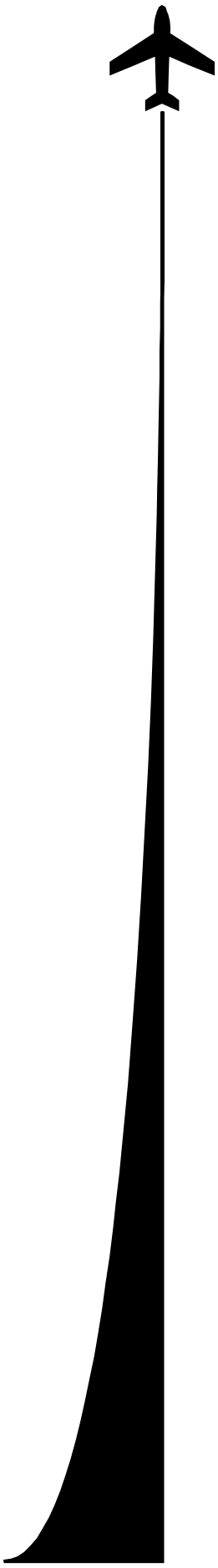

Session 2

Statistical Applications

ПРИМЕНЕНИЕ ПАКЕТНЫХ СРЕДСТВ В ЗАДАЧАХ ОЦЕНКИ ПАРАМЕТРОВ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Виктор Люмкис

Институт транспорта и связи
Ломоносова 1, Рига, LV-1019, Латвия
Тел.: (+371)-7100598. Факс: (+371)-7100660. E-mail: vlyumkis@yahoo.com

Задача исследования регрессионных моделей, в которых зависимость отклика от коэффициентов при предикторных переменных оказывается нелинейной, остается весьма актуальной. В этой связи применение современных пакетных средств таких, как Matchcad 7.0 [2], пакет Statistica [3] могут существенно облегчить проблему оценки параметров нелинейных моделей. Пусть рассматривается регрессионная модель, в которой зависимость отклика от коэффициентов при предикторных переменных оказывается нелинейной вида

$$\mathbf{Y}= f(\mathbf{X},\boldsymbol{\theta}) + \boldsymbol{\varepsilon}\,, \qquad (1)$$

где отклик $\mathbf{Y}$ связан с вектором предикторных переменных $\mathbf{X}$ некоторой нелинейной функцией $f(\mathbf{X},\boldsymbol{\theta})$, а относительно ошибок $\boldsymbol{\varepsilon}$ принимаются обычные предположения о нормальном распределении вида $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$.

Известный метод наименьших квадратов (МНК) требует минимизации суммы квадратов ошибок для нелинейных моделей вида

$$S(\boldsymbol{\theta})=\sum_{i=1}^{n}[\,\mathbf{Y}_i - f(\mathbf{X}_i,\boldsymbol{\theta})]^2, \qquad (2)$$

где суммирование осуществляется по совокупности n наблюдений.

Чтобы найти МНК – оценку $\hat{\theta}$, необходимо минимизировать функцию $S(\boldsymbol{\theta})$, например, продифференцировав (1), получить известную систему p нормальных уравнений [1] (p – размерность вектора $\boldsymbol{\theta}$). Весьма удобным вычислительным средством решения такого рода системы нормальных уравнений является метод линеаризации (или метод разложения в ряд Тейлора). Его суть состоит в многократном применении МНК к функции [1], которая предварительно разложена в ряд Тейлора в окрестности некоторой "разумной" точки $\boldsymbol{\theta}_0=(\theta_{10}, \theta_{20},\ldots, \theta_{P0})$, т. е. к функции вида

$$f(x_i,\boldsymbol{\theta})= f(x_i,\boldsymbol{\theta}_0) + \sum_{j=1}^{p}[\frac{\partial f(x_i,\theta)}{\partial \theta_j}]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}(\theta_i-\theta_{i0}) \qquad (3)$$

Применимость этого метода с последующей численной реализацией в пакете Matchcad оценок МНК реализована для конкретной модели вида

$$Y=\theta_1(1-\exp(-\theta_2*t)) +\varepsilon, \quad \text{где} \qquad (4)$$

выбраны следующие данные t=(1,2,3,4,5,6,7); Y=(82,112,153,163,176,192,200).

