

## ВЫСОКОСИММЕТРИЧНЫЕ НЕЙРОННЫЕ СЕТИ ХОПФИЛДОВА ТИПА

Изучается множество неподвижных точек нейронной сети хопфилдова типа, матрица связей которой строится по правилу Хебба из высокосимметричного набора исходных образов (memorized patterns). Зависящие от внешнего параметра исходные образы трактуются как искаженные экземпляры вектора-эталона, которому хотели обучить сеть. Удаётся аналитически описать зависимость множества неподвижных точек сети от величины параметра искажения. Полученные результаты интерпретируются в терминах нейронных сетей и модели Изинга.

### 1. ВВЕДЕНИЕ

Рассматривается задача максимизации симметричной формы, квадратичной по *спиновым* переменным  $\sigma_i = \{\pm 1\}$ :

$$\sum_{i,j=1}^n J_{ij} \sigma_i \sigma_j \rightarrow \max, \quad (1)$$
$$J_{ij} = J_{ji}, \quad i, j = 1, 2, \dots, n.$$

Эта задача возникает в модели Изинга [1] и в физике поверхности [2], в оптимальном кодировании [3] и в теории нейронных сетей [1, 4], в факторном анализе [5, 6] и в теории оптимизации. В общем случае максимизируемый функционал имеет много локальных экстремумов и проблема состоит, во-первых, в эффективном отыскании его глобального максимума и, во-вторых, в конструктивном описании множества локальных экстремумов.

Используя физическую терминологию, будем называть *конфигурационными*  $n$ -мерные векторы

$$\vec{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n), \quad \sigma_i = \{\pm 1\}, \quad i = 1, 2, \dots, n,$$

задающие  $2^n$  конфигураций, среди которых ищется решение задачи (1). Конфигурационный вектор, решающий задачу (1), будем называть *основным состоянием*, а макси-

---

\*Институт физики высоких давлений РАН, Троицк, Московская обл., Россия

мизируемый функционал перепишем в векторном виде

$$F(\vec{\sigma}) = \sum_{i,j=1}^n J_{ij} \sigma_i \sigma_j = (\mathbf{J} \cdot \vec{\sigma}, \vec{\sigma}).$$

Очевидным образом множество локальных максимумов функционала не зависит от диагональных элементов матрицы  $\mathbf{J}$  (подробнее см. в [7]). Специальным выбором диагонали матрицу  $\mathbf{J}$  можно сделать положительно- (неотрицательно) определенной, а значит, представимой в виде

$$\mathbf{J} = \mathbf{S}^* \cdot \mathbf{S}, \quad (2)$$

где  $(p \times n)$ -матрица  $\mathbf{S}$  имеет вид

$$\mathbf{S} = \begin{pmatrix} s_1^{(1)} & s_2^{(1)} & \dots & s_p^{(1)} & s_{p+1}^{(1)} & \dots & s_n^{(1)} \\ s_1^{(2)} & s_2^{(2)} & \dots & s_p^{(2)} & s_{p+1}^{(2)} & \dots & s_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ s_1^{(p)} & s_2^{(p)} & \dots & s_p^{(p)} & s_{p+1}^{(p)} & \dots & s_n^{(p)} \end{pmatrix}, \quad (3)$$

а  $\mathbf{S}^* - (n \times p)$ -матрица, сопряженная к  $\mathbf{S}$ . Никак не ограничивая общности рассуждений, будем считать, что матрица  $\mathbf{J}$  представлена в виде (2). Число  $p$  строк прямоугольной матрицы  $\mathbf{S}$  определяется по матрице  $\mathbf{J}$  однозначно и не превосходит  $n$ :  $p \leq n$ . Прямоугольная матрица  $\mathbf{S}$  играет существенную роль в содержательной постановке задачи в теории нейронных сетей и факторном анализе. Интерпретация матрицы  $\mathbf{S}$  в терминах модели Изинга пока неизвестна.

**1.1. Модель Хопфилда нейронной сети.** В этой модели [1, 4, 8] в качестве матрицы связей используют так называемую матрицу Хебба, ее получают по рецепту (2), (3), но при том условии, что все элементы матрицы  $\mathbf{S}$  равны  $\{\pm 1\}$ <sup>1)</sup>.

Условимся, что всюду в дальнейшем греческие буквы используются для обозначения величин, принимающих значения  $\{\pm 1\}$ . При таком соглашении для обозначения матрицы (3) с единичными по модулю матричными элементами будем использовать символ  $\Xi$ , а ее  $n$ -мерные векторы-строки обозначим через  $\vec{\xi}^{(l)}$ :

$$\vec{\xi}^{(l)} = (\xi_1^{(l)}, \xi_2^{(l)}, \dots, \xi_n^{(l)}) \in \mathbb{R}^n, \quad l = 1, 2, \dots, p.$$

В теории нейронных сетей  $n$ -мерные конфигурационные векторы  $\{\vec{\xi}^{(l)}\}_1^p$  называются *исходными образами (memorized patterns)*, а  $(n \times n)$ -матрица  $\mathbf{H} = \Xi^* \cdot \Xi$  — *матрицей связей Хебба*. Важная роль, которая принадлежит матрице Хебба, обусловлена ее глубоким нейрофизиологическим обоснованием.

<sup>1)</sup>Обычно еще требуют, чтобы диагональные элементы матрицы Хебба были равны нулю: тогда множество неподвижных точек сети совпадает с множеством локальных максимумов функционала  $F(\vec{\sigma})$ . Однако последние не зависят от диагональных элементов матрицы, следовательно, при изучении локальных максимумов функционала диагональные элементы можно выбирать любым удобным для нас образом. Найденные локальные максимумы будут неподвижными точками сети, диагональ матрицы связей которой равна нулю.

Основная цель модели Хопфилда – *добиться того, чтобы исходные образы  $\vec{\xi}^{(l)}$  были единственными неподвижными точками сети*. Хорошо известно, однако, что, если ограничиться только матрицей  $\mathbf{H}$ , достичь этой цели практически невозможно. В сколько-нибудь общем случае нельзя даже гарантировать, что *все* исходные образы будут неподвижными точками сети, не говоря уж о том, что у сети почти неизбежно возникнут так называемые *фиктивные* неподвижные точки, весьма слабо коррелирующие с исходными образами  $\{\vec{\xi}^{(l)}\}_{l=1}^p$ . Чтобы все-таки навязать сети наперед заданный набор неподвижных точек  $\{\vec{\xi}^{(l)}\}_{l=1}^p$ , сеть *дообучают*: в соответствии с тем или иным рецептом поэтапно перестраивают элементы матрицы связей, добиваясь того, чтобы, во-первых, ушли фиктивные неподвижные точки и, во-вторых, чтобы возможно больше исходных образов стали неподвижными точками сети. При некоторых ограничениях на набор исходных образов такая стратегия приводит, как правило, к успеху [9]. Обзор современного состояния теории обучения нейронных сетей можно найти в [10].

Подчеркнем, что процесс дообучения сети в модели Хопфилда завершается переходом от матрицы связей Хебба  $\mathbf{H}$  с целочисленными коэффициентами к симметричной матрице связей  $\mathbf{J}$  общего вида<sup>2)</sup>. Эту матрицу  $\mathbf{J}$  тоже можно представить в виде (2), но теперь элементы матрицы  $\mathbf{S}$  (3) уже необязательно будут равны  $\{\pm 1\}$ . Следовательно,  $n$ -мерные строки такой матрицы  $\mathbf{S}$  уже не обязаны быть *конфигурационными* векторами и не будут исходными образами в классическом смысле этого понятия.

Итак, решение задачи (1)–(3) с симметричной матрицей связей *общего вида* представляет интерес для модели Хопфилда. Более того, абстрагируясь от понимания исходных образов как непременно конфигурационных векторов, можно утверждать, что любая нейронная сеть с симметричной матрицей связей есть сеть хопфилдова типа: ее матрицу связей можно представить в виде произведения (2). Это замечание оправдывает модель, которая изучается в настоящей работе.

**1.2. Предмет исследования.** Пусть  $p \leq n$  и  $(p \times n)$ -матрица  $\mathbf{S}$  исходных образов имеет вид

$$\mathbf{S} = \begin{pmatrix} 1-x & 1 & \dots & 1 & 1 & \dots & 1 \\ 1 & 1-x & \dots & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & \dots & 1-x & 1 & \dots & 1 \end{pmatrix}, \quad (4)$$

где  $x$  – произвольное вещественное число.

Матрицу связей, полученную из такой матрицы  $\mathbf{S}$  по рецепту (2), мы и будем изучать;

---

<sup>2)</sup>Некоторые алгоритмы дообучения сети приводят и к несимметричной матрице связей [11, 12]; мы ограничимся рассмотрением только симметричных матриц.

введем для нее специальное обозначение  $\mathbf{N}$ :

$$\mathbf{N} = \mathbf{S}^* \cdot \mathbf{S} = \begin{pmatrix} p-2x+x^2 & p-2x & \dots & p-2x & p-x & \dots & p-x \\ p-2x & p-2x+x^2 & \dots & p-2x & p-x & \dots & p-x \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ p-2x & p-2x & \dots & p-2x+x^2 & p-x & \dots & p-x \\ p-x & p-x & \dots & p-x & p & \dots & p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p-x & p-x & \dots & p-x & p & \dots & p \end{pmatrix}. \quad (5)$$

Эту модель естественно интерпретировать, отталкиваясь от устройства обобщенных “исходных образов” – строк матрицы (4). Можно предложить следующую содержательную трактовку: *сеть пытались обучить  $p$  предъявлениями ей вектора-эталона*

$$\vec{\varepsilon}(n) = \left( \underbrace{1, 1, \dots, 1}_n \right),$$

*однако в процесс обучения вкрались ошибки и на деле сеть обучили  $p$  искаженными экземплярами вектора-эталона. Величина искажения  $x$  у всех исходных образов*

$$\vec{s}^{(l)} = \left( 1, \dots, 1, \underbrace{1-x}_l, 1, \dots, 1 \right), \quad l = 1, 2, \dots, p, \quad (6)$$

*одинакова, и искажается всякий раз только один признак.*

Когда искажение  $x$  равно нулю, получаем сеть, обученную  $p$  экземплярами вектора-эталона  $\vec{\varepsilon}(n)$ . В этом случае, как известно, основным состоянием будет сам вектор-эталон  $\vec{\varepsilon}(n)$ , и других локальных максимумов у функционала не будет. Из соображений непрерывности понятно, что эта ситуация сохранится и для достаточно малых значений  $x$ . Однако с увеличением  $x$  основное состояние будет меняться. Для задачи (1), (5) удастся получить *аналитическую* зависимость основного состояния от параметра  $x$ . Чрезвычайно важным оказывается при этом то, что исходные образы  $\vec{s}^{(l)}$  образуют высокосимметричную группу векторов, все они одинаково коррелируют друг с другом. Термин “высокосимметричный” в заглавии статьи продиктован этим свойством исходных образов.

