ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ и МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 2 Ноябрь 1962 Декабрь № 6

К ВОПРОСУ О САМООБУЧЕНИИ В ПЕРСЕПТРОНЕ

В. М. ГЛУШКОВ

(*Киев*)

В предыдущей работе автора [1] было исследовано поведение одного класса дискретных персептронов (так называемых α-персептронов) в режиме *обучения*. Характерной чертой режима обучения является наличие учителя, которому известна правильная классификация изображений. В настоящей статье исследуются некоторые вопросы, связанные с поведением дискретных α-персептронов в режиме *самообучения*. В этом случае учитель отсутствует, а процессы самоорганизации, приводящие к изменению производимой персептроном классификации изображений, определяются введенной в схему персептрона положительной обратной связью.

Хорошо известно, что анализ поведения персептронов в режиме самообучения, сделанный Розенблаттом [2], весьма далек от какой бы то ни было математической строгости. Отсутствие строго доказанных предложений в этой области привело к тому, что некоторые авторы (особенно в публикациях научно-популярного характера) приписывают самообучению персептронов многие свойства, которыми оно в действительности не обладает и не может обладать. Для того чтобы хотя бы частично восполнить этот пробел, в настоящей статье предпринята попытка более строгого рассмотрения некоторых задач, связанных с изучением самообучения в персептронах. На основании такого рассмотрения нетрудно сделать ряд выводов, очерчивающих границы действительных возможностей, заложенных в самообучении персептронов. При дальнейшем изложении мы будем предполагать известными основные определения и результаты работы [1].

Рассмотрим дискретный симметричный α-персептрон, рассчитанный на распознавание двух образов P и Q. Условимся в качестве единого выходного сигнала персептрона рассматривать разность выходных сигналов сумматоров P-го и Q-го образов:

$$V_i(l) = U_i^P(l) - U_i^Q(l). \tag{1}$$

Здесь, как и в [1], индекс i пробегает все изображения (как P-го, так и Q-го образа), а l — любая последовательность изображений, показанная персептрону в процессе его самообучения.

Как и в случае обучения, нетрудно показать, что функционирование симметричного персептрона в режиме самообучения определяется суммой $a + b$ констант поощрения и штрафа, а не этими константами, рассматриваемыми отдельно. Имея в виду также возможность произвольного изменения масштабов, допустимо, не нарушая общности, предположить, что

$a = 1$, а $b = 0$. В дальнейшем мы всегда будем исходить из этого допущения.

Обозначая, как и в [1], через $\| T_{ij} \|$ характеристическую матрицу персептрона и вспоминая определение α-закона поощрения (см. [1]), легко придем к формуле

$$V_i(lj) = V_i(l) + T_{ij} \operatorname{sign} V_j(l). \qquad (2)$$

Формула (2) имеет место для любой пары изображений i, j и для любой последовательности изображений l. Функция $\operatorname{sign} x$, как обычно, для положительных значений x принимается равной плюс единице, а для отрицательных значений x — минус единице. Ясно, что в случае нулевого значения величины $V_j(l)$, по точному смыслу закона поощрения (положительной обратной связи), величина $\operatorname{sign} V_j(l)$ в формуле (2) должна быть не определена. С целью избежать неопределенности в дальнейшем мы будем по определению считать нуль положительной величиной, так что $\operatorname{sign} 0 = + 1$.

Имея в виду указанное видоизменение в определении функции $\operatorname{sign} x$, мы будем рассматривать формулу (2) как способ рекуррентного задания вектора $\mathbf{V}(l) = (V_1(l), \ldots, V_m(l))$, определяющего выходные сигналы персептрона под действием любого изображения $j = 1, \ldots, m$ после подачи на вход персептрона последовательности изображений l. Начальное значение этого вектора $\mathbf{V}(0) = (V_1(0), \ldots, V_m(0))$ предполагается при этом заданным. Изображение j причисляется персептроном к образу P или к образу Q в соответствии с тем, положительна или отрицательна соответствующая компонента $V_j(l)$ рассматриваемого вектора (напомним, что нуль, по принятому соглашению, считается положительным числом).

Поскольку все величины T_{ij} являются целыми (и к тому же неотрицательными) числами, задача расчета персептрона в режиме самообучения сводится по существу к задаче случайного блуждания по дискретной решетке в пространстве с числом измерений $\leqslant m$. Предполагая, что показ изображений в процессе самообучения производится по схеме независимых испытаний, легко заметить, что вероятности переходов из любой точки такой решетки определяются лишь набором знаков координат этой точки.

