***БУТЫРНОВА Татьяна Валерьяновна*** *– кандидат экономических наук, доцент кафедры ми- ровой экономики и налоговых систем. Чебоксарский кооперативный институт (филиал) Россий- ского университета кооперации. Россия. Чебоксары. E-mail:* [*btv2303@mail.ru.*](mailto:btv2303@mail.ru)

***BUTYRNOVA, Tatyana Valeryanovna*** *– Candidate of Economics Sciences, Associate Professor of the World Economy and Tax Systems. Cheboksary Cooperative Institute (branch) of the Russian University of Cooperation. Russia. Cheboksary. E-mail:* [*btv2303@mail.ru.*](mailto:btv2303@mail.ru)

**УДК 65.011.56**

# О ПОДДЕРЖКЕ ПРИНЯТИЯ РЕШЕНИЯ В ТЕХНОЛОГИИ DATA MINING

## В.И. Возяков, В.П. Филиппов

Рассмотрено применение технологии Data Mining для кластерного и дискриминантного ана- лиза инвестиционной привлекательности предприятий потребительской кооперации.

***Ключевые слова*:** Data Mining; классификация; кластеризация; дискриминантный анализ; кластерный анализ.

**V.I. Vozyakov, V.P. Filippov.** ABOUT DECISION-MAKING SUPPORT IN DATA MINING TECHNOLOGIES

The application of data mining technology for cluster and discriminant analysis of investment attractiveness of Consumer Co-operatives is considered in this article.

***Keywords:*** Data Mining; classification; clustering; discriminant analysis; cluster analysis.

В настоящее время деятельность любого предприятия (коммерческого, производственно- го, научного и т.д.) сопровождается регистраци- ей и записью всех подробностей его функцио- нирования. Эффективное использование этой информации связано с продуктивной перера- боткой данных. Специфика современных требо- ваний к такой переработке следующая:

* данные являются разнородными (количе- ственными, качественными, текстовыми);
* результаты должны быть конкретны и по- нятны;
* инструменты для обработки данных долж- ны быть просты в использовании.

Традиционная математическая статистика при обработке данных использует концепцию усреднения по выборке. Практика показывает, что эта процедура недостаточна для принятия решений в различных сферах человеческой де- ятельности, требующей обнаружения в данных ранее неизвестных, нетривиальных, практиче- ски полезных закономерностей. Исследователь- ский аспект изучения закономерностей в данной предметной области объектов и их отношений связан с использованием средств информацион- ных и коммуникационных технологий.

В основу современной технологии Data

Mining (discovery-driven data mining) положена концепция шаблонов, отражающих фрагменты

многоаспектных взаимоотношений в данных. Эти шаблоны представляют собой закономер- ности, свойственные подвыборкам данных, которые могут быть компактно выражены в понятной человеку форме. Поиск шаблонов производится методами, не ограниченными рамками априорных предположений о струк- туре выборки и виде распределений значений анализируемых показателей.

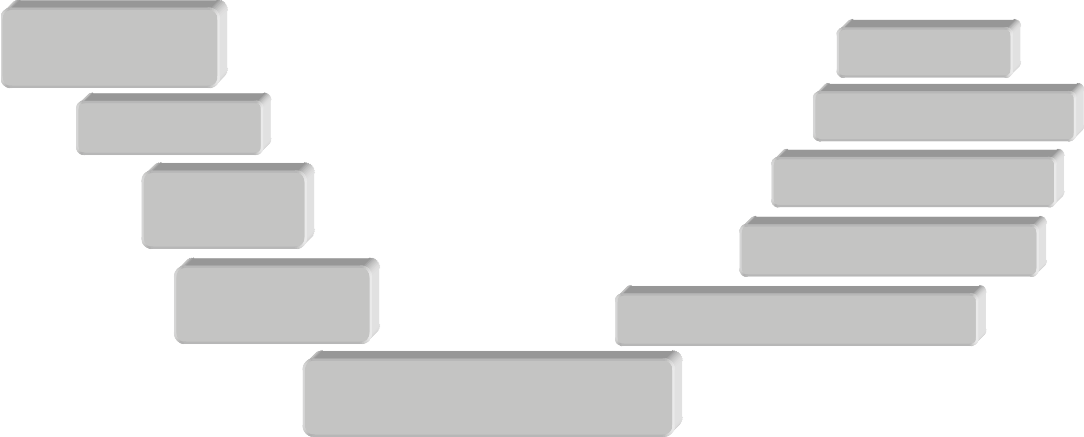
Data Mining является мультидисциплинар- ной областью, возникшей и развивающейся на базе достижений прикладной статистики, рас- познавания образов, методов искусственного интеллекта, теории баз данных и др. (рис. 1). От- сюда обилие методов и алгоритмов, реализован- ных в различных действующих системах Data Mining. Многие из таких систем интегрируют в себе сразу несколько подходов. Как правило, в каждой системе имеется какая-то ключевая ком- понента, на которую делается главная ставка.

Выделяют пять стандартных типов законо- мерностей, которые позволяют выявлять ме- тоды Data Mining: ассоциация, последователь- ность, классификация, кластеризация и про- гнозирование.

Далее остановимся на классификационном анализе, включающем кластерный и дискри- минантный анализ.

*Кластерный анализ* – это совокупность

Распознавание образов



Нейросети

Хранилища данных

Эффективные вычисления

**Data Mining**

Статистика Теория баз данных

Экспертные системы Визуализация данных

Информационный поиск

Оперативная аналитическая обработка

Рис. 1. Data Mining — мультидисциплинарная область

многомерных статистических процедур, кото- рая позволяет упорядочить объекты по одно- родным группам.

Решение задачи кластерного анализа выпол- няется в следующей последовательности:

1. Формируются выборки для анализа.
2. Выбирается совокупность признаков, характеризующих объект.
3. Выбираются меры сходства (расстояния) между объектами и производится их расчет.
4. Формируются кластеры.

