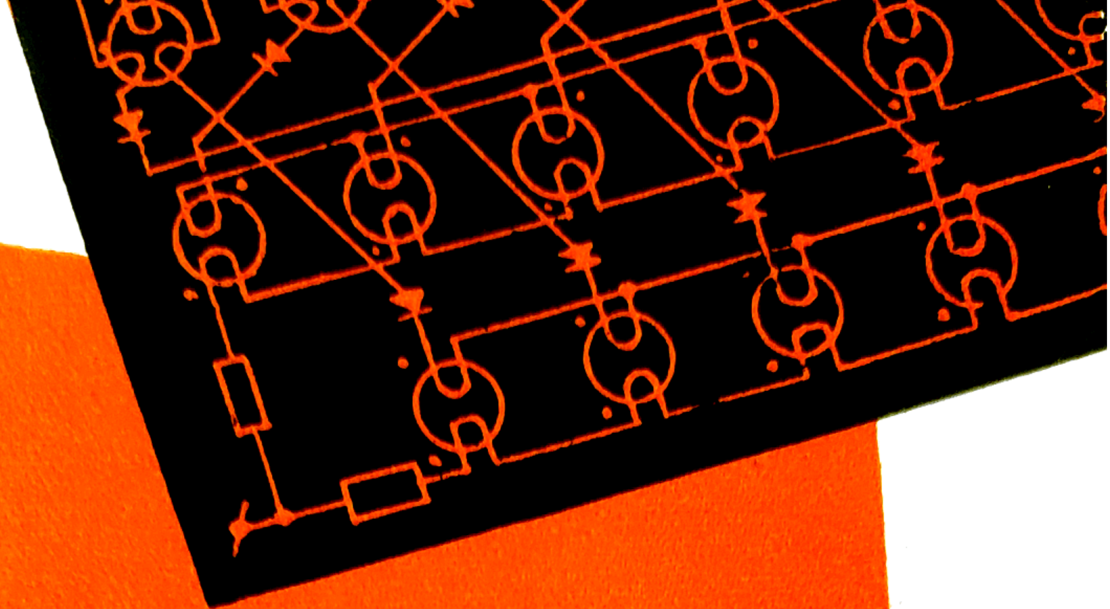

Вычислительная техника *и* вопросы программирования

3

1964

ЛЕНИНГРАДСКИЙ ОРДЕНА ЛЕНИНА ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ имени А. А. ЖДАНОВА

КАФЕДРА ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ВОПРОСЫ ПРОГРАММИРОВАНИЯ

Выпуск 3

ИЗДАТЕЛЬСТВО
ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА
1964

Б. Н. Козинец

ОБ ОДНОМ АЛГОРИФМЕ ОБУЧЕНИЯ ЛИНЕЙНОГО ПЕРСЕПТРОНА

В работе [1], посвященной линейному персептрону, для большого класса реальных понятий (сюда относятся, например, гласные фонемы, различные контурные рисунки и т. д.) строится n-мерное евклидово пространство — пространство образов, в котором этим понятиям соответствуют открытые выпуклые множества. Линейный персептрон с одним суммирующим элементом может распознавать два понятия из этого пространства образов. Процесс обучения состоит в том, что на вход персептрона подается „тренировочная последовательность“ — выбранные наугад представители двух разных понятий с указанием, к какому из понятий принадлежит каждый из них. Персептрон же должен определить в пространстве образов плоскость (здесь и в дальнейшем под „плоскостью“ будем понимать $(n—1)$-мерную гиперплоскость), разделяющую тренировочную последовательность, а следовательно, с каким-то приближением, и множества, соответствующие разным понятиям. Тем самым задача распознавания большого класса реальных понятий сводится к следующей задаче.

В n-мерном евклидовом пространстве заданы два множества точек* $M=\{x^k\}_{k=1}^{m}$ и $N=\{y^l\}_{l=1}^{p}$ (элементы множеств M и N — это образы тренировочной последовательности). Предполагается, что соответствующие выпуклые оболочки $\overline{M}$ и $\overline{N}$ этих множеств не имеют общих точек. Требуется построить плоскость, разделяющую множества M и N, т. е. такую плоскость, которая не имеет общих точек с множествами $\overline{M}$ и $\overline{N}$, и, кроме того, множества $\overline{M}$ и $\overline{N}$ лежат по разные стороны этой плоскости.

---

* Здесь и в дальнейшем через x обозначаем радиус-вектор точки, а также и самую точку; так, например, $\rho(x, y)$ означает расстояние от точки x до точки y.