Статья устроена следующим образом. В разделе 2 для задачи (1), (5) приводятся основные полученные результаты. В разделе 3 матрица связей (5) естественным образом обобщается. В разделе 4 дается содержательная интерпретация полученных результатов в терминах нейронных сетей и модели Изинга. Математические доказательства, несущественные для понимания структуры решения, вынесены в приложения.

**1.3. Обозначения.** На протяжении всей статьи  $n$  – число спиновых переменных,  $p$  – число исходных образов,  $q = n - p$  – число неискажаемых координат вектора-эталона. Векторы снабжаются стрелкой сверху, причем для обозначения конфигурационных векторов используются греческие строчные буквы, а для обозначения векторов с вещественными координатами – строчные латинские. Векторы из  $n$ -мерного пространства  $\mathbb{R}^n$

нумеруются верхним индексом в скобках, а их координаты – нижним. Матрицы обозначаются полужирными прописными латинскими буквами.

Конфигурационный вектор, решающий задачу (1), называется основным состоянием. Конфигурационный вектор, коллинеарный биссектрисе главного угла пространства  $\mathbb{R}^k$ , обозначается через  $\vec{\varepsilon}(k)$ :

$$\vec{\varepsilon}(k) = \left( \underbrace{1, 1, \dots, 1}_k \right).$$

Вектор, искажениями которого являются  $p$  исходных образов  $\vec{s}^{(l)}$  (6), называется *вектором-эталоном*.

## 2. ОСНОВНАЯ МОДЕЛЬ

Использованный для решения задачи метод опирается на аналитические выражения для собственных значений и векторов матрицы связей  $\mathbf{N}$ ; приведем их (доказательство см. в приложении 1).

**ТЕОРЕМА 1.** *Отличные от нуля собственные значения  $a_i$  матрицы  $\mathbf{N}$  (5) и соответствующие собственные векторы  $\vec{h}^{(i)}$  имеют вид*

$$\begin{aligned} a_1 &= x^2 - 2px + pn, & \vec{h}^{(1)} &= \frac{1}{\sqrt{pa_1}}(p - x, \dots, p - x, p, \dots, p), \\ a_2 &= \dots = a_p = x^2, & \vec{h}^{(k)} &= (h_1^{(k)}, \dots, h_p^{(k)}, 0, \dots, 0), \end{aligned}$$

где

$$\sum_{i=1}^p h_i^{(k)} = 0, \quad \sum_{i=1}^p (h_i^{(k)})^2 = 1, \quad k = 2, 3, \dots, p.$$

Из результатов работ [13, 14] следует, что функционал  $F(\vec{\sigma}) = (\mathbf{N} \cdot \vec{\sigma}, \vec{\sigma})$  достигает локального максимума на таких конфигурационных векторах, у которых последние  $q$  координат равны между собой. Следовательно, “подозреваемый на экстремум” конфигурационный вектор имеет вид<sup>3)</sup>

$$\vec{\sigma}^* = \left( \underbrace{\sigma_1, \sigma_2, \dots, \sigma_p}_{\vec{\sigma}'}, \underbrace{1, \dots, 1}_q \right), \quad (7)$$

где через  $\vec{\sigma}'$  обозначен  $p$ -мерный вектор, образованный первыми  $p$  координатами конфигурационного вектора  $\vec{\sigma}^*$ . Используя теорему 1, имеем

$$F(\vec{\sigma}^*) = a_1 \cdot (\vec{\sigma}^*, \vec{h}^{(1)})^2 + a_2 \cdot \sum_{k=2}^p (\vec{\sigma}^*, \vec{h}^{(k)})^2,$$

<sup>3)</sup>Поскольку функционал  $F(\vec{\sigma}) = (\mathbf{N}\vec{\sigma}, \vec{\sigma})$  инвариантен относительно обращения знака у всех координат вектора  $\vec{\sigma}$ , любое его значение дважды вырожденно. В дальнейшем нам придется делать утверждения о числе локальных максимумов функционала  $F(\vec{\sigma})$ ; во избежание путаницы условимся, что локальные максимумы ищутся среди таких конфигурационных векторов, у которых  $n$ -я координата положительна. Вообще же проблема дважды вырожденности функционала  $F(\vec{\sigma})$  снимается добавлением к нему линейного члена  $h \sum_{i=1}^n \sigma_i$ , моделирующего в физике учет магнитного поля [15].

а учитывая выражение (7), после элементарных геометрических рассуждений и приведения подобных получаем

$$F(\vec{\sigma}^*) = p[x^2 - 2x \cos w(q + p \cos w) + (q + p \cos w)^2], \quad (8a)$$

где

$$\cos w = \frac{(\vec{\sigma}', \vec{\varepsilon}(p))}{\|\vec{\sigma}'\| \cdot \|\vec{\varepsilon}(p)\|} = \frac{\sum_{i=1}^p \sigma_i}{p} \quad (8б)$$

– косинус угла между  $p$ -мерным вектором  $\vec{\sigma}'$  и биссектрисой  $\vec{\varepsilon}(p)$  главного угла пространства  $\mathbb{R}^p$ .

В зависимости от того, сколько координат  $\sigma_1, \sigma_2, \dots, \sigma_p$  равны  $-1$ ,  $\cos w$  может принимать значения

$$\cos w_k = 1 - \frac{2k}{p}, \quad k = 0, 1, \dots, p. \quad (9)$$

Следовательно, конфигурационные векторы  $\vec{\sigma}^*$ , которые, возможно, доставляют экстремум, распределятся между  $(p+1)$  классами эквивалентности  $\Sigma_k$ , и на векторах  $\vec{\sigma}^*$  из одного класса функционал  $F(\vec{\sigma}^*)$  принимает одно и то же значение  $F_k(x)$ . Классы эквивалентности  $\Sigma_k$  нумеруются числом  $k$  отрицательных координат  $\sigma_1, \sigma_2, \dots, \sigma_p$ :  $k = 0, 1, \dots, p$ , а количество конфигурационных векторов в классе  $\Sigma_k$  равно  $C_p^k$ :

Класс векторов $\vec{\sigma}^*$	$\cos w_k$	Число векторов в классе
$\Sigma_0$	1	$1 = C_p^0$
$\Sigma_1$	$1 - \frac{2}{p}$	$p = C_p^1$
$\vdots$	$\vdots$	$\vdots$
$\Sigma_k$	$1 - \frac{2k}{p}$	$C_p^k$
$\vdots$	$\vdots$	$\vdots$
$\Sigma_{p-1}$	$-\left(1 - \frac{2}{p}\right)$	$p = C_p^{p-1}$
$\Sigma_p$	-1	$1 = C_p^p$

Суммируем эти рассуждения: при заданном значении  $x$  основное состояние системы достигается на конфигурационном векторе  $\vec{\sigma}^*$  (7), принадлежащем тому классу векторов  $\Sigma_k$  (10), для которого наибольшее значение принимает функция

$$F_k(x) = p[x^2 - 2x \cos w_k(q + p \cos w_k) + (q + p \cos w_k)^2], \quad (11)$$

где значения  $\cos w_k$  даются выражениями (9).

Для того чтобы отыскать основное состояние, необходимо сравнить между собой величины  $F_0(x), F_1(x), \dots, F_p(x)$  и определить среди них наибольшую. Поскольку квадратичный по  $x$  член у всех  $F_k(x)$  одинаков, его при сравнении можно опустить. Следовательно, для выяснения того, как основное состояние зависит от параметра  $x$ , необходимо изучить семейство прямых

$$L_k(x) = (q + p \cos w_k)^2 - 2x \cos w_k(q + p \cos w_k), \quad k = 0, 1, \dots, p. \quad (12)$$

Там, где прямая  $L_k(x)$  мажорирует все остальные прямые, основное состояние принадлежит классу  $\Sigma_k$ . Подробный анализ взаимного расположения прямых  $L_k(x)$  дан в приложении 2. Сформулируем здесь окончательный результат в виде отдельной теоремы.

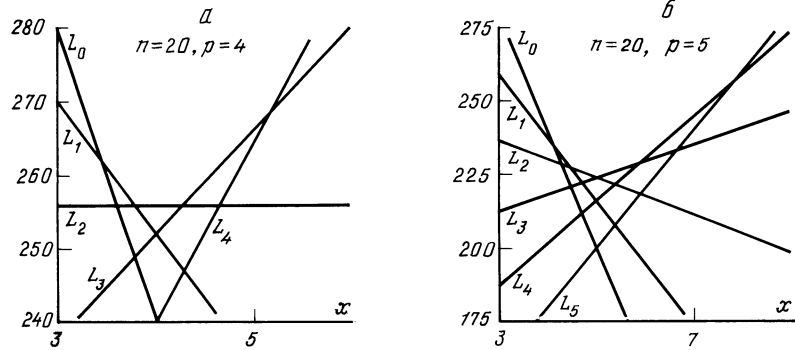


Рис. 1. Типичное поведение прямых  $L_k(x)$ ,  $k = 0, 1, \dots, p$ , в случае  $(p-1)/(n-1) < 1/3$  для четного (а) и нечетного (б) значений  $p$ .

ТЕОРЕМА 2. По мере изменения  $x$  от  $-\infty$  до  $\infty$  основное состояние последовательно принадлежит классам векторов  $\vec{\sigma}^*$ :

$$\Sigma_0, \Sigma_1, \dots, \Sigma_{k_{\max}}.$$

Перескок основного состояния из класса  $\Sigma_{k-1}$  в класс  $\Sigma_k$  происходит в точке  $x_k$ :

$$\begin{aligned} x_k &= \frac{p}{2} \left[ 1 + \frac{q}{q + p(\cos w_{k-1} + \cos w_k)} \right] = \\ &= p \frac{n - (2k - 1)}{n + p - 2(2k - 1)}, \quad k = 1, 2, \dots, k_{\max}. \end{aligned} \tag{13}$$

Если

$$\frac{p-1}{n-1} < \frac{1}{3},$$

то по этой схеме одна за другой произойдут все  $p$  перестроек основного состояния,  $p = k_{\max}$ . Если же

$$\frac{p-1}{n-1} > \frac{1}{3},$$

то последней будет перестройка основного состояния с номером

$$k_{\max} = \left[ \frac{n + p + 2}{4} \right]. \tag{14}$$

Других локальных максимумов у функционала  $F(\vec{\sigma})$  не будет.