при изоморфном – в однородные группы будут включаться те компании, в которых структуры распределения прибыли сходны.

В обоих способах признаки сначала преобра- зуют таким образом, чтобы не было единиц из- мерения и размах шкалы был одинаковым.

Для нормирования шкал необходимо выпол- нить следующие преобразования. При изотони- ческом разбиении сначала каждое значение при- знака заменяется на вычисленное по формуле:

*n*

1. Проводится анализ полученной инфор- мации.

*Vij*  *xij*

 *xij* ,

*i*1

Большинство алгоритмов кластерного ана- лиза относят к так называемым агломератив- ным процедурам, которые сначала объединяют в группы самые близкие объекты, а затем к ним

где *xij –* значение *j*-го признака для *i*-го объекта.

После этого каждому объекту ставится в соответствие одно число, вычисленное по фор- муле:

*m*

присоединяют более дальние.

Рассмотрим процедуру кластерного ана- лиза, предлагаемую для обработки данных. Существуют две разновидности, которые мо- гут давать различное разбиение на кластеры. Выбирать подходящую разновидность следует исходя из постановки задачи. Если это невоз- можно, необходимо провести разбиение двумя

*i*  *Vij* .

*j* 1

Расстояния между двумя объектами опре-

деляют по формуле:

*dij*  *i*   *j* .

При изоморфном разбиении сначала вы- полняют нормирование шкал по формуле:

способами и попытаться определить, какой из них более соответствует фактически сущест-

*Zij*

 *Vij*

*Vij* .

*j* 1

*m*

вующим структурам данных. При изотониче- ском разбиении группы объектов состоят из

однородных по уровню значений, а при изо-

Расстояния между двумя объектами опре- деляют по формуле:

*m*

морфном в группы включаются объекты, близ- кие по структуре, т.е. те, в которых пропорции

*dik* 

*Zij*

*j* 1

 *Zik*

2 .

признаков мало отличаются. Это означает, что различные способы разбиения могут давать различное объединение по группам. Напри- мер, у нас есть данные, которые характеризу- ют распределение прибыли фирм на расши- рение производства, научные исследования, социальные выплаты и пр. Тогда при изото- ническом разбиении группы будут состоять из фирм, в которых уровни прибыли близки, а

В изоморфном преобразовании расстояние будет минимальным в том случае, когда век- торы коллинеарны, и максимальным, если они перпендикулярны.

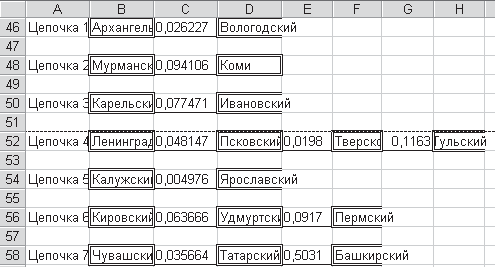
После определения расстояний возможно разбиение на группы с помощью метода ближай- шего соседа*.* Метод заключается в следующем: сначала для первого объекта находится ближай- ший и объединяется с ним в один кластер. Затем

для второго объекта ищется ближайший сосед и объединяется с ним в другой кластер и т.д. В случае, если ближайший сосед уже находится в одном из кластеров, происходит присоединение исследуемого объекта к нему. Таким образом образуются начальные кластеры.

На следующем этапе обычно выполняют построение дендритов и определение связно- сти в системе кластеров. Это позволяет объ- единить первоначальные кластеры в более сложные структуры, которые в большей степе- ни отвечают их реальной форме.

*Дискриминантный анализ* относится к мето- дам классификаций многомерных наблюдений при наличии обучающих выборок (в отличие от кластерного анализа, осуществляющего класси- фикацию автоматически – без обучения). Его цель состоит в идентификации новых объектов и отнесении их к уже имеющимся группам или совокупностям.

Пусть имеется *п* наблюдений, характери- зующих набор из *к* признаков. Тогда каждое наблюдение представляет собой случайный

вектор *x =* (*x* , *x* , … *x* )Т*.* Задача дискримина-

без дорогостоящих специализированных про- грамм для обработки данных.

Далее в качестве примера проведем класси- фикационный анализ прогноза развития потре- бительской кооперации на 20xx год (таблица) с использованием программ для кластерного (Приложение 1) и дискриминантного анализа (Приложение 2).

*Пример кластерного анализа.* Первым эта- пом работы программы является вычисление расстояний между объектами (рис. 5).

На следующем этапе, с использованием найденных расстояний, происходит объеди- нение объектов в кластеры (рис. 2):

1 2 *k*

ции состоит в разбиении всего множества ре- ализаций рассматриваемой многомерной вели- чины на некоторое число групп (областей) *Ri*

(*i*=1, 2, ..., *l*) и последующем отнесении нового

наблюдения к одной из них, используя некое решающее правило. При этом информация об истинной принадлежности объекта считается недоступной или требует чрезмерных матери- альных и временных затрат.

В ходе процедуры автоматически вычи- сляются функции классификации, предназна- ченные для определения той группы, к кото- рой наиболее вероятно принадлежит новый объект. Количество функций классификации равно числу имеющихся групп. Наблюде- ние считается принадлежащим той группе (совокупности), для которой получено наи- высшее значение функции классификации или наивысшее значение апостериорной ве- роятности – вероятности, с которой новое наблюдение принадлежит к этому классу, вычисленное на основе априорной вероятно- сти и расстояния Махаланобиса. Расстоянием махаланобисского типа называют естествен- ную меру отдаленности двух объектов друг от друга, если наблюдения извлекаются из нор- мальных генеральных совокупностей с одной и той же матрицей ковариации [4].

