

MULTIDIMENSIONAL
GEOMETRY
AND PATTERN
RECOGNITION

Yu. I. NEIMARK

Mathematical models of pattern recognition, models of an object, a pattern, a recognition and recognition training are discussed.

Рассказывается о математических моделях распознавания образов, о моделях объекта, образа, распознавания и обучения распознаванию.

МНОГОМЕРНАЯ ГЕОМЕТРИЯ И РАСПОЗНАВАНИЕ ОБРАЗОВ

Ю. И. НЕЙМАРК

Нижегородский государственный университет

ВВЕДЕНИЕ И ПОСТАНОВКА ПРОБЛЕМЫ

В предыдущей статье [1] я говорил о математике как об операционной системе и моделях. Сейчас я постараюсь рассказать о замечательных простых математических моделях очень сложных интеллектуальных способностей человека, моделях, которые послужили основой и трамплином для бурного триумфального шествия нового раздела науки — распознавания образов. В 1957 году в новой, только в 1948 году провозглашенной Н. Винером науке кибернетике произошло ошеломляющее, получившее благодаря журналистам широкую огласку, событие: демонстрация устройства со звучным названием перцептрон [2, 3]. Перцептрон распознавал (узнавал, различал, опознавал) зрительные образы и обучался этому распознаванию. Он мог узнавать новые образы, если его, как и нас, людей, этому учили путем показов. Ранее казалось, что такое умение — бесспорная прерогатива человека, что только человек с его высоким интеллектом способен к обобщению и формированию абстрактного понятия образа, только он это может и только его можно этому научить. Оказалось, что это может делать устройство, бездушная машина.

Перцептрон имел не только звучное название, он был на вид прост и загадочен. Его создал Розенблатт, физиолог по профессии, исходя из современных ему представлений о конструкции мозга [4], связей между слоями его нервных клеток и о самих нервных клетках, исходя из того, что он сам назвал принципами нейродинамики [5]. Перцептрон имел фасетчатый глаз, как у стрекозы, состоящий из десяти тысяч фотоэлементов. Далее шел ящик с лампочками, отвечающими на показ, и кнопками для сообщения ему, что показывают при обучении, состоящем в указаниях, правильно ли он, перцептрон, распознает то, что ему показывают. Что внутри ящика, не видно, ящик черный, как в игре “Что? Где? Когда?”.

Но вскоре содержимое ящика стало понятным, и сотни групп ученых во всем мире стали экспериментировать с перцептронами. Уже в начале 60-х годов в нашей стране применили машинное распознавание образов к опознанию нефтеносной местности, местности, в недрах которой скрыта нефть. Затем с помощью ЭВМ начали распознавать болезни человека и делать это подчас лучше, чем сам человек. ЭВМ научилась “читать” печатный и даже рукописный текст и воспринимать текст с голоса,

обнаруживать неисправности, предсказывать аварийные ситуации и многое, многое другое.

Отдельный конкретный стул мы можем в деталях описать и научить ЭВМ по этому описанию его узнавать. Но как научить ЭВМ узнавать любой стул, в том числе и тот, которого мы сами не видели, то есть каким образом описать образ “стул”? Стул предназначен для сидения, имеет ножки и спинку. Но что значит для ЭВМ: “предназначен для сидения, ножка и спинка”? Это столь же неясно, как и то, что такое сам стул.

Итак, что такое “объект”, что такое “образ” и как можно ЭВМ научить “распознаванию образов”, как можно распознавание использовать для перечисленных выше вопросов? Могущество математики состоит в том, что ответы на эти вопросы могут быть получены, если найдены соответствующие этим вопросам математические модели, то есть математические модели объекта, образа, распознавания и обучения распознаванию. Как только такие модели станут общим достоянием, ЭВМ будут распознавать не хуже, а может быть, и лучше человека. Приложения, и самые разнообразные, посыплются, как горох из мешка, причем не только ожидаемые, но и совершенно неожиданные. Ведь персептрон распознавал зрительные образы, а затем ЭВМ ставила диагнозы болезней, находила скрытые неисправности и осуществляла геологическую разведку.