Теорема 2 позволяет решить целый ряд практических задач. Например, при фиксированных  $n$  и  $p$  для любого наперед заданного  $x$  достаточно вычислить, между какими  $x_k$  и  $x_{k+1}$  оно попадает, чтобы знать, конфигурации какого класса  $\Sigma_k$  доставляют глобальный максимум функционалу  $F(\vec{\sigma})$ , а также каковы будут его значение и вырожденность основного состояния. Можно, наоборот, по заданному  $x$  подбирать такие  $n$  и  $p$ ,

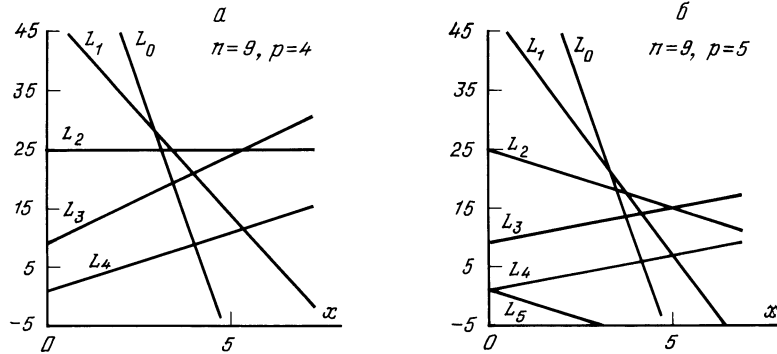


Рис. 2. Типичное поведение прямых  $L_k(x)$ ,  $k = 0, 1, \dots, p$ , в случае  $(p-1)/(n-1) > 1/3$  для четного (а) и нечетного (б) значений  $p$ .

чтобы основное состояние системы принадлежало к наперед заданному классу  $\Sigma_k$ . Мы не будем подробно останавливаться на подобных естественных приложениях теоремы 2.

На рис. 1 и 2 даны типичные примеры взаимного расположения прямых  $L_k(x)$  для четного (а) и нечетного (б) значений  $p$ ; рис. 1 отвечает случаю

$$\frac{p-1}{n-1} < \frac{1}{3},$$

рис. 2 – случаю

$$\frac{p-1}{n-1} > \frac{1}{3}.$$

Пока  $x$  меняется от  $-\infty$  до  $x_1$ , основным состоянием является вектор-эталон  $\vec{\varepsilon}^{\prime}(n)$ , поскольку им исчерпывается класс векторов  $\Sigma_0$ . В точке  $x_1$  происходит перескок основного состояния из класса  $\Sigma_0$  в класс  $\Sigma_1$  и основное состояние становится  $p$  раз вырожденным. Когда  $x$  достигает значения  $x_2$ , основное состояние перескакивает из класса  $\Sigma_1$  в класс  $\Sigma_2$  и становится  $C_p^2$  раз вырожденным и т.д. По мере увеличения  $x$  значение функционала для основного состояния сначала монотонно падает, а затем, пройдя через минимум, монотонно возрастает. Номер  $k_{\max}$  последней перестройки основного состояния, которая произойдет при неограниченном увеличении  $x$ , не меньше  $[(p+1)/2]$ :

$$k_{\max} = \left\lceil \frac{\frac{n+1}{2} + \frac{p+1}{2}}{2} \right\rceil \geq \left\lceil \frac{p+1}{2} \right\rceil.$$

Случай  $p = n$  имеет смысл обсудить отдельно. Подставив  $p = n$  в выражения (13), получаем, что при этом все абсциссы  $x_k$  слипаются в одну:

$$x_k \equiv x' = \frac{n}{2}, \quad k = 1, 2, 3, \dots \quad (15a)$$



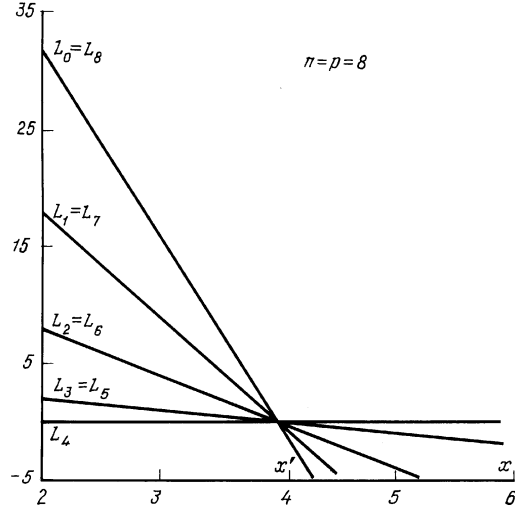


Рис. 3. Типичное поведение прямых  $L_k(x)$ ,  $k = 0, 1, \dots, p$ , для  $p = n$ .

Действительно, все прямые  $L_k(x)$  пересекаются теперь в одной точке  $x'$  (рис. 3). Таким образом, при  $p = n$  для любого  $x$  слева от  $x'$  основным состоянием будет вектор-эталон  $\vec{\varepsilon}(n)$ , а справа от  $x'$  основное состояние будет  $C_n^{[(n+1)/2]}$  раз вырожденным, поскольку

$$k_{\max} = \left[ \frac{n+1}{2} \right].$$

Интервал от  $x_1$  до  $x_{k_{\max}}$  – область перестройки основного состояния – подробно исследован в приложении 2. Здесь лишь отметим, что при  $p = \text{const}$  и  $n \rightarrow \infty$  область перестройки стягивается в точку

$$x'' = p. \tag{15б}$$

Эта точка на оси значений  $x$  во многих отношениях замечательна.

### 3. ОБОБЩЕНИЕ ОСНОВНОЙ МОДЕЛИ

**3.1. Нормированные исходные образы.** По мере увеличения  $x$  корреляции между исходными образами  $\vec{s}^{(l)}$  (6) уменьшаются, набор исходных образов представляет собой постепенно разжимающийся с увеличением  $x$  пучок векторов. Некоторая нефизичность исходных образов (6) связана с тем, что их длина зависит от варьируемого параметра  $x$ :  $s^2 = \|\vec{s}^{(l)}\|^2 = n - 2x + x^2$ . Эту зависимость легко устранить, отнормировав исходные образы на единицу:

$$\vec{\bar{s}}^{(l)} = \frac{\vec{s}^{(l)}}{s}, \quad l = 1, 2, \dots, p.$$

Все элементы матрицы связей окажутся при этом поделенными на  $s^2$ , равно как и выражение (8а) для максимизируемого функционала:

$$\tilde{F}(\vec{\sigma}^*) = p \frac{x^2 - 2x \cos w(q + p \cos w) + (q + p \cos w)^2}{n - 2x + x^2}.$$

Ясно, что для функционала  $\tilde{F}$  остаются в силе все основные результаты предыдущего раздела, в частности распределение “подозреваемых на экстремум” конфигурационных векторов  $\vec{\sigma}^*$  по классам  $\Sigma_k$  и теорема 2. Можно показать, что значение функционала для *основного состояния* как функция от  $x$  будет при этом монотонно убывать.

### 3.2. Произвольный вектор-эталон.

1. Если в качестве вектора-эталона вместо  $\vec{\varepsilon}(n)$  взять  $\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_p, \alpha_{p+1}, \dots, \alpha_n)$  – произвольный конфигурационный вектор, а в качестве исходных образов –  $p$  его искажений по типу (6), то матрица связей – обозначим ее  $\mathbf{N}^{(\alpha)}$  – будет иметь вид  $N_{ij}^{(\alpha)} = N_{ij} \alpha_i \alpha_j$ ,  $i, j = 1, 2, \dots, n$ , где  $N_{ij}$  – элементы матрицы связей  $\mathbf{N}$  (5). Если ввести диагональную  $(n \times n)$ -матрицу

$$\mathbf{A} = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_p, \alpha_{p+1}, \dots, \alpha_n),$$

то  $\mathbf{N}^{(\alpha)} = \mathbf{A} \cdot \mathbf{N} \cdot \mathbf{A}$  и максимизируемый функционал  $F^{(\alpha)}(\vec{\sigma})$  принимает вид

$$F^{(\alpha)}(\vec{\sigma}) = (\mathbf{N}^{(\alpha)} \vec{\sigma}, \vec{\sigma}) = (\mathbf{N} \cdot \mathbf{A} \vec{\sigma}, \mathbf{A} \vec{\sigma}) = (\mathbf{N} \vec{\xi}, \vec{\xi}) = F(\vec{\xi}),$$

где  $\vec{\xi} = \mathbf{A} \vec{\sigma}$ , а функционал  $F(\vec{\xi})$  изучен выше. Таким образом, все результаты предыдущего раздела остаются в силе, но формулируются теперь относительно конфигурационных векторов  $\vec{\xi} = \mathbf{A} \vec{\sigma}$ . В дальнейшем в качестве вектора-эталона будем использовать биссектрису  $\vec{\varepsilon}(n)$  главного угла в пространстве  $\mathbb{R}^n$ , памятуя о том, что получающиеся результаты легко переформулируются для произвольного конфигурационного вектора-эталона.

2. Исходные образы  $\vec{s}^{(l)}$  (6) представляют собой симметричный пучок векторов, порожденных вектором-эталонем  $\vec{\varepsilon}(n)$ . Можно ожидать, что вращение исходных образов как целого не отразится кардинально на свойствах максимизируемого функционала. Результат, однако, сильно зависит от того, смещается ли при таком вращении вектор-эталон или нет.

