

МАТЕМАТИЧЕСКИЙ АРХИВ

В. А. Якубович

МАШИНЫ, ОБУЧАЮЩИЕСЯ РАСПОЗНАВАНИЮ ОБРАЗОВ. II¹§ 5. ОБ ОДНОМ АЛГОРИФМЕ РАЗДЕЛЕНИЯ
ВЫПУКЛЫХ МНОЖЕСТВ

Простейший алгоритм разделения выпуклых множеств описывается формулами (3.5), (3.7). Этот алгоритм эффективен лишь в том случае, если разделяемые множества находятся в пространстве образов достаточно далеко друг от друга и тренировочная последовательность достаточно велика. В противном случае, как показывают элементарные геометрические примеры, вырабатываемая этим алгоритмом гиперплоскость может не разделять даже образы тренировочной последовательности (рис. 3).

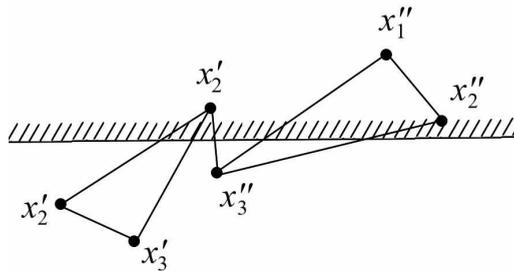


Рис. 3.

Основная часть алгоритма обучения состоит в определении по заданным $x'_j, x''_h \in X$ числа α и вектора a так, чтобы выполнялись неравенства

$$\psi(x'_j) \equiv \alpha + (x'_j, a) > 0, \quad \psi(x''_h) \equiv \alpha + (x''_h, a) < 0$$

$$(j = 1, \dots, m_1, \quad h = 1, \dots, m_2).$$
(5.1)

Этот алгоритм должен доставлять не многообразие всех решений неравенств (5.1), а какое-либо одно решение, притом такое, чтобы имела место теорема, аналогичная

¹ Первая часть статьи опубликована в предыдущем выпуске журнала: Якубович В. А. Машины, обучающиеся распознаванию образов. I // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2021. Т. 8 (66). Вып. 4. С. 625–638. <https://doi.org/10.21638/spbu01.2021.408>. Нумерация глав, теорем, формул, рисунков и пр. продолжается из первой части статьи.

© Санкт-Петербургский государственный университет, 2022

теореме 1 (см. § 3). Стандартные методы динамического программирования здесь, по-видимому, мало пригодны.

Предлагаемый ниже алгоритм обладает тем свойством, что вырабатываемая им гиперплоскость разделяет выпуклые замкнутые оболочки S'_0, S''_0 множеств $\{x'_j\}$ и $\{x''_h\}$. Другим его достоинством является возможность относительно простого технического осуществления.

Положим

$$x(\sigma)' = \sum_{j=1}^{m_1} \sigma_j x'_j, \quad x''(\tau) = \sum_{h=1}^{m_2} \tau_h x''_h, \quad (5.2)$$

где числа τ_h, σ_j подчинены условиям

$$0 \leq \sigma_j \leq 1, \quad 0 \leq \tau_h \leq 1, \quad (5.3)$$

$$\sum_{j=1}^{m_1} \sigma_j = 1, \quad \sum_{h=1}^{m_2} \tau_h = 1. \quad (5.4)$$

При изменении чисел $\sigma_j(\tau_h)$ с сохранением соотношений (5.3), (5.4) вектор $x'(\sigma)(x''(\tau))$ описывает выпуклую замкнутую оболочку, натянутую на точки $x'_j(x''_h)$. Пусть²

$$\min_j |x' - x''| = |x'_0 - x''_0|, \quad (5.5)$$

где

$$x''_0 = \sum_j \sigma_j^0 x''_j, \quad x'_0 = \sum_h \tau_h^0 x'_h, \quad (5.6)$$

и числа σ_j^0, τ_h^0 удовлетворяют условиям (5.3), (5.4). (Числа σ_j^0, τ_h^0 и векторы x''_0, x'_0 могут определяться не единственным образом, даже если $S'_0 \cap S''_0 = \Lambda$.)

Если множества $S' \supset S'_0 = \{x'_j\}$ и $S'' \supset S''_0 = \{x''_h\}$ могут быть разделены перцептроном с одним суммирующим элементом (с.э.), то, очевидно, $S'_0 \cap S''_0 = \Lambda$. Полагая, как и выше,

$$\begin{aligned} \alpha &= -\frac{1}{2}(x'_0 + x''_0, x'_0 - x''_0), \\ \alpha &= \sum_{j=1}^{N_a} \alpha_j a_j = x'_0 - x''_0, \\ \psi(x) &= \alpha + (a, x), \end{aligned} \quad (5.7)$$

получим, что будут выполнены соотношения

$$\begin{aligned} \psi(x) &> 0 \text{ при } x \in S'_0, \\ \psi(x) &< 0 \text{ при } x \in S''_0 \end{aligned} \quad (5.8)$$

и тем более соотношения (5.1). Задача, таким образом, состоит в отыскании чисел σ_j^0, τ_j^0 .

Рассмотрим вначале случай $m_1 + m_2 \leq n + 1$. Обозначим L' и L'' гиперплоскости, натянутые на векторы x'_j, x''_h соответственно. Когда числа σ_j (τ_h) изменяются произвольным образом с сохранением первого (второго) соотношения (5.4) (но не соотношения (5.3)), вектор $x'(\sigma)$ ($x''(\tau)$) пробегает многообразие L' (L''). Хорошо известно (ниже это будет показано), что при $m_1 + m_2 \leq n - 1$ многообразия L' и L'' , как правило, не пересекаются³. В этом случае трудную задачу об отыскании чисел

² По-видимому, левая часть (5.5) должна иметь вид $\min_{\sigma, \tau} |x'(\sigma) - x''(\tau)|$ (Прим. А. Л. Фрадкова).

³ Они располагаются примерно как две прямые в трехмерном пространстве, когда $m_1 = 2, m_2 = 2, n = 3$.

σ_j^0, τ_h^0 , когда минимум в (5.5) берется по множеству (5.3), (5.4), можно заменить на значительно более легкую задачу отыскания чисел σ_j^0, τ_h^0 , реализующих минимум между многообразиями L', L'' .

Составим систему дифференциальных уравнений

$$\frac{d\sigma_j}{dt} = \lambda_j, \quad \frac{d\tau_h}{dt} = \mu_h \quad (j = 1, \dots, m_1, \quad h = 1, \dots, m_2), \quad (5.9)$$

где

$$\begin{aligned} \lambda_j &= -\frac{\partial \rho}{\partial \sigma_j} + \frac{1}{m_1} \sum_{k=1}^{m_1} \frac{\partial \rho}{\partial \sigma_k}, \\ \mu_h &= -\frac{\partial \rho}{\partial \tau_h} + \frac{1}{m_2} \sum_{k=1}^{m_2} \frac{\partial \rho}{\partial \tau_k}, \\ \rho &= \left| \sum_{j=1}^{m_1} \sigma_j x'_j - \sum_{h=1}^{m_2} \tau_h x''_h \right|^2. \end{aligned} \quad (5.10)$$

Будем решать эту систему в предположении, что начальные условия удовлетворяют соотношениям (5.4). Поскольку из (5.9), (5.10) следует $\sum \sigma_j = 0, \sum \tau_h = 0$, то соотношения (5.4) будут выполнены тождественно.

