

222

МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ БССР  
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ имени В.И.ЛЕНИНА

БЕЛОРУССКИЙ РЕСПУБЛИКАНСКИЙ СОВЕТ НАУЧНО-ТЕХНИЧЕСКИХ ОБЩЕСТВ

КОМИТЕТ ПО ПРИКЛАДНОЙ МАТЕМАТИКЕ И ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКЕ БРСНТО

2 85-56  
98-17

МЕТОДЫ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ОБРАБОТКИ  
ИНФОРМАЦИИ И ПРИКЛАДНОГО СТАТИСТИЧЕСКОГО  
АНАЛИЗА ДАННЫХ НА ЗЕМ

ТЕЗИСЫ ДОКЛАДОВ

Минск, декабрь 1985



ФБ-2  
732/1

Минск 1985

ЗАКОНОМЕРНОСТИ, ИССЛЕДУЕМЫЕ МЕТОДОМ СОГЛАСОВАННЫХ ОЦЕНОК

При решении ряда задач естествознания требуется выявлять взаимосвязи между исследуемыми объектами и характеризующими их признаками. В ряде случаев эти задачи решаются методом согласованных оценок [1]. Согласованные оценки строк и столбцов таблиц данных представляют собой неподвижную точку некоторого оператора, для нахождения которой применяется итерационная процедура. Метод основан на подобном тестовому [2] способе вычисления оценок строк через оценки столбцов

$$\lambda \cdot \pi_i = \sum_{j=1}^n t_{ij} \cdot \omega_j, \mu \cdot \omega_j = \sum_{i=1}^m t_{ij} \cdot \pi_i \quad (1)$$

но отличается от тестового тем, что оценки столбцов так же точно вычисляются через оценки строк. Нагрузки строк ( $\pi_i$ ) и столбцов ( $\omega_j$ ) представляют собой первую пару собственных векторов, соответствующую максимальному собственному числу, в сингулярном разложении таблицы данных. Учетывание  $\pi_i$  и  $\omega_j$ , как показала практика, приводит к выявлению закономерностей, содержащихся в таблицах исходных данных.

Пусть совокупность эмпирических описаний объектов задана таблицей исходных данных, тогда:

Предположение 1. Исходные табличные данные, организованные в соответствии с целью исследования объектов, содержат указания на существующую закономерность, подлежащую выявлению анализирующими алгоритмами.

Предположение 2. Согласно целеуказанию, профилирующему вид искомой закономерности, анализирующие алгоритмы могут выявить скрытую закономерность, содержащуюся в таблице эмпирических данных. Предположения 1,2 накладывают ограничения на поиск возможных закономерностей целеуказанием. Поэтому введем:

Определение 1. Скрытые закономерности, выявляемые анализирующими алгоритмами на таблицах эмпирических данных, называются целеориентированными закономерностями (ЦЗ).

Существуют альтернативы:

- а) признак  $x_{i1}$  имеет отношение к исследуемой закономерности - гипотеза  $A_1$  и
- б) признак  $x_{j2}$  не имеет отношения к исследуемой закономерности - гипотеза  $A_2$ .

"Иметь отношение" суть участие данного признака в законообразовании. Законообразование признака измеряется долей участия данного призна-

нака в общем количестве элементарных актов обнаружения закономерностей. В методе согласованных оценок элементарными актами являются процедуры (I), которые устанавливают вес каждого столбца и каждой строки во взаимосвязи при оценке их согласованности. Тогда:

Предположение 3. Максимальным значениям нагрузок строк ( $\pi_i$ ) и столбцов ( $\omega_j$ ) таблиц исходных данных соответствуют максимальная согласованность объектов (строк) и признаков (столбцов) между собой и с другими объектами и признаками, имеющими меньшие веса. Следовательно, нулевые значения  $\pi_i$  и  $\omega_j$  можно считать подпадающими под гипотезу  $A_2$ , а если  $\pi_i \neq 0$  и  $\omega_j \neq 0$ , то, - под гипотезу  $A_1$ . Объекты и признаки с максимальными значениями нагрузок являются носителями силы взаимосвязи объектов и признаков.

Конкретизируем, несколько, предположение 3. Формулы (1), если сделать их подстановки друг в друга, приводят к следующим соотношениям:

$$\lambda \cdot \mu \cdot \pi_i = \sum_{\alpha=1}^m h_{\alpha i} \pi_{\alpha}, \lambda \cdot \mu \cdot \omega_j = \sum_{\beta=1}^n q_{\beta j} \omega_{\beta}, \quad (2)$$

где  $h_{\alpha i}, q_{\beta j}$  - коэффициенты, отражающие собой силу взаимосвязи строк  $S_{\alpha}, S_i$  и соответственно столбцов  $x_{\beta}, x_j$  между собой. Как видно из формул (2), чем больше эта сила взаимосвязи, тем с большим коэффициентом вес соответствующей строки (или столбца) вносит вклад в вес вычисляемой строки (столбца).

Формулы (1) - (2), как и предположения 1-3, относятся к так называемым закономерностям выраженности исследуемых свойств.

Но метод может быть применяем и для поиска закономерностей невыраженности исследуемых свойств, для чего употребляются следующие формулы вычисления нагрузок строк и столбцов:

$$\lambda \cdot \pi_i = \sum_{j=1}^n (1-t_{ij}) \cdot \omega_j, \mu \cdot \omega_j = \sum_{i=1}^m (1-t_{ij}) \cdot \pi_i \quad (3)$$

Может иметь значение также отыскание связи между закономерностями выраженности и невыраженности, которое проводится по формуле:

$$\lambda \cdot \pi_i = \sum_{j=1}^n t_{ij} \cdot \omega_j, \mu \cdot \omega_j = \sum_{i=1}^m (1-t_{ij}) \cdot \pi_i \quad (4)$$

Формулы (1) - (4) имеют программное обеспечение.

Л и т е р а т у р а

1. Метод согласованных оценок: Метод. рекомендации. /Составители: А. Н. Дмитриев, С. В. Макаров, Е. А. Смертин и др./ - Новосибирск, 1982, с.133.
2. А. Н. Дмитриев, Ю. И. Журавлев, Ф. П. Кренделев. О математических принципах классификации предметов и явлений. - В кн.: Дискретный анализ. Вып. 7. Новосибирск, Наука, 1966, с.3 - 15.