Кажется, что построить эти математические модели невозможно и уж, во всяком случае, очень трудно. Будущий врач учится полжизни, чтобы верно ставить диагнозы заболевания, только талантливый химик догадывается о свойствах вещества по его структурной химической формуле, а мы хотим, чтобы это делали ЭВМ. Конечно, трудно, даже очень трудно, потому что эти модели существенно новые, отличные от привычных нам уравнений, и мы еще не знаем, как их построить. Нужен взлет, скачок мысли, и он был бы невозможен, не обладая человек изобретением своего воображения, называемым многомерным пространством. Изобретением потому, что многомерного геометрического пространства нет в природе, его изобрел математик. Изобрести это многомерное пространство в связи с моделированием распознавания образов очень маловероятно, пожалуй, невозможно. Идеи многомерного пространства высказывались еще в XVIII веке И. Кантом и Ж. Д’Аламбером, а многомерная геометрия построена А. Кэли, Г. Грассманом и Л. Шлефли в прошлом веке. После этого первоначальные сомнения и связанная с ними мистика были преодолены, и многомерное пространство стало плодотворным математическим понятием. Многомерное пространство такое же, как наше трехмерное, но только у него не три координатных оси, а много, любое число n . Представить его в нашем трехмерном пространстве нельзя, но вообразить можно: n осей координат Ox_1, Ox_2, \dots, Ox_n , каждая точка x

имеет n координат x_1, x_2, \dots, x_n ; плоскость, точнее гиперплоскость, делит все пространство на две части; через любые две точки можно провести единственную прямую; конечный отрезок прямой имеет длину и т.д. Для физиков уже давно стало привычным рабочим инструментом четырехмерное пространство с тремя пространственными координатами и четвертой — временем. Говорят, наш замечательный геометр Делоне легко представлял себе четырехмерное пространство так же, как мы трехмерное, и мог в нем производить геометрические построения. В конце концов, мы же изображаем наше трехмерное пространство на двумерных картинах, рисунках и чертежах.

Но вернемся к математическим моделям. Мне очень трудно перед Вами хитрить, ведь я знаю эти модели, и они просты, как дважды два — четыре. Помогают мне хитрить только прошлые воспоминания, когда они были неизвестны и всем казались недостижимыми и загадочными. Чувство удивления и затем восхищения не прошли для меня даром: я занялся разработкой и приложением идей распознавания образов в медицине. В 1972 году вышла книга “Распознавание образов и медицинская диагностика” [6], плод труда и энтузиазма математиков, врачей и инженеров. Сейчас возросшие возможности персональных ЭВМ побудили нас на основе приобретенного опыта к разработке самообучающегося, автоматизированного консультанта врача в постановке диагноза, прогноза течения болезни, показанности оперативных вмешательств, выборе метода лечения и др. [7]. Естественно, что мне хочется, чтобы сейчас эти же чувства удивления и восхищения посетили и Вас. Итак, требуется описать на математическом языке, то есть построить, математические модели объекта, образа, распознавания и обучения.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ РАСПОЗНАВАНИЯ

Что такое изображение? Для простоты черно-белое, такое, как в газете, полученное путем нанесения или не нанесения на бумагу в определенных ее местах черных точек. Перенумеруем от 1 до n все места, где можно ставить черные точки. В s -м месте может стоять или не стоять черная точка, это опишем тем, что $x_s = 1$ или, соответственно, $x_s = 0$. В результате любое изображение описывается последовательностью чисел x_1, x_2, \dots, x_n . Аналогично сетчатка нашего глаза, состоящая из колбочек и палочек, воспринимает любое изображение как некоторую последовательность чисел. Еще отчетливей это для фасетчатого глаза стрекозы. Теперь введем n -мерное пространство X с координатами x_1, x_2, \dots, x_n . Любое изображение представляется в нем некоторой точкой x . Если бы изображение состояло только из трех черных и белых точек, то пространство X и точку — изображение x можно было бы наглядно нарисовать. Но координат x_1, x_2, \dots, x_n тысячи и сотни ты-

сяч. Ну и что? Изобразим его так же, только с n осями и такой же точкой x — изображением. Таким образом, конкретный объект-рисунок — это точка x в многомерном пространстве X всевозможных рисунков.