Один из методов автоматизированного

классификационного анализа данных основан на использовании программной среды Excel [4]. Достоинством такого подхода является возможность получения искомого результата

Рис. 2. Начальные кластеры

При наличии большего числа показателей или более сложной структуры данных возмож- но образование большего числа кластеров. Путем расчета межкластерных расстояний возможно их дальнейшее объединение в более крупные структуры (рис. 3).

Таким образом, исходя из рис. 3. с учетом значений, приведенных на рис. 2, получаем следующие кластеры:

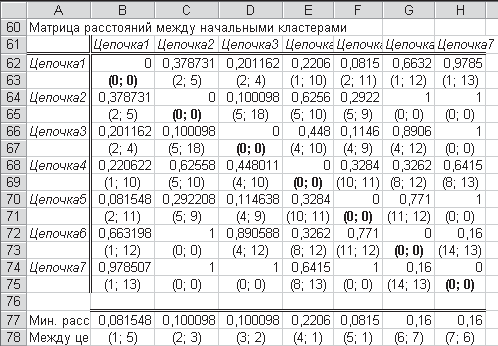
1. Архангельский, Вологодский, Мурман- ский, Карельский, Коми, Ярославский, Калуж- ский, Ивановский.
2. Ленинградский, Псковский, Тверской, Тульский.
3. Кировский, Чувашский, Пермский, Башкирский, Удмуртский, Татарский.

Программа позволяет провести анализ вли- яния показателей деятельности потребитель- ских союзов на состав получаемых кластеров (аналог факторного анализа).

*Пример дискриминантного анализа.* Рас- смотрим пример классификации Ивановско- го и Татарского потребсоюзов при наличии обучающей выборки из предыдущего разде- ла. Заполним рабочий лист Microsoft Excel согласно рис. 6. Столбец *Группа принадлеж- ности наблюдения* относит наблюдения к тому или иному кластеру (использовано де- ление на кластеры, полученное в кластерном анализе).

**Основные показатели прогноза развития потребительской кооперации на 20xx год**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Наимено- вание потреб- союзов** | **Весь розничный**  **товарооборот** | | **Объем бытовых услуг, млн руб.** | **Производство основных видов промышленной продукции** | | | | | | **Закупки основных видов сельскохозяйственной продукции** | | | | | |
| **млн руб.** | **в т. ч. оборот обще-**  **ственного питания, млн руб.** | **хлеб, тыс., т** | **колбасные изделия, т** | **кондитерские изделия, т** | **консервы, туб** | **безалкогольные напитки, тыс. дкл** | **пиво, тыс. дкл** | **Мясо, т** | **Молоко т** | **Яйца, млн шт.** | **Картофель, т** | **Овощи, т** | **Плоды, т** |
| 1 | Архангель-  ский | 1900 | 130 | 7 | 18,7 | 610 | 900 | 130 | 20 | 12 | 1800 | 1000 | 11,7 | 5000 | 1830 | 650 |
| 2 | Вологодский | 2550 | 140 | 5 | 22,8 | 600 | 1100 | 15 | 110 |  | 2400 | 4000 |  | 8000 | 1000 | 80 |
| 3 | Мурманский | 145 | 12 | 1 | 1,3 |  | 15 |  |  |  | 110 | 40 | 0,9 | 700 | 540 | 100 |
| 4 | Карельский | 630 | 30 | 2 | 3,5 | 250 | 200 |  | 8 |  | 440 | 500 | 40 | 1000 | 400 | 380 |
| 5 | Коми | 530 | 28 | 2 | 2,4 |  | 75 |  | 7 |  | 500 | 1500 | 2,5 | 5800 | 620 | 100 |
| 6 | Ленинград-  ский | 1900 | 150 | 1,5 | 4 | 1100 | 300 | 9500 | 10 |  | 5000 | 2000 | 12 | 5000 | 2000 | 1100 |
| 7 | Псковский | 3150 | 225 | 10 | 32 | 160 | 1600 |  | 240 |  | 2300 | 7000 | 23 | 8400 | 2300 | 1200 |
| 8 | Тверской | 1950 | 150 | 6 | 21 | 240 | 1200 | 1000 | 300 | 45 | 2000 | 8000 | 18 | 8000 | 4000 | 800 |
| 9 | Калужский | 860 | 68 | 7 | 12 | 410 | 600 | 170 | 40 |  | 2200 | 4000 | 5 | 7500 | 2400 | 520 |
| 10 | Тульский | 850 | 45 | 3 | 12,5 | 1000 | 400 | 7750 | 60 |  | 3750 | 3000 | 8 | 8000 | 4500 | 950 |
| 11 | Ярославский | 1600 | 95 | 6 | 8 |  | 800 |  | 8 |  | 2200 | 2000 | 11 | 3000 | 3000 | 1300 |
| 12 | Кировский | 3800 | 450 | 8 | 44 | 730 | 1730 | 700 | 400 | 33 | 3000 | 5000 | 22 | 8000 | 4000 | 1100 |
| 13 | Чувашский | 2550 | 310 | 10 | 23,5 | 2000 | 2700 | 1900 | 920 | 12 | 6300 | 20000 | 12,1 | 11000 | 5000 | 1200 |
| 14 | Пермский | 1850 | 115 | 5 | 25 | 140 | 900 | 3900 | 100 | 200 | 3000 | 5000 | 5,9 | 17000 | 8000 | 450 |
| 15 | Башкирский | 3000 | 300 | 20 | 33 | 1000 | 1600 | 1500 | 600 | 60 | 11000 | 25000 | 15 | 20000 | 5000 | 8000 |
| 16 | Удмуртский | 2150 | 240 | 15 | 22 | 1500 | 3000 | 3000 | 560 |  | 6200 | 13000 | 10,5 | 13500 | 3500 | 250 |
| 17 | Татарский | 3100 | 300 | 29 | 40 | 1500 | 1900 | 1900 | 510 |  | 6500 | 20000 | 15 | 19000 | 5000 | 1000 |
| 18 | Ивановский | 560 | 30 | 1,5 | 6,7 |  | 200 | 2000 | 30 |  | 900 | 1000 | 4,5 | 1000 | 1000 | 700 |

а Татарский – к третьей. Данный результат соответствует полученному при кластерном анализе, а также согласуется с выводами ра- боты, выполненной в программной среде Sta- tistica.

