рассмотрим участок сети, изображенный на рис.9. Для этого участка нам необходимо вычислить промежуточное среднеквадратическое отклонение погрешности  $\sigma^{part}$ .

Пусть собственное среднеквадратическое отклонение погрешности сумматора  $\Sigma_1$  равно  $\sigma_{\Sigma_1}$ , собственное среднеквадратическое отклонение погрешности нелинейного преобразователя равно  $\sigma_{\varphi}$  и  $\sigma_{tv}$  – собственное среднеквадратическое отклонение погрешности точки ветвления.

Рассмотрим сначала вариант, когда собственные погрешности элементов добавляются к выходным сигналам этих элементов. В этом случае среднеквадратическое отклонение погрешности входного сигнала нелинейного преобразователя вычисляется по формуле  $\sigma = \sigma_{own}/|\varphi'(A)|$ , где  $\sigma_{own} = \sqrt{\sigma_1^2 - \sigma_{\varphi}^2}$ ,  $\sigma_1$  – общая погрешность выходного сигнала нелинейного преобразователя. Для точки ветвления среднеквадратическое отклонение погрешности входного сигнала определяется как  $\min\{D_i\}_{i=1}^k - \sigma_{tv}^2$ .

Среднеквадратическое отклонение погрешности  $\sigma^{part}$ , которое придет к входу сумматора  $\Sigma_2$  при прямом функционировании сети, начиная от выходного сигнала сумматора  $\Sigma_1$  (рис.9), равно

$$\sigma^{part} = \sqrt{\sigma_{\Sigma_1}^2 \cdot \varphi'(A)^2 + \sigma_{\varphi}^2 + \sigma_{tv}^2}.$$

Среднеквадратические отклонения погрешностей  $\sigma_i^{part}$  придут к каждому входу сумматора  $\Sigma_2$ . Если сумма квадратов среднеквадратических отклонений  $\sigma_i^{part}$  с коэффициентами  $\alpha_i$  меньше квадрата среднеквадратического отклонения погрешности выходного сигнала сумматора ( $\sum_{i=1}^n \alpha_i^2 \cdot (\sigma_i^{part})^2 < \sigma^2$ ), то вычисляем разность  $\sum_{i=1}^n \alpha_i^2 \cdot (\sigma_i^{part})^2 - \sigma^2$ . Оставшуюся часть квадрата среднеквадратического отклонения погрешности выходного сигнала сумматора  $\sigma$  распределяем равномерно по всем входам, чтобы среднеквадратические отклонения погрешностей входов превышали собственные среднеквадратические отклонения погрешностей элементов на одну и ту же величину  $\xi$ . Тогда получаем следующую формулу

$$\sum_{i=1}^n \alpha_i^2 \cdot ((\sigma_i^{part})^2 + \xi^2) = \sigma^2 \Rightarrow \xi = \sqrt{(\sigma^2 - \sum_{i=1}^n \alpha_i^2 \cdot (\sigma_i^{part})^2) / \sum_{i=1}^n \alpha_i^2}.$$

Среднеквадратические отклонения погрешностей по входам сумматора будут равны  $\sigma_i = \sqrt{(\sigma_i^{part})^2 + \xi^2}$ .

Пусть теперь собственные погрешности элементов добавляются к входным сигналам этих элементов. В этом случае среднеквадратическое отклонение погрешности входного сигнала нелинейного преобразователя вычисляется по формуле  $\sigma = \sqrt{\sigma_1^2 / \varphi'(A)^2 - \sigma_{\varphi}^2}$ , где  $\sigma_1$  – погрешность выходного сигнала нелинейного преобразователя. Для точки ветвления среднеквадратическое отклонение погрешности входного сигнала определяется как было указано выше.

Среднеквадратическое отклонение погрешности  $\sigma^{part}$  в этом случае равно

$$\sigma^{part} = \sqrt{\left(\sum_{i=1}^n \alpha_i^2 \cdot (\sigma_{\Sigma_1}^i)^2 + \sigma_{\varphi}^2\right) \cdot |\varphi'(A)|^2 + \sigma_{tv}^2}.$$

Величины  $\xi$  для вычисления среднеквадратических отклонений погрешностей входных сигналов  $\sigma_i$  вычисляются как было показано выше.

Промежуточные среднеквадратические отклонения погрешностей  $\sigma^{part}$  можно вычислять как для участков сети, так и для сети в целом.

Таким образом, мы получили формулы для вычисления среднеквадратических отклонений погрешностей сигналов нейронной сети с собственными погрешностями элементов.