Исходные образы  $\vec{s}^{(l)}$  будем вращать так, чтобы последние  $q$  координат вектора-эталона не менялись. Нетривиальная часть матрицы вращения имеет размерность  $p \times p$  и действует на первые  $p$  координат  $n$ -мерных векторов, переводя искажаемую часть вектора-эталона в  $p$ -мерный вектор  $\vec{u}$ :

$$\mathbf{U} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1p} \\ u_{21} & u_{22} & \dots & u_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1} & u_{p2} & \dots & u_{pp} \end{pmatrix}, \quad \mathbf{U} \cdot \vec{\varepsilon}(p) = \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{pmatrix}, \quad \|\vec{u}\|^2 = p. \quad (16)$$

Играющие в дальнейшем существенную роль координаты  $u_l$  вектора  $\vec{u}$  суть

$$u_l = \sum_{i=1}^p u_{li}, \quad l = 1, 2, \dots, p. \quad (17)$$

В результате вращения исходные образы переходят в векторы

$$\vec{s}^{(l)}(u) = (u_1 - xu_{1l}, u_2 - xu_{2l}, \dots, u_p - xu_{pl}, 1, \dots, 1), \quad l = 1, 2, \dots, p, \quad (18)$$

а отвечающей им матрице связей  $\mathbf{N}^{(U)}$  удобно придать вид

$$\mathbf{N}^{(U)} = \begin{pmatrix} (p-2x+x^2)u_1^2 & \dots & (p-2x)u_1u_p & (p-x)u_1 & \dots & (p-x)u_1 \\ (p-2x)u_2u_1 & \dots & (p-2x)u_2u_p & (p-x)u_2 & \dots & (p-x)u_2 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ (p-2x)u_pu_1 & \dots & (p-2x+x^2)u_p^2 & (p-x)u_p & \dots & (p-x)u_p \\ (p-x)u_1 & \dots & (p-x)u_p & p & \dots & p \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (p-x)u_1 & \dots & (p-x)u_p & p & \dots & p \end{pmatrix}. \quad (19)$$

Отсюда сразу видно, если в результате вращения вектор-эталон  $\vec{\varepsilon}(n)$  остается на месте, т.е. если  $u_1 = u_2 = \dots = u_p = 1$ , то матрица связей  $\mathbf{N}^{(U)}$  совпадает с матрицей  $\mathbf{N}$  (5). Следовательно, несмотря на то что в этом случае исходные образы  $\vec{s}^{(l)}(u)$  (18) получаются из вектора-эталона искажением сразу  $p$  его координат, все результаты предыдущего раздела относительно зависимости основного состояния от параметра  $x$  остаются в силе.

Если же величины  $u_i$  не равны тождественно 1, то удобно ввести диагональную  $(n \times n)$ -матрицу  $\mathbf{D} = \text{diag}(u_1, u_2, \dots, u_p, 1, \dots, 1)$  и представить матрицу связей  $\mathbf{N}^{(U)}$  в виде тройного произведения:  $\mathbf{N}^{(U)} = \mathbf{D} \cdot \mathbf{N} \cdot \mathbf{D}$ , где  $\mathbf{N}$  – уже изученная нами матрица связей (5). Рассуждая, как в предыдущем пункте, и используя теорему 1, можно показать, что, как и раньше:

1) для функционала  $F^{(U)}(\vec{\sigma}) = (\mathbf{N}^{(U)}\vec{\sigma}, \vec{\sigma})$  “подозреваемыми на экстремум” являются только конфигурационные векторы  $\vec{\sigma}^*$  вида (7);

2) значение функционала  $F^{(U)}$  на векторах  $\vec{\sigma}^*$  равно

$$F^{(U)}(\vec{\sigma}^*) = p \cdot [x^2 - 2x \cos w (q + p \cos w) + (q + p \cos w)^2], \quad (20a)$$

где

$$\cos w = \frac{\sum_{i=1}^p \sigma_i \cdot u_i}{p} = \frac{(\vec{\sigma}', \vec{u})}{\|\vec{\sigma}'\| \cdot \|\vec{u}\|} \quad (20б)$$

есть косинус угла между  $p$ -мерными векторами  $\vec{\sigma}' = (\sigma_1, \sigma_2, \dots, \sigma_p)$  и  $\vec{u}$ . Выражения (20) обобщают аналогичные равенства (8) из предыдущего раздела.

Итак, значение функционала  $F^{(U)}(\vec{\sigma}^*)$  полностью определяется величиной искажения  $x$  и косинусом угла между  $p$ -мерными векторами  $\vec{\sigma}'$  и  $\vec{u}$ . Если несколько векторов  $\vec{\sigma}'$  равноудалено от  $\vec{u}$  (они образуют с ним одинаковые углы), то значение функционала

$F^{(U)}(\vec{\sigma}^*)$  на соответствующих  $n$ -мерных векторах  $\vec{\sigma}^*$  одинаково. Следовательно, как и в предыдущем разделе, множество “подозреваемых на экстремум” конфигурационных векторов  $\vec{\sigma}^*$  распадается на классы такие, что на векторах из одного класса функционал  $F^{(U)}(\vec{\sigma}^*)$  принимает одно и то же значение. По аналогии с предыдущим разделом обозначим эти классы через  $\Sigma_k^{(U)}$ . Число классов  $\Sigma_k^{(U)}$  равно числу *различных* значений  $\cos w$ , а их состав определяется тем, на *какие* группы распадается совокупность  $p$ -мерных векторов  $\vec{\sigma}'$ , равноудаленных от  $\vec{u}$ .

Таким образом, необходимо уметь упорядочивать  $p$ -мерные векторы  $\vec{\sigma}'$  по степени близости к вектору  $\vec{u}$ . Подробнее эту задачу мы обсудим в отдельной публикации, а сейчас положим, что требуемое упорядочение векторов  $\vec{\sigma}'$  проведено и получена убывающая последовательность *различных* значений  $\cos w$ :

$$\cos w_0 > \cos w_1 > \dots > \cos w_t. \quad (21)$$

Число классов эквивалентности  $\Sigma_k^{(U)}$  равно  $t + 1$ , а количество конфигурационных векторов  $\vec{\sigma}^*$  в  $k$ -м классе обозначим через  $v_k$ :

$$v_0 + v_1 + \dots + v_t = 2^p.$$

Поскольку половина  $p$ -мерных векторов  $\vec{\sigma}'$  получается из другой половины обращением знака у всех координат, для всякого  $k = 0, 1, \dots, t$  выполняются очевидные равенства:

$$\begin{aligned} \cos w_k &= -\cos w_{t-k}, \\ v_k &= v_{t-k}. \end{aligned}$$

Отметим, что с ростом  $k$  значения последовательных сумм  $\cos w_{k-1} + \cos w_k$  не только убывают, но и начиная с некоторого  $k$  становятся отрицательными (этот результат нам вскоре потребуется).

Итак, аналогично предыдущему разделу совокупность  $2^p$  “подозреваемых на экстремум” конфигурационных векторов  $\vec{\sigma}^*$  распадается на классы эквивалентности  $\Sigma_k^{(U)}$ ,  $k = 0, 1, \dots, t$ . Для определения при заданном  $x$  основного состояния необходимо сравнить между собой величины

$$F_k^{(U)}(x) = p \cdot [x^2 - 2x \cos w_k (q + p \cos w_k) + (q + p \cos w_k)^2], \quad k = 0, 1, \dots, t,$$

и отыскать среди них наибольшую. Поскольку квадратичный по  $x$  член у всех  $F_k^{(U)}(x)$  одинаков, его при сравнении можно опустить. Следовательно, для выяснения зависимости основного состояния от параметра  $x$  необходимо изучить семейство прямых

$$L_k^{(U)}(x) = (q + p \cos w_k)^2 - 2x \cos w_k (q + p \cos w_k), \quad k = 0, 1, \dots, t,$$

во многом аналогичное семейству прямых  $L_k(x)$  (12). Анализ взаимного расположения прямых  $L_k^{(U)}(x)$  дает следующий результат, обобщающий теорему 2.

ТЕОРЕМА 3.. По мере изменения  $x$  от  $-\infty$  до  $\infty$  основное состояние системы последовательно исчерпывается классами конфигурационных векторов  $\vec{\sigma}^*$

$$\Sigma_0^{(U)}, \Sigma_1^{(U)}, \dots, \Sigma_{k_{\max}}^{(U)}.$$

Перескок основного состояния из класса  $\Sigma_{k-1}^{(U)}$  в класс  $\Sigma_k^{(U)}$  происходит в точке  $x_k$ :

$$x_k = \frac{p}{2} \left[ 1 + \frac{q}{q + p(\cos w_{k-1} + \cos w_k)} \right], \quad k = 1, 2, \dots, k_{\max}, \quad (22)$$

где  $\cos w_k$  принадлежат упорядоченной последовательности (21). Если  $x_1 > 3p/4$ , то по этой схеме одна за другой произойдут все  $t$  перестроек основного состояния:  $k_{\max} = t$ . Если же  $x_1 < 3p/4$ , то перестройки основного состояния прекратятся, когда знаменатель в (22) станет отрицательным. Других локальных максимумов у функционала  $F^{(U)}(\vec{\sigma})$  не будет.

Сделаем в заключение два замечания.

ЗАМЕЧАНИЕ 1. “Основная модель” предыдущего раздела обобщена нами на случай, когда в качестве вектора-эталона используется  $n$ -мерный вектор

$$(u_1, u_2, \dots, u_p, 1, \dots, 1),$$

а исходные образы получаются искажением не одной, но сразу  $p$  его координат (см. выражения (18)). И хотя на исходные образы налагается требование, чтобы они одинаково коррелировали друг с другом, отвечающая им матрица связей  $\mathbf{N}^{(U)}$  (19) заметно расширяет (по сравнению с  $\mathbf{N}$ ) класс изученных симметричных матриц связей.

Когда  $u_i \equiv 1$ , теорема 3 переходит в теорему 2. При этом оказывается, что  $t = p$ , а  $v_k = C_p^k$ . Небольшого пояснения требует тот факт, что в теореме 2 номер  $k_{\max}$  последней перестройки основного состояния зависит от соотношения между  $(p-1)/(n-1)$  и  $1/3$ , а в теореме 3 – от соотношения между  $x_1$  и  $3p/4$ : при  $u_i \equiv 1$  обе формулировки эквивалентны.

ЗАМЕЧАНИЕ 2. Когда искажение  $x$  принадлежит интервалу  $(x_k, x_{k+1})$ , класс  $\Sigma_k^{(U)}$  конфигурационных векторов  $\vec{\sigma}^*$  представляет собой множество локальных максимумов функционала  $F^{(U)}(x)$ . Иными словами,  $\Sigma_k^{(U)}$  – это множество неподвижных точек сети хопфилдова типа с матрицей связей  $\mathbf{N}^{(U)}(x)$  (19) (диагональ у которой, напомним, равна нулю – см. введение). В то же время состав каждого класса  $\Sigma_k^{(U)}$  определяется исключительно величинами  $\{u_i\}_{i=1}^p$ , выбор которых находится полностью в руках исследователя. Следовательно, подбирая специальным образом величины  $\{u_i\}_{i=1}^p$  и искажение  $x$ , можно создавать сети хопфилдова типа с наперед заданным набором неподвижных точек.

