

УДК 62-506.2:62I.39I.I63:5I9.24.00I.5

ПОСЛЕДОВАТЕЛЬНАЯ ПРОЦЕДУРА РАСПОЗНАВАНИЯ
ДЛЯ СЛУЧАЯ КОРРЕЛИРОВАННЫХ ИЗМЕРЕНИЙ

Г.С.Лбов, В.И.Котюков

1. В ряде работ [1,2,3] рассматривается последовательная решающая процедура распознавания образов. Учет соотношения между потерями от ошибок распознавания и стоимостью измерения значений признаков приводит к последовательному процессу их измерения. Естественно, что признаки должны быть измерены у объекта, подлежащего распознаванию, в таком порядке и таком количестве, чтобы требуемая надежность его классификации была обеспечена при минимальной стоимости измерений.

Обычно в литературе предлагаются процедуры измерения, не учитывающие статистическую взаимосвязь признаков, хотя, очевидно, что создание легко реализуемого на ЭВМ метода, использующего эту дополнительную информацию, привело бы лишь к уменьшению стоимости измерений в процессе распознавания.

В данной работе разработан алгоритм последовательного измерения для случая зависимых признаков.

Задача решается при следующих ограничениях:

- а) для простоты полагается, что стоимости измерения каждого из N признаков $\{X_1, \dots, X_N\}$ равны;
- б) повторное измерение у объекта какого-либо из признаков X_i не несет дополнительной информации;
- в) для каждого i -го класса ($i = 1, \dots, m$) известна его априорная вероятность P_i^0 , а также в результате предварительного процесса обучения определены вид и параметры функции плотности распределения его вероятности $P_i(x_1, \dots, x_N)^*$

* Ясно, что распределение $P_i(x_1, \dots, x_N)$ задаёт распределение для любого подпространства признаков.

Пусть на n -м шаге процедуры распознавания объекта у него замеры лишь n каких-то признаков ($n < N$), значения которых обозначим $(\tilde{x}^1, \dots, \tilde{x}^n)$. Незмеренные признаки, в свою очередь, обозначим $\{X^1, \dots, X^{\omega}, \dots, X^{N-n}\}$.

После проведенных измерений $(\tilde{x}^1, \dots, \tilde{x}^n)$ апостериорные вероятности классов для данного неизвестного объекта равны:

$$q_i(n) = \frac{P_i^0 \cdot P_i(\tilde{x}^1, \dots, \tilde{x}^n)}{\sum_{i=1}^m P_i^0 \cdot P_i(\tilde{x}^1, \dots, \tilde{x}^n)}; \quad i = 1, \dots, m. \quad (I)$$

Заметим, что до начала измерений имели $\{q_i(0) = P_i^0\}$.

В качестве критерия останова процедуры измерения выбрано следующее правило. Последовательное измерение значений признаков заканчивается, и выносится решение о принадлежности объекта к S -му классу, если

$$F(n) = \frac{q_S(n)}{q_Z(n)} \geq B, \quad (2)$$

где $q_S(n) = \max_i \{q_i(n)\}$;

$$q_Z(n) = \max_{i, i \neq S} \{q_i(n)\};$$

B — величина некоторого порога, связанная с требуемыми вероятностями ошибочной классификации I-го и 2-го рода [3].

Обозначим через n_0 число проведенных измерений, при котором уже выполняется соотношение (2).

Таким образом, в рассматриваемой постановке задача заключается в том, чтобы достичь выполнения соотношения (2) за минимальное (в среднем) число измерений

$$M(n_0) \rightarrow \min \quad (3)$$

(M — знак математического ожидания).

Предлагаемая стратегия состоит в том, что если на n -м шаге имеем $F(n) < B$, то на следующем $(n+1)$ -м шаге процедуры распознавания у объекта производится измерение значения того признака, для которого при имеющейся информации $(\tilde{x}^1, \dots, \tilde{x}^n)$ ожидаемое значение величины $F(n+1)$ максимально.

Заметим, что рассмотренный ниже алгоритм с небольшими изменениями пригоден и для некоторых иных критериев [3], отличных от (2).

II. Рассмотрим алгоритм. Для каждого из еще неизмеренных признаков X^{ω} определение ожидаемого значения $F(n+1)$, обозначаемого через $F_{ож}^{\omega}(n+1)$, производится по такой схеме.