А что такое образ? Это обобщенное представление обо всех его конкретных изображениях или нечто объемлющее все эти изображения, то есть множество точек x в многомерном пространстве X , соответствующих всем этим изображениям. Будем мыслить это множество точек в виде некоторого облака, назвав его, например, A (рис. 1). Другому образу будет отвечать другое облако, пусть B . Например, A — всевозможные стулья, а B — ведра.

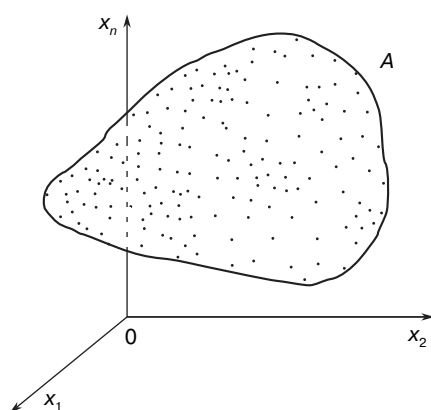


Рис. 1.

Таким образом, строятся математические модели конкретного изображения и множества изображений, составляющих образы. Теперь мы можем ответить на вопрос “что такое распознавание образов?” Это умение выяснить, к какому из облаков A или B принадлежит точка x , отвечающая показываемому изображению — объекту. Как узнать, лежит ли точка x в облаке A или B ? Ведь нарисовать на самом деле их нельзя, и пространство X воображаемое, воображаемые и облака A и B в нем, да и точка x тоже, но сделать это все же можно, например, так. Разделим облака A и B гиперплоскостью или, если не выходит, гиперповерхностью S (рис. 2). С одной ее стороны облако A , с другой — B . Как это сделать? Фантазия? Вот и нет! Гиперповерхность S — это множество точек, на которых “зануляется” некоторая функция $f(x_1, x_2, \dots, x_n)$. С одной стороны поверхности S $f > 0$, с другой $f < 0$. Так что, например, $x \in A$, если $f < 0$, и $x \in B$, если $f > 0$. Это не единственный способ, но он очень хороший и дает нам определенную математическую модель различия, распознавания образов. Надо только еще научиться находить эту поверхность S и ее уравнение

$$f(x_1, x_2, \dots, x_n) = 0. \quad (1)$$

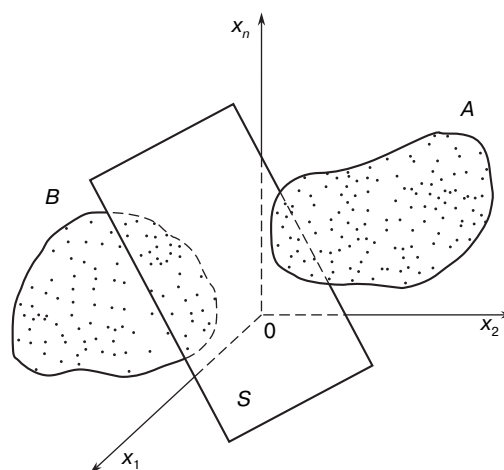


Рис. 2.

Ясно, что найти эту поверхность можно, только “показывая” точки облаков A и B . Показать сразу все облако A или B мы не можем, они воображаемые, и точек в них может быть слишком много. Но показать не очень большие количества точек множеств облаков A и B можно, просто перечислив их.

Теперь можно сформулировать, в чем состоит модель обучения распознаванию образов. Это нахождение разделяющей поверхности S по конечным ограниченными показам точек облаков A и B . (Все время речь шла о распознавании двух образов, но, умея распознать два, нетрудно распознать любое их конечное число. Достаточно, например, каждое из множеств отличать от всех остальных. Это не лучший способ, даже плохой, и можно придумать лучший, при котором разделяющих поверхностей требуется много меньше.)