На самом деле, чтобы лучше разделялись сами понятия, желательно среди всех плоскостей, разделяющих множества M и N, найти такую плоскость P, расстояние от которой до множества $M \cup N$ имеет максимальную величину. Очевидно, что такой плоскостью будет плоскость, проходящая через середину отрезка, соединяющего какие-нибудь две ближайшие точки множеств $\overline{M}$ и $\overline{N}$, перпендикулярно к нему. В настоящей заметке дается простой алгорифм, удобный для ЭВМ, который позволяет находить какие-нибудь две ближайшие точки множеств $\overline{M}$ и $\overline{N}$.

Описание алгорифма

а) Берем наугад две точки из разных множеств $x_1 \in M$ и $y_1 \in N$.

б) Предположим, что точки $x_i \in \overline{M}$ и $y_i \in \overline{N}$ найдены. Ищем

$$\max_{x^k \in M,\, y^l \in N} \{(y_i - x_i,\ x^k - x_i);\ (x_i - y_i,\ y^l - y_i)\} \equiv m_i. \quad (1)$$

Если $m_i \leqslant 0$, то, очевидно, $\rho(x_i, y_i) = \rho(\overline{M}, \overline{N})$, т. е. x_i и y_i — искомые точки.

Пусть же $m_i > 0$. Покажем, как строить точки x_{i+1} и y_{i+1}. Для этого в качестве вспомогательной точки берем точку, на которой реализуется максимум (1) (если таких точек несколько, то берем любую из них). Эта точка может принадлежать как множеству M, так и множеству N. Пусть для определенности вспомогательная точка принадлежит множеству N. Обозначим ее y_i^*. Тогда сразу же полагаем $x_{i+1} = x_i$. y_{i+1} вычисляется по формулам:*

$$y_{i+1} = \begin{cases} y_i^*, \text{ если } (x_i - y_i^*,\ y_i - y_i^*) \leqslant 0, \\ y_i^* + \dfrac{(y_i^* - y_i,\ y_i^* - x_i)}{(y_i^* - y_i,\ y_i^* - y_i)}(y_i - y_i^*), \\ \quad \text{если } (x_i - y_i^*,\ y_i - y_i^*) > 0. \end{cases} \quad (2)$$

Замечание. Геометрически $(i+1)$-й шаг описанного выше алгорифма означает следующее. Среди всех векторов

---

* Аналогично поступаем, если вспомогательная точка принадлежит множеству M. Обозначая ее x_i^*, имеем

$$y_{i+1} = y_i, \quad x_{i+1} = \begin{cases} x_i^*, \text{ если } (y_i - x_i^*,\ x_i - x_i^*) \leqslant 0, \\ x_i^* + \dfrac{(x_i^* - x_i,\ x_i^* - y_i)}{(x_i^* - x_i,\ x_i^* - x_i)}(x_i - x_i^*), \\ \quad \text{если } (y_i - x_i^*,\ x_i - x_i^*) > 0. \end{cases}$$

$x_i x^k$ $(k=1, \ldots, m)$, $y_i y^l$ $(l=1, \ldots, p)$ выбирается такой, проекция которого соответственно на вектор $x_i y_i$ или $y_i x_i$ имеет наибольшую величину. Пусть для определенности это вектор $y_i y^{l'}$. В качестве вспомогательной точки берем $y_i^* = y^{l'}$. Тогда y_{i+1} — точка на отрезке $y_i y_i^*$, ближайшая к x_i.

Теорема. Действуя согласно указанному алгорифму, мы получим последовательность точек $x_i \in \overline{M}$, $y_i \in \overline{N}$ таких, что $\rho(x_i, y_i) \to \rho(\overline{M}, \overline{N})$ при $i \to \infty$.

Доказательство. Если величина m_i (см. (1)), начиная с некоторого i_0, становится меньше любого наперед заданного $\varepsilon > 0$, то для всех $i > i_0$ выполнено неравенство

$$0 \leqslant \rho(x_i, y_i) - \rho(\overline{M}, \overline{N}) \leqslant 2 \frac{\varepsilon}{\rho(x_i, y_i)}$$

и теорема становится очевидной.

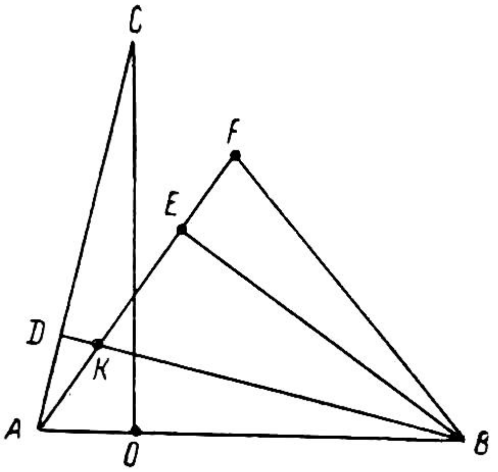

Рис. 1.