Выяснение границ возникающих на этом пути возможностей требует дополнительных исследований. Добавим, что удастся обобщить “основную модель” предыдущего раздела на случай, когда исходные образы получаются из вектора-эталона одновременным искажением сразу целой группы его координат, причем внутри группы параметры

искажения независимы. К примеру,  $(p \times n)$ -матрица исходных образов  $\mathbf{S}$  может выглядеть так:

$$\begin{pmatrix} 1 - x_1 & \dots & 1 - x_m & 1 & \dots & 1 & \dots & 1 & \dots & 1 & 1 & \dots & 1 \\ 1 & \dots & 1 & 1 - x_1 & \dots & 1 - x_m & \dots & 1 & \dots & 1 & 1 & \dots & 1 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & \dots & 1 & 1 & \dots & 1 & \dots & 1 - x_1 & \dots & 1 - x_m & 1 & \dots & 1 \end{pmatrix}.$$

Здесь  $n = p \times m + q$ . С помощью варьирования параметров искажения  $x_1, x_2, \dots, x_m$  независимо друг от друга можно охватить довольно широкий круг симметричных матриц связей. Эти результаты готовятся к публикации.

#### 4. ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

**4.1. Нейронные сети.** Полученные результаты естественно интерпретируются в рамках содержательной постановки задачи (см. введение): *качество восстановления сетью “истины” (вектора-эталона) зависит от величины искажения  $x$  на стадии обучения и от длины  $p$  обучающего набора.* Отметим некоторые характерные моменты.

Когда искажение  $x$  принадлежит интервалу  $(x_k, x_{k+1})$ , “истина” воспроизводится сетью с ошибкой: неподвижные точки отличаются от вектора-эталона  $k$  координатами. Чем больше искажение на стадии обучения, тем больше ошибка сети.

Согласуется со здравым смыслом и то, что левая граница  $x_1$  области перестройки является возрастающей функцией как  $p$ , так и  $n$ . Например, зафиксируем  $n$  и  $x$  (не очень большое, не больше чем  $n/2$  – см. ниже). Тогда одним только увеличением числа исходных образов  $p$  можно заставить  $x_1(p)$  превзойти  $x$ . В результате  $x$  окажется слева от  $x_1(p)$  в области, где единственной неподвижной точкой является вектор-эталон. Иначе говоря, при не очень большом  $x$  только лишь за счет увеличения  $p$  можно добиться, чтобы сеть правильно воспроизвела “истину”. Это согласуется с привычным опытом, согласно которому большая длина обучающей выборки позволяет вернее выделить зашумленный истинный сигнал (если уровень шумов не очень высок). Конечно, с увеличением  $p$  растет не только длина обучающей выборки, но и число искаженных координат вектора-эталона. Однако все они искажены не очень сильно, а большая обучающая выборка передает вектор-эталон более полно и всесторонне. С этой точки зрения наиболее надежной сетью рассматриваемого типа является та, у которой длина обучающей выборки  $p = n$ : при любых искажениях  $x < n/2$  она устойчиво воспроизводит вектор-эталон. Чем меньше  $p$ , тем меньше область устойчивого воспроизведения вектора-эталона. В том же духе интерпретируется и возрастание левой границы области перестройки  $x_1$  как функции от  $n$ .

В заключение обсудим еще один аспект, возможно, более дискуссионный. Когда величина искажения  $x \in (x_k, x_{k+1})$ , число неподвижных точек сети равно  $C_p^k$ . Следовательно, пока  $k$  увеличивается от 1 до  $p/2$ , количество воспроизводимых сетью “искаженных истин” тоже растет, это выглядит вполне правдоподобно. Однако стоит параметру искажения  $x$  превзойти  $x'' = p$  ( $k$  при этом станет больше  $p/2$ ), как количество

воспроизводимых сетью искаженных истин начнет уменьшаться! Это тем более необычно, что степень искажения сетью “истины” по-прежнему увеличивается: неподвижные точки будут все большим числом координат отличаться от вектора-эталона. Этот парадокс следует обсудить.

Нам представляется разумным следующее объяснение. Когда величина искажения  $x$  очень велика, исходные образы (6)

$$\vec{s}^{(l)} = \left( 1, \dots, 1, \underbrace{1-x}_l, 1, \dots, 1 \right), \quad l = 1, 2, \dots, p,$$

следует трактовать уже не как искажения вектора-эталона  $\vec{\varepsilon}^{\pm}(n)$ , обозначим его здесь иначе, чем прежде:

$$\vec{\varepsilon}^{(+)}(n) = \left( \underbrace{1, 1, \dots, 1}_p, 1, \dots, 1 \right), \quad (23)$$

а как искажения другого конфигурационного вектора –

$$\vec{\varepsilon}^{(-)}(n) = \left( \underbrace{-1, -1, \dots, -1}_p, 1, \dots, 1 \right). \quad (24)$$

В своей несовпадающей части векторы  $\vec{\varepsilon}^{(+)}(n)$  и  $\vec{\varepsilon}^{(-)}(n)$  являются противоположностью друг друга, двумя противостоящими друг другу высказываниями. Любой конфигурационный вектор  $\vec{\sigma}^*$  вида (7), у которого среди первых  $p$  координат ровно  $k$  равны  $-1$ , можно трактовать как некое промежуточное между векторами  $\vec{\varepsilon}^{(+)}(n)$  и  $\vec{\varepsilon}^{(-)}(n)$  высказывание, тяготеющее либо к одному краю шкалы, либо к другому. В зависимости от величины искажения  $x$  сеть в качестве усвоенной ею истины будет воспроизводить высказывание, тяготеющее либо к вектору  $\vec{\varepsilon}^{(+)}(n)$ , либо к его противоположности – вектору  $\vec{\varepsilon}^{(-)}(n)$ . А точка  $x'' = p$  является той границей в области изменения параметра  $x$ , по разные стороны от которой реализуется та или другая возможность.

Проанализируем с этой точки зрения то, как множество неподвижных точек зависит от параметра искажения  $x$  при  $k_{\max} = p$  (см. теорему 2). По мере все большего *увеличения*  $x$  за  $x''$  число неподвижных точек сети убывает как  $C_p^k$  и они группируются вокруг одного полюса шкалы – вектора  $\vec{\varepsilon}^{(-)}(n)$ , который сеть будет воспроизводить в качестве усвоенной ею истины *при любом значении  $x$  из полубесконечного интервала*  $(x_p, \infty)$ . И наоборот, по мере все большего *уменьшения*  $x$  *вниз от  $x''$*  убывающие теми же темпами неподвижные точки группируются вокруг другого полюса шкалы – вектора  $\vec{\varepsilon}^{(+)}(n)$ , который сеть станет воспроизводить *при любом значении  $x$  из противоположного полубесконечного интервала*  $(-\infty, x_1)$ . Картина выглядит достаточно симметричной относительно  $x'' = p$ , однако это имеет место только при

$$\frac{p-1}{n-1} < \frac{1}{3} \iff p < \frac{q}{2} + 1, \quad (25)$$

поскольку именно в этом случае  $k_{\max} = p$ . Если же  $p$  больше, чем это разрешено неравенствами (25), то

$$\left\lceil \frac{p+1}{2} \right\rceil < k_{\max} = \left\lfloor \frac{n+p+2}{4} \right\rfloor < p$$

и симметрия разрушается: какое бы большое  $x$  из интервала  $(x_{k_{\max}}, \infty)$  при этом ни взять, сеть будет воспроизводить в качестве усвоенной ею истины не противоположный  $\vec{\varepsilon}^{(+)}(n)$  вектор  $\vec{\varepsilon}^{(-)}(n)$ , а всего лишь некоторое промежуточное между  $\vec{\varepsilon}^{(+)}(n)$  и  $\vec{\varepsilon}^{(-)}(n)$  высказывание, хотя и тяготеющее к  $\vec{\varepsilon}^{(-)}(n)$ , поскольку

$$k_{\max} > \left[ \frac{p+1}{2} \right].$$

Возможно, этот эффект является математическим отражением трудностей, сопровождающих процесс “переобучения” (ведь хорошо известно, что чем крепче ученик усвоил “неправильную” истину, тем труднее – и даже не всегда возможно! – научить его “правильной”). В самом деле, отнесемся к  $x$  как к внешнему управляющему параметру, варьированием которого можно регулировать усваиваемую сеть “истину”. Пока  $x$  невелико, обучающие векторы  $\vec{s}^{(l)}$  мало отличаются от  $\vec{\varepsilon}^{(+)}(n)$  и сеть надежно его усваивает в качестве единственной истины. Пусть мы считаем эту истину неверной и хотим переобучить сеть противоположной истине – вектору  $\vec{\varepsilon}^{(-)}(n)$ . Станем для этого увеличивать параметр искажения  $x$  так сильно, чтобы исходные образы  $\vec{s}^{(l)}$  воспринимались уже как искажения “правильной” истины  $\vec{\varepsilon}^{(-)}(n)$ . Иначе говоря, станем увеличивать  $x$  далеко за  $x'' = p$ . Переобучится ли сеть противоположному  $\vec{\varepsilon}^{(+)}(n)$  вектору  $\vec{\varepsilon}^{(-)}(n)$ ? Неравенства (25) дают ответ: если число  $p$  подлежащих исправлению признаков невелико, то переобучится. Если же  $p$  сравнительно велико, то демонстрацией сети набора противоположных (“правильных”) обучающих примеров не обойтись. Конечно, поначалу усвоенная сеть “неправильная” истина в результате такого переобучения изменится, и даже станет тяготеть к цели переобучения – вектору  $\vec{\varepsilon}^{(-)}(n)$ , но не больше. Впрочем без подкрепляющих примеров рассуждения этого абзаца достаточно спорны и высказываются нами только в порядке дискуссии.

**4.2. Модель Изинга.** Как уже отмечалось, интерпретация матрицы  $\mathbf{S}$  исходных образов в терминах модели Изинга пока неизвестна. Поэтому полученные результаты будем интерпретировать исходя из устройства матрицы связей  $\mathbf{N}$ . Перепишем ее в более традиционном для физических рассуждений виде:

$$\mathbf{N} \propto \begin{pmatrix} 0 & 1 - 2\frac{x}{p} & \dots & 1 - 2\frac{x}{p} & 1 - \frac{x}{p} & \dots & 1 - \frac{x}{p} \\ 1 - 2\frac{x}{p} & 0 & \dots & 1 - 2\frac{x}{p} & 1 - \frac{x}{p} & \dots & 1 - \frac{x}{p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 1 - 2\frac{x}{p} & 1 - 2\frac{x}{p} & \dots & 0 & 1 - \frac{x}{p} & \dots & 1 - \frac{x}{p} \\ 1 - \frac{x}{p} & 1 - \frac{x}{p} & \dots & 1 - \frac{x}{p} & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 - \frac{x}{p} & 1 - \frac{x}{p} & \dots & 1 - \frac{x}{p} & 1 & \dots & 0 \end{pmatrix}.$$

По сравнению с (5) здесь диагональные элементы равны нулю и вынесен общий множитель  $p$ .