Теорема 3. *Предположим, что $m_1 + m_2 \leq n + 1$ и что при $m_1 + m_2 \leq n$ векторы $x'_1, \dots, x'_{m_1}, x''_1, \dots, x''_{m_2}$ линейно независимы и при $m_1 + m_2 = n + 1$ отличен от нуля определитель $(n + 2) \times (n + 2)$ -матрицы*

$$A = \begin{vmatrix} x'_1 & \dots & x'_{m_1} & x''_1 & \dots & x''_{m_2} & 0 \\ & & & & & & \vdots \\ & & & & & & 0 \\ 1 & \dots & 1 & 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & -1 & \dots & -1 & -1 \end{vmatrix}$$

Тогда для любого решения системы (5.9) такого, что начальные значения удовлетворяют (5.4), существуют пределы

$$\lim_{t \rightarrow \infty} \sigma_j(t) = \sigma_j^0, \quad \lim_{t \rightarrow \infty} \tau_h(t) = \tau_h^0, \quad (5.11)$$

где σ_j^0, τ_h^0 — значения, удовлетворяющие (5.4), (5.5), (5.6). При этом выполнены соотношения (5.8), где $\psi(x)$ определяется по числам σ_j^0, τ_h^0 формулами (5.6), (5.7).

ДОКАЗАТЕЛЬСТВО. Покажем, что многообразия L' и L'' не пересекаются и что существует единственный набор значений σ_j^0, τ_h^0 , реализующий минимальное расстояние

$$\min_{\sigma_j, \tau_h} \sqrt{\rho} = |x'(\sigma_1^0, \dots, \sigma_{m_1}^0) - x''(\tau_1^0, \dots, \tau_{m_2}^0)|$$

между многообразиями L' и L'' . При $m_1 + m_2 \leq n$ из равенства

$$\sum_{j=1}^{m_1} \sigma_j x'_j = \sum_{h=1}^{m_2} \tau_h x''_h \quad (5.12)$$

в силу линейной независимости векторов x'_j, x''_h следует $\sigma_j = 0, \tau_h = 0$, что противоречит (5.4). При $m_1 + m_2 = n + 1$ соотношения (5.12), (5.4) могут быть записаны в виде $Az = 0$, где z — вектор с компонентами

$$\sigma_1, \dots, \sigma_{m_1}, \tau_1, \dots, \tau_{m_2}, 1.$$

Так как $\det A \neq 0, z \neq 0$, то равенство $Az = 0$ не может быть выполнено. Таким образом, многообразия L' и L'' не пересекаются.

Если

$$\sum_{j=1}^{m_1} |\sigma_j| \rightarrow \infty \text{ или } \sum_{h=1}^{m_2} |\tau_h| \rightarrow \infty,$$

то $\rho \rightarrow \infty$. Следовательно, существуют значения σ_j^0, τ_h^0 , для которых ρ достигает минимума. Эти значения должны, согласно правилу Лагранжа, удовлетворять уравнениям

$$\frac{\partial \rho}{\partial \sigma_j} = \lambda, \quad \frac{\partial \rho}{\partial \tau_k} = \mu \quad (j = 1, \dots, m_1, \quad h = 1, \dots, m_2), \quad (5.13)$$

или в развернутой форме уравнениям

$$\begin{aligned} \sum_{k=1}^{m_1} \sigma_k(x'_k, x'_j) - \sum_{k=1}^{m_2} \tau_k(x''_k, x'_j) &= \lambda, \\ \sum_{k=1}^{m_1} \sigma_k(x'_k, x''_h) - \sum_{k=1}^{m_2} \tau_k(x''_k, x''_h) &= \mu \end{aligned} \quad (5.14)$$

и уравнениям (5.4). Определителем системы (5.14) относительно величин $\sigma_j(-\tau_h)$ является определитель Грамма векторов x'_j, x''_h . При $m_1 + m_2 \leq n$ в силу линейной независимости векторов x'_j, x''_h этот определитель отличен от нуля. При $m_1 + m_2 = n + 1$ систему (5.14) можно, воспользовавшись соотношениями (5.4), преобразовать к виду

$$\begin{aligned} \sum_{k=1}^{m_1} \sigma_k(a'_k, a'_j) - \sum_{k=1}^{m_2} \tau_k(a''_k, a'_j) &= \lambda, \\ \sum_{k=1}^{m_1} \sigma_k(a'_k, a''_h) - \sum_{k=1}^{m_2} \tau_k(a''_k, a''_h) &= \mu, \end{aligned}$$

где $a'_1, \dots, a'_{m_1}, a''_1, \dots, a''_{m_2}$ — первые $m_1 + m_2$ столбцов матрицы A . Определитель последней системы также отличен от нуля в силу линейной независимости векторов a'_j, a''_h . Таким образом, значения σ_j^0, τ_h^0 определяются единственным образом.

Уравнения (5.9) задают на многообразии (5.4) в пространстве $[\sigma_1, \dots, \sigma_{m_1}, \tau_1, \dots, \tau_{m_2}]$ динамическую систему, для которой точка $\sigma_1^0, \dots, \sigma_{m_1}^0, \tau_1^0, \dots, \tau_{m_2}^0$ является, согласно соотношениям (5.13), точкой покоя. Взяв в качестве функции Ляпунова функцию $Y = \rho - \rho_{\min}$, найдем, что

$$\dot{Y} = - \sum_{j=1}^{m_1} \left(\frac{\partial \rho}{\partial \sigma_j} \right)^2 + \frac{1}{m_1} \left(\sum_{j=1}^{m_1} \frac{\partial \rho}{\partial \sigma_j} \right)^2 - \sum_{h=1}^{m_2} \left(\frac{\partial \rho}{\partial \tau_h} \right)^2 + \frac{1}{m_2} \left(\sum_{h=1}^{m_2} \frac{\partial \rho}{\partial \tau_h} \right)^2 \leq 0$$

по неравенству Шварца — Буняковского. При этом из $\dot{Y} \equiv 0$ следует, что выполнены соотношения (5.13) для некоторых λ, μ и, в силу единственности значений σ_j^0, τ_h^0 , что

$\sigma_j(t) \equiv \sigma_j^0, \tau_h(t) \equiv \tau_h^0$. Применяя, например, теорему 3 (см. [11]), получим утверждение теоремы относительно существования пределов (5.11).

Пусть $r > 0$ настолько велико, что

$$|x'_j| \leq r, \quad |x''_h| \leq r \quad (j = 0, 1, \dots, m_1; \quad h = 0, 1, \dots, m_2),$$

где, напомним, x'_0, x''_0 определяются формулами (5.6). Мы имеем

$$\min |x' - x''| = |x'_0 - x''_0|,$$

где минимум берется по всем

$$x' \in L'_0 = L' \cap E(|x| \leq r), \quad x'' \in L''_0 = L'' \cap E(|x| \leq r).$$

Множества L'_0, L''_0 выпуклы. Из доказательства теоремы 1 § 3 следует, что

$$\psi(x) > 0 \text{ при } x \in L'_0,$$

$$\psi(x) < 0 \text{ при } x \in L''_0.$$

Так как $S'_0 \subset L'_0, S''_0 \subset L''_0$, то тем более выполнено (5.8). Теорема доказана.

Предположение теоремы $m_1 + m_2 \leq n$ не является серьезным ограничением, так как размерность пространства образов n обычно очень велика. Поскольку образы x'_j, x''_h тренировочной последовательности выбираются наугад, то практически всегда выполнены предположения относительно определителя матрицы A или относительно линейной независимости векторов x'_j, x''_h . Тем не менее алгоритм, доставляемый теоремой 3, не является достаточно удовлетворительным. Дело в том, что при случайном выборе векторов x'_j, x''_h может оказаться, что многообразия L' и L'' проходят близко друг от друга. В этом случае разделяющая плоскость $\psi(x) = 0$ будет также близка к L' и L'' и весьма вероятно, что эта плоскость пересечет множества S', S'' . Таким образом, хотя плоскость $\psi(x) = 0$ по теореме 1 всегда разделит тренировочные последовательности, довольно вероятен случай, когда эта плоскость не будет разделять множества S' и S'' ⁴. Этот плохой случай не будет иметь место, когда

$$|x'_0 - x''_0| \ll |x'_j - x''_h| \quad (j = 1, \dots, m_1; \quad h = 1, \dots, m_2).$$

Проверив это неравенство, например, для выбранных наугад j и h , можно (разумеется, очень приближенно) установить, имеет ли место указанный плохой случай.