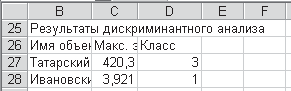


Рис. 3. Матрица расстояний между начальными кластерами

В результате вычислений с использовани- ем программы (рис. 4) получаем, что Иванов- ский потребсоюз относится к первой группе,

Рис. 4. Результат работы программы дискриминантного анализа

В заключение отметим, что настоящая схема классификации может быть использо- вана для анализа инвестиционной привлека- тельности предприятий потребительской ко- операции.

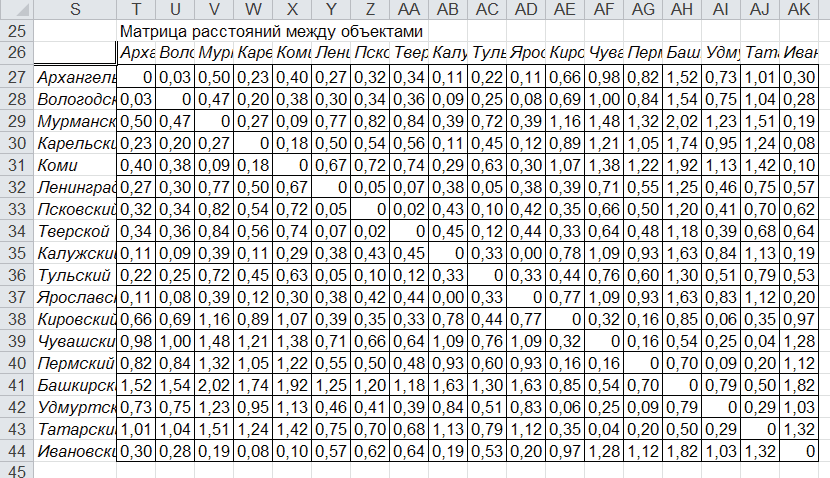


Рис. 5. Матрица расстояний между объектами

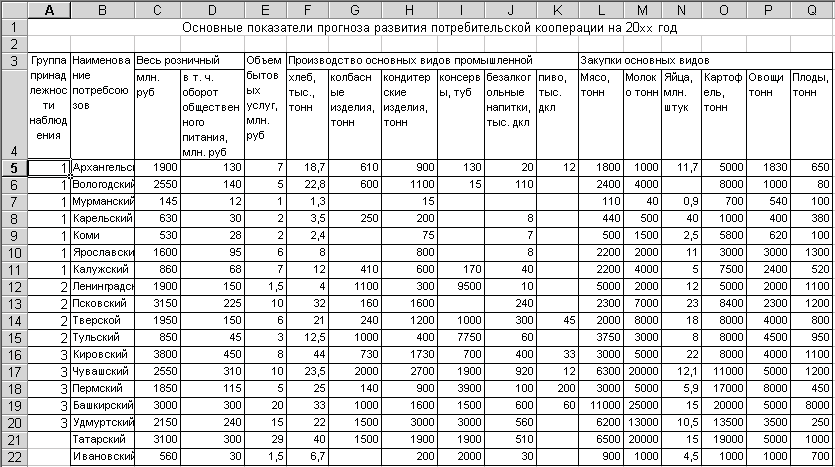


Рис. 6. Начальные данные для дискриминантного анализа

### Список литературы

1. *Возяков В.И., Филиппов В.П.* Инструмен- тальная поддержка управленческих решений в условиях экономического соперничества // Обо- зрение прикладной и промышленной математи- ки. 2008. Т. 15. № 3. С. 456-457.
2. *Возяков В.И., Филиппов В.П.* Модель рын- ка с прогнозируемыми ценами // Обозрение при- кладной и промышленной математики. 2007. Т. 14. № 4. С. 694-695.
3. *Пичужкин А.Б., Романов Ю.А., Филиппов В.П.* Об управлении бизнес-процессами торгово-

го предприятия // Вестник Российского универ- ситета кооперации. 2013. № 2 (12). С. 144-148.

1. *Филиппов В.П., Возяков В.И.* О поддержке принятия решения в технологии Data Mining // Обозрение прикладной и промышленной мате- матики. 2006. Т. 13, вып. 6. С. 1063-1064.
2. *Филиппов В.П., Шульдяшева Е.О., Ярчен- ков Н.А.* Классификационный анализ деятель- ности потребительских организаций за 2008- 2011 гг. // Актуальные вопросы науки и образо- вания: сб. науч. ст. Чебоксары: ЧКИ РУК, 2013. С. 325-330.

### Приложение 1

**Программа для кластерного анализа**

'Разбиение на кластеры методом изотонических рас- стояний

Option Base 1

Dim s\_1, s\_2(), s\_3(), RowCount, ColCount Dim mat\_res()

Dim arr\_Max\_Min() As Double

Dim arr\_min(), Max\_Min As Double

Dim arr\_clast\_dist() As Double 'матрица расстояний между кластерами

Dim arr\_num() As Integer Dim Ind\_arr() As Integer Dim myCell3, myCELL2

'Разбиение на кластеры методом изотонических рас- стояний

Sub clast\_izoton()

Set myCELL = Application.InputBox(Prompt:="", Title:="Выберите исходную матрицу данных", Default:=Cells(1, 255).Value, Type:=8)

Set myCell3 = Application.InputBox(Prompt:="", Title:="Выберите ячейки, содержащие имена объектов", Default:=Cells(2, 255).Value, Type:=8)

