

КРАТКИЕ СООБЩЕНИЯ

УДК 621.391.19.3

М. И. МОГИЛЬНИЦКИЙ, А. Г. СЕНИН
 (Новосибирск)

О КРИТЕРИЯХ ОПТИМАЛЬНОСТИ В ЗАДАЧАХ ОБУЧЕНИЯ

В практике исследований по распознаванию образов широко используются рекуррентные алгоритмы обучения [1]. Такие алгоритмы минимизируют функционал ошибки между реакцией оператора и выходным сигналом разделяющей поверхности. Ограничимся примерами наиболее простых критериев среднего квадрата ошибки

$$I_{\text{кв}} = \left(y - \sum_{j=1}^n c_j \varphi_j(x) \right)^2 \quad (1)$$

и модульного

$$I_{\text{м}} = \left| y - \sum_{j=1}^n c_j \varphi_j(x) \right|, \quad (2)$$

где y — реакция оператора, т. е. величина, характеризующая класс принадлежности образа; c_1, c_2, \dots, c_n — коэффициенты разложения разделяющей поверхности; $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ — система линейно независимых функций. Заметим, что в общем случае минимизация их не обеспечивает минимума ошибок распознавания. Так, для двух одномерных нормальных распределений $f_A(x)$ и $f_B(x)$ с математическими ожиданиями m_A, m_B и среднеквадратическими отклонениями σ_A и σ_B и разделяющей поверхности вида

$$v(x) = x - c$$

оптимальное значение порога, которое определяем из выражения

$$\frac{d I_{\text{кв}}}{d c} = \int_{-\infty}^{\infty} (-1 - x + c) f_A(x) dx + \int_{-\infty}^{\infty} (1 - x + c) f_B(x) dx = 0, \quad (3)$$

равно

$$c_{\text{кв}} = \frac{m_A + m_B}{2}. \quad (4)$$

Для модульного критерия значение порога, вычисляемое по формуле

$$\frac{d I_{\text{м}}}{d c} = \int_{-\infty}^{\infty} \text{sign}(-1 - x + c) f_A(x) dx + \int_{-\infty}^{\infty} \text{sign}(1 - x + c) f_B(x) dx = 0, \quad (5)$$

равно

$$c_{\text{м}} = \frac{(1 + m_B) \sigma_A - (1 - m_A) \sigma_B}{\sigma_A + \sigma_B}. \quad (6)$$

При этом

$$y = \begin{cases} -1 & \text{для класса А;} \\ +1 & \text{для класса В.} \end{cases}$$

Лишь в случае $c_A = c_B$ эти значения совпадают с оптимальным, определяемым из условия

$$f_A(c^*) = f_B(c^*). \quad (7)$$

По этой причине предложено много иных критериев и алгоритмов обучения, однако и они страдают отмеченными недостатками.

Минимум ошибок распознавания будет обеспечен, если минимизировать критерий вида

$$I = F\left(y - \text{sign} \sum_{j=1}^n c_j \varphi_j(x)\right), \quad (8)$$

где $F(\cdot)$ — выпуклая функция, причем вид функции не влияет на окончательный результат обучения. К сожалению, минимизировать его алгоритмами стохастической аппроксимации не представляется возможным, поскольку производная функционала по искомым параметрам везде равна нулю, за исключением отдельных значений. Для возможности дифференцирования критерия с целью определения градиента функционала несущественно изменим его аргумент:

$$I = F\left(y - \sum_{j=1}^n c_j \varphi_j(x) - \alpha \text{sign} \sum_{j=1}^n c_j \varphi_j(x)\right). \quad (9)$$

Введение пороговой функции позволяет придать больший вес ошибкам распознавания, причем величина его определяется коэффициентом α . На рис. 1 представлена зависимость модульного функционала ошибки (9)

$$I = f\left(y - \sum_{j=1}^n c_j \varphi_j(x)\right) \quad (10)$$

при разных значениях коэффициента α и

$$y = \begin{cases} +1; & x \in A; \\ -1; & x \in B. \end{cases}$$

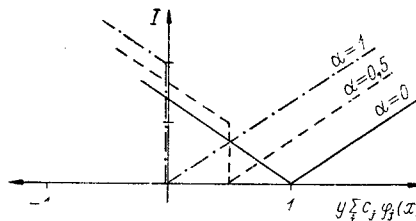


Рис. 1.