Этот недостаток явился следствием того обстоятельства, что мы заменили задачу разделения выпуклых замкнутых оболочек S'_0, S''_0 более простой задачей разделения многообразий L', L'' .

Рассмотрим теперь задачу разделения множеств S'_0, S''_0 .

Пусть α — какое-нибудь подмножество во множестве чисел $1, \dots, m_1$ и β — какое-нибудь подмножество во множестве чисел $1, \dots, m_2$. Обозначим $m(\alpha)$ ($m(\beta)$) число элементов множества α (β), и пусть

$$\begin{aligned} \mu_j^\alpha &= -\frac{\partial \rho}{\partial \sigma_j} + \frac{1}{m(\alpha)} \sum_{k \in \alpha} \frac{\partial \rho}{\partial \sigma_k}, \\ \nu_j^\beta &= -\frac{\partial \rho}{\partial \tau_h} + \frac{1}{m(\beta)} \sum_{k \in \beta} \frac{\partial \rho}{\partial \tau_k}. \end{aligned} \tag{5.15}$$

⁴ Персептрон, блок обучения которого работает по алгоритму, порождаемому теоремой 3, будет безошибочно разделять образы, близкие к тренировочной последовательности, но он плохо «обобщает» — он может ошибаться, когда ему показывают образы, далекие от образов тренировочной последовательности.

Пусть $P(\sigma_1, \dots, \sigma_{m_1}; \tau_1, \dots, \tau_{m_2})$ — какая-либо точка в многообразии R , определяемом формулами (5.3), (5.4), α — множество индексов координат $\sigma_j \neq 0$ и β — множество индексов координат $\tau_h \neq 0$. Точку P будем называть особой, если

$$\mu_j^\alpha = 0, \quad \nu_h^\beta = 0 \quad \text{для всех } j \in \alpha, \quad h \in \beta,$$

и неособой в противном случае. Отметим, что в число особых точек входят все $m_1 + m_2$ вершин

$$\begin{aligned} \sigma_1 = 1, \sigma_2 = 0, \dots, \sigma_{m_1} = 0, \tau_1 = 0, \dots, \tau_{m_2} = 0, \\ \dots \\ \sigma_1 = 0, \sigma_2 = 0, \dots, \sigma_{m_1} = 0, \tau_1 = 0, \dots, \tau_{m_2} = 1. \end{aligned}$$

Во множестве неособых точек определим систему дифференциальных уравнений формулами⁵

$$\begin{aligned} \dot{\sigma}_j &= \frac{1}{\xi_{\alpha\beta}} \mu_j^\alpha \quad \text{при } j \in \alpha, \quad \dot{\sigma}_j = 0 \quad \text{при } j \in \bar{\alpha}, \\ \dot{\tau}_h &= \frac{1}{\xi_{\alpha\beta}} \nu_h^\beta \quad \text{при } h \in \beta, \quad \dot{\tau}_h = 0 \quad \text{при } h \in \bar{\beta}, \end{aligned} \quad (5.16)$$

где

$$\xi_{\alpha\beta} = \left[\sum_{j \in \alpha} (\mu_j^\alpha)^2 + \sum_{h \in \beta} (\nu_h^\beta)^2 \right]^{1/2}. \quad (5.17)$$

В особых точках определим $\dot{\sigma}_j, \dot{\tau}_h$ согласно следующему алгоритму.

Пусть

$$\alpha_1 = (1, \dots, m_1), \alpha_2, \alpha_3, \dots, \alpha_{k_1} = \alpha \quad (5.18)$$

— всевозможные подмножества множества $(1, \dots, m_1)$, содержащие множество α , упорядоченные произвольным образом с сохранением условия

$$m(\alpha_1) \geq m(\alpha_2) > \dots \geq m(\alpha_{k_1}),$$

и

$$\beta_1 = (1, \dots, m_2), \beta_2, \beta_3, \dots, \beta_{k_2} = \beta \quad (5.19)$$

— аналогичным образом определенные подмножества, содержащие множества β .

Пусть α_0 — первое из подмножеств в ряду (5.18) такое, что все $\mu_j^{\alpha_0} \geq 0$ при $j \in \alpha_0$ и β_0 — первое из подмножеств в ряду (5.19) такое, что все $\nu_h^{\beta_0} \geq 0$ при $h \in \beta_0$. Отметим, что $\alpha_0 \supseteq \alpha, \beta_0 \supseteq \beta$. В особых точках заменим систему (5.16) следующей:

$$\begin{aligned} \dot{\sigma}_j &= \frac{1}{\xi_{\alpha_0\beta_0}} \mu_j^{\alpha_0} \quad \text{при } j \in \alpha_0, \quad \dot{\sigma}_j = 0 \quad \text{при } j \in \bar{\alpha}_0, \\ \dot{\tau}_h &= \frac{1}{\xi_{\alpha_0\beta_0}} \nu_h^{\beta_0} \quad \text{при } h \in \beta_0, \quad \dot{\tau}_h = 0 \quad \text{при } h \in \bar{\beta}_0, \end{aligned} \quad (5.20)$$

⁵ Отметим, что если бы мы определили систему дифференциальных уравнений формулами (5.16) во всем множестве R , то особые точки были бы точками покоя этой системы.

Первые части системы (5.16)–(5.20) разрывны в каждой точке $P_0 \subset R$, для которой хотя бы один из σ_j или τ_h равен нулю.

Тем не менее легко убедиться, что для каждой такой точки P_0 существует кривая, входящая в эту точку, такая, что при стремлении точки P по кривой к точке P_0 правые части системы (5.16)–(5.20) меняются непрерывно.

Под решением систем (5.16)–(5.20) будем понимать абсолютно-непрерывные функции $\tau_j(t), \sigma_h(t)$, удовлетворяющие системам (5.16)–(5.20) почти всюду⁶.

Теорема 4. *Предположим, что существует единственная точка $(\sigma_1^0, \dots, \sigma_{m_1}^0; \tau_1^0, \dots, \tau_{m_2}^0)$ такая, что*

$$\min_R \rho = \rho(\sigma_1^0, \dots, \sigma_{m_1}^0, \tau_1^0, \dots, \tau_{m_2}^0). \quad (5.21)$$

Для произвольных начальных значений

$$\sigma_1(0), \dots, \sigma_{m_1}(0), \tau_1(0), \dots, \tau_{m_2}(0),$$

удовлетворяющих соотношениям (5.3), (5.4), существует решение систем (5.16)–(5.20), удовлетворяющее при любых значениях t соотношениям (5.3), (5.4), и число T такое, что

$$\begin{aligned} \sigma_j(t) = \sigma_j^0, \tau_h(t) = \tau_h^0 \text{ при } t \geq T \\ (j = 1, \dots, m_1, h = 1, \dots, m_2). \end{aligned} \quad (5.22)$$

Если $\psi(x)$ определяется по числам σ_j^0, τ_h^0 формулами (5.7), (5.6), то будут выполнены соотношения (5.8).

Замечание. Предположение теоремы о единственности точки $(\sigma_1^0, \dots, \tau_{m_2}^0)$ практически всегда выполнено, так как векторы x'_j, x''_h выбираются случайным образом. Если точка $(\sigma_1^0, \dots, \tau_{m_2}^0)$ не единственна, то справедливо утверждение теоремы с заменой (5.22) на

$$r[(\sigma_1(t), \dots, \tau_{m_2}(t)); M_0] \rightarrow 0 \text{ при } t \rightarrow \infty,$$

где M_0 — множество всех точек $(\sigma_1^0, \dots, \tau_{m_2}^0)$, удовлетворяющих (5.21), и $r[(\sigma_1(t), \dots, \tau_{m_2}(t)); M_0]$ — расстояние от точки $P(\sigma_1(t), \dots, \tau_{m_2}(t))$ до множества M_0 .