Set myCELL2 = Application.InputBox(Prompt:="", Title:="Выберите ячейку, с которой будут выводиться результаты", Default:=Cells(3, 255).Value, Type:=8) Cells(1, 255) = myCELL.Address

Cells(2, 255) = myCell3.Address Cells(3, 255) = myCELL2.Address

Num\_row = myCELL.Rows.count 'Вычисление количества строк

Num\_col = myCELL.Columns.count 'Вычисление количества столбцов

' Выполним нормировку исходных данных ReDim s\_1(Num\_col)

ReDim s\_2(1 To Num\_col, 1 To Num\_row) 'Вычисление суммы по столбцам и помещение ее в массив s\_1

For i = 1 To Num\_col

s\_1(i) = Application.Sum(myCELL.Columns(i))

For j = 1 To Num\_row

s\_2(i, j) = myCELL.Columns(i).Cells(j) / s\_1(i) Next j

Next i

ReDim s\_3(Num\_row) ' Расчет длин векторов

ReDim mas\_min(Num\_row) For i = 1 To Num\_row s\_3(i) = 0

mas\_min(i) = 1

For j = 1 To Num\_col s\_3(i) = s\_3(i) + s\_2(j, i) Next j

Next i

ReDim mat\_res(1 To Num\_row, 1 To Num\_row) ' Расчет матрицы расстояний

For i = 1 To Num\_row For j = 1 To Num\_row

mat\_res(i, j) = Abs(s\_3(i) - s\_3(j))

If mat\_res(i, j) < 1.1E-15 Then mat\_res(i, j) = 0 Next j

Next i

'Формирование и вывод изотонической матрицы рас- стояний

myCELL2.Offset(0, 1).Value = "Матрица расстояний между объектами"

For i = 1 To Num\_row

myCELL2.Offset(1, i).Value = myCell3.Cells(i) myCELL2.Offset(1, i).Font.Italic = True myCELL2.Offset(1, i).Borders.Item(xlEdgeBottom).Lin- eStyle = xlDouble

myCELL2.Offset(i + 1, 0).Value = myCell3.Cells(i) myCELL2.Offset(i + 1, 0).Font.Italic = True myCELL2.Offset(i + 1, 0).Borders(xlEdgeRight).Lin- eStyle = xlDouble

Next i

myCELL2.Offset(1, 0).Borders.Item(xlEdgeBottom).

LineStyle = xlDouble

myCELL2.Offset(1, 0).Borders(xlEdgeRight).LineStyle

= xlDouble

For i = 1 To Num\_row For j = 1 To Num\_row

myCELL2.Offset(i + 1, j).Value = mat\_res(i, j) Next j

Next i

'Разбиение на кластеры и поиск следов

Dim Sum\_dist As Double ' Суммарное расстояние Dim Num\_Links As Integer ' Количество связей ReDim arr\_num(1 To RowCount) As Integer ReDim Ind\_arr(1 To RowCount) As Integer Num\_of\_trek = 0 ' Количество кластеров

Num\_in\_treks = 0 ' Счетчик уже классифицирован- ных объектов

shift1 = 0

'Создание индексного массива для запоминания, включен ли

'рассматриваемый объект в какой - либо кластер или нет: 0-еще не включен;

'n-включен в n-й кластер.

Dim k\_A As Integer 'Номер объекта А Dim k\_B As Integer 'Номер объекта В

Dim Ind\_arr\_order() As Integer 'Массив порядка вклю- чения объектов в кластер

ReDim Ind\_arr\_order(RowCount) Dim Num\_obj\_in\_trek() As Integer For i = 1 To RowCount

Ind\_arr(i) = 0

Ind\_arr\_order(i) = 0

arr\_min(i) = 1 Next i

'Max\_Min = 0.088977

Max\_Min = InputBox("", "Введите значение критиче-

ского растояния", Max\_Min)

'Формирование вектора минимальных значений For i = 1 To RowCount

If Ind\_arr(i) = 0 Then k\_A = i

For j = 1 To RowCount If i <> j Then

'If arr\_min(i) > mat\_res(i, j) Then If Max\_Min > mat\_res(i, j) Then 'arr\_min(i) = mat\_res(i, j)

k\_B = j

Else: GoTo 111000 End If

Else: GoTo 111000 End If

'Next j 'End If

If ((Ind\_arr(k\_B) <> 0) And (Ind\_arr(k\_A) <> 0)) Then GoTo 111000

If (Ind\_arr(k\_B) <> 0) Or ((Ind\_arr(k\_B) = 0) And (Ind\_ arr(k\_A) <> 0)) Then 'Запись объекта в уже существу- ющий кластер

If ((Ind\_arr(k\_B) = 0) And (Ind\_arr(k\_A) <> 0)) Then Ind\_arr(k\_B) = Ind\_arr(k\_A)

Else

Ind\_arr(k\_A) = Ind\_arr(k\_B) End If

Num\_obj\_in\_trek(Ind\_arr(k\_B)) = Num\_obj\_in\_ trek(Ind\_arr(k\_B)) + 1

Ind\_arr\_order(k\_A) = Num\_obj\_in\_trek(Ind\_arr(k\_B)) Sum\_dist = Sum\_dist + arr\_min(i)

Else 'Создание нового кластера Num\_of\_trek = Num\_of\_trek + 1

ReDim Preserve Num\_obj\_in\_trek(Num\_of\_trek) Ind\_arr(k\_B) = Num\_of\_trek

Ind\_arr(k\_A) = Num\_of\_trek Num\_obj\_in\_trek(Num\_of\_trek) = 2

Ind\_arr\_order(k\_A) = 1

Ind\_arr\_order(k\_B) = 2

Sum\_dist = Sum\_dist + arr\_min(i) End If

111000

Next j End If Next i

Dim Tmp\_arr() As Integer Dim ind As Integer

Dim tp\_ind As Integer For i = 1 To Num\_of\_trek Num\_in\_treks = 0

myCELL2.Offset(RowCount + i \* 2 + 1, 0).Value = "Кла- стер" & i & " (" & Num\_obj\_in\_trek(i) & ")" & ":" ReDim Tmp\_arr(2, Num\_obj\_in\_trek(i))

