УДК 519.95

МЕТОДЫ И АЛГОРИТМЫ СЕЛЕКЦИИ ПРИЗНАКОВ С ПОМОЩЬЮ КРИТЕРИЕВ ЗАДАННЫХ В ВИДЕ ФУНКЦИОНАЛОВ k-ПОРЯДКА

Фазылов Ш.Х.
д.т.н., проф., зам.директора Центра разработки программных продуктов и аппаратно-программных комплексов при Ташкентском университете информационных технологий,
тел.: +(99890) 927-72-13, e-mail: sh.fozilov@mail.ru

Маматов Н.С.
к.т.н., старший научный сотрудник Центра разработки программных продуктов и аппаратно-программных комплексов при Ташкентском университете информационных технологий,
тел.: +(99893) 506-24-57, e-mail: m_narzullo@mail.ru

На основе критериев, заданных в виде однородных функционалов нулевого порядка, разработаны почти 50 методов формирования информативных наборов признаков. Однако в настоящее время отсутствуют подобные методы для критериев, заданных в виде однородных функционалов k - порядка. Поэтому для восполнения данного пробела в статье предлагаются методы и алгоритмы выбора информативных признаков.

Ключевые слова: признак, объект, класс, критерий, информативный признак.

METHODS AND ALGORITHMS OF SELECTION FEATURES BY THE CRITERIA OF FUNCTIONALS OF K-TYPE

Fazilov Sh.X., Mamatov N.S.

Based on the given criteria in the form of homogeneous zero-order functionals, almost 50 methods for the formation of informative character sets have been developed. However, at the present time there are no such methods for the criteria given in the form of homogeneous k -type functionals. Therefore, to fill this gap, we propose methods for selecting informative features.

Keywords: feature, object, class, criterion, informative feature.

k-TARTIBLI FUNKSIONALLAR KO'RINISHIDA BERILGAN MEZONLAR ASOSIDA BELGILARNI AJRATISH USUL VA ALGORITMLARI

Fozilov Sh.X., Mamatov N.S.

Nolinchi tartibli bir jinsli funksionallar ko'rinishida berilgan mezonlar asosida informativ belgilar majmuasini shakllantirishning 50 ga yaqin usullari ishlab chiqilgan. Biroq hozirgi kunda k -tartibli bir jinsli funksionallar ko'rinishida berilgan mezonlar asosida informativ belgilar majmuasini shakllantirishning usullari mavjud emas. Mazkur kamchilikni bartaraf etish uchun maqolada mazkur mezonlar asosida informativ belgilarni ajratishning usul va algoritmlari taklif etilgan.

Tayanch iboralar: belgi, obyekt, sinf, mezon, informativ belgi.

1. Введение

В настоящее время системы распознавания образов широко применяются для решения задач управления. Это вызвано тем, что современные сложные технические объекты часто не могут быть описаны достаточно корректно и всесторонне математическими уравнениями. Управление такими объектами, неподдающимися строгому математическому описанию, является важной первостепенной задачей.

Опыт создания систем управления сложными объектами показал, что необходимо использовать все возможности классической теории управления, но там, где это невозможно, следует применять адаптивные, обучающиеся и самообучающиеся системы, в том числе обучающиеся системы распознавания.

Разработка системы распознавания образов требует решения ряда сложных задач, в том числе определения на базе исходного множества признаков такого информативного признака, который обеспечивает максимальную эффективность решений, принимаемых системой управления на основании результатов решения задачи распознавания [1-4].

В статье рассматривается задача выбора набора информативных признаков с помощью однородного критерия с k-й степенью.

2. Постановка и решение задачи

ОПРЕДЕЛЕНИЕ. Если для $\forall \alpha \in R, (\alpha \neq 0)$ выполняется равенство $f(\alpha x_1, \alpha x_2, ..., \alpha x_k) = \alpha^k f(x_1, x_2, ..., x_k)$, то f называется однородной функцией *k*-порядка.

Пусть дан критерий

$$I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)(c,\lambda)}, \quad (1)$$

который является однородным критерием «-1» порядка. Рассмотрим следующую задачу:

$$\begin{cases} I(\lambda) = \dfrac{(a,\lambda)}{(b,\lambda)(c,\lambda)} \to \max, \\ \lambda \in \Lambda^l, \lambda_i = \{0,1\}, i = \overline{1,N}, \\ a,b \in R^N, a_i \geq 0, b_i > 0, i = \overline{1,N}. \end{cases} \quad (2)$$

Введем обозначения:

$$A = \sum_{i=1}^{l} a_i, \; B = \sum_{i=1}^{l} b_i, \; C = \sum_{i=1}^{l} c_i,$$

$$\begin{cases} \Delta a_{ij} = a_j - a_i, \\ \Delta b_{ij} = b_j - b_i, i = \overline{1,l}, j = \overline{l+1,N}. \end{cases}$$

Пусть даны действительные числа x, y, z, а также $w, e, f > 0$, удовлетворяющие условиям $y + e > 0, z + f > 0$. Тогда имеет место одна из следующих лемм.