Данная матрица связей отвечает спиновой системе, состоящей из двух подсистем размерами  $p$  и  $q = n - p$  таких, что взаимодействие между  $p$  спинами первой подсистемы одинаково и равно  $1 - 2x/p$ ; взаимодействие между  $q$  спинами второй подсистемы одинаково и равно 1; перекрестное взаимодействие между спинами обеих подсистем тоже одинаково и равно  $1 - x/p$ . Таким образом, речь идет о спиновой системе, состоящей из двух различных, но как бы однородных (в смысле взаимодействия между спинами) подсистем<sup>4)</sup>. Варьирование параметра  $x$  не повлияет на взаимодействия между  $q$  спинами второй подсистемы, но отразится на остальных взаимодействиях. Теорема 2 позволяет проследить за тем, что при этом будет происходить с основным состоянием.

Пока  $x/p \in (-\infty, 1/2)$ , все матричные элементы положительны, спины взаимодействуют только ферромагнитным образом. Основное состояние при этом должно быть ферромагнитным (спины должны смотреть в одну сторону). Так оно и происходит в нашей модели, поскольку, пока  $x < x_1$ , основным состоянием является вектор  $\vec{\varepsilon}^{(+)}(n)$  (23), а  $1/2$  меньше  $x_1/p$ :

$$\frac{x_1}{p} = \frac{n-1}{n+p-2} \geq \frac{1}{2}.$$

Равенство здесь имеет место только при  $p = n$ ; этот случай мы рассмотрим отдельно, а пока будем считать, что  $p < n$ .

После того как  $x/p$  превзойдет  $1/2$ , левый диагональный  $(p \times p)$ -блок матрицы  $\mathbf{N}$  станет отрицательным. Однако до тех пор, пока  $x/p$  не превзойдет  $x_1/p$ , основное состояние будет оставаться ферромагнитным! Иначе говоря, *существует “зазор” между тем значением параметра  $x/p$ , при котором разрушается ферромагнитный тип взаимодействия, и тем его значением, при котором разрушается ферромагнитное основное состояние*. Только после того как во взаимодействиях спинов накопится достаточно “антиферромагнитности”, ферромагнитное основное состояние разрушится и в системе произойдет фазовый переход. Затем после накопления следующей критической порции антиферромагнитности произойдет следующий фазовый переход (когда  $x/p$  превзойдет  $x_2/p$ ) и т.д. После того как  $x/p$  превзойдет значение  $1 = x''/p$  ( $x'' = p - \text{см. (15б)}$ ), антиферромагнитным станет и перекрестное взаимодействие между спинами двух подсистем. Фазовые переходы будут продолжаться.

В физике интересуются зависимостью энергии

$$E = - \sum_{i=1}^n N_{ij} \sigma_i \sigma_j$$

основного состояния от внешнего параметра, каковым в нашем случае является  $x$ . Можно показать, что, пока  $x < x'' = p$ , энергия основного состояния с ростом  $x$  уве-

<sup>4)</sup> При  $p = n$  получаем  $n$  одинаково взаимодействующих друг с другом спинов. Напомним, кстати, что использование в качестве вектора-эталоны произвольного конфигурационного вектора  $\vec{a}$  приведет к появлению у матричных элементов  $N_{ij}$  множителей  $\alpha_i \alpha_j$ , а вращение исходных образцов как целого отразится на  $N_{ij}$  домножением их на  $u_i u_j$  (см. (19)). Однако мы будем интерпретировать только результаты, относящиеся к “основной модели”.

личивается, а после перехода  $x$  через  $x''$  энергия станет уменьшаться<sup>5)</sup>. В точках  $x_k$  энергия как функция параметра  $x$  имеет изломы.

Когда  $p$  мало по сравнению с  $q$  (см. неравенства (25)), постепенным увеличением параметра  $x$  можно перевести основное состояние из начального  $\vec{\varepsilon}^{(+)}(n)$  в конечное  $\vec{\varepsilon}^{(-)}(n)$ ; оба эти основных состояния невырождены. Если же  $p$  велико по сравнению с  $q$ , то никаким увеличением  $x$  достичь состояния  $\vec{\varepsilon}^{(-)}(n)$  не удастся: основное состояние, в котором система окончательно застрянет и перестанет реагировать на увеличение внешнего параметра, будет  $C_p^{k_{\max}}$  раз вырожденным. Осмысление этого результата в терминах модели Изинга нам неизвестно.

Все эти фазовые переходы обычно ускользают от внимания физиков, потому что, как правило, они рассматривают ситуацию  $p = \text{const}$ ,  $n \rightarrow \infty$ . В этом случае согласно нашим выкладкам (см. (15б)) все перескоковые абсциссы  $x_k$  слипаются в одну  $x'' = p$ . Слева от  $x''$  основным состоянием будет вектор  $\vec{\varepsilon}^{(+)}(n)$ , а справа от  $x''$  — вектор  $\vec{\varepsilon}^{(-)}(n)$ . Именно эта картина хорошо знакома физикам по их рассмотрениям. Оказывается, однако, что при  $0 < p/n < 1$  существует целая область значений внешнего параметра, насыщенная фазовыми переходами.

Рассмотрим, наконец, случай  $p = n$ . Все перескоковые абсциссы слипаются при этом в  $x' = n/2$  (см. соотношения (15а)). Поскольку  $p/2 \equiv n/2 = x_1$ , исчезает “зазор” между тем значением параметра  $x$ , при котором разрушается ферромагнитный тип взаимодействия, и тем его значением, при котором основное состояние перестает быть ферромагнитным. При любом  $x$ , меньшем  $x'$ , основным состоянием будет вектор  $\vec{\varepsilon}^{(+)}(n)$ . При любом  $x$  справа от  $x'$  основным состоянием будет  $C_n^{\lfloor (n+1)/2 \rfloor}$  раз вырожденным. Состояние системы справа от  $x'$  естественно ассоциировать с фазой спинового стекла.

**4.3. Факторный анализ.** Во введении отмечалось, что матрица  $\mathbf{S}$  исходных образов играет существенную роль в содержательной постановке задачи в факторном анализе [5]. Действительно, матрица  $\mathbf{S}$  в этом случае не что иное, как эмпирическая матрица “объектов-параметров”, которая собственно и содержит исчерпывающее количественное описание изучаемого явления. Ее строки суть  $p$  объектов, заданных значениями  $n$  параметров (доступных исследователю для измерения):

$$\vec{s}^{(l)} = (s_1^{(l)}, s_2^{(l)}, \dots, s_n^{(l)}) \in \mathbb{R}^n, \quad l = 1, 2, \dots, p, \quad \text{— векторы-объекты,}$$

$$\vec{s}_i = \begin{pmatrix} s_i^{(1)} \\ s_i^{(2)} \\ \vdots \\ s_i^{(p)} \end{pmatrix} \in \mathbb{R}^p, \quad i = 1, 2, \dots, n, \quad \text{— векторы-параметры.}$$

Основная идея факторного анализа состоит в конструировании из  $n$   $p$ -мерных векторов-параметров  $\vec{s}_i$  *значительно меньшего* числа  $t$  ( $t \ll n$ ) искусственных характеристик-признаков (их называют *факторами*) таких, чтобы переход к описанию объектов в

<sup>5)</sup>См., однако, конец п. 3.1.

терминах факторов не сопровождался потерей существенной информации об объектах. Факторы  $\vec{f}_j$  строятся как линейные комбинации векторов-параметров  $\vec{s}_i$ :

$$\vec{f}_j = \sum_{i=1}^n a_i^{(j)} \vec{s}_i,$$

минимизирующие (в смысле той или иной меры близости) суммарное отклонение фактора от всей совокупности векторов-параметров. В работе [6] показано, что для широко используемого в факторном анализе *центроидного метода* задача построения первого фактора эквивалентна задаче (1)–(3). Интерпретация в этих терминах полученных у нас результатов будет представлена в дальнейшем.

**Благодарности.** Автор признателен А. А. Ежову за целый ряд советов как по существу настоящей работы, так и по поводу рукописи.

### ПРИЛОЖЕНИЕ 1

Докажем теорему 1. Пусть  $\mathbf{S}$  – прямоугольная  $(p \times n)$ -матрица. Ее строки суть  $p$   $n$ -мерных векторов  $\vec{s}^{(l)}$  ( $l = 1, \dots, p$ ), а столбцы –  $n$   $p$ -мерных векторов  $\vec{s}_i$  ( $i = 1, \dots, n$ ). С помощью сопряженной к  $\mathbf{S}$  и тоже прямоугольной  $(n \times p)$ -матрицы  $\mathbf{S}^*$  можно построить две симметричные матрицы:

- $(n \times n)$ -матрицу  $\mathbf{N} = \mathbf{S}^* \cdot \mathbf{S}$ , матричные элементы которой – скалярные произведения  $p$ -мерных векторов-столбцов  $\vec{s}_i$  матрицы  $\mathbf{S}$ ;
- $(p \times p)$ -матрицу  $\mathbf{P} = \mathbf{S} \cdot \mathbf{S}^*$ , матричные элементы которой – скалярные произведения  $n$ -мерных векторов-строк  $\vec{s}^{(l)}$  матрицы  $\mathbf{S}$ .

Хорошо известно, что ненулевые собственные значения симметричных матриц  $\mathbf{N}$  и  $\mathbf{P}$  совпадают, а их собственные векторы весьма просто связаны друг с другом; если

$$\begin{aligned} \mathbf{N} \cdot \vec{h}^{(j)} &= a_j \cdot \vec{h}^{(j)}, & \|\vec{h}^{(j)}\| &= 1, & j &= 1, \dots, n, \\ \mathbf{P} \cdot \vec{f}_i &= b_i \cdot \vec{f}_i, & \|\vec{f}_i\| &= 1, & i &= 1, \dots, p, \end{aligned}$$

то справедливы соотношения

$$a_i = b_i, \quad i = 1, \dots, p; \quad a_{p+1} = a_{p+2} = \dots = a_n = 0, \quad (\text{П.1})$$

$$\vec{h}^{(i)} \sim \mathbf{S}^* \cdot \vec{f}_i, \quad \vec{f}_i \sim \mathbf{S} \cdot \vec{h}^{(i)}, \quad i = 1, \dots, p, \quad (\text{П.2})$$

которые проверяются непосредственно (либо так, как, например, в работе [16]).