I. Во всем пространстве неизмеренных признаков $\{X^1, \dots, X^{\omega}, \dots, X^{N-n}\}$ для каждого i -го класса находится вектор условных матожиданий значений соответствующих признаков:

$$x_{iож} = (x_{iож}^1, \dots, x_{iож}^{\omega}, \dots, x_{iож}^{N-n})$$

при уже имеющихся измерениях $(\tilde{x}^1, \dots, \tilde{x}^n)$. Простой численный метод нахождения $x_{iож}^{\omega}$ ($i = 1, \dots, m$) будет описан в п. III, стр. 40.

2. Полагая $\tilde{x}^{n+1} = x_{iож}^{\omega}$, найдем ожидаемые апостериорные вероятности классов:

$$q_{jож}(n+1)_i^{\omega} = \frac{P_j^0 \cdot P_j(\tilde{x}^1, \dots, \tilde{x}^n, x_{iож}^{\omega})}{\sum_{j=1}^m P_j^0 \cdot P_j(\tilde{x}^1, \dots, \tilde{x}^n, x_{iож}^{\omega})}; \quad j = 1, \dots, m. \quad (4)$$

3. Вычислим

$$F_{iож}^{\omega}(n+1) = \frac{q_{Sож}(n+1)_i^{\omega}}{q_{Zож}(n+1)_i^{\omega}}, \quad (5)$$

где $q_{Sож}(n+1)_i^{\omega} = \max_j \{q_{jож}(n+1)_i^{\omega}\}$,

$$q_{Zож}(n+1)_i^{\omega} = \max_{j, j \neq S} \{q_{jож}(n+1)_i^{\omega}\}.$$

Другими словами, $F_{iож}^{\omega}(n+1)$ есть значение критерия F для вектора $(\tilde{x}^1, \dots, \tilde{x}^n, x_{iож}^{\omega})$.

4. Аналогично определяем значения F для всех векторов $(\tilde{x}^1, \dots, \tilde{x}^n, x_{iож}^{\omega}), \dots, (\tilde{x}^1, \dots, \tilde{x}^n, x_{mож}^{\omega})$, т.е.

$$F_{1ож}^{\omega}(n+1), \dots, F_{mож}^{\omega}(n+1).$$

5. Для признака X^ω вычисляем искомую величину матожидания критерия F

$$F_{ож}^\omega(n+1) = \sum_{i=1}^m q_i(n) \cdot F_{iож}^\omega(n+1). \quad (6)$$

Точно таким же образом определяются величины $F_{ож}^\omega(n+1)$ для всех еще неизмеренных признаков ($\omega = 1, \dots, N-n$).

Находится величина

$$F_{ож}^k(n+1) = \max_{\omega} \{F_{ож}^\omega(n+1)\}. \quad (7)$$

На $(n+1)$ -м шаге измеряется k -й признак.

Необходимо заметить, что указанная последовательная процедура измерения признаков, вообще говоря, не является оптимальной с точки зрения минимума матожидания числа необходимых измерений n_0 . Известно, что в последовательных методах планирования экспериментов оптимизация процедуры на каждом шаге не всегда приводит к оптимизации общей процедуры.

Кроме того, для определения ожидаемого значения критерия F на каждом шаге необходимо усреднение по всему множеству возможных реализаций признака X^ω ($\omega = 1, \dots, N-n$). В нашем случае такое усреднение не делается, что также приводит в общем случае к отклонению от оптимальной стратегии.

Однако, усреднение по всему множеству возможных реализаций приводит к резкому увеличению времени вычислений на ЭВМ, что делает процедуру планирования трудоемкой даже для случая независимых признаков [3].

Рассматриваемая в данной работе последовательная процедура реализована для ЭВМ "Минск-22" для случая многомерных нормальных распределений с неравными матрицами ковариаций.

III. Определим вектор $x_{iож}$ для случая нормального распределения. Пусть для i -го класса $P_i(x_1, \dots, x_N) = N(\mu^i, A^i)$, где $\mu^i = (\mu_1^i, \dots, \mu_N^i)$ - вектор матожиданий признаков, а $A^i = \| \sigma_{jk}^i \|$ - матрица ковариаций.