For j = 1 To RowCount If Ind\_arr(j) = i Then

Num\_in\_treks = Num\_in\_treks + 1 Tmp\_arr(1, Num\_in\_treks) = j

Tmp\_arr(2, Num\_in\_treks) = Ind\_arr\_order(j) '(1,7)(1,6) End If

Next j

'Сортировка кластеров для вывода ind = 1

While ind = 1 ind = 0

For i1 = 1 To Num\_obj\_in\_trek(i) - 1

If Tmp\_arr(2, i1) > Tmp\_arr(2, i1 + 1) Then ind = 1

tp\_ind = Tmp\_arr(2, i1)

Tmp\_arr(2, i1) = Tmp\_arr(2, i1 + 1)

Tmp\_arr(2, i1 + 1) = tp\_ind tp\_ind = Tmp\_arr(1, i1)

Tmp\_arr(1, i1) = Tmp\_arr(1, i1 + 1) Tmp\_arr(1, i1 + 1) = tp\_ind

End If Next i1 Wend

'Выведение кластеров

For j = 1 To Num\_obj\_in\_trek(i) myCELL2.Offset(RowCount + i \* 2 + 1, j \* 2 - 1).Value

= myCell3(Tmp\_arr(1, j))

myCELL2.Offset(RowCount + i \* 2 + 1, j \* 2 - 1).Hori- zontalAlignment = xlCenter myCELL2.Offset(RowCount + i \* 2 + 1, j \* 2 - 1).Bor- ders.LineStyle = xlDouble

If j < Num\_obj\_in\_trek(i) Then myCELL2.Offset(RowCount + i \* 2 + 1, j \* 2).Value = mat\_res(Tmp\_arr(1, j), Tmp\_arr(1, j + 1)) myCELL2.Offset(RowCount + i \* 2 + 1, j \* 2).Borders(xlEdgeBottom).LineStyle = xlContinuous End If

Next j Next i

'Формирование матрицы расстояний между кластера- ми

Dim arr1() As Integer Dim arr2() As Integer

Dim Arr\_Links() As Integer

Dim arr\_trek\_dist() As Double 'матрица расстояний между кластерами

ReDim arr\_trek\_dist(Num\_of\_trek, Num\_of\_trek) ReDim arr1(RowCount)

ReDim arr2(RowCount)

ReDim Arr\_Links(2, Num\_of\_trek, Num\_of\_trek)

For i = 1 To Num\_of\_trek - 1

' Формирование массива номеров объектов і-го кла- стера

k\_A = 0

For j = 1 To RowCount If Ind\_arr(j) = i Then k\_A = k\_A + 1 arr1(k\_A) = j

End If Next j

For j = i + 1 To Num\_of\_trek

' Формирование массива номеров объектов j-го кла- стера

k\_B = 0

For k = 1 To RowCount

If Ind\_arr(k) = j Then k\_B = k\_B + 1 arr2(k\_B) = k

End If Next k

' Поиск наименьшего расстояния между объектами, находящимися в массивах arr1 и arr2

arr\_trek\_dist(i, j) = 1 For i1 = 1 To k\_A For j1 = 1 To k\_B

If arr\_trek\_dist(i, j) > mat\_res(arr1(i1), arr2(j1)) Then arr\_trek\_dist(i, j) = mat\_res(arr1(i1), arr2(j1)) Arr\_Links(1, i, j) = arr1(i1)

Arr\_Links(2, i, j) = arr2(j1) End If

Next j1 Next i1

arr\_trek\_dist(j, i) = arr\_trek\_dist(i, j) Arr\_Links(1, j, i) = Arr\_Links(1, i, j) Arr\_Links(2, j, i) = Arr\_Links(2, i, j) Next j

Next i

' Выведение матрицы расстояний между кластерами shift1 = Num\_of\_trek \* 2 + 2 + 1 myCELL2.Offset(RowCount + shift1, 0).Value = " Ма- трица межкластерных расстояний "

For i = 1 To Num\_of\_trek myCELL2.Offset(RowCount + shift1 + 1, i).Value = "Кластер" & i

myCELL2.Offset(RowCount + shift1 + 1, i).Font.Italic

= True

myCELL2.Offset(RowCount + shift1 + 1, i).Borders. Item(xlEdgeBottom).LineStyle = xlDouble myCELL2.Offset(RowCount + shift1 + 2 \* i, 0).Value = "Кластер" & i

myCELL2.Offset(RowCount + shift1 + 2 \* i, 0).Font.

Italic = True

myCELL2.Offset(RowCount + shift1 + 2 \* i, 0).Borders(xlEdgeRight).LineStyle = xlDouble myCELL2.Offset(RowCount + shift1 + 2 \* i + 1, 0).Borders(xlEdgeRight).LineStyle = xlDouble Next i

myCELL2.Offset(RowCount + shift1 + 1, 0).Borders. Item(xlEdgeBottom).LineStyle = xlDouble myCELL2.Offset(RowCount + shift1 + 1, 0).Borders(xlEdgeRight).LineStyle = xlDouble