Если в качестве  $\mathbf{S}$  взять матрицу исходных образов (4), то  $(n \times n)$ -матрица  $\mathbf{N}$  – это матрица связей (5), которую мы изучаем, а матрица  $\mathbf{P}$  имеет весьма специальный вид:

$$\mathbf{P} = s^2 \cdot \begin{pmatrix} 1 & r & \dots & r \\ r & 1 & \dots & r \\ \vdots & \vdots & \ddots & \vdots \\ r & r & \dots & 1 \end{pmatrix} = s^2 \cdot \mathbf{R}, \quad (\text{П.3})$$

здесь  $s^2 = \|\vec{s}^{(l)}\|^2 = n - 2 \cdot x + x^2$ , а  $r$  – коэффициент корреляции между  $n$ -мерными векторами-строками  $\vec{s}^{(l)}$  и  $\vec{s}^{(l')}$  матрицы  $\mathbf{S}$ :

$$r = \frac{(\vec{s}^{(l)}, \vec{s}^{(l')})}{\|\vec{s}^{(l)}\|^2} = \frac{n - 2x}{n - 2x + x^2}, \quad l, l' = 1, 2, \dots, p.$$

Таким образом, если отыскать собственные значения и векторы достаточно простой матрицы  $\mathbf{R}$  из выражения (П.3), то с помощью (П.1) и (П.2) получим собственные значения и векторы интересующей нас матрицы  $\mathbf{N}$ . Обозначим собственные значения матрицы  $\mathbf{R}$  через  $r_i$ ; ее собственные векторы, естественно, совпадают с собственными векторами  $\vec{f}_i$  матрицы  $\mathbf{P}$ . Легко проверяется, что

$$r_1 = 1 + (p - 1) \cdot r, \quad r_2 = r_3 = \dots = r_p = 1 - r,$$

а

$$\vec{f}_1 = \frac{1}{\sqrt{p}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \vec{f}_k = \begin{pmatrix} f_k^{(1)} \\ f_k^{(2)} \\ \vdots \\ f_k^{(p)} \end{pmatrix}, \quad k = 2, \dots, p,$$

где собственные векторы  $\vec{f}_k$  для  $k > 1$  ортогональны биссектрисе  $\vec{\varepsilon}(p)$  главного угла пространства  $\mathbb{R}^p$ :

$$(\vec{f}_k, \vec{\varepsilon}(p)) = \sum_{i=1}^p f_k^{(i)} = 0, \quad k = 2, \dots, p. \quad (\text{П.4})$$

Домножив  $r_i$  на  $s^2$ , получим ненулевые собственные значения  $a_i$  матрицы  $\mathbf{N}$ :

$$a_1 = x^2 - 2pn + pn, \quad a_2 = a_3 = \dots = a_p = x^2.$$

Для получения соответствующих собственных векторов  $\vec{h}^{(i)}$  подействуем на векторы  $\vec{f}_i$  матрицей  $\mathbf{S}^*$ :

$$\mathbf{S}^* = \begin{pmatrix} 1-x & 1 & \dots & 1 \\ 1 & 1-x & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1-x \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

Легко видеть, что

$$\mathbf{S}^* \cdot \vec{f}_1 \sim \left( \underbrace{p-x, p-x, \dots, p-x}_p, \underbrace{p, p, \dots, p}_q \right),$$

а

$$\mathbf{S}^* \cdot \vec{f}_k = -x \cdot \left( f_k^{(1)}, f_k^{(2)}, \dots, f_k^{(p)}, \underbrace{0, 0, \dots, 0}_q \right), \quad k = 2, \dots, p;$$

для получения последнего равенства мы воспользовались соотношениями (П.4).

Отнормировав векторы  $\mathbf{S}^* \vec{f}_1$  и  $\mathbf{S}^* \vec{f}_k$  на единицу, получим выражения для собственных векторов  $\vec{h}^{(i)}$  из теоремы 1.

**ПРИЛОЖЕНИЕ 2**

1. Изучим семейство прямых  $L_k(x)$  (12):

$$L_k(x) = (n - 2k)^2 - 2x \frac{p - 2k}{p} (n - 2k), \quad k = 0, 1, \dots, p.$$

При  $x = 0$  имеем  $L_k(0) = (n - 2k)^2$ , и в случае  $p \leq n/2$  величины  $L_k(0)$  строго убывают с ростом  $k$ :

$$L_0(0) > L_1(0) > L_2(0) > \dots > L_p(0),$$

а при  $p > n/2$  начиная с номера  $n - p$  величины  $L_k(0)$  становятся дважды вырожденными:

$$L_0(0) > L_1(0) > \dots > L_{n-p}(0) = L_p(0) > L_{n-p+1}(0) = L_{p-1}(0) > \dots$$

Как функции от  $x$  прямые  $L_k(x)$  ведут себя по-разному в зависимости от номера  $k$ : для  $k < p/2$  и  $k > n/2$  прямые  $L_k(x)$  с ростом  $x$  убывают, а для  $p/2 < k < n/2$  возрастают; прямые с номерами  $k = p/2$  и  $k' = n/2$  (если такие  $k$  и  $k'$  найдутся) параллельны оси  $OX$ .

Абсцисса пересечения прямых  $L_k(x)$  и  $L_{k'}(x)$  – обозначим ее  $x_{(k,k')}$  – дается выражением

$$\begin{aligned} x_{(k,k')} &= \frac{p}{2} \left[ 1 + \frac{q}{q + p(\cos w_k + \cos w_{k'})} \right] = \\ &= p \frac{n - (k' + k)}{n + p - 2(k' + k)}, \quad 0 \leq k < k' \leq p. \end{aligned} \tag{П.5}$$

Обратим внимание на то, что абсцисса пересечения зависит только от суммы  $k + k'$  номеров прямых. При  $n + p$  нечетном число *различных* абсцисс  $x_{(k,k')}$  равно в точности  $2p - 1$ . При  $n + p$  четном возможны комбинации номеров  $k'$  и  $k$  такие, что  $n + p - 2(k' + k) = 0$ ; соответствующие прямые  $L_k(x)$  и  $L_{k'}(x)$  параллельны (пересекаются на бесконечности).

С увеличением  $k + k'$  абсцисса  $x_{(k,k')}$ , вообще говоря, растет:

$$x_{(0,1)} < x_{(0,2)} < x_{(0,3)} = x_{(1,2)} < x_{(0,4)} = x_{(1,3)} < \dots$$

Однако нетрудно видеть, что только при<sup>6)</sup>

$$\alpha = \frac{p - 1}{n - 1} \leq \frac{1}{3} \tag{П.6}$$

---

<sup>6)</sup> В теории нейронных сетей отношение числа исходных образов к числу нейронов традиционно обозначают символом  $\alpha$ , и в нижеприведенном определении мы следуем данной традиции, хотя это и противоречит договоренности использовать греческие буквы только для обозначения величин, равных  $\pm 1$ .

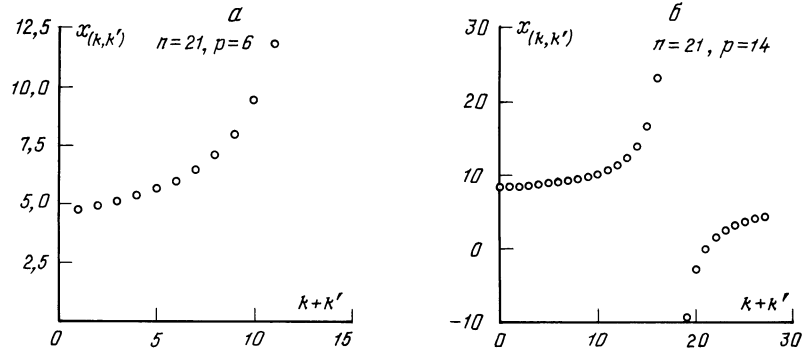


Рис. 4. Зависимость абсциссы  $x_{(k, k')}$  пересечения прямых  $L_k(x)$  и  $L_{k'}(x)$  от суммы номеров  $k + k'$ .

это возрастание будет строго монотонным: в этом случае знаменатель в выражении (П.5) неотрицателен при любых  $k$  и  $k'$ , и, следовательно, все пересечения прямых происходят в области положительности  $x$ . Если же

$$\alpha > \frac{1}{3},$$

то монотонность возрастания  $x_{(k, k')}$  один раз нарушается: при переходе  $k + k'$  через значение  $(n + p)/2$  абсцисса  $x_{(k, k')}$  становится отрицательной, после чего опять монотонно возрастает. Нетрудно получить выражение для абсциссы  $x_{(p-1, p)}$ , отвечающей максимальному значению суммы  $k + k' = 2p - 1$ :

$$x_{(p-1, p)} = p \frac{1 - 2\alpha}{1 - 3\alpha}.$$

Видно, что при  $1/3 < \alpha \leq 1/2$  все абсциссы  $x_{(k, k')}$ , отвечающие значениям  $k + k'$ , большим чем  $(n + p)/2$ , останутся отрицательными. Если же  $\alpha > 1/2$ , то, по крайней мере,  $x_{(p-1, p)}$  опять попадет в область положительности  $x$ . Тем не менее легко проверяется, что и в этом случае  $x_{(p-1, p)}$  не может превзойти  $x_{(0, 1)}$ :

$$x_{(p-1, p)} < x_{(0, 1)}.$$

На рис. 4 приведены графики зависимости  $x_{(k, k')}$  от  $k + k'$  для значений параметра  $\alpha$ , меньших и больших  $1/3$ :

а) для  $\alpha \leq 1/3$  кривая  $x_{(k, k')}$  монотонно возрастает, оставаясь все время в области положительности  $x$ ;

б) для  $1/3 < \alpha$  монотонно возрастающая кривая  $x_{(k, k')}$  терпит разрыв при  $k + k' = (n + p)/2$ , после чего монотонное возрастание продолжается.

**2.** Для доказательства теоремы 2 проследим за изменением основного состояния системы по мере увеличения  $x$ . При  $x = 0$  основным является состояние с  $k = 0$  (см. цепочку неравенств для  $L_k(0)$ ), и это состояние невырождено. С увеличением  $x$  состояние

с  $k = 0$  будет оставаться основным до тех пор, пока прямая  $L_0(x)$  не пересечется с какой-либо из прямых  $L_{k'}$ . Очевидно, что ближайшее пересечение произойдет с прямой  $L_1(x)$  в точке

$$x_1 \equiv x_{(0,1)} = p \frac{n-1}{n+p-2},$$

после чего основным станет состояние с  $k = 1$  из класса  $\Sigma_1$ . Это состояние будет  $p$  раз вырожденным.