ЛЕММА 1. Если $\frac{a}{d} \geq \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$, тогда имеет место $\frac{d}{ef} \leq \frac{d+a}{(b+e)(c+f)}$.

Доказательство:

$$\frac{a}{d} \geq \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef} \Rightarrow aef \geq dbf + dce + dbc \Rightarrow$$

$$\Rightarrow aef + def \geq dbf + dce + dbc + def \Rightarrow$$

$$\Rightarrow ef(a+d) \geq d(b+e)(c+f) \Rightarrow$$

$$\Rightarrow \frac{a+d}{(b+e)(c+f)} \geq \frac{d}{ef}.$$

ЛЕММА 2. Если $\frac{a}{d} < \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$, тогда имеет место $\frac{d}{ef} > \frac{d+a}{(b+e)(c+f)}$.

Доказательство:

$$\frac{a}{d} < \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef} \Rightarrow aef < dbf + dce + dbc \Rightarrow$$

$$\Rightarrow aef + def < dbf + dce + dbc + def \Rightarrow$$

$$\Rightarrow ef(a+d) < d(b+e)(c+f) \Rightarrow$$

$$\Rightarrow \frac{a+d}{(b+e)(c+f)} < \frac{d}{ef}.$$

Если в леммах 1 и 2 $a = \Delta a_{ij}, b = \Delta b_{ij}, c = \Delta c_{ij}$, $d = A, \; e = B, \; f = C$, тогда для $\forall i, j$ одновременно выполняются неравенства $A + \Delta a_{ij} \geq 0; \; B + \Delta b_{ij} \geq 0;$ $C + \Delta c_{ij} \geq 0$, и одна из указанных выше лемм (1 или 2) будет верной.

Пусть выбран $\forall \lambda \in \Lambda^\ell$.

ТЕОРЕМА 1. Для того чтобы выбранный вектор λ обеспечил оптимальное решение задачи (2), необходимо и достаточно отсутствие $a = \Delta a_{ij}, b = \Delta b_{ij}$ и $c = \Delta c_{ij}$ $\left(i = \overline{1,l}, \; j = \overline{l+1,N}\right)$, удовлетворяющих условиям леммы 1.

Доказательство. ***Достаточность.*** Пусть дано $\forall \lambda \in \Lambda^\ell$. Тогда выражения (произвольные суммы)

$$\begin{cases} A^* = (a,\lambda) = \sum_{i=1}^{N} a_i \lambda_i, \\ B^* = (b,\lambda) = \sum_{i=1}^{N} b_i \lambda_i, \\ C^* = (c,\lambda) = \sum_{i=1}^{N} c_i \lambda_i \end{cases}$$

можно привести в следующий вид:

$$\begin{cases} A^* = A + \sum_{k=1}^{p} \Delta a_{fk}, \\ B^* = B + \sum_{k=1}^{p} \Delta b_{fk}, \\ C^* = C + \sum_{k=1}^{p} \Delta c_{fk}. \end{cases}$$

Сохраняя l, информативность вектора λ, f и k определяется следующим образом:

1. Если $\lambda_i = 0$ и $\lambda_j = 1$, то $f = j$ и $k = i$ $\left(i = \overline{1,l}, \; j = \overline{l+1, N}\right)$.

2. Если $\lambda_i = 1$ и $\lambda_j = 0$, то $f = i$ и $k = j$ $\left(i = \overline{1,l}, \; j = \overline{l+1, N}\right)$.

Для A^*, B^*, C^* имеет место равенство

$$\begin{cases} A^* = A + A_1 + A_2, \\ B^* = B + B_1 + B_2, \\ C^* = C + C_1 + C_2, \end{cases}$$

где A_m и B_m - сумма $\Delta a_{ij}, \Delta b_{ij}$ и Δc_{ij}, удовлетворяющая условиям *m*-леммы $\left(m = \overline{1, 2}\right)$.