Строгое доказательство теоремы достаточно сложно; ограничимся следующим пояснением. Введение выше множества α, β определяют грань $\sigma_j = 0, \tau_h = 0, j \in \alpha, h \in \beta$ многогранника R . Можно показать, что оптимальное движение по этой грани, при котором ρ уменьшается более всего, задается системой (5.16). Перемещаясь согласно системе (5.16), точка P придет за конечное время (конечность обуславливается введением множителя $\xi_{\alpha\beta}$) в одну из особых точек P_0 . Поскольку оптимальное движение по грани α, β может быть не оптимальным движением по какой-либо объемлющей грани $\alpha_1 \supseteq \alpha, \beta_1 \supseteq \beta$, то в точке P_0 , вообще говоря, может не реализоваться даже локальный в R минимум функции ρ . Если локального минимума нет, то система (5.20) выведет точку P из особой и далее вступит в силу снова система (5.16). При таком движении постоянно уменьшается ρ . Поэтому точка P приходит в конце концов к точке P_0 , реализующей локальный минимум, который в силу свойств функции ρ будет общим минимумом в R .

⁶ Поскольку правые части систем (5.16)–(5.20) — разрывные функции, производные $\dot{\sigma}_j, \dot{\tau}_h$, как правило, всюду не могут существовать. Отметим, что приведенное определение решения существенно отличается от определения [12].

Из этих рассуждений можно сделать еще следующий вывод. Неудобную для реализации систему (5.20) можно исключить, введя взамен следующее правило. Если точка $P_0(\sigma_{10}, \dots, \tau_{m20})$ — особая, то ее следует заменить точкой $P(\sigma_1, \dots, \tau_{m2})$ с координатами

$$\sigma_j = \sigma_{j0} + \Delta'_j, \quad \tau_h = \tau_{h0} + \Delta''_h,$$

где Δ'_j, Δ''_h — какие-либо случайные, достаточно малые числа, такие, что выполнены соотношения (5.3), (5.4). Дальнейшее движение продолжается в силу системы (5.16). Если система (5.16) приведет точку P снова в ту же точку P_0 и так будет некоторое фиксированное число q раз, то полагаем

$$\sigma_j^0 = \sigma_{j0}, \quad \tau_h^0 = \tau_{h0}.$$

Система (5.9), а также (5.16)–(5.20) или (5.16) с указанным правилом могут быть промоделированы. На вход соответствующего блока должны подаваться параметры тренировочной последовательности $x'_1, x'_2, \dots, x'_m, x''_1, x''_2, \dots, x''_m$; выходом являются веса ассоциативных элементов (а. э.), связанных с фиксированным с. э. Этот блок является блоком обучения персептрона с одним с. э. и составной частью блока обучения для персептрона со многими с. э.

При моделировании процесса обучения на ЭВМ системы (5.9), (5.16)–(5.20) должны быть заменены какой-либо конечно-разностной схемой. Поскольку точность в этих расчетах не играет роли, можно ограничиться простейшим методом Эйлера. Система (5.9) породит тогда следующий алгоритм.

1⁰. По заданным $x'_1, x'_2, \dots, x'_m, x''_1, x''_2, \dots, x''_m$ находим

$$x' = \sum_{j=1}^m \tau_j x'_j, \quad x'' = \sum_{j=1}^m \sigma_j x''_j,$$

где

$$\tau_j = \frac{1}{m}, \quad \sigma_j = \frac{1}{m},$$

или определяются каким-либо иным способом, удовлетворяющим условию (5.4)⁷.

2⁰. Определяем числа

$$\alpha'_j = (x'_j, x' - x''), \quad \alpha''_j = (x''_j, x' - x''),$$

$$\alpha_0 = -\frac{1}{2}(x' + x'', x' - x''),$$

$$\psi(x'_j) = \alpha_0 + \alpha'_j, \quad \psi(x''_j) = \alpha_0 + \alpha''_j.$$

Если

$$\psi(x'_j) > 0, \quad \psi(x''_j) < 0 \quad (j = 1, 2, \dots, m), \quad (5.23)$$

то x', x'' найдены удовлетворительно и вычисления на этом заканчиваются.

3⁰. Если хотя бы одно из неравенств (5.23) не выполнено, то определяем

$$\alpha' = \frac{1}{m} \sum_{j=1}^m \alpha'_j, \quad \alpha'' = \frac{1}{m} \sum_{j=1}^m \alpha''_j,$$

⁷ Полученные x', x'' соответствуют начальным значениям решения системы (5.9); они могут быть выбраны произвольно, лишь бы $x' \in \{x'_j\}, x'' \in \{x''_j\}$.

$$\begin{aligned}\bar{x}' &= x' + \delta \sum_{j=1}^m (\alpha'_j - \alpha') x'_j, \\ \bar{x}'' &= x'' + \delta \sum_{j=1}^m (\alpha''_j - \alpha'') x''_j,\end{aligned}\tag{5.24}$$

где $\delta > 0$ — некоторое достаточно малое число. (Контролем может служить неравенство

$$|\bar{x}' - \bar{x}''|^2 \leq |x' - x''|^2.$$

Если это неравенство не выполнено, то число δ нужно уменьшить.)

4⁰. Возвращаемся к п. 2⁰, беря вместо x', x'' векторы \bar{x}', \bar{x}'' .

Для систем (5.16)–(5.20) получается аналогичный алгоритм.

Можно привести и другие, связанные с дифференциальными уравнениями алгоритмы разделения выпуклых множеств.

§ 6. РАЗДЕЛЕНИЕ ПРОИЗВОЛЬНОГО ЧИСЛА, ВООБЩЕ ГОВОРЯ, НЕВЫПУКЛЫХ МНОЖЕСТВ. БЛОК ОБУЧЕНИЯ ПЕРСЕПТРОНА СО МНОГИМИ С. Э.

Будем исходить из представления о том, что понятиям соответствуют достаточно «хорошие» множества в пространстве образов, именно множества, которые можно разбить на конечное число выпуклых ограниченных множеств⁸. Разумеется, это представление формально не используется в дальнейшем изложении.

Будем называть сейчас тренировочной последовательностью произвольную последовательность $k_1 + k_2$ образов, разбитых на две группы:

$$M = x'_1, x'_2, \dots, x'_{k_1}, \quad N = x''_1, x''_2, \dots, x''_{k_2}.$$

При разделении двух множеств S_1 и S_2 имеем $M \subset S_1, N \subset S_2$.

Предположим, что имеется алгоритм A , вырабатывающий по произвольной тренировочной последовательности линейный неоднородный функционал

$$\psi(x) = \alpha + (a, x),$$

такой, что множества

$$E[\psi(x) > 0], \quad E[\psi(x) < 0]$$

не пусты. (В идеальном случае первое множество содержит M , а вторая последовательность — N или наоборот.) Этим алгоритмом может быть любой из описанных выше алгоритмов. Можно, например, проводить плоскость $\psi(x) = 0$ через середину отрезка, соединяющего центры тяжести этих множеств, ортогонально к этому отрезку, т. е. определять α, a формулами

$$\begin{aligned}x' &= \frac{1}{k_1} \sum_{j=1}^{k_1} x'_j, & x'' &= \frac{1}{k_2} \sum_{j=1}^{k_2} x''_j, \\ \alpha &= \frac{1}{2}(x'' - x', x'' + x'), & a &= x' - x''\end{aligned}$$

и т. п.

⁸ Более точная формулировка: элементарному понятию соответствует множество, которое можно представить как сумму выпуклых тел, пересекающихся, может быть, лишь по границе.

Для любого конечного множества N будем обозначать

$$N_0 = N \cap E[\varphi(x) < 0], \quad N_1 = N \cap E[\varphi(x) \geq 0],$$

$N_{\alpha\beta} = (N_\alpha)_\beta$; $\alpha, \beta = 0, 1$ и через mN будем обозначать число элементов множества N .