Раскладывая правую часть (4) (без учета ε) в ряд Тейлора в некоторой точке $\boldsymbol{\theta}_0=(\theta_{10}, \theta_{20})$ (ее выбор оговаривается всегда особо), получим

$$y_i = \theta_{10}(1-e^{=\theta_{20}t_i}) + (1-e^{-\theta_{20}t_i})(\theta_1-\theta_{10}) + \theta_{10}e^{-\theta_{20}t_i}t_i(\theta_2-\theta_{20}),\ \ i=1,2,...,n\,,$$

или

$$y_i - \theta_{10}(1-e^{=\theta_{20}t_i}) = z_{1i}\beta_1 + z_{2i}\beta_2\,, \quad \text{где}$$

$$\beta_1 = \theta_1 - \theta_{10}, \quad \beta_2 = \theta_2 - \theta_{20}, \quad z_{1i} = 1-e^{-\theta_2 t_i}, \; z_{2i} = \theta_{10}e^{-\theta_2 t_i}t_i, i = 1,2,...,n.$$

Используя для полученной линейной системы обычный МНК, получим решение в векторной форме

$$\boldsymbol{\beta}=(\mathbf{Z}^{\mathbf{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathbf{T}}(\mathbf{Y}-\boldsymbol{\varphi}), \text{ где} \qquad (5)$$

$\boldsymbol{\varphi}_j = \theta_{10}(1-e^{=\theta_{20}t_i})$, i=1,2,…n .

Осуществим далее итерационный процесс вида

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + \boldsymbol{\beta}_j \quad \text{или}$$

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + (\mathbf{Z}_j^T\mathbf{Z}_j)^{-1}\mathbf{Z}_j^T(\mathrm{Y}-\varphi_j). \qquad (6)$$

Процесс будет продолжаться до тех пор, пока для последовательных итераций $\boldsymbol{\theta}_j$, $\boldsymbol{\theta}_{j+1}$ не будет иметь место соотношение $|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j| < \varepsilon$.

В нашем случае компактная запись процесса итерирования в виде программы пакета Matchcad для конкретной системы (4) позволила получить устойчивое решение при θ_1=205.27, θ_2 =0.431. Интересно отметить, что использование раздела Nonlinear Statistics пакета Statistica [3] дало тот же результат.

Рассмотрим далее применимость пакета Matchcad для нахождения оценок параметров следующей регрессионной модели

$$\mathbf{Y}=\alpha-\beta*\rho^{\mathrm{X}}+\varepsilon,$$

где зависимость переменной **Y** от **X** задается соотношениями **X** = (0,1,2,3,4); **Y** = (44.4, 14.6, 63.8, 65.7, 68.9), при этом параметр $0 < \rho < 1$, а распределение **ε** нормальное, как и в модели (1).

Подход, который реализован в этом случае, базируется на весьма простом методе перебора. Замечая, что параметр ρ ограничен, осуществим перебор значений ρ, двигаясь с некоторым шагом h. Обозначая $\rho^{\mathrm{X}}_{\mathrm{i}}$, приходим к линейной зависимости

$$y_i = \alpha \quad + \quad (-\beta)z_i \quad +\varepsilon, \quad i=1,2,...,n.$$

Обычное решение МНК будет задаваться формулой вида

$$\boldsymbol{\theta}=(\mathbf{Z}^{\mathbf{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathbf{T}}\mathbf{Y},$$

где

$$\boldsymbol{\theta}=\begin{pmatrix}\alpha\\ -\beta\end{pmatrix}, \quad \mathbf{Z}=\begin{pmatrix}1 & z_1\\ 1 & z_2\\ \dots & \dots\\ 1 & z_n\end{pmatrix}.$$

Выбирая то ρ, при котором невязка $S=\sum_{j=1}^{n}(y_j-(\theta_0+\theta_1\rho^{x_j}))^2$ минимальна, получим для нашего случая α=72.5, β=28.3, ρ=0.6. Программа в пакете Matchcad, реализующая весь процесс вычислений весьма компактна и составляет порядка 15 операторов.