Набор знаков координат любой точки решетки определяет, как нетрудно понять, также и классификацию изображений, производимую персептроном, который имеет в качестве вектора $\mathbf{V}(l)$ своих выходных сигналов радиус-вектор этой точки. С точки зрения теории персептрона представляет интерес прежде всего предельное распределение знаков координат вектора $\mathbf{V}(l)$ при неограниченном увеличении длины обучающей последовательности l. Проведенные выше рассмотрения показывают, что требуемое распределение получается из предельного распределения для марковской цепи, соответствующей описанному выше блужданию по дискретной решетке.

Поскольку указанная цепь имеет бесконечное число состояний, нахождение предельного распределения в общем случае является достаточно сложным. Можно, однако, указать ряд случаев, когда задача нахождения предельного распределения легко сводится к исследованию марковской цепи с конечным числом состояний.

Рассмотрим в качестве примера дискретный симметричный α-персептрон A, рассчитанный на распознавание $2n$ изображений, первые n из которых принадлежат образу P, а последние n — образу Q. Пусть, далее, для элементов характеристической матрицы персептрона A имеют место соотношения $T_{ij} = a > 0$, если i и j принадлежат одному и тому же образу и $T_{ij} = 0$, если i и j принадлежат различным образам. Ввиду теоремы 4 из [1], рассматриваемый персептрон обладает абсолютной способностью к экстраполяции и, следовательно, наилучшим образом ведет себя в режиме обучения (обучается правильному распознаванию в результате показа хотя бы одного изображения из каждого образа). Предположим, что начальными условиями будут условия $V_i(0) = b > 0$ для $i = 1, 2, \ldots$ $\ldots, m \leqslant n$ и $V_j(0) = -b$ для $j = m + 1, \ldots, n, \ldots, 2n$.

Из формулы (2) непосредственно следует, что $V_j(l) < 0$ для любой последовательности l при всех $j = n + 1, n + 2, \ldots, 2n$. Что же касается остальных компонент, то они будут выражаться формулами $V_i(l) = b + ka$ для $i = 1, \ldots, m$ и $V_i(l) = -b + ka$ для $i = m + 1, \ldots, n$, где k — разность между числом появления изображений, которым соответствуют положительные компоненты $V_i(l')$, и числом изображений, которым соответствуют отрицательные компоненты $V_i(l')$ (l' — соответствующая подпоследовательность последовательности l).

Предположим, что процесс самообучения совершается по схеме независимых испытаний с равными вероятностями появления всех изображений. Поскольку показ изображений одного образа в рассматриваемом случае никак не влияет на распознавание изображений второго образа, можно, не нарушая общности, предполагать, что в процессе самообучения участвуют только изображения образа P (изображениям образа Q всегда отвечает отрицательный выходной сигнал независимо от того, включаются они в процесс самообучения или нет).

Допустим, что b не делится на a, и обозначим через t целое число $[b/a] + 1$. Нетрудно понять, что для изучения функционирования персептрона A представляют интерес лишь те значения параметра k, которые заключены в замкнутом интервале $[-t, t]$. Действительно, если в процессе самообучения хотя бы раз величина k достигнет значения t, то в дальнейшем в силу формулы (2) она может только возрастать, причем персептрон, начиная с этого момента, будет давать положительный выходной сигнал для всех изображений образа P (что соответствует правильной классификации). Аналогично, если параметр k примет значение $-t$, то в дальнейшем он может только убывать, а персептрон будет давать отрицательный выходной сигнал для всех изображений (что фактически означает отсутствие какой-либо классификации изображений, поскольку все изображения причисляются персептроном к одному и тому же образу).

Теперь, как легко видеть, предельное поведение персептрона A определяется марковской цепью с $2t + 1$ состояниями $k = -t, -t + 1, \ldots$ $\ldots, -1, 0, 1, \ldots, t - 1, t$. В силу принятого предположения о вероятностях появления изображений в процессе самообучения, для любого k, отличного от t или $-t$, вероятность перехода в состояние $k + 1$ равна m/n, а вероятность перехода в состояние $k - 1$ равна $(n - m)/n$. Из состояния t (равным образом как и из состояния $-t$) возможен переход только в это же самое состояние, поскольку с точки зрения функционирования персептрона все состояния при $k > t$ (соответственно при $k < -t$) не отличаются от состояния $k = t$ (соответственно от состояния $k = -t$).