For i = 1 To Num\_of\_trek For j = 1 To Num\_of\_trek

If (i = j) Then myCELL2.Offset(RowCount + shift1 + 2

\* i + 1, i).Font.Bold = True myCELL2.Offset(RowCount + shift1 + 2 \* i, j).Value = arr\_trek\_dist(i, j)

myCELL2.Offset(RowCount + shift1 + 2 \* i + 1, j).Value = "(" & Arr\_Links(1, i, j) & "; " & Arr\_Links(2, i, j) & ")"

myCELL2.Offset(RowCount + shift1 + 2 \* i + 1, j).HorizontalAlignment = xlCenter

Next j Next i

'Поиск минимальных расстояний между кластерами и их связей

ReDim arr\_min(Num\_of\_trek) Dim arr\_trek\_num() As Integer

ReDim arr\_trek\_num(Num\_of\_trek) For i = 1 To Num\_of\_trek arr\_min(i) = 1

For j = 1 To Num\_of\_trek If i <> j Then

If arr\_min(i) > arr\_trek\_dist(i, j) Then arr\_min(i) = arr\_trek\_dist(i, j) arr\_trek\_num(i) = j

End If End If Next j Next i

shift1 = shift1 + 2 \* Num\_of\_trek + 3 myCELL2.Offset(RowCount + shift1, 0).Value = "Мин. расст."

myCELL2.Offset(RowCount + shift1 + 1, 0).Value = "Между класт."

For i = 1 To Num\_of\_trek myCELL2.Offset(RowCount + shift1, i).Borders(xlEdgeTop).LineStyle = xlDouble myCELL2.Offset(RowCount + shift1, i).Value = arr\_ min(i)

myCELL2.Offset(RowCount + shift1 + 1, i).Value = "(" & i & "; " & arr\_trek\_num(i) & ")" myCELL2.Offset(RowCount + shift1 + 1, i).HorizontalAlignment = xlCenter

Next i End Sub

### Приложение 2

Sub diskrim\_n\_m()

### Программа для дискриминантного анализа

Dim mas\_min() As Double

'Подпрограмма дискриминантного анализа для n со- вокупностей объектов, позволяющая сразу классифи- цировать целую матрицу новых объектов

Dim Mean\_n() As Double

Dim Mean\_n\_tmp() As Double Dim Cov\_mat\_n() As Double Dim Cov\_all() As Double

Dim Cof\_disk() As Double Dim mat\_res()

Dim myfactor\_n() As Double

Dim Obj\_count As Integer 'Счетчик совокупностей (классов) объектов

Dim Obj\_mas() As Object 'Массив ссылок на совокуп- ности (классы)объектов

Dim Row\_n() As Integer Dim Col\_n() As Integer Dim All\_Row As Double

Dim Count\_New\_Obj As Integer 'Количество новых

объектов, предназначенных для классификации

'Ввод ссылки на ячейку, в которой хранится количест- во 'совокупностей объектов

Set CELL\_count = Application.InputBox(Prompt:="", Title:="Выберите ячейку, в которой содержится ко- личество совокупностей объектов", Default:=Cells(1, 255).Value, Type:=8)

Cells(1, 255) = CELL\_count.Address Obj\_count = CELL\_count.Cells.Value ReDim Obj\_mas(Obj\_count)

ReDim Row\_n(Obj\_count) As Integer ReDim Col\_n(Obj\_count) As Integer 'Ввод выборок совокупностей объектов For i = 1 To Obj\_count

Set Obj\_mas(i) = Application.InputBox(Prompt:="", Title:="Выберите " & i & "-ю совокупность объектов", Default:=Cells(i + 1, 255).Value, Type:=8)

Cells(i + 1, 255) = Obj\_mas(i).Address Next i

'Ввод ссылки на вторую совокупность объектов

Set myObj = Application.InputBox(Prompt:="", Title:="Выберите группу объектов, которые предназ- начены для классификации", Default:=Cells(10, 255). Value, Type:=8)

Cells(10, 255) = myObj.Address Count\_New\_Obj = myObj.Rows.count

Set Name\_myObj = Application.InputBox(Prompt:="", Title:="Выберите названия объектов, которые пред- назначены для классификации", Default:=Cells(11, 255).Value, Type:=8)

Cells(11, 255) = Name\_myObj.Address

Set myCell3 = Application.InputBox(Prompt:="", Title:="Выберите ячейку, с которой будут выводиться результаты", Default:=Cells(12, 255).Value, Type:=8) Cells(12, 255) = myCell3.Address

All\_Row = 0

For i = 1 To Obj\_count

Row\_n(i) = Obj\_mas(i).Rows.count 'Вычисление коли- чества строк

All\_Row = All\_Row + Row\_n(i)

Col\_n(i) = Obj\_mas(i).Columns.count 'Вычисление ко- личества столбцов

Next i

'1. Вычисление векторов средних значений для каждо- го класса (совокупности) объектов

ReDim Mean\_n(Col\_n(1), Obj\_count) ReDim Mean\_n\_tmp(Col\_n(1), 1)

For j = 1 To Obj\_count For i = 1 To Col\_n(1)

Mean\_n(i, j) = Application.Average(Obj\_mas(j). Columns(i))

Next i Next j

'2. Вычисление оценок ковариационных матриц для каждого класса

ReDim Cov\_mat\_n(1 To Obj\_count, 1 To Col\_n(1), 1 To Col\_n(1))

For k = 1 To Obj\_count For i = 1 To Col\_n(1)

For j = i To Col\_n(1)

Cov\_mat\_n(k, i, j) = Application.Covar(Obj\_mas(k). Columns(i), Obj\_mas(k).Columns(j))

Next j Next i Next k

'3. Вычисление несмещенной оценки объединенной ковариационной матрицы

ReDim Cov\_all(1 To Col\_n(1), 1 To Col\_n(1)) For k = 1 To Obj\_count

For i = 1 To Col\_n(1) For j = i To Col\_n(1)

Cov\_all(i, j) = Cov\_all(i, j) + (1 / (All\_Row - Obj\_count

- 1))

\* (Row\_n(k) \* Cov\_mat\_n(k, i, j))