Важной задачей построения любой многомерной регрессии является выбор "наилучшей" модели. Смысл выбора "наилучшей" модели сводится к включению не слишком большого числа предикторов и при этом модель должна иметь сравнительно малую сумму предсказываемых расхождений.

Пакетные средства типа Matchcad могут послужить хорошим инструментом реализации подобных алгоритмов. Перед непосредственной реализацией алгоритма распишем его суть более подробно [1], (построение алгоритма касается многомерной линейной модели, которая содержит p параметров и имеется n измерений). В алгоритме выделяются следующие шаги:

1. Исключим первый опыт и построим все возможные регрессионные модели по данным оставшихся n-1 опытов.

2. По каждой модели подсчитываем предсказываемое значение отклика $\hat{Y}_{1p}$ в условия первого опыта и найдем разность ($\hat{Y}_{1p} - Y_1$).
3. Повторим шаги 1,2, исключая последовательно результаты второго опыта и получая ($\hat{Y}_{2p} - Y_2$), исключаем третий опыт, получаем ($\hat{Y}_{3p} - Y_3$) и т. д. до исключения последнего n-ого опыта.
4. Выбор наилучшей модели сводится к выбору такой модели, для которой сумма квадратов предсказываемых расхождений $\sum_{i=1}^{n} (\hat{Y}_{ip} - Y_{ip})^2$ мала.

Разработанная в пакете Matchcad программа была применена к данным Хальда [1]. Сама программа для конкретного случая выбора двух предикторных переменных при общем числе наблюдений n=13 представлена на рис. 1.

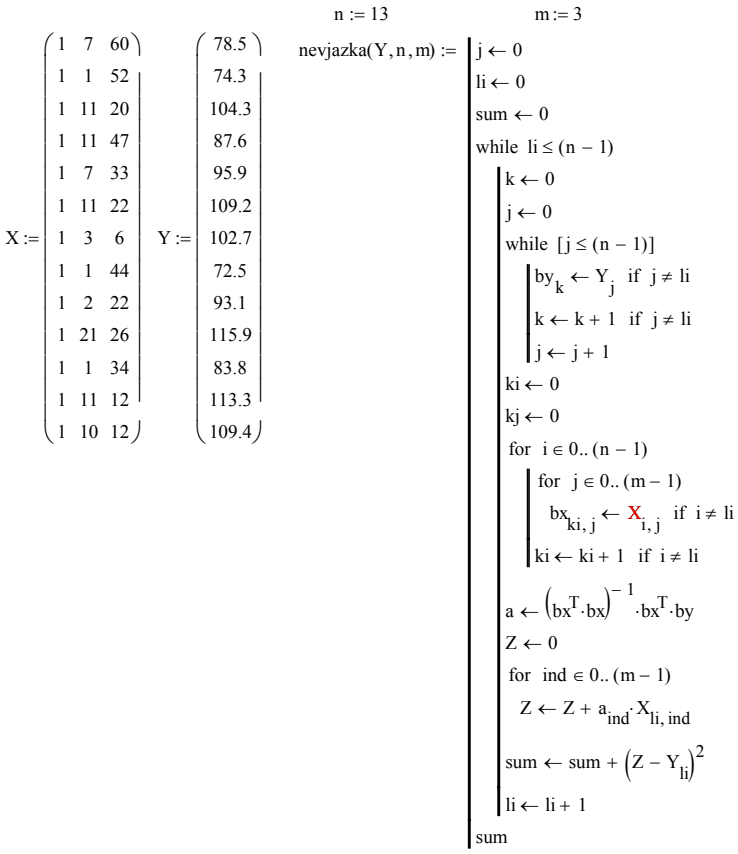

Рис.1. *Реализация алгоритма выбора минимальной невязки в пакете Matchcad*

Результат, который получается при использовании данной программы следующий Nevjazka(Y,n,m)=121.2. Прочие просчеты выбора предикторных переменных с расчетом соответствующих невязок дают возможность выбрать “наилучшую” регрессионную модель.

Литература

1. Дрейпер Н., Смит Г. Прикладной регрессионный анализ, Кн. 2, М.,Финансы и статистика, 1987, 351 с.
2. Кудрявцев Е.М. Matchcad 2000, М., ДМК, 2001, 571 с.
3. Боровиков В. Программа Statistica для студентов и инженеров, М., 2001, 300 с.