Вводя обозначения $p = m/n$ и $q = (n-m)/n$, получим для рассматриваемой марковской цепи матрицу переходных вероятностей

$$\mathbf{M} = \begin{Vmatrix} 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ p & 0 & q & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & p & 0 & q & \ldots & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & p & 0 & q & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & p & 0 & q \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \end{Vmatrix}$$

Эта матрица имеет единицу в качестве своего двукратного собственного значения. Вероятности перехода цепи в состояния t и $-t$ равны предельным переходным вероятностям $p_{t+1,1}^{\infty}$ и $p_{t+1,2t+1}^{\infty}$. Для вероятности $p_{t+1,i}^{\infty}$ по известной формуле Перрона получаем выражение

$$p_{t+1,i}^{\infty} - \lim_{s\to\infty} \frac{d}{d\lambda} \frac{\lambda^s M_{i,t+1}(\lambda)}{\psi(\lambda)} \bigg|_{\lambda=1}. \tag{3}$$

Легко видеть, что $M_{i,t+1}(\lambda)$ для i, отличного от 1 и от $2t+1$, делится на $(\lambda-1)^2$ и потому $p_{t+1,i}^{\infty} = 0$ для этих значений i. Для $i = 1$ имеем $M_{1,t+1}(\lambda) = (\lambda-1)N_1(\lambda)$, а для $i = 2t+1$ $M_{2t+1,t+1}(\lambda) = (\lambda-1)N_2(\lambda)$, где $N_1(\lambda) = p^tQ(\lambda)$, $N_2(\lambda) = q^tR(\lambda)$.

Из формулы (3) теперь легко получаем

$$p_{t+1,1}^{\infty} = cp^t, \qquad p_{t+1,2t+1}^{\infty} = cq^t.$$

Поскольку все остальные предельные переходные вероятности в $(t+1)$-й строке равны нулю, из условия стохастичности матрицы предельных переходных вероятностей находим значение c: $c = 1/(p^t + q^t)$. Тем самым нами доказано следующее предложение:

При неограниченном продолжении процесса самообучения описанный выше персептрон A с вероятностью $p^t/(p^t + q^t)$ устанавливает правильную классификацию изображений и с вероятностью $q^t/(p^t + q^t)$ относит все изображения к одному и тому же образу.

Рассмотренный пример, как легко заметит внимательный читатель, строго говоря, не может быть осуществлен в реальном персептроне, исключая тривиальные случаи $m = n$, $p = 1$, $q = 0$ и $m = 0$, $p = 0$, $q = 1$. Причина заключается в том, что при сделанных предположениях относительно характеристической матрицы все изображения одного и того же образа возбуждают одно и то же множество нейронов. Поэтому выходные сигналы, индуцируемые изображениями одного и того же образа, всегда, в том числе и в начальный момент, должны быть равны между собой.

Нетрудно, однако, заметить, что, полагая $T_{ii} = a + \delta$ для всех $i = 1, \ldots, 2n$ $(\delta > 0)$ и не изменяя остальных элементов характеристической матрицы, мы получаем возможность удовлетворить введенным в примере начальным условиям. Вместе с тем, если δ существенно меньше, чем a, а t относительно невелико, то описанное в примере поведение персептрона будет служить хорошим приближением для его истинного поведения.

Рассмотрим теперь полный дискретный α-персептрон B с $(1, 1, 1)$-нейронами, с квадратной $(n \times n)$-сетчаткой, рассчитанный на распознавание двух образов P и Q. Образ P состоит из n горизонтальных линий, а образ Q — из n вертикальных линий; каждая из этих $2n$ линий составляет отдельное изображение. Как уже отмечалось в работе [1], персептрон B можно рассматривать как наиболее характерный представитель класса персептронов со случайными связями нейронов с сетчаткой. Согласно теореме 1 и следующему за ней следствию из работы Розенблатта [2], такие персептроны, будучи построены на непрерывных нейтронах, с вероятностью сколь угодно близкой к единице, при самообучении должны стремиться к состоянию, в котором все изображения относятся к одному и тому же образу. Покажем, что для персептрона B подобное утверждение не имеет места.