Наша цель достигнута: построены математические модели объекта (изображения), образа, распознавания образов и обучения распознаванию образов. Но все же, как конкретно находить требуемую разделяющую поверхность, мы не знаем. Тут возможны разные способы, но прежде давайте немного осмыслим полученные математические модели. Математические модели — это описания на абстрактном математическом языке. Куда уж дальше? Модели не только абстрактные, но и фантастические, воображаемые. Но это не плохо, а, оказывается, очень хорошо. Хорошо потому, что под эти модели подходят не только зрительные объекты, но и звуковые, обонятельные и осязательные, и те представления об объектах, которые дают не только наши органы чувств, но и приборы, проще, всё, что может быть описано или, как говорят, закодировано числами, а числами может быть закодировано все, что можно описать словами и любыми значками, то есть всё, что мы можем знать. Такова всеобъемлющая мощь этих математических моделей. Нужен

диагноз болезни – пожалуйста, только опишите больного и дайте достаточно случаев этой болезни, которые я использую как представителей ее образа. Нужно узнать, есть ли в глубине земли золото, – пожалуйста; в чем неисправность машин – пожалуйста; не раковая ли это клетка – пожалуйста; чей это почерк – пожалуйста; не жулик ли этот человек – пожалуйста; будет ли этот человек хорошим водителем автомобиля – пожалуйста; какие свойства у еще не полученного химического вещества – пожалуйста ... Такова необычайная широта приложений математических моделей распознавания образов.

Теперь, я думаю, Вы готовы узнать, как же найти разделяющую поверхность, во всяком случае, важность этого Вам ясна.

АЛГОРИТМ ОБУЧЕНИЯ

Выше были описаны математические модели объекта и образа. Объект – это точка многомерного пространства его признаков, образ – это множество, облако таких точек. Распознавать образы – это указывать, к какому из облаков-образов принадлежит заданная точка. Обучение показами – это когда предъявляются какие-то точки x и указывается, к какому из множеств-образов они принадлежат, а в результате вырабатывается способность определять, к какому из множеств-образов принадлежит новая, в том числе и ранее не показываемая, точка.

Перейдем теперь к вопросу о том, как осуществить обучение показами, с помощью какого алгоритма его достигнуть. Как уже говорилось, этого можно достигнуть путем построения разделяющей поверхности. Но как ее найти? Об этом и пойдет речь. Будем искать, например, разделяющую поверхность в виде

$$f(x, a) = a_1\varphi_1(x_1, \dots, x_n) + \dots + a_m\varphi_m(x_1, \dots, x_n) = 0, \quad (2)$$

где x – точка многомерного пространства X с координатами-признаками x_1, \dots, x_n , а a_1, \dots, a_m – неизвестные параметры, которые мы хотим подобрать так, чтобы она была искомой разделяющей, то есть чтобы для всех точек $x \in A$ $f(x, a) < 0$, а для всех $x \in B$ $f(x, a) > 0$. Функции $\varphi_s(x_1, \dots, x_n)$ известны и подобраны так, что мы можем надеяться найти требуемую разделяющую поверхность. Согласно сказанному, искомые значения $a_1^*, a_2^*, \dots, a_m^*$ параметров можно найти из системы неравенств

$$\begin{aligned} f(x, a^*) < 0 & \quad \text{при } x \in A, \\ f(x, a^*) > 0 & \quad \text{при } x \in B. \end{aligned} \quad (3)$$

Ясная математическая задача поставлена. Однако как ее решить? Точек x очень много, все не перечислишь, параметров a_s также не мало. Персептрон Розенблатта делал это следующим образом. Возьмем произвольные начальные значения параметров a_s , $s = 1, 2, \dots, m$, и будем последовательно показывать точки множеств-облаков A и B , указывая каждый раз с помощью величины ξ , к какому из образов

A или B она принадлежит. Если $\xi = -1$, то показываемое x принадлежит A , если $\xi = +1$, то $x \in B$. При этих показах одновременно выясняется, а верно ли, в соответствии с условиями (3), распознает объект x персептрон. Если верно, его не трогают. Если неверно, меняют параметры a_s в соответствии с формулами

$$\bar{a}_s = a_s + \xi\varphi_s(x_1, \dots, x_n), \quad s = 1, 2, \dots, m. \quad (4)$$