Будем говорить, что алгоритм A удовлетворяет условию разделения, если из $mM_0 \leq mM_1$ следует, что $mN_0 \geq mN_1$ и из $mM_0 \geq mM_1$ следует $mN_0 \leq mN_1$. (Здесь множества M и N равноправны.) Любой алгоритм можно дополнить так, чтобы полученный алгоритм удовлетворял условию разделения. Алгоритм, который мы рассматриваем ниже, может не удовлетворять условию разделения.

Рассмотрим сначала случай, когда должны быть разделены два множества $S_1 \supset M$ и $S_2 \supset N$.

Алгоритм разделения множеств M и N состоит в следующем.

1⁰. Применяем ко множеству M, N алгоритм A , порождающий функцию $\psi(x)$ и множества M_0, N_0, M_1, N_1 ($M = M_0 \cup M_1, M_0 \cap M_1 = A$), и аналогично для множества N .

2⁰. Если одно из множеств в паре M_0, N_0 или в паре M_1, N_1 пусто, то запоминаем название соответствующего непустого множества.

3⁰. К каждой паре M_μ, N_μ с одинаковым индексом, в которой оба множества не пусты, применяем алгоритм A , порождающий функцию $\psi_\mu(x)$ и множества $M_{\mu 0}, N_{\mu 0}, M_{\mu 1}, N_{\mu 1}$.

4⁰. Обозначая через μ полученные в п. 3⁰ индексы ($\mu 0$) или ($\mu 1$), запоминаем те множества $M_\mu(N_\mu)$, для которых множество $N_\mu(M_\mu)$ с тем же индексом пусто, и переходим к п. 3⁰.

Если этот процесс заканчивается, то мы получим в результате запомненные множества

$$M'_{\mu 1}, M'_{\mu 2}, \dots, M'_{\mu r}, \tag{6.1}$$

$$N''_{\mu 1}, N''_{\mu 2}, \dots, N''_{\mu s}, \tag{6.2}$$

причем ни один из индексов ряда (6.1) не встречается в ряду (6.2). Процесс обязательно заканчивается, если алгоритм A удовлетворяет условию разделения. Действительно, в этом случае

$$\min[mM_\mu, mN_\mu] \leq \frac{\max(k_1, k_2)}{2^q},$$

где q — число цифр индекса μ и процесс разделения потребует не более

$$\log_2 \max(k_1, k_2)$$

возвращений к п. 3⁰.

Сопоставим каждой функции $\psi_\mu(x)$ логическую переменную

$$\pi_\mu(x) = \begin{cases} 0 & \text{при } \psi_\mu(x) < 0, \\ 1 & \text{при } \psi_\mu(x) \geq 0 \end{cases}$$

и аналогично для функции $\psi(x)$ — логическую переменную $\pi(x)$. Сопоставим далее индексу

$$\mu = j_1, j_2, \dots, j_{q-1}, j_q \quad (j_i = 0, 1, i = 1, 2, \dots, q)$$

функцию алгебры логики

$$\xi_\mu = \pi^{j_1} \pi_{j_1}^{j_2} \pi_{j_1, j_2}^{j_3}, \dots, \pi_{j_1, j_2, \dots, j_{q-1}}^{j_q}.$$

Очевидно, $\xi_\mu = \xi_\mu(x)$.

Обозначим через $[M]$ множество индексов в ряду (6.1), через $[N]$ — множество индексов в ряду (6.2) и построим функции

$$\xi_M(x) = \bigvee_{\mu \in [M]} \xi_\mu(x), \quad \xi_N(x) = \bigvee_{\mu \in [N]} \xi_\mu(x).$$

Из изложенного следует, что

$$\xi_M(x)\xi_N(x) \equiv 0, \quad 1 = \xi_M(x) \vee \xi_N(x) \quad (6.3)$$

и

$$\begin{aligned} \xi_M(x) &= 1 \text{ при } x \in M, \\ \xi_N(x) &= 1 \text{ при } x \in N. \end{aligned}$$

Уравнения $\xi_M(x) = 1$, $\xi_N(x) = 1$ выделяют в пространстве образов непересекающиеся полиэдры, аппроксимирующие множества $S_1 \supset M$ и $S_2 \supset N$, соответствующие разделяемым понятиям. Поэтому естественно считать, что перцептрон отвечает « $x \in S_1$ », если $\xi_M(x) = 1$, и « $x \in S_2$ », если $\xi_N(x) = 1$. Согласно (6.3) любой образ будет отнесен к понятию S_1 и S_2 и противоречивых ответов быть не может. Разумеется, ответ может быть и неверным. Интуитивно ясно, однако, что при достаточно «хороших» множествах S_1 , S_2 , достаточно длинной тренировочной последовательности и достаточно «хорошем» обучении (хотя бы в смысле условий теоремы 1) вероятность неправильного ответа будет сколь угодно мала. Описанный алгоритм, порожденный алгоритмом A , обозначим A^* .

Если допускается неполное обучение, то приведенный выше алгоритм можно закончить на некотором шаге. Кроме рядов (6.1), (6.2), мы получим ряд неразделенных множеств

$$M_{\mu_1}, N_{\mu_1}, M_{\mu_2}, N_{\mu_2}, \dots, M_{\mu_r}, N_{\mu_r}. \quad (6.5)$$

Множества соответствующих индексов обозначим через $[M \cup N]$. Полагая

$$\xi_{M \cup N}(x) = \bigvee_{\mu \in [M \cup N]} \xi_\mu(x),$$

получим вместо (6.3), (6.4)

$$\xi_M(x)\xi_N(x)\xi_{M \cup N}(x) \equiv 0, \quad 1 = \xi_M(x) \vee \xi_N(x) \vee \xi_{M \cup N}(x). \quad (6.6)$$

При этом, если $x \in M$, то $\xi_M(x) = 1$; если $x \in N$, то $\xi_N(x) = 1$.

Для произвольного x ответы $\xi_M(x) = 1$, $\xi_N(x) = 1$ истолковываются, как и выше. Однако теперь возможен ответ $\xi_{M \cup N}(x) = 1$, что можно трактовать, как «не знаю». Соотношения (6.6) означают, что указанные три ответа являются взаимно исключающими и один из них обязательно последует.

Приведенный алгоритм, по-видимому, далек от оптимального. Его достоинством является, возможно, простота технической реализации. Именно, построим релейно-контактную схему, изображенную для случая шестнадцати с. э. на рис. 4.

Если соединить вершины, помеченные индексами $\mu \in [M]$, с вершиной $\xi_M(x)$, то между вершинами $*$ и $\xi_M(x)$ реализуется функция $\xi_M(x)$. Аналогично реализуются функции $\xi_N(x), \xi_{M \cup N}(x)$.

Таким образом, блок обучения персептрона со многими с. э. должен:

1) по заданной тренировочной последовательности выдавать, осуществляя описанный алгоритм, веса а. э., так что реакциями с. э. будут функции

$$\psi(x), \psi_0(x), \psi_1(x), \psi_{01}(x), \dots, \psi_\mu(x), \dots$$

(нумерация двоичным кодом);

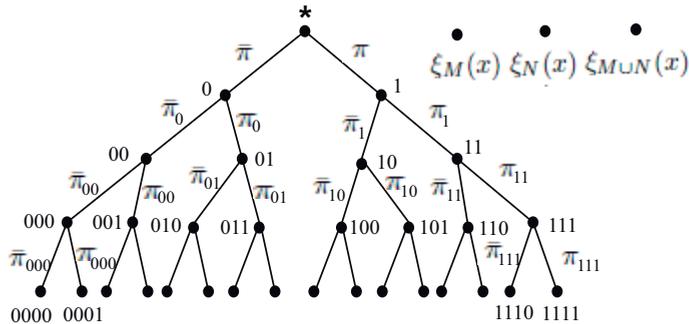


Рис. 4.