По условию теоремы сумма A_1, B_1 и C_1 равна нулю. Значит,

$$\begin{cases} A^* = A + A_2, \\ B^* = B + B_2, \\ C^* = C + C_2. \end{cases}$$

Так как каждое слагаемое сумм $A_2 = \sum_{z=1}^{p} \Delta a_{ji}^{(z)}$, $B_2 = \sum_{z=1}^{p} \Delta b_{ji}^{(z)}$ и $C_2 = \sum_{z=1}^{p} \Delta c_{ji}^{(z)}$

удовлетворяет условиям леммы 2, то $\frac{\sum_{z=1}^{p}\Delta a_{ji}^{(z)}}{A} < \frac{\sum_{z=1}^{p}\Delta b_{ji}^{(z)}}{B} + \frac{\sum_{z=1}^{p}\Delta c_{ji}^{(z)}}{C} + \frac{\sum_{z=1}^{p}\Delta b_{ji}^{(z)}\Delta c_{ji}^{(z)}}{BC}$, отсюда следует $\frac{A+\sum_{z=1}^{p}\Delta a_{ji}^{(z)}}{\left(B+\sum_{z=1}^{p}\Delta b_{ji}^{(z)}\right)\left(C+\sum_{z=1}^{p}\Delta c_{ji}^{(z)}\right)} < \frac{A}{BC}$ и $\frac{A+A_2}{(B+B_2)(C+C_2)} < \frac{A}{BC}$.

Необходимость. Пусть существуют $\Delta a_{ij}, \Delta b_{ij}$ и Δc_{ij}, удовлетворяющие условиям леммы 1. Тогда в соответствии с этой леммой получаем $\frac{A+\Delta a_{ij}}{\left(B+\Delta b_{ij}\right)\left(C+\Delta b_{ij}\right)} \geq \frac{A}{B}$.

Так как каждое слагаемое сумм

$$A_1 = \sum_{z=1}^{p}\Delta a_{ji}^{(z)},\ B_1 = \sum_{z=1}^{p}\Delta b_{ji}^{(z)} \text{ и } C_1 = \sum_{z=1}^{p}\Delta c_{ji}^{(z)}$$

удовлетворяет условиям леммы 1, то $\frac{\sum_{z=1}^{p}\Delta a_{ji}^{(z)}}{A} \geq \frac{\sum_{z=1}^{p}\Delta b_{ji}^{(z)}}{B} + \frac{\sum_{z=1}^{p}\Delta c_{ji}^{(z)}}{C} + \frac{\sum_{z=1}^{p}\Delta b_{ji}^{(z)}\Delta c_{ji}^{(z)}}{BC}$, отсюда следует

$$\frac{A+\sum_{z=1}^{p}\Delta a_{ji}^{(z)}}{\left(B+\sum_{z=1}^{p}\Delta b_{ji}^{(z)}\right)\left(C+\sum_{z=1}^{p}\Delta c_{ji}^{(z)}\right)} \geq \frac{A}{BC}.$$

Таким образом, имеем $\frac{A+A_1}{(B+B_1)(C+C_1)} > \frac{A}{B}$, и значение $I(\lambda) = \frac{A}{BC}$, соответствующее выбранному λ не является оптимальным.

Теорема доказана.

Если вектор λ не является оптимальным решением задачи (2), то осуществляются замены на основе леммы 1.

Процесс замены продолжается до момента, когда исчерпаны все $\Delta a_{ij}, \Delta b_{ij}$ и Δc_{ij}, удовлетворяющие условиям леммы 1 и при этом в соответствии с теоремой 1 найденное решение будет оптимальным.

В данном методе значения функционала и компоненты вектора λ определяются следующим образом.

Допустим, для $\Delta a_{ij}, \Delta b_{ij}$ и Δc_{ij} справедлива лемма 1. В этом случае, в соответствии с этой леммой имеет место соотношение $\frac{A+\Delta a_{ij}}{\left(B+\Delta b_{ij}\right)\left(C+\Delta b_{ij}\right)} \geq \frac{A}{B}$ и значения компонент i и j вектора λ взаимоменяются.

Процесс последовательной взаимозамены продолжается до тех пор, пока не будут выполнены условия теоремы 1.

Метод, основанный на этой теореме, обозначен нами как метод «Дельта-4». Его алгоритм состоит из следующих шагов.

Шаг 1. Выбирается $\lambda^{(0)} = \{\underbrace{1,\ 1,\ \dots,\ 1}_{\ell}, 0, 0, \dots 0\}$.

Шаг 2. Вычисляется начальное значение $I\left(\lambda^{(0)}\right)$.

Шаг 3. Осуществляется присвоение $\lambda = \lambda^{(0)}$ и $j = \ell$.

Шаг 4. Осуществляется сравнение j *и* N. Если $j = N$, то $j = \ell + 1$, в противном случае $j = j + 1$.