Дальше за параметры персептрона a_s принимают \bar{a}_s , и показывают следующий объект, и опять либо не меняют, либо меняют параметры согласно формулам (4). Процесс очень простой и смысл его тоже, так как при каждой ошибке неверное значение функции $f(x, a)$ за счет изменения параметров становится “менее” неверным. Именно:

$$\begin{aligned} f(x, a) \longrightarrow f(x, \bar{a}) &= \sum (a_s + \xi\varphi_s(x))\varphi_s(x) = \\ &= f(x, a) + \xi \sum \varphi_s^2(x). \end{aligned}$$

Поэтому, если $f(x, a) > 0$ и $\xi = -1$ ($x \in A$), то $f(x, \bar{a}) < f(x, a)$, а если $f(x, a) > 0$ и $\xi = 1$ ($x \in B$), то $f(x, \bar{a}) > f(x, a)$. Трудно поверить, что такой простой алгоритм ведет к цели, но это так при весьма общих условиях и, самое важное, удачности выбора общего вида (2) разделяющей функции $f(x, a)$. Удачность означает принципиальную разрешимость с запасом неравенств (3). С запасом – это значит, они выполняются не только для некоторого a^* , но и всех a , достаточно близких к a^* .

Достигаемая цель состоит в том, что в описанном процессе персептрон совершает не более, чем некоторое конечное число ошибок N . Если персептрон уже их совершил, то после этого он уже никогда не ошибается, узнавая правильно даже те объекты, которые он никогда ранее не видел и которые ему не показывали.

Идея доказательства очень проста [8]. Рассмотрим изменение величины V , равной

$$V = \sum (a_s - \lambda a_s^*)^2 \geq 0, \quad \lambda > 0, \quad (5)$$

при совершении персептроном ошибки

$$\begin{aligned} \Delta V &= \sum (a_s + \xi\varphi_s - \lambda a_s^*)^2 - \sum (a_s - \lambda a_s^*)^2 = \\ &= 2\lambda\xi \left\{ \sum a_s\varphi_s - \sum a_s^*\varphi_s \right\} + \sum \varphi_s^2. \end{aligned}$$

Первый член этой суммы отрицателен, так как в силу ошибки персептрона знаки ξ и $\sum a_s\varphi_s = f(x, a)$ противоположны, а знаки ξ и $\sum a_s^*\varphi_s = f(x, a^*)$ одинаковы, и по величине неограниченно растет вместе с λ . Это позволяет выбрать λ так, что убывание величины V при каждой ошибке персептрона не менее, чем на $q > 0$. Из этого и следует конечность числа ошибок, так как величина V не может стать отрицательной.

ЗАКЛЮЧЕНИЕ

Есть поговорка: “За одного битого двух небитых дают”. Раз побиты, значит, в чем-то ошибся. Поэтому поговорку можно перефразировать и так: “За одного ошибавшегося двух неопытных дают”. Ученик, чтобы познать новое, должен ошибиться положенное ему число раз. Поэтому ученику мы стараемся давать постепенно задачи с разумными подвохами, так, чтобы, ошибаясь и осмысливая ошибки, он быстрее поумнел. Высказанным я хочу обратить Ваше внимание на то, что законы обучения перцептрона и ученика схожи, что в основе обучения обоих, возможно, лежат общие закономерности обучения, порождаемые общностью их математических моделей, тех самых, о которых я Вам рассказал. Розенблатт шел от принципов работы мозга, нейродинамики, и придумал перцептрон. Мы же от перцептрона, от его принципа обучения и общих математических моделей вернулись к человеческому мозгу. Круг замкнулся, обогатившись новыми знаниями, удивительным фактом о конечности числа ошибок при обучении, о связи этого предельного числа ошибок с выбором обучаемой функции $f(x, a)$ и влиянии на обучение порядка показа представитель образ.