ПРОВЕРКА ОБОСНОВАННОСТИ КЛАСТЕРНОГО РЕШЕНИЯ

Лада Гусарова, Ирина Яцкив

Институт транспорта и связи
ул. Ломоносова 1, LV-1019, Рига, Латвия
Тел.: (+371)-7100650. Факс: (+371)-7100660. E-mail: ivl@tsi.lv

Ключевые слова: кластерный анализ, кластерная структура, тесты проверки адекватности

Введение

Кластерный анализ, являясь одним методов моделирования, также как и другие его разновидности требует проверки адекватности (валидации) своего итогового решения, а именно той структуры (модели), которую он вносит в данные. Необходимость этого обуславливается также большим количеством методов кластерного анализа и возможностью получения различных результатов для одних и тех же данных. Следовательно, оценка результатов после завершения процедуры кластеризации является обязательным, но на практике не всегда выполняемым шагом.

1. Классификация критериев валидации кластерного решения

При анализе результатов кластеризации необходимо проанализировать свойства кластеров, наиболее важные из которых – плотность, дисперсия, размеры, форма и отделимость [1,2]. Однозначные количественные характеристики этих свойств в литературе отсутствуют. Можно лишь сказать, что в качественной кластерной структуре кластеры должны быть значимо отделены друг от друга и объекты внутри кластеров должны располагаться как можно ближе друг к другу.

В настоящее время существует три группы критериев, отражающих три подхода к исследованию валидации кластеров в зависимости от используемых данных:

- *Внешние критерии* - сравнивают результаты классификации исходных данных с результатами классификации данных, не участвующих в исходной классификации.
- *Внутренние критерии* - используют только исходные данные для оценки качества классификации.
- *Относительные критерии* - сравнивают несколько различных классификаций одного набора данных.

Тесты валидации также различаются для пересекающихся и непересекающихся классов.

В данной статье рассматриваются только внешние и внутренние критерии для проверки кластерного решения и часть из них применяется к реальным данным.

2. Внешние критерии

При использовании внешних критериев существует два следующих подхода:

- Кластерная структура $C=\{C_1, C_2, \ldots, C_m\}$ сравнивается со структурой, полученной по другим данным (например, экспертно $P=\{P_1, P_2, \ldots, P_s\}$).
- Матрица близости исходных данных A сравнивается с матрицей, построенной на основе структуры, полученной по сгенерированным данным Y.

В нашем обзоре выберем первый вариант применения. Будем рассматривать все возможные пары объектов (x_v, x_u) и модифицировать четыре переменных: SS, SD, DS, DD. Добавим 1 к:

- SS = a, если оба объекта принадлежат одному и тому же кластеру в C и P;
- SD = b, если оба объекта принадлежат одному кластеру в C и разным в P;
- DS = c, если оба объекта принадлежат разным кластерам в C и одному кластеру в P;
- DD = d, если оба объекта принадлежат различным кластерам в C и P.

Далее, на основе полученных коэффициентов ассоциативности, высчитываются следующие различные метрики [3].

- Коэффициент Рэнда определяется по формуле

$$R = \frac{a+d}{a+b+c+d}.$$

Его значение лежит в пределах от 0 до 1. Чем ближе к 1, тем более вероятна полученная кластерная структура.

- Коэффициент Жаккарда определяется по формуле

$$J = \frac{a}{a+b+c}.$$

Данный коэффициент аналогичен коэффициенту Рэнда только без учета случая, когда оба объекта принадлежат разным классам (DD = d).

- Коэффициент Фолка и Мэллоу определяется по формуле

$$FM = \sqrt{\frac{a}{a+b} \cdot \frac{a}{a+c}}.$$

Высокое значение данного критерия говорит о высокой вероятности полученного разбиения.

Для применения вышеперечисленных критериев в статистических тестах необходимо знать их плотность распределения. Для этого используются методы Монте-Карло.