Шаг 5. Проверяется выполнение условий леммы 1. Если они выполняются, то в соответствии с правой частью этой леммой осуществляются преобразования и переход на очередной шаг, в противном случае – возврат на шаг 4.

Шаг 6. Если $I(\lambda) = I(\lambda^*)$, то λ - оптимальное решение и процесс останавливается, в противном случае осуществляется переход на шаг 4.

Данный алгоритм обозначен нами как A_4.

Теперь рассмотрим метод «Градиент» для решения задачи (2).

Для этого вводим следующие обозначения:

$$A = (a,\lambda),\ B = (b,\lambda),\ D = (d,\lambda),$$

$$A_1 = (a,\mu),\ B_1 = (b,\mu),\ D_1 = (d,\mu).$$

Тогда соответствующие значения функционала (1) равны $I(\lambda) = \frac{A}{BD}$ и $I(\mu) = \frac{A_1}{B_1 D_1}$.

Для того чтобы решить задачу (2), воспользуемся вспомогательным направляющим вектором $C(c_1, c_2, ..., c_n)$, и его координаты определим следующим образом:

$$c_i = \sqrt{\frac{1}{\ell}\cdot\frac{a_i}{A} - \frac{b_i}{B} - \frac{c_i}{C}}.$$

Для выбранного вектора λ и формированного на его основе вектора справедлива следующая теорема.

ТЕОРЕМА 2. Для того чтобы выбранный вектор λ обеспечил оптимальное решение задачи (2), необходимо и достаточно отсутствие вектора μ, удовлетворяющего условию $(C,\mu) > 0$.

Доказательство. ***Достаточность.*** По условию теоремы не существует вектора $\mu \in \Lambda^{\ell}$, удовлетворяющего условию $(D,\mu) > 0$. Тогда от $\frac{A}{BC} \geq \frac{A_1}{B_1 C_1}$ следует

$$\frac{B_1 C_1}{BC} \geq \frac{A_1}{A}, \tag{3}$$

$$\left|\frac{t_1 + t_2 + ... + t_n}{n}\right| \leq \sqrt{\frac{t_1^2 + t_2^2 + ... + t_n^2}{n}}, \tag{4}$$

где $t_i \in R\left(i=\overline{1,n}\right)$.

Из неравенств (3) и (4) получаем следующее неравенство:

$$\left(\frac{B_1}{B}+\frac{C_1}{C}\right) \geq \sqrt{\frac{B_1 C_1}{BC}} \geq \sqrt{\frac{A_1}{A}} \geq \geq \frac{1}{\ell}\sqrt{\frac{A_1}{A}} \geq \sum_{i=1}^{1}\sqrt{\frac{1}{\ell}\cdot\frac{a_{m_i}}{A}}. \quad (5)$$

Из (5) следует $(D,\ \mu) \leq 0$.

Необходимость. Допустим, существует μ, удовлетворяющий условию $(D,\mu) > 0$.

Из $(D,\mu)=\sum_{j=1}^{\ell} d_{m_j} > 0$ следует

$$\sum_{j=1}^{\ell}\sqrt{\frac{1}{\ell}\cdot\frac{a_{m_j}}{A}} > \sum_{j=1}^{\ell}\frac{b_{m_j}}{B}+\sum_{j=1}^{\ell}\frac{c_{m_j}}{C}.$$

Из неравенства Коши выражение (3) приобретает следующий вид:

$$\sum_{j=1}^{\ell}\sqrt{\frac{1}{\ell}\cdot\frac{a_{m_j}}{A}} > \sum_{j=1}^{\ell}\frac{b_{m_j}}{B}+\sum_{j=1}^{\ell}\frac{c_{m_j}}{C} \geq \geq \sqrt{\sum_{j=1}^{\ell}\frac{b_{m_j}}{B}\cdot\sum_{j=1}^{\ell}\frac{c_{m_j}}{C}}. \quad (6)$$

Из $a_{m_j} \geq 0$ и неравенства (4) выражение (6) приводится к следующему виду:

$$\sqrt{\frac{\sum_{j=1}^{\ell} a_{m_j}}{A}} \geq \sum_{j=1}^{\ell}\sqrt{\frac{1}{\ell}\cdot\frac{a_{m_j}}{A}} > > \sum_{j=1}^{\ell}\frac{b_{m_j}}{B}+\sum_{j=1}^{\ell}\frac{c_{m_j}}{C} \geq \sqrt{\sum_{j=1}^{\ell}\frac{b_{m_j}}{B}\cdot\sum_{j=1}^{\ell}\frac{c_{m_j}}{C}}. \quad (7)$$

Из (7) получаем следующее:

$$\frac{\sum_{j=1}^{\ell} a_{m_j}}{A} \geq \sum_{j=1}^{\ell}\frac{b_{m_j}}{B}\cdot\sum_{j=1}^{\ell}\frac{c_{m_j}}{C} \Rightarrow$$

$$\Rightarrow \frac{\sum_{j=1}^{\ell} a_{m_j}}{\sum_{j=1}^{\ell} b_{m_j}\cdot\sum_{j=1}^{\ell} c_{m_j}} \geq \frac{A}{BC},$$

т.е. $\frac{A_1}{B_1 C_1} \geq \frac{A}{BC}$.