Очевидно, что при α , близких к единице, коэффициенты, минимизирующие критерий ошибки (9), обеспечивают также и минимум ошибок распознавания, а вид выпуклой функции не меняет существенно искомые значения настраиваемых параметров.

Для ранее рассмотренного примера нормальных распределений и видоизмененного модульного критерия с добавлением в аргумент пороговой функции оптимальное значение порога определяется из условия

$$\alpha (f_A(c) - f_B(c)) + \int_{-\infty}^{c+1-\alpha} f_A(x) dx - \int_{c-1+\alpha}^{\infty} f_B(x) dx = 0. \quad (11)$$

Из (11) можно видеть, что при достаточно больших значениях коэффициента α , когда значения двух последних интегралов в выражении (11) близки к нулю, минимум критерия достигается при пороговом значении α , для которого

$$f_A(c) = f_B(c), \quad (12)$$

что обеспечивает также и оптимальное распознавание.

Рис. 2 отражает зависимость расчетного значения порога c для квадратичного, модульного и предлагаемого критерия при различении двух распределений:

$$f_A = \frac{1}{1,37 \sqrt{2\pi}} e^{-\frac{(x-4,7)^2}{2 \cdot 1,37^2}}; \quad f_B = \delta(x-6). \quad (13)$$

Там же представлена экспериментальная зависимость

$$c = f(\alpha), \quad (14)$$

удовлетворительно согласующаяся с расчетными результатами.

Эксперименты проводились на макете классификатора шумов, разработанного в ИАЭ СО АН СССР [2]. Упрощенная схема, использовавшаяся при моделировании, пред-

ставлена на рис. 3. Алгоритм обучения рекуррентный. Рис. 4 иллюстрирует процесс настройки искомого параметра c и сходимости к оптимальному значению.

Введение пороговой функции незначительно усложняет алгоритм адаптации и его реализацию в специализированном устройстве, но позволяет оптимизировать разделяющую поверхность из условия минимума ошибок распознавания.

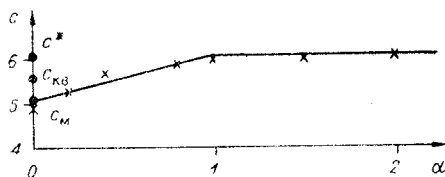


Рис. 2.

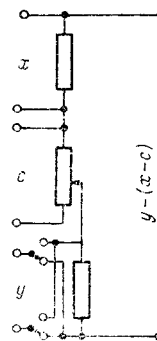


Рис. 3.

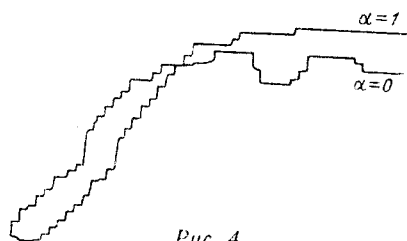


Рис. 4.

ЛИТЕРАТУРА

1. Я. З. Цыпкин. Адаптация и обучение в автоматических системах. М., «Наука», 1968.
2. Б. Д. Борисов, М. И. Могильницкий, А. Г. Сенин. К анализу работы классификатора шумов. — Автометрия, 1970, № 3.

Поступило в редакцию
26 февраля 1970 г.

УДК 621.317.3 : 531.717

А. С. БЕРШТЕЙН, Н. Н. КАРИМОВ, С. Ф. КОРНДОРФ
(Москва)

ОБ ОДНОМ МЕТОДЕ СОКРАЩЕНИЯ ОБЪЕМА ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ

Сущность разработанного метода [1, 2], позволяющего существенно сокращать объем измерительной информации, можно проиллюстрировать на примере бесконтактного измерения толщины диэлектрических покрытий на металлической подложке. В этом случае могут быть использованы различные типы датчиков: емкостные, индуктивные, оптические, радиоактивные, ультразвуковые. При контроле толщины покрытия, нанесенного на движущуюся немагнитную металлическую подложку, на выходные параметры датчиков кроме контролируемой толщины x_1 влияют: x_2 — величина воздушного зазора между датчиком и поверхностью покрытия, которая может изменяться при движении подложки как вследствие неравномерности толщины последней, так и вследствие загрязнения поверхности опоры, по которой перемещается подложка; x_3 — перекос поверхности подложки относительно датчика, изменяющийся при ее перемещении; x_4 — толщина подложки, изменяющаяся в пределах допуска. Поэтому при составлении уравнений,