Всеобъемлющая значимость общих абстрактных моделей математики не только в их конкретных приложениях, но еще более, пожалуй, в их эвристических возможностях. В своей научной деятельности я столкнулся с рассматриваемыми общими принципами обучения, разрабатывая адаптивные (самообучающиеся) системы управления. Так вот, для ускорения их адаптации и улучшения их работы лучше, если они в самом начале покажут все свои “фокусы” и нарушения хорошего управления, затем, после этого, они будут работать надежнее и лучше.

Только столкнув при обучении будущего оператора атомной станции с возможными аварийными ситуациями, его можно допустить к самостоятельной работе: ошибаться нужно при обучении, а не на работе. Для этого, конечно, необходимо провести предварительно большую исследовательскую работу не только технической стороны, но и динамики атомного реактора и станции (создав достаточно хорошую их математическую модель). И еще: в плане подчеркивания важнейшей эвристической роли общих абстрактных математических моделей я хотел бы обратить внимание на то, что понимание общих моделей обучения наталкивает на мысль о желательности к своему процессу обучения относиться критически, оценивая как бы со стороны его результаты и, в зависимости от этой оценки, корректировать свои дальнейшие действия. Именно эта общая идея побудила нас в разрабатываемой самообучающейся системе принятия решений в трудно формализуемых условиях ввести отдельный блок критики. Этот блок анализирует ошибки и подсказывает коррективы, которые нужно внести в алгоритм распозна-

вания: расширить описание объекта и как именно, запросить дополнительные обучающие данные и какие именно, изменить кодировку признаков, в каких случаях к ответам алгоритма можно относиться с полным доверием и в каких нельзя и др.

После высказанных общих соображений о процессе обучения и распознавания укажем, что же распознавание позволяет получить, например, для диагностики рака желудка, отделяющей его от других, сходных по симптомам, заболеваний: язвы, полипоза, гастрита, доброкачественной опухоли. Оказывается, что только по результатам иммунологического анализа крови, дающего девять чисел x_1, x_2, \dots, x_9 — количество в ней разных типов клеток (лейкоцитов, лимфоцитов, В-лимфоцитов и др.), можно с достоверностью порядка 98% выяснить, рак ли это. Диагноз “рак” или “не рак”, зависит от знака функции $f(x, a)$, равной

$$f(x, a) = 0,02x_1 + 0,207x_2 + 0,31x_3 + 0,75x_4 + 0,618x_5 + 0,154x_6 + 0,212x_7 + 0,403x_8 - 0,492x_9 - 0,5,$$

который легко вычисляется по данным анализа крови.

ЛИТЕРАТУРА

1. *Неймарк Ю.И.* Математика как операционная система и модели // Соросовский Образовательный Журнал. 1996. № 1. С. 82 – 85.
2. *Rosenblatt F.* The Perceptron, A Perceiving and Recognizing Automation / Project PARA, Cornell Aeronaut. Lab. Rep. № 85-460-1. Jan. 1957.
3. *Rosenblatt F.* Two Theorems of Statistical Separability in the Perceptron, Proceedings of Symposium on the Mechanization of Thought Processes, London, 1959. PP. 421 – 456.
4. *Эшби У.Р.* Конструкция мозга. М.: ИЛ, 1962.
5. *Розенблатт Ф.* Принцип нейродинамики. М.: Мир, 1966.
6. *Неймарк Ю.И., Баталова З.С., Васин Ю.Г., Брейдо М.Д.* Распознавание образов и медицинская диагностика. М.: Наука, 1972.
7. *Неймарк Ю.И., Теклина Л.Г., Таранова Н.Н., Котельников И.В.* Обучающаяся статистическая консультативная система. Межвузовский сборник “Динамика систем”, Изд. Нижегородского ун-та, 1995. С. 3 – 28.
8. *Novikoff A.* On Convergence Proofs for Perceptrons. Proceedings of Symposium on Mathematical Theory of Automata. Polytechnic Institute of Brooklyn, 1963. V. XII.

* * *

Юрий Исаакович Неймарк, доктор технических наук, профессор Нижегородского государственного университета, академик Российской академии естественных наук. Лауреат Международной премии Винера, лауреат премии АН СССР им. А.А. Андропова. Автор более 400 научных работ и 8 монографий.