Легко видеть, что для персептрона B начальные условия могут быть выбраны любыми. *Нижней границей модулей начальных условий* будем называть наименьшее из чисел $|V_i(0)|$ $(i = 1, 2, \ldots, 2n)$. При сделанных предположениях имеет место следующий результат:

Теорема 1. *Для любого сколь угодно малого положительного числа ε найдется такое число s, что в случае, когда нижняя граница модулей начальных условий превосходит s, персептрон B в режиме самообучения (с равновероятным появлением всех изображений) с вероятностью $p \geqslant 1 - \varepsilon$ сохраняет начальную классификацию изображений.*

Доказательство. Обозначим через N длину обучающей последовательности l, а через ν_i — число появлений i-го изображения $(i = 1, 2, \ldots, 2n)$ в этой последовательности. Пусть $V_i(0) = x_i$, $(i = 1, 2, \ldots, 2n)$, K_i — множество всех индексов j (изображений), относящихся к противоположному по сравнению с i образу и таких, что знак x_j совпадает со знаком x_i, L_i — множество всех индексов j, относящихся к противоположному по сравнению с i образу и таких, что знак x_j противоположен знаку x_i (нуль, как и раньше, считается при этом положительным числом).

Как было показано в [1], произвольный элемент T_{ij} характеристической матрицы персептрона B равняется $n^2(n-1)$, 0 или $(n-1)^2$ в зависимости от того, совпадают ли индексы i и j, не совпадают, но относятся к одному и тому же образу, или не совпадают и относятся к различным образам. Используя это обстоятельство, с помощью формулы (2) легко получаем, что исходная классификация изображений сохранится в процессе самообучения, если для всех $N = 1, 2, \ldots$ будут выполняться неравенства

$$n^2(n-1)\nu_i + (n-1)^2\Big(\sum_{j \in K_i} \nu_j - \sum_{j \in L_i} \nu_j\Big) + |x_i| > 0 \qquad (i = 1, 2, \ldots, 2n),$$

и тем более, если будут выполнены неравенства

$$n^2(n-1)\nu_i - (n-1)^2 \sum_{j \in K_i \cup L_i} \nu_j + x > 0, \tag{4}$$

где через x обозначено минимальное из чисел $|x_i|$ $(i = 1, 2, \ldots, 2n.)$

В свою очередь, нетрудно проверить, что неравенства (4) выполняются, если выполнены неравенства

$$\left|\frac{\nu_i}{N}-\frac{1}{2n}\right|<\frac{1}{4n^2}\left(1+\frac{2x}{N(n-1)}\right) \qquad (i=1, 2, \ldots, 2n), \tag{5}$$

Величины $z_1=\nu_1/N-\frac{1}{2n}, \ldots, z_{2n-1}=\nu_{2n-1}/N-\frac{1}{2n}$ распределены по закону, асимптотически стремящемуся к нормальному закону с плотностью вероятности вида

$$a\, N^{n-1/2}e^{-NQ(z_1,\ldots,\, z_{2n-1})},$$

где a — некоторая положительная константа, а Q — положительно определенная квадратичная форма. Легко показать также, что неравенства (5) выполняются, если выполнены неравенства

$$\left|\frac{\nu_i}{N}-\frac{1}{2n}\right|<\frac{1}{8n^3} \qquad (i=1, 2, \ldots, 2n-1). \tag{6}$$

Для вероятности $q(N)$ неудовлетворения хотя бы одного из неравенств (6) получается оценка сверху вида

$$q(N)<cN^{n-1/2}\left(2\int\limits_d^\infty e^{-bNx^2}dx\right)^{2n-1}<cN^{n-1/2}\left(-\frac{1}{bNd}\int\limits_d^\infty e^{-bNx^2}d(-bNx^2)\right)^{2n-1}=\frac{f}{N^{n-1/2}}\,e^{-gN},$$

где c, d, f и g — положительные величины, не зависящие от N ($d=1/8n^3$, $f=c/(bd)^{2n-1}$, $g=(2n-1)\,bd^2$). Константы c и b выбираются таким образом, чтобы в результате использования формулы Муавра — Лапласа получить для вычисляемой вероятности приближение с избытком.

Вероятность невыполнения хотя бы одного из неравенств (6) для значений N от M до ∞ не превосходит суммы ряда

$$\sum_{N=M}^{\infty}\frac{f}{N^{n-1/2}}\,e^{-gN},$$

а эта последняя сумма, очевидно, меньше, чем

$$R(M)=\frac{f}{M^{n-1/2}}\,\frac{e^{-gM}}{1-e^{-g}}\,.$$

При $M\to\infty$ величина $R(M)$ стремится к нулю. Выберем M так, чтобы $R(M)<\varepsilon$.