2) в процессе осуществления описанного алгоритма соединять на релейно-контактной схеме вершины с индексами $\mu \in [M]$ с клеммой $\xi_M(x)$, вершины с индексами $\mu \in [N]$ с клеммой $\xi_N(x)$ и в конце процесса вершины $\mu \in [M \cup N]$ с клеммой $\xi_{M \cup N}(x)$.

Перейдем теперь к случаю разделения $q > 2$ множеств S_1, S_2, \dots, S_q . Пусть тренировочную последовательность образуют образы

$$\begin{aligned} M_1 &= x'_1, x'_2, \dots, x'_{k_1} \subset S_1, \\ M_2 &= x''_1, x''_2, \dots, x''_{k_2} \subset S_2, \\ &\dots \\ M_q &= x_1^{(q)}, x_2^{(q)}, \dots, x_{k_1}^{(q)} \subset S_q. \end{aligned}$$

Будем для простоты предполагать, что алгоритм A удовлетворяет условию разделения.

Из различных возможных обобщений приведенного выше алгоритма рассмотрим обучение по плану:

- « S_1 или S_2 ?»,
- « S_1 или S_3 ?», « S_2 или S_3 ?»,
- ...
- « S_1 или S_q ?», « S_2 или S_q ?», ..., « S_{q-1} или S_q ?».

Применяя алгоритм A^* к паре множеств M_1 и M_2 , получим функции ξ_{M_1} и $\xi_{M_2}(x) = \xi_{M_1}(x)$. Нам удобно теперь обозначить $\xi_{12}(x) = \xi_{M_1}(x)$, $\xi_{21}(x) = \xi_{M_2}(x)$. Точно также по множествам M_1 и M_3 построим функции $\xi_{13}(x)$ и $\xi_{31}(x)$ и так далее.

Всего мы получим $q(q-1)$ функций $\xi_{jh}(x)$ таких, что $\xi_{hj}(x) = \overline{\xi_{jh}(x)}$ и

$$\begin{aligned}\xi_{jh}(x) &= 1 \text{ при } x \in M_j, \\ \xi_{jh}(x) &= 0 \text{ при } x \in M_h.\end{aligned}$$

Обозначая

$$\xi_j(x) = \prod_{h \neq j} \xi_{jh}(x),$$

получим, что для любого $x \in M_1 \cup M_2 \cup \dots \cup M_q$ равенство равносильно соотношению

$$x \in M_j.$$

При этом

$$1 = \xi_1(x) \vee \dots \vee \xi_1(x), \quad (6.8)$$

$$\xi_j(x) \cdot \xi_h(x) = 0 \text{ при } j \neq h. \quad (6.9)$$

Формулы (6.8), (6.9) показывают, что для любого x найдется единственное j такое, что $\xi_j(x) = 1$. В этом случае ответ персептрона интерпретируется как « $x \in S_j$ ».

§ 7. ОПИСАНИЕ ПРОВЕДЕННЫХ ЭКСПЕРИМЕНТОВ⁹

Во всех описанных ниже экспериментах обучение проводилось по простейшему алгоритму § 3.

1⁰. *Распознавание выпуклых и вогнутых функций.* Ретиной является система пятидесяти равноотстоящих точек на числовой оси («рецепторные нейроны»). Образуем систему 50 соответствующих ординат («кривая»). Таким образом, размерность пространства образов $n = 50$. Ординаты кривой брались с точностью трех восьмеричных знаков (в других терминах: каждый рецепторный нейрон ретины может находиться в 128 состояниях). Разделяемыми понятиями были «выпуклые кривые» и «вогнутые кривые». Очевидно, им соответствуют выпуклые множества S_1 и S_2 в пространстве образов. Тренировочная последовательность состояла из выбранных наугад десяти выпуклых и десяти вогнутых кривых, показанных на рис. 5, 6. Для распознавания была взята последовательность также выбранных наугад 15 выпуклых и 15 вогнутых кривых, показанных на рис. 7, 8. Все ответы оказались верными. Тренировочная последовательность также разделилась полностью.

2⁰. *Распознавание рукописных цифр.* Ретиной является квадрат 10×10 см и образом — упорядоченная система 25 точек на нем. Размерность пространства образов $n = 50$. Координаты точек брались с точностью трех восьмеричных знаков (≈ 0.8 мм), что соответствует размеру ретины 128×128 . Разделяемыми понятиями были изображения цифр 2 и 3. Как следует из теоремы 2 § 4, соответствующие множества являются выпуклыми. Тренировочная последовательность состояла из пяти изображений цифры 2 и из пяти изображений цифры 3, различных по величине, форме и расположению на экране. Из четырнадцати образов (семь двоек и семь троек) машина верно распознала двенадцать (85.8%). Для двух ошибочных ответов

⁹ Автор пользуется случаем поблагодарить студентов М. Персиянову, Т. Богдарину и Н. Беляеву, проделавших большую, описанную в этом параграфе работу.

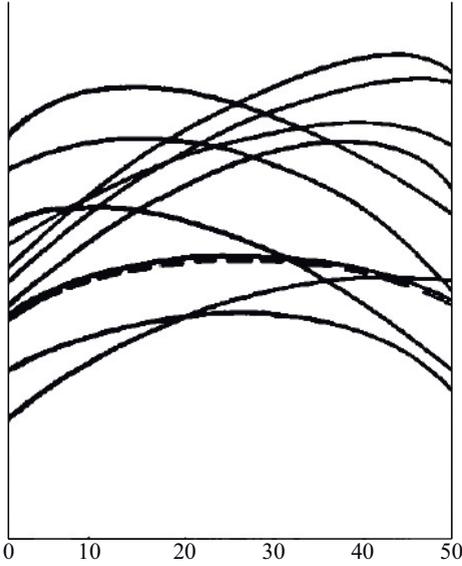


Рис. 5.

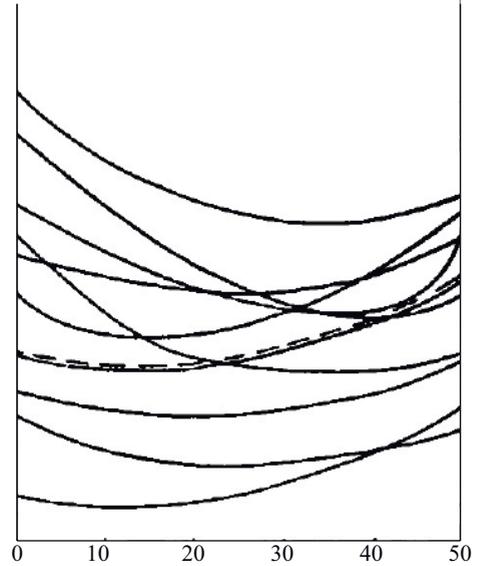


Рис. 6.

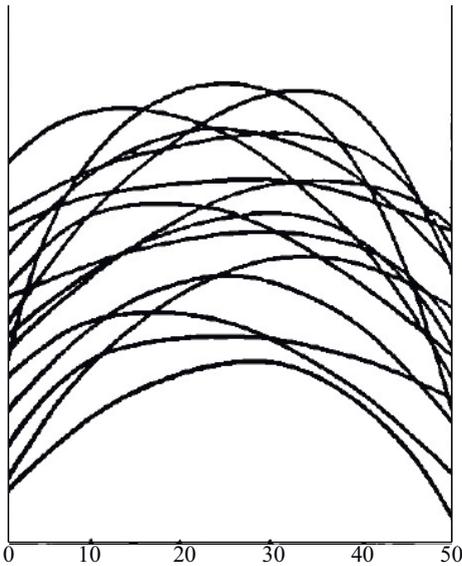


Рис. 7.