3. Внутренние критерии

Внутренние тесты можно разделить по типу исследуемой классовой структуры:

- *полное отсутствие классовой структуры;*
- *наличие единого кластера;*
- *наличие нескольких кластеров;*
- *наличие иерархической кластеризации.*

3.1. Для проверки на полное отсутствие классовой структуры используются следующие методы, которые носят название правил остановки [2,5].

- Критерий Duda и Hart (1973). Этот критерий относится к локальным правилам остановки и рассчитывается по формуле

$$F_1 = \frac{w[2]}{w[1]},$$

где w[2] – сумма квадратов внутрикластерных расстояний в случае, когда данные распределены по двум кластерам;

w[1] – сумма квадратов внутрикластерных расстояний в случае одного кластера.

Основная гипотеза состоит в существовании единого кластера однородных данных. Она отвергается, если

$$F_1 < 1 - 2/(\pi p) - z_{1-\alpha}\sqrt{2(1-8/(\pi^2/p))/(np)},$$

где α – заданный уровень значимости;
p – число переменных;
n – число объектов.

- Критерий Beale (1969) рассчитывается по формуле

$$F_2 = \left(\frac{w[1]-w[2]}{w[2]}\right) \Big/ \left(\frac{n-1}{n-2}\cdot 2^{2/p} - 1\right).$$

Это также локальное правило остановки. Основная гипотеза состоит в существовании единого кластера однородных данных. Она отвергается, если

$$F_2 > F_{1-\alpha}(p, p(n-2)),$$

где $F_{1-\alpha}(p, p(n-2))$ – квантиль распределения Фишера уровня $(1-\alpha)$

- Критерий Calinski и Harabasz (1974) относится к глобальным правилам остановки и рассчитывается по формуле

$$F_3 = \frac{trace(B)/(k-1)}{trace(W)/(n-k)}, \qquad (1)$$

где **B** – матрица межкластерных расстояний,
W – матрица внутрикластерных расстояний.

Максимальное значение критерия указывает на наиболее вероятное число кластеров.

- Кофенетический коэффициент корреляции. Этот коэффициент применяется для иерархической кластеризации и использует кофенетическую матрицу, представляющую иерархическую диаграмму.

3.2. Для проверки валидности отдельного кластера используется понятие идеального кластера С. Существует несколько определений идеального кластера [2]:

- «Comprehansive type» или "ball" – шар:

$$\max_j \{d_{ij} \mid j \in C\} < \min_k \{d_{ik} \mid k \notin C\}. \qquad (2)$$

- L*-cluster (van Rijbergen, 1970):

$$\max\{d_{ij} \mid i, j \in C\} < \min_k \{d_{kl} \mid k \in C, l \notin C\}. \qquad (3)$$

где d_{ij} – расстояние между i-м и j-м объектами.

Для определения, является ли некоторая группа объектов кластером или сгущением применяются условия, рассматриваемые в [1]:

- *кластером* называется группа объектов, удовлетворяющая условию:

$$\bar{d}_i^2 = \frac{W_i}{n} < \frac{S}{n}, \tag{4}$$

где W_i – сумма квадратов расстояний объектов до центра i-го класса;

S – общее рассеивание;

n – число объектов.

- *сгущением* называется группа объектов, для которой выполняется условие:

$$d_{i\max}^2 = \max_{X_j \in G_i} \left(d^2{}_{ij} \right) < \bar{d}^2 = \frac{S}{n}. \tag{5}$$

где X_j – j-й объект выборки;

G_i – i-я группа объектов;

d_{ij} – расстояние между i-м и j-м объектами;

$\bar{d}$ – среднее расстояние до центра класса.

На основе введенных критериев можно построить следующую методику проверки валидности отдельного кластера:

1 шаг - проверяется, удовлетворяет ли кластер свойствам идеального кластера;

2 шаг - проверяется гипотеза о его единой структуре с помощью одного из статистических критериев. Для этого может быть использована процедура Монте-Карло для оценки функции плотности выбранного критерия.