Теорема доказана.

Введем оператор (следования)

$$\mu : \Lambda^{\ell} \rightarrow \Lambda^{\ell}$$

такой, при котором

$$\left(\phi(\lambda),\mu(\lambda)\right)=\max_{\eta\in\Lambda^{\ell}}\left(\phi(\lambda),\eta\right).$$

Оператор μ имеет очевидное конструктивное представление. Если упорядочить компоненты вектора $\phi(\lambda)$, т.е. найти набор попарно различных индексов $j_1, j_2, ..., j_N$ таких, при которых

$$\phi^{j_1}(\lambda) \geq \phi^{j_2}(\lambda) \geq ...\phi^{j_N}(\lambda),$$

то компоненты вектора $\mu(\lambda)$ будут определены как

$$\mu^{j_1}(\lambda)=1, \mu^{j_2}(\lambda)=1, ...,$$
$$\mu^{j_\ell}(\lambda)=1, \mu^{j_{\ell+1}}(\lambda)=0,$$
$$\mu^{j_{\ell+2}}(\lambda)=0, ..., \mu^{j_N}(\lambda)=0.$$

Иначе говоря, компоненты вектора $\mu(\lambda)$, соответствующие первым ℓ - максимальным компонентам вектора $\phi(\lambda)$, равны единице, остальные - нулю.

Очевидно, что $\mu(\lambda)$ также является ℓ - информативным вектором, причем

$$\left(\phi(\lambda),\mu(\lambda)\right)= = \max\left\{\left(\phi(\lambda),\eta\right) \middle| \eta \in \Lambda^{\ell}\right\}.$$

Нахождение оптимального решения осуществляется аналогично методу «Градиент».

Алгоритм, решающий задачу (2), обозначен как A_5 и состоит из следующих шагов.

Шаг 1. Выбирается $\lambda=\{\underbrace{1,\ 1,\ ...\ ,\ 1}_{\ell},\ 0,\ 0,\ ...\ 0\}$.

Шаг 2. Вычисляются A и B_j, т.е. $A=(a,\lambda), B_j=\left(b^{(j)},\lambda\right)$.

Шаг 3. Определяются компоненты вектора С: $c_i=\sqrt{\frac{1}{\ell}\cdot\frac{a_i}{A}}-\frac{b_i}{B}-\frac{c_i}{C}$.

Шаг 4. Определение вектора μ. Основываясь на (9) и (10), выполняется процедура формирования вектора μ.

Шаг 5. Вычисляется скалярное произведение векторов C и μ. Если $(C,\mu)>0$, то $\lambda=\mu$ и переход на шаг 2, в противном случае λ - оптимальное решение и процесс останавливается.

3. Заключение

Задача выбора информативных признаков является в проблеме распознавания образов традиционной. Для выбора информативных признаков с помощью однородного критерия с нулевой степенью разработано большое число методов и алгоритмов.

В практике задачу выбора информативных признаков требуется исследовать с помощью однородного критерия с *k*-й степенью.

Литература

[1] *Фозилов Ш.Х., Маматов Н.С.* l-информатив белгиларни танлаб олиш усули // «Информатика ва энергетика

муаммолари» Ўзбекистон журнали. - Тошкент, 2014. - 1-сон.

[2] *Фозилов Ш.Х., Маматов Н.С.* Информатив белгилар фазосини қуришда «Дельталар усули» // «Информатика ва энергетика муаммолари» Ўзбекистон журнали. - Тошкент, 2005. - 5-сон.

[3] *Камилов М.М., Фазылов Ш.Х., Нишанов А.Х.* Метод выбора признаков с использованием критерия информативности фишеровского типа // Узбекский журнал «Проблемы информатики и энергетики». – Ташкент, 1992. - № 2. – С. 9-12.

[4] *Камилов М.М., Фазылов Ш.Х., Нишанов А.Х.* Эффективный метод выделения информативных подсистем признаков в распознавания образов // Деп. в ВИНИТИ, 03.08.89. № 5218-В89. Ред. журн. Изв. АН УзССР. СТН. – 7 с.