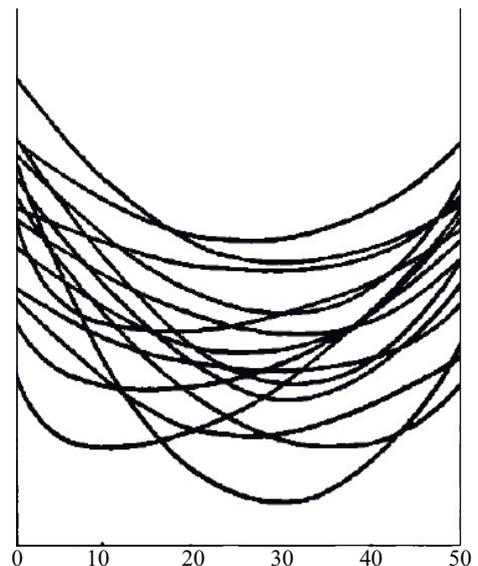


Рис. 8.

соответствующие значения $\psi(x_j)$ были близки к нулю. На рис. 9, 10 показано десять изображений цифр 2 и 3 из последовательности для распознавания.

3⁰. *Распознавание профилей.* Образом является упорядоченная система 50 точек на квадрате 128 мм × 128 мм, координаты которых брались с точностью трех восьмичных знаков (≈ 1.0 мм). (Размер ретины снова 128 × 128, размерность пространства образов $n = 100$.) Разделяемыми понятиями являлись «профиль курного

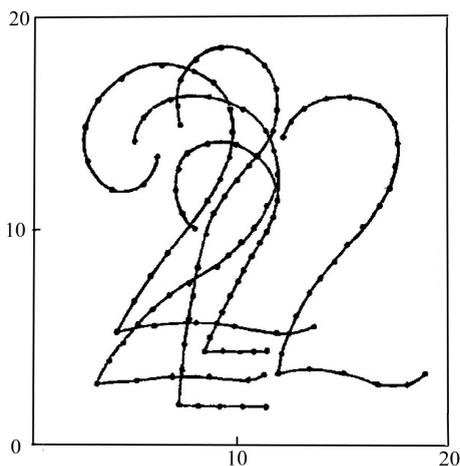


Рис. 9.

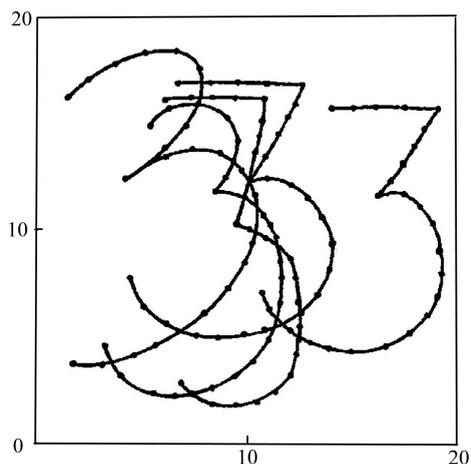


Рис. 10.

лица» и «профиль лица с длинным носом». (Им соответствуют согласно § 4 приблизительно выпуклые множества.) На рис. 11, 12 показаны десять типичных профилей из последовательности для распознавания. Тренировочная последовательность состояла из семи образов каждого типа и последовательность для распознавания — из 14 образов каждого типа. Машина верно отгадала 28 образов последовательности для распознавания. Тренировочная последовательность также разделилась полностью.

4⁰. *Распознавание гласных фонем.* С физической точки зрения образом является звук и более точно — кривая $\xi(\lambda)$ (см. (4.1)). Разделяемыми понятиями были фонемы «у», произносимые различными дикторами-женщинами¹⁰.

Первоначально на каждой кривой было взято 11 точек ($n = 11$) с одинаковыми для всех кривых абсциссами. Это равносильно (см. § 4) распознаванию звука по одиннадцати фиксированным гармоникам. Тренировочная последовательность состояла из восьми образов «у» и восьми образов «ы». При проверке тренировочной последовательности на распознавание машина сделала три ошибки (81% распознавания), опознав три фонемы «у» как «ы». В последовательности для распознавания, состоящей из десяти образов «у» и десяти образов «ы», машина сделала 9 ошибок, что свидетельствовало о том, что машина не обучилась распознаванию. После этой неудачи было взято 40 точек на каждой кривой. На обучение и распознавание были поданы те же последовательности. После обучения все образы обеих последовательностей были опознаны верно. На рис. 13, 14 показаны типичные кривые из последовательности для распознавания.

Были проделаны аналогичные эксперименты по распознаванию фонем «о» и «ы», «а» и «у», «о» и «а», «а» и «ы», «у» и «о».

¹⁰ Автор пользуется случаем, чтобы поблагодарить проф. Г. В. Гершуни и проф. Л. А. Чистович, любезно предоставивших соответствующие графики.

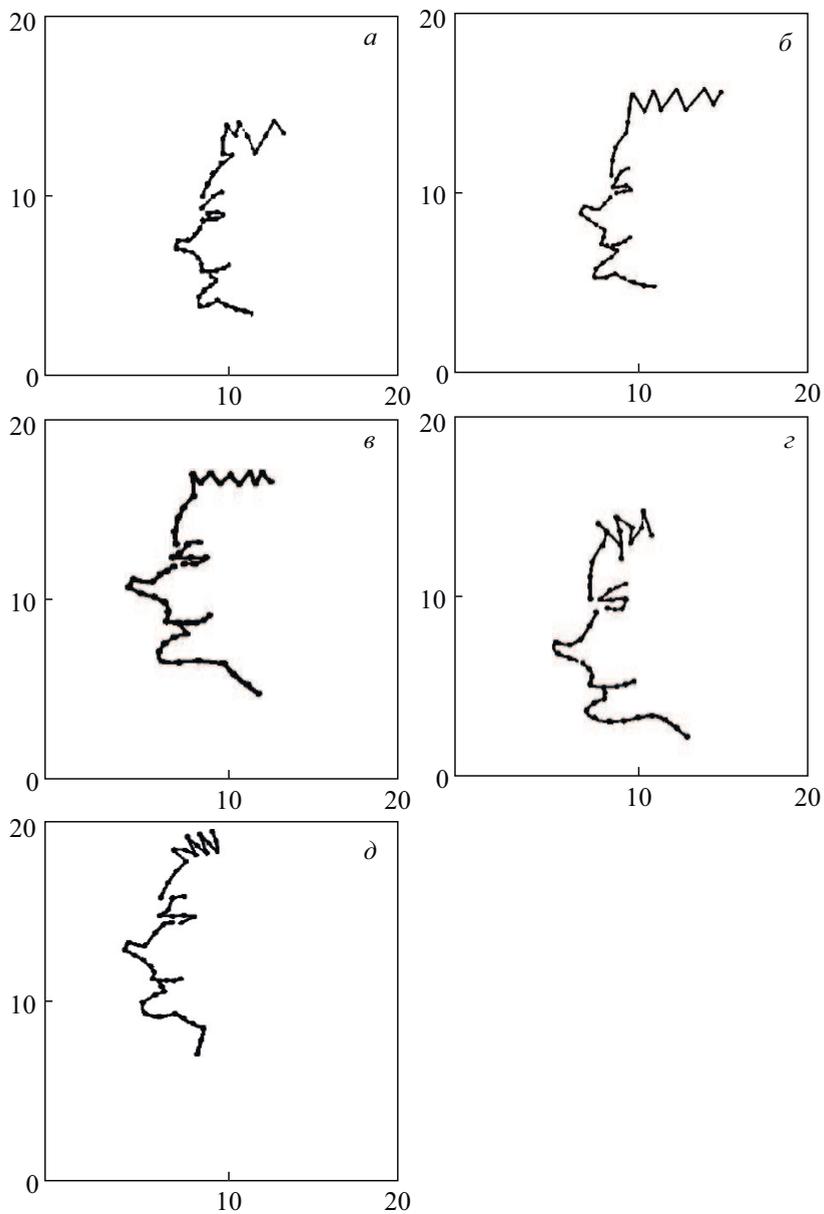


Рис. 11.