В качестве примера рассмотрим подход, описанный Гордоном в 1994 г.[2]

1) Постулируется основная гипотеза H_0, заключающаяся в отсутствии кластерной структуры.
2) Генерируется набор данных, того же объема n, что и оригинальный, в предположении, что кластерная структура отсутствует.
3) Классифицируем сгенерированный набор тем же методом, что и оригинальный набор данных.
4) Обозначаем результирующие кластерные структуры для оригинального набора P и для сгенерированного S.
5) Если в структурах P и S более 1 класса имеют одинаковое число объектов, оставляется один случайно выбранный кластер C.
6) Для кластера C, содержащего r объектов, рассчитывается статистика Манна-Уитни $U(r)$ по формуле:

$$U_{ijkl} = \begin{cases} 0, & \text{если } d_{ij} < d_{kl} \\ 1/2, & \text{если } d_{ij} = d_{kl} \\ 1, & \text{если } d_{ij} > d_{kl} \end{cases}, \tag{6}$$

где d_{ij} – расстояние между i и j объектами.

Можно рассчитать глобальный критерий:

$$U_G = \sum_{(i,j) \in W} \sum_{(k,l) \in B} U_{ijkl}$$

или локальный критерий:

$$U_L = \sum_{i \in C} \sum_{j \in C} \sum_{k \in C} U_{ijkl}. \tag{7}$$

где $W \equiv \left\{ (i,j) \middle| i,j \in C, i < j \right\}$,

$$B \equiv \{(k,l) | k \in C, l \notin C\}.$$

7) Шаги 2) – 6) повторяются до тех пор, пока не будет получено (m–1) значений статистики $U(r)$.
8) Из значений статистики $U(r)$ составляется вариационный ряд.
9) Если значение статистики $U(r)$ для P меньше, чем j-е значение статистики в вариационном ряду для сгенерированных S, то гипотеза H_0 отвергается при уровне значимости $100(j/m)\%$.

Недостатком данного метода является то, что для кластеров малого объема j-е значение статистики $U(r)$ может оказаться равным 0.

Рассмотрим применение предложенной методики на практическом примере.

4. Пример проверки валидности кластерного решения

Рассмотрим классификацию 31 региона России по уровню развития малого бизнеса. В качестве признаков классификации использовались показатели развития региона: доля экономически активного населения, количество студентов Вузов, инвестиции, расходы бюджета на производство, капитальные вложения, количество банков в регионе. В пакете Statistica/Win методом Уорда (Евклидово расстояние в качестве метрики) была получена дендрограмма, представленная на рис. 1.