Выбирая теперь $s=2\,(M-1)\,n^2\,(n-1)$, получим, что при $x>s$ неравенства (5) выполнены для всех значений $N=1, 2, \ldots, M-1$. В силу же выбора M, для всех остальных значений N неравенства (5) выполняются с вероятностью, большей, чем $1-\varepsilon$. Поскольку выполнение неравенств (5) для всех значений N от 1 до ∞ означает сохранение исходной классификации изображений, то теорема доказана.

Теорема 1 показывает, что при достаточно больших начальных значениях выходных сигналов для всех изображений рассмотренный персептрон фактически почти лишен способности не только самообучаться, но даже просто изменять изначально задаваемую ему классификацию изображений. Из доказательства теоремы легко усмотреть, что остающаяся при этом слабая способность к самоизменению имеет наибольшую величину в случае правильной исходной классификации. Иными словами, персептрон наименее склонен к сохранению именно правильного способа функционирования.

Переходя к рассмотрению β-закона поощрения, фиксируем произвольное число β, заключенное между нулем и единицей, и рассмотрим произвольный симметричный персептрон C с β-законом поощрения, характеристическая матрица которого диагональна, т. е., иными словами, имеет отличные от нуля элементы только на главной диагонали. Как показано в [1], таким свойством обладает симметричный персептрон C_1 с (1, 1, 1)-нейронами, рассчитанный на распознавание горизонтальных и вертикальных линий, у которого входы каждого нейрона подсоединяются к элементам сетчатки, расположенным на одной горизонтали или на одной вертикали.

В случае β-закона поощрения основное рекуррентное соотношение для определения выходных сигналов запишется в виде

$$V_i(lj) = (1-\beta)\,(V_i(l) + T_{ij}\operatorname{sign} V_j(l)). \tag{7}$$

Обозначения здесь точно такие же, как и в формуле (2), причем это соотношение (как и формула (2)) имеет место для произвольных дискретных симметричных персептронов. В случае персептронов с диагональной характеристической матрицей оба слагаемых в правой части формул (2) и (7) имеют всегда один и тот же знак (случай, когда второе слагаемое равно нулю, исключаем из рассмотрения). Отсюда непосредственно вытекает справедливость следующего предложения:

Т е о р е м а 2. *Дискретный симметричный персептрон C с диагональной характеристической матрицей полностью лишен способности к самообучению (т. е. сохраняет неизменной любую задаваемую ему исходную классификацию изображений) как в случае α-закона поощрения, так и в случае β-закона поощрения.*

Легко видеть также, что имеет место следующее предложение:

Т е о р е м а 3. *Никакой дискретный симметричный персептрон (как с α-, так и с β-законом поощрения), работая в режиме самообучения, не может изменить исходной классификации изображений, если эта классификация относит все изображения к одному и тому же образу.*

Нетрудно заметить, что полученные нами результаты можно рассматривать как контрпримеры к результатам Розенблатта [2] в той мере, в какой примененные им рассуждения относятся не только к непрерывным, но и к дискретным нейронам. Во всяком случае, результаты настоящей статьи свидетельствуют о том, что асимптотическое поведение персептронов в режиме самообучения гораздо более сложно и требует значительно более тонких приемов для своего изучения по сравнению с приемами чисто качественного характера, употребляемыми Розенблаттом [2].

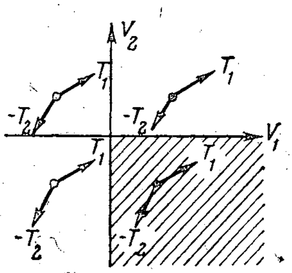
Для наглядного представления особенностей поведения персептронов в режиме самообучения по сравнению с режимом обучения рассмотрим случай, когда число изображений равно двум (каждый образ состоит из одного единственного изображения). Этот случай допускает простую графическую интерпретацию.