### Типы погрешностей

В методе обратного распространения точности приведены формулы для вычисления погрешностей сигналов сети. Эти формулы рассчитаны для сигналов, полученных при прямом функционировании сети с одним примером из обучающей выборки в качестве входных сигналов сети. Вообще говоря, допустимые погрешности сигналов зависят от вида входных сигналов сети. Исходя из этого, для метода обратного распространения точности можно выделить четыре типа допустимых погрешностей:

- 1) погрешности, вычисленные для одного примера;
- 2) погрешности, вычисленные для всей обучающей выборки;
- 3) погрешности, вычисленные для примеров, компоненты которых принадлежат области  $A_i \leq x_i \leq B_i, i = 1, n$ , где  $n$  – размерность области,  $A_i, B_i$  – действительные числа;

4) погрешности, вычисленные для примеров, компоненты которых принадлежат области  $\sum_{i=1}^n x_i^2 \leq R^2$ , где  $R$  – действительное число.

Погрешности первого типа вычисляются по формулам, описанным в методе обратного распространения точности.

Для того, чтобы вычислить погрешности второго типа, вычисляем погрешности для каждого примера из обучающей выборки. Затем в качестве допустимой погрешности для каждого элемента сети выбирается минимум допустимых погрешностей этого элемента, вычисленных для каждого примера из обучающей выборки.

Рассмотрим, как вычисляются допустимые погрешности третьего и четвертого типов. В формулах для вычисления допустимых погрешностей входной сигнал используется только у нелинейного преобразователя. Допустимые погрешности остальных элементов сети от входных сигналов не зависят. Поэтому для вычисления этих типов погрешностей следует выяснить, какие сигналы будут входными для нелинейных преобразователей, если входные сигналы сети принадлежат области, которая является прямоугольником или шаром.

Для начала рассмотрим допустимые погрешности третьего типа, то есть те допустимые погрешности элементов сети, которые получаются при входных сигналах, принадлежащих прямоугольной области. Нам известны интервалы, в которых изменяются входные сигналы сети. Требуется вычислить интервалы для входных сигналов каждого элемента сети. Будем вычислять их следующим образом. При прохождении интервалов через сумматор концы интервалов соответствующих входов умножаются на веса синапсов  $\alpha_i$  и затем складываются. Предположим, что функция активации нелинейного преобразователя непрерывна и монотонна. Тогда в качестве концов интервала его выходного сигнала берутся значения характеристической функции нелинейного преобразователя от концов интервала входного сигнала. Точка ветвления посылает приходящий к ней интервал на входы следующих элементов.

Таким образом, для каждого элемента сети мы можем вычислить интервал, в котором изменяются его входные сигналы. Нас интересуют интервалы, в которых изменяются входные сигналы нелинейных преобразователей. Для того, чтобы вычислить допустимые погрешности входного сигнала нелинейного преобразователя, необходимо вычислить максимум производной функции активации нелинейного преобразователя на интервале изменения его входных сигналов и затем разделить на эту величину допустимую погрешность выходного сигнала нелинейного преобразователя.

Таким образом вычисляются допустимые погрешности сигналов сети для прямоугольной области входных сигналов сети.

Рассмотрим пример, в котором будем вычислять погрешности третьего типа. Воспользуемся нейросетью, изображенной на рис.9. Нейросеть имеет такие же веса синапсов, но входные сигналы принадлежат прямоугольной области:

$$I_1=[-2,6]; I_2=[3,11]; I_3=[1.5,4]; I_4=[2,7]; I_5=[4,10]; I_6=[-5,5];$$

$$I_7=[1,8]; I_8=[-4,6].$$

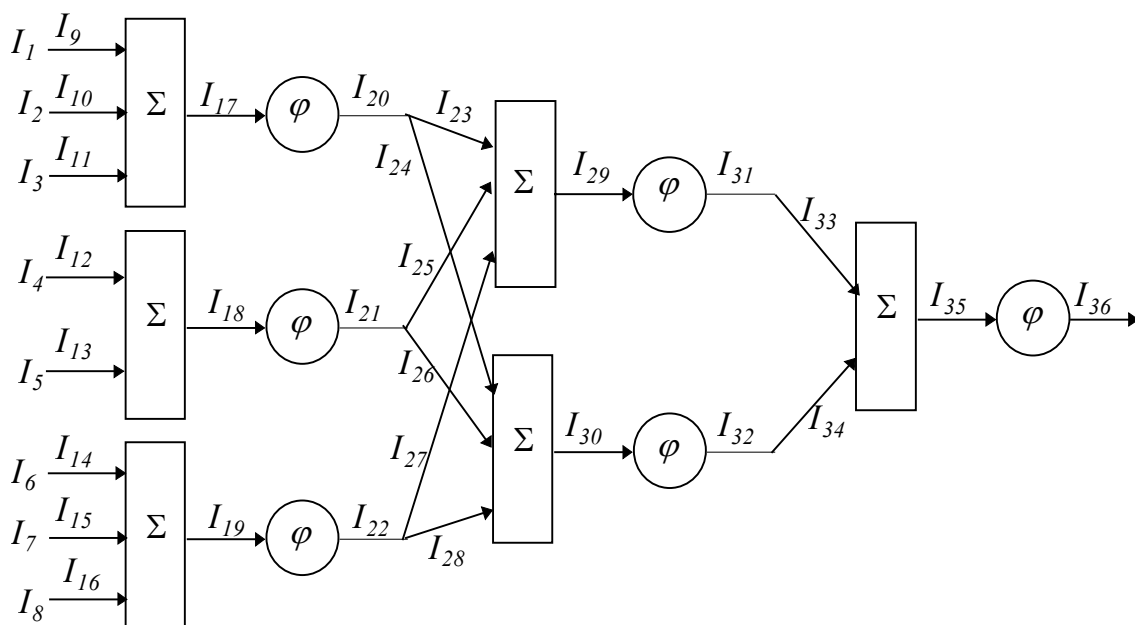


Рис. 12

Сигналы сети изменяются в следующих интервалах:

$$\begin{aligned}
 &I_9=[-0.6,3.6]; I_{10}=[2.85,10.45]; I_{11}=[0.96,2.56]; I_{12}=[0.56,1.96]; \\
 &I_{13}=[3,7.5]; I_{14}=[-2.7,2.7]; I_{15}=[0.41,3.28]; I_{16}=[-1.12,1.68]; \\
 &I_{17}=[3.21,16.61]; I_{18}=[3.56,9.46]; I_{19}=[-3.41,7.66]; I_{20}=[0.62,0.89]; \\
 &I_{21}=[0.64,0.83]; I_{22}=[-0.63,0.79]; I_{23}=[0.07,0.1]; I_{24}=[0.45,0.65]; \\
 &I_{25}=[0.39,0.51]; I_{26}=[0.34,0.44]; I_{27}=[-0.2,0.26]; I_{28}=[-0.55,0.69]; \\
 &I_{29}=[0.26,0.86]; I_{30}=[0.24,1.78]; I_{31}=[0.12,0.3]; I_{32}=[0.1,0.47]; \\
 &I_{33}=[0.04,0.1]; I_{34}=[0.07,0.32]; I_{35}=[0.11,0.42]; I_{36}=[0.05,0.17].
 \end{aligned}$$

Мы можем вычислить максимум производной функции активации на интервале изменения входного сигнала нелинейного преобразователя.

$$\max \varphi'_{[3.21,16.61]} = 0.07; \max \varphi'_{[3.56,9.46]} = 0.06;$$

$$\max \varphi'_{[-3.41,7.66]} = 0.07; \max \varphi'_{[0.26,0.86]} = 0.39;$$

$$\max \varphi'_{[0.24,1.78]} = 0.4; \max \varphi'_{[0.11,0.42]} = 0.45.$$

Зная эти величины, можно вычислить допустимые погрешности третьего типа.

Выясним теперь, как вычисляются допустимые погрешности сигналов четвертого типа, то есть погрешности, получающиеся, когда область входных сигналов сети является шаром.

Рассуждения, приведенные выше для допустимых погрешностей третьего типа, справедливы и для допустимых погрешностей четвертого типа. Отличие состоит в том, что нам необходимо "развернуть" шаровую область таким образом, чтобы получить интервалы, в которых изменяются входные сигналы элементов.

Рассмотрим для этого квадраты выходных сигналов сумматоров входного слоя сети. Используя неравенство Коши, получаем

$$A^2 = (\sum_{i=1}^k \alpha_i x_i)^2 \leq \sum_{i=1}^k \alpha_i^2 \cdot \sum_{i=1}^k x_i^2 \leq R^2 \cdot \sum_{i=1}^k \alpha_i^2 \Rightarrow |A| \leq R \sqrt{\sum_{i=1}^k \alpha_i^2},$$

где  $k$  – число входных сигналов сумматора. Получили интервалы, в которых изменяются выходные сигналы сумматоров входного слоя нейронной сети. Используя эти интервалы, можем вычислить интервалы, в которых изменяются входные сигналы элементов сети. Затем, как уже было описано выше, вычисляем допустимые погрешности входных сигналов нелинейных преобразователей.