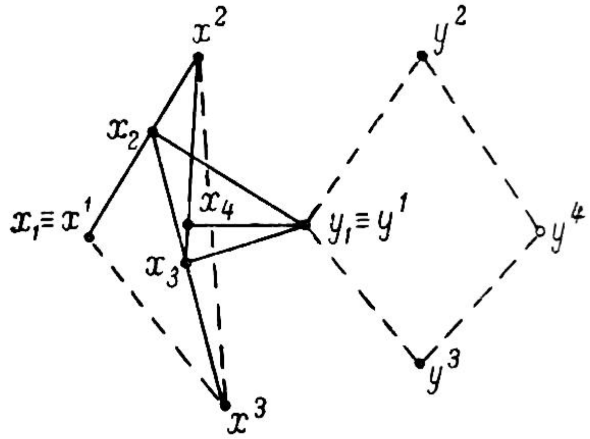

Рис. 2.

Допустим противное, т. е. существует $\varepsilon_0 > 0$ и последовательность m_i $(i=1, 2, \ldots)$ такая, что $m_i > \varepsilon_0$ для всех i. Если мы докажем, что в этом случае для всех i

$$\rho(x_i, y_i) - \rho(x_{i+1}, y_{i+1}) \geqslant \delta > 0, \tag{3}$$

то теорема будет доказана полностью.

Пусть b_1 — диаметр множества $\overline{M}$, b_2 — диаметр множества $\overline{N}$, $d = \rho(\overline{M}, \overline{N})$, $b = \max\{b_1, b_2\, d\}$. Сделаем следующее построение. Пусть $AB = d$ (рис. 1), $AO = \frac{\varepsilon_0 d}{D^2}$, где D — диаметр множества $M \cup N$. Точка C удовлетворяет двум условиям $CA = b$ и $CO \perp AB$. Проведем $BD \perp AC$. Докажем, что

$$\rho(y_i, x_i) - \rho(y_{i+1}, x_{i+1}) > AB - BD > 0, \tag{4}$$

тогда будет доказано и (3) $(\delta \equiv AB - BD)$.

При построении пары точек x_{i+1}, y_{i+1} мы использовали описанную выше вспомогательную точку. Пусть для определенности получилось, что вспомогательная точка принадлежит множеству $\overline{M}$. Обозначим ее $\overset{*}{x_i}$. Тогда $y_{i+1}=y_i$. Построим $\Delta AFB \sim \Delta x_i \overset{*}{x_i} y_i$ (см. рис. 1) так, что точке x_i соответствует точка A, точке $\overset{*}{x_i}$ — точка F, точке y_i — точка B и точке x_{i+1} — точка E. При этом точки F и A будут лежать по разные стороны от прямой OC, что следует из равенства $AO=\frac{\varepsilon_0 d}{D^2}$. Из описания алгорифма ясно, что либо $\angle AFB \geqslant 90^\circ$ и тогда $\angle AEB=90^\circ$, либо $\angle AFB > 90^\circ$ и тогда точка E совпадает с точкой F. Во всяком случае $\angle AEB \geqslant 90^\circ$, т. е.

$$BE \leqslant BK. \tag{5}$$

Из того, что $\rho(x_i, y_i) > AB$ и из подобия следует, что

$$\rho(x_i, y_i) - \rho(x_{i+1}, y_{i+1}) > AB - BE, \tag{6}$$

а так как $AF < AC$ и точки A и F лежат по разные стороны от прямой CO, то $\angle FAB < \angle CAB$, т. е.

$$BK < BD. \tag{7}$$

Из (5), (6) и (7) следует (4). Что и требовалось доказать.

Замечание. Алгорифм может не сходиться в конечное число шагов. Это видно из рис. 2, если взять $x_1=x^1$, $y_1=y^1$. Однако справедлива следующая оценка:

$$0 \leqslant \rho(x_i, y_i) - \rho(\overline{M}, \overline{N}) \leqslant 2\frac{m_i}{\rho(x_i, y_i)},$$

где m_i определяется по формуле (1).

ЛИТЕРАТУРА

В. А. Якубович. Машины, обучающиеся распознаванию образов. Сб. «Методы вычислений», 2. Изд. ЛГУ, 1963.

Статья поступила в редакцию 16 IV 1963 г.