По мере дальнейшего увеличения  $x$  состояние с  $k = 1$  будет оставаться основным до тех пор, пока прямая  $L_1(x)$  не пересечется с какой-либо из прямых  $L_{k'}(x)$ . Очевидно, что ближайшее пересечение произойдет с прямой  $L_2(x)$  в точке

$$x_2 \equiv x_{(1,2)} = p \frac{n-3}{n+p-6},$$

после чего основным станет состояние с  $k = 2$  из класса  $\Sigma_2$ . Это состояние будет  $C_p^2$  раз вырожденным.

Учитывая свойства прямых  $L_k(x)$  и рассуждая аналогично, получаем, что по мере дальнейшего увеличения  $x$  перестройки основного состояния будут происходить в точках

$$x_k \equiv x_{(k-1,k)} = \frac{p}{2} \left[ 1 + \frac{q}{q + p(\cos w_{k-1} + \cos w_k)} \right] = p \frac{n - (2k - 1)}{n + p - 2(2k - 1)}$$

для  $k = 1, 2, 3, \dots$ , причем в точке  $k$ -й перестройки  $C_p^{k-1}$  раз вырожденное состояние из класса  $\Sigma_{k-1}$  сменяется  $C_p^k$  раз вырожденным состоянием из класса  $\Sigma_k$ .

Перестройки основного состояния будут происходить до тех пор, пока увеличению  $k$  будет отвечать возрастание  $x_k \equiv x_{(k-1,k)}$ . Если  $\alpha \leq 1/3$ , то по мере увеличения  $x$  одна за другой произойдут все  $p$  возможных перестроек основного состояния. Так что при дальнейшем увеличении  $x$  сколь угодно далеко за точку последней перестройки  $x_p \equiv x_{(p-1,p)}$  основным состоянием будет конфигурационный вектор

$$\left( \underbrace{-1, -1, \dots, -1}_p, 1, 1, \dots, 1 \right)$$

из класса  $\Sigma_p$ <sup>7)</sup>.

Если же  $\alpha > 1/3$ , то перестройки основного состояния будут происходить до тех пор, пока сумма  $2k - 1$  номеров прямых  $L_{k-1}(x)$  и  $L_k(x)$  будет оставаться меньше  $(n + p)/2$ ; как только  $2k - 1$  превзойдет  $(n + p)/2$ , абсцисса  $x_{(k-1,k)}$  пересечения прямых  $L_{k-1}$  и  $L_k$  станет отрицательной и, следовательно, дальнейшее увеличение  $x$  уже не будет сопровождаться пересечениями прямой  $L_{k-1}$  с какой-либо из прямых  $L_{k'}$ ; иначе говоря, состояние с индексом  $k - 1$  будет оставаться основным при неограниченном увеличении  $x$ .

<sup>7)</sup>Строго говоря, если  $\alpha = 1/3$ , то последняя перестройка основного состояния произойдет на бесконечности.

В этом случае номер  $k_{\max}$  последней перестройки основного состояния определяется из уравнения

$$2k_{\max} - 1 = \frac{n+p}{2} \implies k_{\max} = \left[ \frac{n+p+2}{4} \right]. \quad (\text{П.7})$$

Итак, в области положительных значений  $x$  последняя перестройка основного состояния произойдет при  $x_{k_{\max}}$ <sup>8)</sup>. При отрицательных значениях  $x$  никаких перестроек основного состояния быть не может, поскольку при  $x = 0$  прямая  $L_0(x)$  мажорирует все остальные прямые и имеет наибольшую по модулю производную.

**3.** Для завершения доказательства теоремы 2 покажем, что других локальных максимумов у функционала не будет. Конфигурационный вектор, являющийся локальным максимумом, с необходимостью имеет вид (7). Пусть он принадлежит классу  $\Sigma_k$  (см. (10)). Тогда при некотором  $x$  должны выполняться неравенства

$$L_k(x) \geq L_{k-1}(x), L_{k+1}(x). \quad (\text{П.8})$$

Раскрывая скобки и приводя подобные, получаем

$$x[2 \pm (n+p-4k)] \geq p[1 \pm (n-2k)]. \quad (\text{П.9})$$

Рассмотрим три случая.

А. Пусть  $n+p-4k \geq 2$ . Для номера  $k$  это означает, что

$$k+1 \leq \frac{n+p+2}{4},$$

иначе говоря, что  $k \in [1, k_{\max} - 1]$ . Знак плюс в (П.9) дает неравенство

$$x \geq p \frac{n-(2k-1)}{n+p-2(2k-1)} = x_k,$$

а знак минус – неравенство

$$x \leq p \frac{n-(2k+1)}{n+p-2(2k+1)} = x_{k+1}.$$

Таким образом, при  $k \in [1, k_{\max} - 1]$  состояние с номером  $k$  будет локальным максимумом только тогда, когда  $x$  принадлежит интервалу  $[x_k, x_{k+1}]$ , на котором это состояние является глобальным максимумом.

Б. Пусть теперь  $-2 \leq n+p-4k < 2$ . Для  $k$  это означает, что  $k > k_{\max} - 1$  и  $k \leq k_{\max}$ , иными словами, что  $k = k_{\max}$ . Знак плюс в (П.9) дает неравенство  $x \geq x_{k_{\max}}$ , а знак минус – неравенство  $x \geq x_{k_{\max}+1}$ . Однако мы знаем, что  $x_{k_{\max}+1}$  уже отрицательно

<sup>8)</sup> Если

$$\frac{n+p+2}{4} = \left[ \frac{n+p+2}{4} \right],$$

то последняя перестройка произойдет на бесконечности.



(см. выше п. 1), и, следовательно, одновременное выполнение двух последних неравенств приходится на область  $x \geq x_{k_{\max}}$ . А это и означает, что состояние с номером  $k_{\max}$  будет локальным максимумом функционала только тогда, когда оно является его глобальным максимумом.

В. Наконец, пусть  $n + p - 4k < -2$ , иными словами, пусть  $k > k_{\max}$ . Знак плюс в (П.9) дает неравенство  $x \leq x_k$ , а знак минус – неравенство  $x \geq x_{k+1}$ . Однако из п. 1 мы знаем, что даже при  $k > k_{\max}$  сохраняется упорядочение  $x_k < x_{k+1}$ . Следовательно, для  $k > k_{\max}$  одновременное выполнение неравенств (П.9) невозможно. Доказательство теоремы 2 закончено.

4. Область изменения параметра  $x$  от  $x_{\min} = x_1$  до  $x_{\max} = x_{k_{\max}}$  естественно называть *областью перестройки* основного состояния. Изучим ее.

С помощью  $\alpha$  из неравенства (П.6) легко получить выражение для левой границы области перестройки

$$x_{\min} = p \frac{n-1}{n+p-2} = \frac{p}{1+\alpha}$$

и убедиться в том, что  $x_{\min}$  – монотонно возрастающая функция как  $p$ , так и  $n$ . Выражение для правой границы зависит от соотношения между  $p$  и  $n$ . Если  $\alpha < 1/3$ , то

$$x_{\max} = p \frac{n+1-2p}{n+2-3p} = p \frac{1-2\alpha}{1-3\alpha}.$$

Как функция от  $n$  правая граница области перестройки убывает. Следовательно, при  $n \rightarrow \infty$  и постоянном  $p$  область перестройки стягивается в точку  $x'' = p$  (15б). Если же  $\alpha > 1/3$ , то

$$x_{\max} = p \frac{n-p+4\kappa}{8\kappa}, \quad \text{где } 4\kappa = 1, 2 \text{ или } 3.$$

При  $n$  постоянном и  $p \rightarrow n$  область перестройки и здесь стягивается в точку. Нетрудно видеть, что этой точкой будет  $x' = n/2$  (15а).

Точка  $x'' = p$  интересна еще и тем, что в ней меняет свою структуру наибольший собственный вектор матрицы  $\mathbf{N}$ : слева от  $x''$  все координаты собственного вектора положительны, а справа от  $x''$  первые  $p$  его координат становятся отрицательными (см. выражение для вектора  $\mathbf{S}^* \vec{f}_1$  в конце приложения 1). Можно определить, сколько перестроек основного состояния произойдет к моменту, когда  $x$  станет равным  $x'' = p$ . Несложные вычисления показывают, что максимальный номер перестройки при этом

$$k = \left[ \frac{p+1}{2} \right].$$

## Список литературы

- [1] Models of neural networks. Eds. E. Domany, J. L. van Hemmen, K. Schulten. Berlin: Springer-Verlag, 1991.
- [2] *U. Blasum, W. Hochstattler, C. Moll, H. Rieger*. Using network-flow techniques to solve an optimization problem from surface-physics. Cond-mat/9604052, 1996; *H. Rieger, U. Blasum*. Ground state properties of solid-on-solid models with disordered substrates. Cond-mat/9608136, 1996.
- [3] *N. Sourlas*. Europhys. Lett. 1994. V. 25. P. 159.
- [4] *G. Parisi*. Attractor Neural Networks. Cond-mat/9412030, 1994.
- [5] *Э. М. Браверман, И. Б. Мучник*. Структурные методы обработки эмпирических данных. М.: Наука, 1983.
- [6] *Л. Б. Литинский*. Математическое моделирование. 1996. Т. 8. С. 11.
- [7] *L. B. Litinsky*. Energy functional and fixed points of neural network. Cond-mat/9706280, 1997.
- [8] *J. J. Hopfield*. Proc. Nat. Acad. Sci. USA. 1982. V. 79. P. 2554.
- [9] *P. Baldi*. IEEE Trans. on Inform. Theory. 1988. V. 34. P. 523.
- [10] *T. L. H. Watkin, A. Rau, M. Biehl*. Rev. Mod. Phys. 1993. V. 65. P. 499.
- [11] *B. Muller, J. Reinhardt, M. T. Strikland*. Neural Networks. An Introduction. Sec. edit. Berlin: Springer-Verlag, 1995.
- [12] *S. Diderich, M. Oppen*. Phys. Rev. Lett. 1987. V. 58. P. 949.
- [13] *Л. Б. Литинский*. ТМФ. 1994. Т. 101. С. 467.
- [14] *A. A. Vedenov, A. A. Ezhov et al.* Structure of patterns in associative memory models. In: Neural Networks—Theory and Architecture. Eds. A. V. Holden, V. I. Kryukov. Manchester: Manch. Univ. Press, 1991. P. 169.
- [15] *Л. Б. Литинский*. Учет магнитного поля для модели Изинга с высокосимметричной матрицей связей. (Готовится к публикации.)
- [16] *L. B. Litinsky*. Solid State Commun. 1991. V. 79. P. 879.

Поступила в редакцию 4.VI.1998 г.