## 8. Обсуждение

Как метод обратного распространения точности, так и метод обратного распространения среднеквадратических отклонений погрешностей можно применять к сетям не только слоистой структуры, но также к циклическим и полносвязным сетям. Рассматривая такт функционирования сети как слой, "разворачиваем" циклические и полносвязные сети в сети слоистой структуры. Вычисляем допустимые погрешности (среднеквадратические отклонения погрешностей) для сигналов стандартных нейронов каждого слоя. Затем "сворачиваем" слоистую сеть в исходную. Так как каждый слой полученной сети на самом деле является тактом функционирования, то для каждого сигнала сети на разных тактах получаем разные допустимые погрешности

(среднеквадратические отклонения погрешностей). В качестве допустимой погрешности (среднеквадратического отклонения погрешности) для каждого сигнала сети выбирается минимум этих величин по всем тактам.

Идея этих методов возникла при решении задачи бинаризации нейронной сети. Бинаризация состоит в построении такой сети, которая функционирует так же, как и исходная, но имеет веса синапсов, равные 0 или 1 (вариант: +1 или -1).

Но метод обратного распространения точности и метод обратного распространения среднеквадратических отклонений погрешностей сигналов сети интересен не только и не столько в приложении к задаче бинаризации. Их можно применять при решении ряда других задач. Например, вычислив допустимые погрешности (среднеквадратические отклонения погрешностей) для всей сети, можно выяснить, в каких пределах можно варьировать входные данные и сигналы на любом участке сети, чтобы вектор выходных сигналов при этом изменился не более, чем на заданную величину.

Метод обратного распространения точности для среднеквадратических оценок погрешности позволяет получать формулы для вычисления погрешностей сигналов сети, налагающие менее жесткие ограничения на величину погрешностей по сравнению с гарантированными интервальными оценками погрешностями. Если для гарантированных интервальных оценок при обратном прохождении слоев допустимые погрешности сигналов уменьшаются, то для среднеквадратических оценок погрешностей есть ситуации, когда погрешности увеличиваются от последнего слоя к первому. Если погрешности сигналов являются независимыми случайными величинами, то, как показано в примере, даже при больших погрешностях входных сигналов получают достаточно точные выходные сигналы.

## **9. Заключение.**

Таким образом, решение задачи вычисления допустимых погрешностей (среднеквадратических отклонений погрешностей) для каждого сигнала сети методом обратного распространения точности удивительно похоже на метод обратного распространения ошибки, но с другими правилами прохождения элементов. Метод позволяет формулировать требования к точности вычисления и реализации технических устройств, если известны требования к точности выходных сигналов сети.

## **Литература.**

1. Горбань А.Н. Обучение нейронных сетей. –М.: СП “ПараГраф”, 1990. – 159 с.
2. Нейроинформатика и ее приложения // Материалы 3 Всероссийского семинара, 6-8 октября 1995 г. Ч. 1 /Под редакцией А.Н.Горбаня. Отв. за выпуск Г.М.Цибульский; Красноярск: изд. КГТУ. – 1995. – 229 с.
3. Kimura T., Shima T. Synapse weight accuracy of analog neuro chip // Proceedings of International Joint Conference on Neural Networks. – Japan, Nagoya, October 25-29, 1993. – Vol.1. – P. 891-894.
4. Anguita D., Ridella S., Rovetta S. Limiting the effects of weight errors in feed forward networks using interval arithmetic // Proceedings of International Conference on Neural Networks (ICNN'96). – USA, Washington, June 3-6, 1996. – Vol.1. – P. 414-417.
5. Edwards P., Murray A. Modelling weight- and input-noise in MLP learning // Proceedings Of International Conference on Neural Networks (ICNN'96). – USA, Washington, June 3-6, 1996. – Vol.1. – P. 78-83.

6. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. Новосибирск: Наука. Сибирская издательская фирма РАН, 1996. – 276 с.
7. Боровков А.А. Теория вероятностей. М.: Наука, 1976. – 352 с.
8. Гнеденко Б.В. Курс теории вероятностей. М.: Наука, 1969.– 400 с.