If (i <> j) Then Cov\_all(j, i) = Cov\_all(i, j) Next j

Next i Next k

'4. Вычисление матрицы, обратной объединенной ко- вариационной матрице

ReDim Cov\_inv(1 To Col\_n(1), 1 To Col\_n(1)) Cov\_inv() = Application.MInverse(Cov\_all()) Dim Cof\_tmp() As Double

ReDim Cof\_tmp(1 To Col\_n(1), 1 To 1)

'Вычисление векторов оценок коэффициентов дискри- минации А(i)

ReDim Cof\_disk(1 To Col\_n(1), 1 To Obj\_count) For i = 1 To Obj\_count

For j = 1 To Col\_n(1) Cof\_disk(j, i) = 0

For l = 1 To Col\_n(1)

Cof\_disk(j, i) = Cof\_disk(j, i) + Cov\_inv(j, l) \* Mean\_n(l, i)

Next l Next j Next i

'Вычисление дискриминантных констант лямда Dim lamda\_n() As Double

ReDim lamda\_n(1 To Obj\_count) Dim Xl\_tmp() As Double

ReDim Xl\_tmp(1, Col\_n(1)) Dim Res\_tmp() As Double ReDim Res\_tmp(1, Col\_n(1)) For i = 1 To Obj\_count

For j = 1 To Col\_n(1) Xl\_tmp(1, j) = Mean\_n(j, i)

Mean\_n\_tmp(j, 1) = Mean\_n(j, i) Next j

For j = 1 To Col\_n(1) Res\_tmp(1, j) = 0 For l = 1 To Col\_n(1)

Res\_tmp(1, j) = Res\_tmp(1, j) + Xl\_tmp(1, l) \* Cov\_inv(l, j)

Next l

lamda\_n(i) = lamda\_n(i) + Res\_tmp(1, j) \* Mean\_n\_ tmp(j, 1)

Next j

lamda\_n(i) = lamda\_n(i) / 2

Next i

'Определение, к какому классу относится каждый из новых объектов

Dim cof\_prinadl() As Double

ReDim cof\_prinadl(Obj\_count) As Double myCell3.Offset(0, 0).Value = "Результаты дискрими- нантного анализа"

'формирование шапки таблицы результатов myCell3.Offset(1, 0).Value = "Имя объекта" myCell3.Offset(1, 1).Value = "Макс. знач." myCell3.Offset(1, 2).Value = "Класс"

For m = 1 To Count\_New\_Obj For k = 1 To Obj\_count cof\_prinadl(k) = 0

For i = 1 To Col\_n(1)

cof\_prinadl(k) = cof\_prinadl(k) + myObj.Cells(m, i) \* Cof\_disk(i, k)

Next i

cof\_prinadl(k) = cof\_prinadl(k) - lamda\_n(k)

Next k k = 1

Max\_val = cof\_prinadl(1) For i = 2 To Obj\_count

If Max\_val < cof\_prinadl(i) Then Max\_val = cof\_ prinadl(i): k = i

Next i

'Вывод результатов дискриминантного анализа для одного объекта

myCell3.Offset(1 + m, 0).Value = Name\_myObj.Cells(m) myCell3.Offset(1 + m, 1).Value = Max\_val myCell3.Offset(1 + m, 2).Value = k

Next m End Sub

***ВОЗЯКОВ Владимир Иванович*** *– доктор физико-математических наук, профессор, зав. кафедрой математических и инструментальных методов экономики. Чебоксарский коопера- тивный институт (филиал) Российского университета кооперации. Россия. Чебоксары. E-mail:* [*vvozyakov@rucoop.ru*](mailto:vvozyakov@rucoop.ru)

***ФИЛИППОВ Владимир Петрович*** *– магистрант. Чебоксарский кооперативный инсти- тут (филиал) Российского университета кооперации. Россия. Чебоксары. E-mail: filippov\_v\_p@ rambler.ru*

***VOZYAKOV, Vladimir Ivanovich*** *– Doctor of Physical and Mathematical Sciences, Professor, Department Chair of Mathematical and Tool Methods of Economics. Cheboksary Cooperative Institute (branch) of Russian University of Cooperation. Russia. Cheboksary. E-mail:* [*vvozyakov@rucoop.ru*](mailto:vvozyakov@rucoop.ru)

***FILIPPOV, Vladimir Petrovich*** *– Undergraduate. Cheboksary Cooperative Institute (branch) of Russian University of Cooperation. Russia. Cheboksary. E-mail:* [*filippov\_v\_p@rambler.ru*](mailto:filippov_v_p@rambler.ru)

# ПОСТРОЕНИЕ ТРЕХУРОВНЕВОЙ СИСТЕМЫ

**УДК 334.7**

# КРЕДИТНОЙ КООПЕРАЦИИ В РОССИЙСКОЙ ФЕДЕРАЦИИ ПУТЕМ УЧРЕЖДЕНИЯ КРЕДИТНОЙ ОРГАНИЗАЦИИ

## А.И. Воробьев

Посвящено исследованию трехуровневой системы кредитной кооперации в Российской Федерации. Проанализировано текущее состояние развития и предложены направления совер- шенствования. Разработана модель системы кредитной кооперации с небанковской кредитной организацией на верхнем уровне.

***Ключевые слова:*** кредитный потребительский кооператив; система кредитной кооперации; кооперативный банк; кредитный кооператив второго уровня.

**A.I. Vorobyev.** THREE-LEVEL SYSTEM CONSTRUCTION OF CREDIT COOPERATIVES IN THE RUSSIAN FEDERATION THROUGH THE ESTABLISHMENT OF A CREDIT ORGANIZATION

The article investigates the three-level system of credit cooperatives in the Russian Federation. The analyze the current state of development is made and suggested ways of its improving. The model system of credit cooperatives with non-bank credit organization at the top level has been worked out.

***Keywords:*** credit consumer cooperative; system of credit cooperatives; Cooperative Bank; credit cooperative second level.