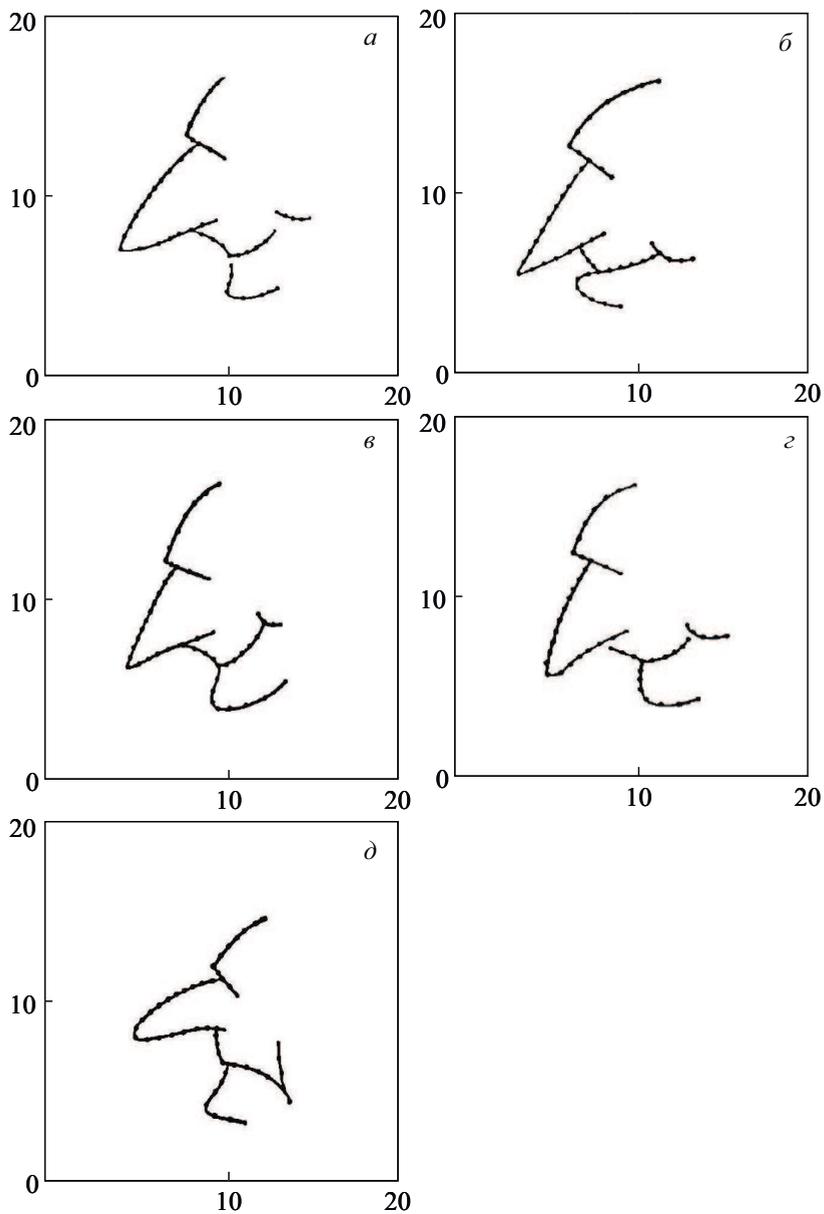


Рис. 12.

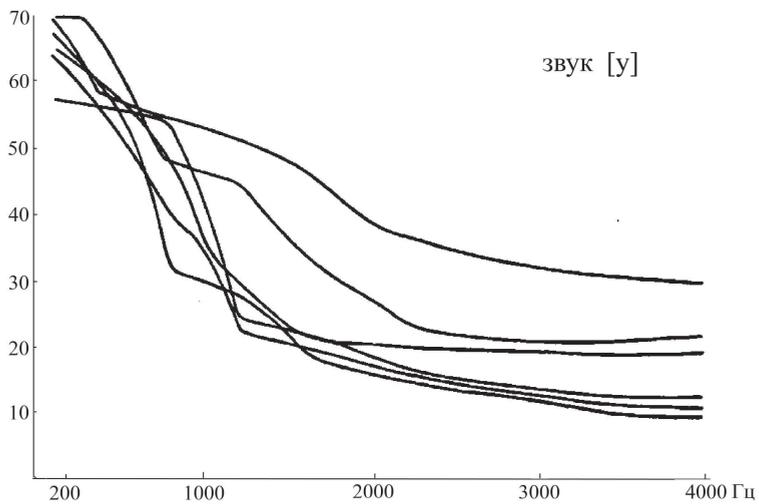


Рис. 13.

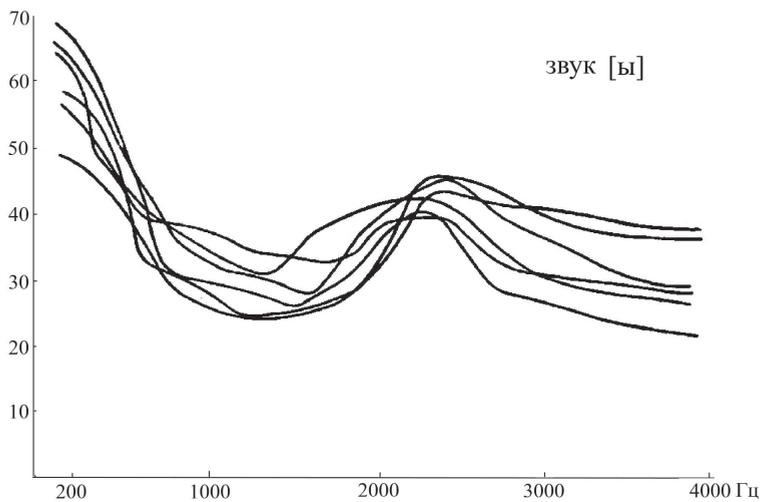


Рис. 14.

ЛИТЕРАТУРА¹¹

1. Трапезников В. А. Кибернетика и автоматическое управление // Автоматика и телемеханика. Т. XXIII, № 3. 1962.
2. Браверман Э. М. Опыты по обучению машины распознаванию зрительных образов // Автоматика и телемеханика. Т. XXIII, № 3. 1962.
3. Браверман Э. М. Некоторые вопросы построения машин, классифицирующих объекты по не заданному заранее признаку // Автоматика и телемеханика. Т. XXI, № 10. 1960.
4. Rosenblatt F. Two Theorems of Statistical Separability in the Perceptron // Symposium of the Mechanisation of Thought Processes. England, November, 1958.
5. Rosenblatt F. Perceptron Simulation Experiments // Proc. IRE. Vol. 48, № 3. 1960.
6. Selfridge O. G. Pandemonium: a Paradigm for Learning // Symposium of the Mechanisation of Thought Processes. England, November, 1958.
7. Харкевич А. А. Опознавание образов // Радиотехника. Т. 14, № 5. 1959.
8. Joseph R. D. On Predicting Perceptron Performance // IRE Intern. Convent. Rec. Vol. 8, № 2. 1960.
9. Глезер В. Д., Цукерман И. И. Информация и зрение. М.: Изд. АН СССР, 1961.
10. Варшавский Л. А., Чистович Л. А. Средние спектры русских гласных фонем // Проблемы физиологической акустики. Т. IV. 1959.
11. Якубович В. А. Об ограниченности и устойчивости в целом решений некоторых нелинейных дифференциальных уравнений // ДАНН СССР. 1958. Т. 121, № 6.
12. Филиппов А. Ф. Дифференциальные уравнения с разрывной частью // Мат. сб. 1960. Т. 51 (93). Вып. 1.

¹¹ Примечание при корректуре: см. также недавно опубликованные работы М. А. Айзермана, М. М. Бонгарда и других авторов в сборнике «Биологические аспекты кибернетики» (Изд. АН СССР, 1962) и книгу М. Глушкова «Введение в теорию самосовершенствующихся систем» (Киев, 1962).