Предварительно заметим, что в случае наличия двух образов (но при произвольном числе изображений) функционирование персептрона в режиме обучения, так же как и в режиме самообучения, удобно характеризовать вектором с компонентами V_i (l) (см. выше). Основное рекуррентное соотношение для этих компонент будет иметь, очевидно, следующий вид:

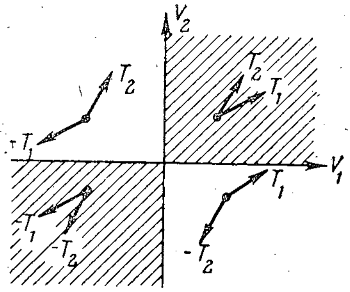
$$V_i(lj) = V_i(l) \pm T_{ij}. \tag{8}$$

Это соотношение справедливо для любой пары изображений i, j и для любой обучающей последовательности l. Второе слагаемое в правой части берется со знаком плюс, если изображению j в правильной классификации соответствует положительный выходной сигнал, и со знаком минус, если соответствующий выходной сигнал должен быть отрицательным.

Рассмотрим дискретный симметричный персептрон с α-законом поощрения, характеристической матрицей которого служит матрица $\mathrm{T} = \left\|\begin{smallmatrix} a, b \\ b, a \end{smallmatrix}\right\|$, где $a > b > 0$.



Фиг. 1



Фиг. 2

Предположим, что при правильной классификации первое изображение должно индуцировать положительный выходной сигнал, а второе изображение — отрицательный выходной сигнал. Откладывая по горизонтальной оси координату V_1 (l), а по вертикальной оси — координату V_2 (l), мы отнесем каждому вектору $(V_1$ (l), V_2 $(l))$ некоторую точку плоскости. Выбирая по одной точке в каждом квадранте, мы получим наглядное представление действия формулы (8) на фиг. 1.

На этой фигуре через T_1 обозначен вектор (a, b), а через T_2 — вектор (b, a). Характерной особенностью режима обучения является то, что направления векторов (определяющих случайные блуждания точки по решетке) не зависят от расположения точек на плоскости. Равнодействующая этих векторов всегда направлена в сторону того квадранта, в котором знаки координат (выходных сигналов персептрона) соответствуют правильной классификации изображений (в данном случае таким квадрантом является заштрихованный — четвертый — квадрант).

В случае режима самообучения интерпретация соответствующей формулы 2 дается фиг. 2. В отличие от предыдущего случая, направления векторов, определяющих случайное блуждание, будут различными в различных квадрантах.

Обозначения векторов — такие же, как и на фиг. 1.

Нетрудно заметить качественные отличия ситуации, изображенной на фиг. 2, от ситуации на фиг. 1. Прежде всего, первый и третий квадранты (на фиг. 2 они заштрихованы) теперь обладают ловушечным свойством: точка, попавшая в процессе случайного блуждания в один из этих квадрантов, уже не может никогда выбраться из него.

Попадание в эти квадранты означает фактически отсутствие какой-либо классификации (оба изображения причисляются к одному и тому же образу). Вместе с тем из квадранта, соответствующего правильной классификации (четвертый квадрант), равно как и из квадранта, соответствующего правильной классификации с точностью до знака выходного сигнала (третий квадрант), всегда имеется ненулевая вероятность выхода в соседние квадранты.

Рассматривая возникшую ситуацию в чисто качественном плане, подобно тому, как это делает Розенблатт [2], мы должны были бы прийти к выводу, что изучаемый нами персептрон асимптотически стремится к состоянию, в котором на все изображения даются выходные сигналы одного и того же знака (отсутствие какой-либо классификации). Более точное рассмотрение (повторяющее выкладки, произведенные при доказательстве теоремы 1) приводит, однако, к совершенно другому выводу: как и в случае теоремы 1, при достаточном удалении начальной точки от границ квадранта вероятность продолжения случайного блуждания без выхода из этого квадранта во все последующие моменты времени (вплоть до бесконечности) может быть сделана сколь угодно близкой к единице.

Тем самым еще раз подтверждается опасность, возникающая в том случае, когда общие выводы об асимптотическом поведении персептронов в режиме самообучения основываются на рассуждениях чисто качественного характера, не подтвержденных точными расчетами и оценками.

Поступила в редакцию
25.05.1962

Цитированная литература

1. В. М. Глушков. Теория обучения одного класса дискретных персептронов. Ж. вычисл. матем. и матем. физ., 1962, 2, № 2, 317—335.
2. F. Rosenblatt. Two theorems of statistical separability in the perceptron. Symp. Mecaniz. Thought Proc., Teddington, England, 1958, Paper 1—3, 3—32.