Индекс i далее для простоты будет опущен. Так как признаки статистически зависимы, то матожидание случайной величины X_j является некоторой функцией переменных $\{X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_N\}$:

$$\mu_j(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N) =$$

$$\int x_j P(x_j/x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N) dx_j, \quad (8)$$

где $x_j \in D_j$; D_j - область значений признака X_j ;

$$P(x_j/x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N) = \frac{P(x_1, \dots, x_N)}{P(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N)}.$$

Для нормального распределения функция (8) имеет вид [4]:

$$\mu_j(x) = b_{j1}x_1 + \dots + b_{jj}x_j + b_{j,j+1}x_{j+1} + \dots + b_{jN}x_N + b_{j0}, \quad (9)$$

где

$$b_{jk} = \beta_{jk} \cdot \frac{\sigma_{jj}}{\sigma_{kk}}; \quad b_{j0} = \mu_j - \sum_{k=1}^N b_{jk} \mu_k; \quad k = 1, \dots, N; \quad k \neq j.$$

Заметим, что если $\{x_1 = \mu_1, \dots, x_{j-1} = \mu_{j-1}, x_{j+1} = \mu_{j+1}, \dots, x_N = \mu_N\}$, то $\mu_j(x) = \mu_j$.

Коэффициенты β_{jk} находятся путем решения системы линейных алгебраических уравнений [4]:

$$\left. \begin{aligned} z_{11}\beta_{j1} + \dots + z_{1,j-1}\beta_{jj-1} + z_{1,j+1}\beta_{jj+1} + \dots + z_{1N}\beta_{jN} &= z_{j1} \\ &\vdots \\ z_{N1}\beta_{j1} + \dots + z_{N,j-1}\beta_{jj-1} + z_{N,j+1}\beta_{jj+1} + \dots + z_{NN}\beta_{jN} &= z_{jN} \end{aligned} \right\}, \quad (10)$$

$$\text{где} \quad z_{jk} = \frac{\sigma_{jk}}{\sqrt{\sigma_{jj}} \sqrt{\sigma_{kk}}}.$$

Таким образом, для i -го класса на основе вектора μ и матрицы A еще на этапе обучения может быть определена (найденны коэффициенты $\{\beta_{jk}\}$) система линейных алгебраических уравнений:

$$\left. \begin{aligned} \mu_1(x) &= \sum_{k=2}^N b_{1k} x_k + b_{10}, \\ &\vdots \\ \mu_j(x) &= \sum_{k=1, k \neq j}^N b_{jk} x_k + b_{j0}, \\ &\vdots \\ \mu_N(x) &= \sum_{k=1}^{N-1} b_{Nk} x_k + b_{N0}. \end{aligned} \right\} \quad (11)$$

Поскольку мы в качестве искоемых величин x_j будем пользоваться их условными матожиданиями $\mu_j(x)$, то в системе уравнений (II) левые части $\{\mu_j(x)\}$ можно заменить на $\{x_j\}$ соответственно.

Допустим на n -м шаге процедуры распознавания мы имеем замеренный вектор значений $(\tilde{x}^1, \dots, \tilde{x}^n)$. Для того, чтобы определить вектор условных матожиданий $\alpha_{i_{ож}} = (\alpha_{i_{ож}}^1, \dots, \alpha_{i_{ож}}^n)$, необходимо полученные значения $(\tilde{x}^1, \dots, \tilde{x}^n)$ подставить на соответствующие места в систему (II). Заметим, что уравнения, у которых $\{\mu_j(x)\}$ соответствуют замеренным признакам $(\tilde{x}^1, \dots, \tilde{x}^n)$, не рассматриваются. После простых преобразований будем иметь систему $(N-n)$ уравнений с $(N-n)$ неизвестными $\{\alpha_{i_{ож}}^1, \alpha_{i_{ож}}^2, \dots, \alpha_{i_{ож}}^{N-n}\}$, решая которую получим искомый числовой вектор $(\alpha_{i_{ож}}^1, \alpha_{i_{ож}}^2, \dots, \alpha_{i_{ож}}^{N-n})$.

Л и т е р а т у р а

1. ФУ К. Последовательные методы в распознавании образов и обучении машин. М., Изд-во "Наука", 1971.
2. ХАЗЕН Э.М. Методы оптимальных статистических решений и задачи оптимального управления. М., Изд-во "Сов.радио", 1968.
3. ВОЛОШИН Г.С. Некоторые вопросы теории иерархических распознающих систем в применении к обработке гидроакустических сигналов. -В кн.: Труды Второй Всесоюзной школы-семинара по статистической гидроакустике. Новосибирск, Изд-во "Наука", 1971, с. 229-246.
4. ЛУКОМСКИЙ Я.И. Теория корреляции и её применение к анализу производства. М., Гос.статист. изд-во, 1958.

Поступила в ред.-изд.отд.
26 декабря 1973 г.