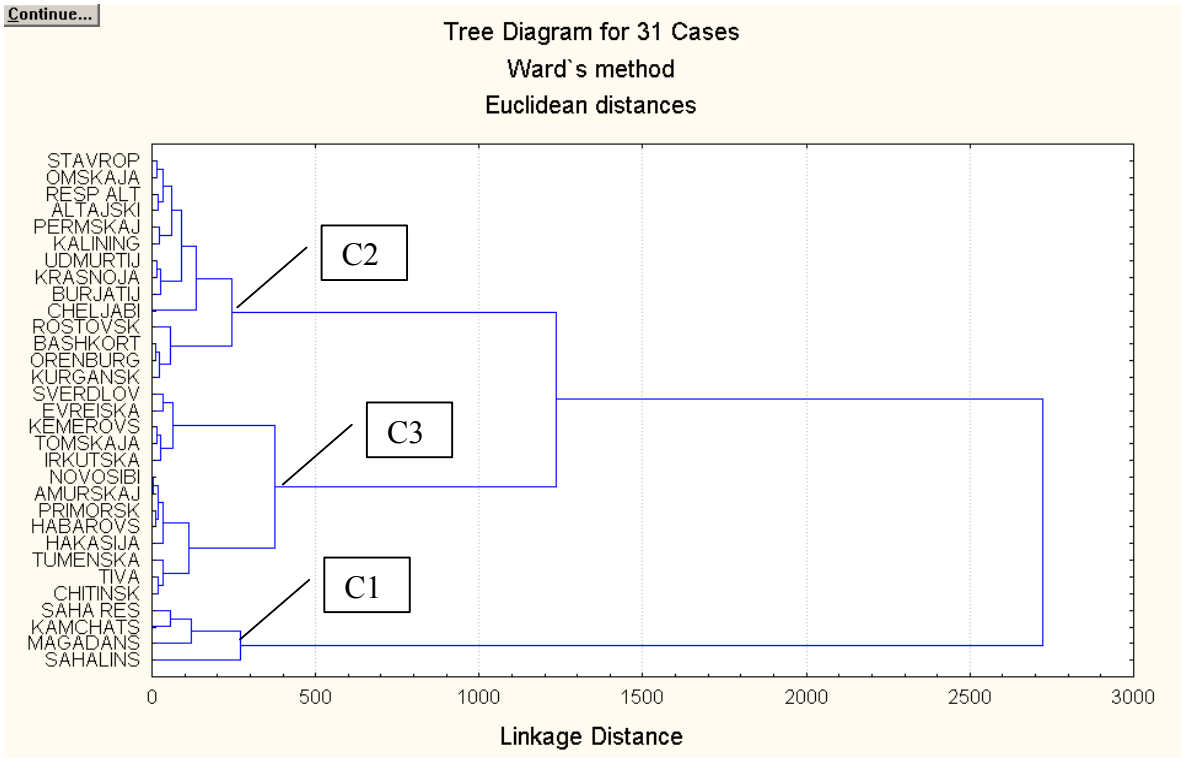

Рис. 1. *Дендрограмма кластеризации, полученная с помощью пакета Statistica/Win*

На дендрограмме четко выделяются 3 класса: C1, C2, C3. Проверим валидность разбиения структуры на данные три класса.

По глобальному правилу (1) было определено оптимальное количество кластеров (Рис. 2). Максимальное значение критерия на графика приходится на количество кластеров равное 3 (число кластеров более 15 не рассматривается).

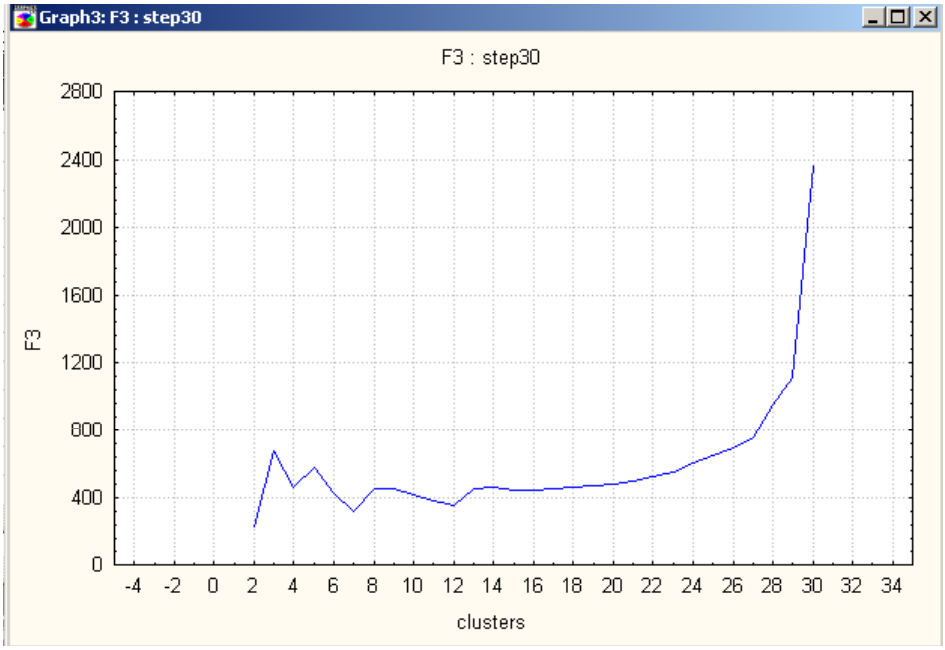

Рис. 2. *Определение оптимального количества кластеров*

При проверки кластеров на соответствие идеальному (рис. 3-5) оказалось, что ни один из кластеров не удовлетворяет данному условию. По результатам, приведенным в таблицах, описывающих кластеры C1, C2, C3 видно, что существуют объекты, для которых минимальное межклассовое расстояние меньше максимального внутриклассового расстояния (выделены в таблицах красным цветом). Очевидно, для выполнения этих условий компактность классов недостаточная.

Min-Max for C1 (new.sta)

CLUSTER ANALYSIS	Within max	Between min
Saha respub	203,9555	395,1519
Kamchatskaja	148,6436	340,0914
Magadanskaja	279,1883	470,2786
Sahalinskaja	279,1883	192,4387

Рис. 3*. Внутриклассовые и межклассовые расстояния для кластера C1*

Проверка по условиям (4, 5), показала, что все группы объектов можно назвать как кластерами, так и сгущениями (результаты на рис. 6). Каждая строка в таблице описывает один кластер C1, C2 и C3. В первой колонке приводится среднее внутриклассовое расстояние, во второй – максимальный квадрат расстояния между объектами, в третьей – средний разброс относительно общего центра. Условия для кластера и сгущения, выполняются по всем классам.

Расчет U_L статистики (рис. 7) для каждого кластера показал, что наиболее близким к идеальному кластеру является 1-й кластер C1, так как значение U_L статистики для этого класса минимально.

Max-min for C2 (new.sta)

CLUSTER ANALYSIS	Within max	Between min
Kaliningradskaja	115,0521	63,1997
Krasnojarskij	121,5246	34,6839
Omskaja	96,4184	51,0173
Stavrop kraj	83,7453	65,2464
Rostovskaja	83,8011	73,2586
Bashkortostan	121,5246	137,0662
Udmurtija	98,5672	109,1272
Kurganskaja	108,9089	31,8978
Saha respub	94,3042	109,1486
Orenburgskaja	95,6530	102,1151
Permskaja	95,4421	59,0277
Cheljabinskaja	115,0521	107,3403
Resp Altaj	96,5493	46,0918
Altajskij kraj	89,5746	48,2748

Рис. 4. *Внутриклассовые и межклассовые расстояния для кластера C2*

Max-min for C3 (new.sta)

CLUSTER ANALYSIS	Within max	Between min
Sverdlovskaja	112,5826	68,6645
Kemerovskaja	146,8505	26,4312
Novosibirskaja	96,5800	117,1345
Tomskaja	149,0001	26,1712
Tumenskaja	149,0001	167,1753
Tiva	128,0145	146,0823
Hakasija	77,5473	99,2660
Irkutskaja	128,0733	43,9725
Chitinskaja	129,8715	148,3485
Evreiskaja	103,7870	67,2946
Primorskij kr	97,1311	116,7696
Habarovskaja	103,6268	122,3687
Amurskaja	94,5104	114,8228

Рис. 5. *Внутриклассовые и межклассовые расстояния для кластера C3*

Rezult (Aivazan) (statistiva4_bak.sta)

CLUSTER ANALYSIS	Wi / ni	max di	S / n
Cluster 1	4237,562	14580,56	35277,37
Cluster 2	1698,599	6443,50	35277,37
Cluster 3	2336,543	5671,25	35277,37

Рис. 6. *Результаты проверки условий кластера и сгущения*


U (statistiva4_bak.sta)

CLAN	COL1
1	9,0000
2	140,0000
3	266,0000

Рис. 7. *Значения U_L статистики для каждого кластера*

Выводы

Большинство универсальных пакетов статистического анализа, предлагая широкий спектр методов кластерного анализа, не имеют в своем арсенале процедур поверки качества полученного решения. В работе сделан обзор критериев, которые являются наиболее подходящими для этой цели и могут быть реализованы в виде макросов или скриптов. Часть из критериев были реализованы в виде скриптов для пакета Statistica/Win.

Литература

1. Яцкив И. В. Многомерный статистический анализ: Классификация и снижение размерности. Рига, ИТС, 2003, 200p.
2. Gordon A.D. Classification, 2nd edn. Chapman & Hall, NewYork, 1999, 256 p.
3. Maria Halkidi, Yannis Batistakis, Michalis Vazirgiannis. Cluster Validity Methods: Part I. Department of Informatics, Athens University of Economics & Business. 2001.
4. Maria Halkidi, Yannis Batistakis, Michalis Vazirgiannis. Cluster Validity Methods: Part II. Department of Informatics, Athens University of Economics & Business. 2001.
5. Яцкив И., Гусарова Л. Методы определения количества кластеров при классификации без обучения. \\ Transport and TeleCommunication, 2002, Volume 4, No.1. Rīga, TSI, 2003. p.23-28.