

Алгоритм кластеризация средств траекторных измерений в задачах программно-целевого планирования развития средств обеспечения испытаний

Ведущий научный сотрудник 46 ЦНИИ МО РФ доктор технических наук старший научный сотрудник В.Г. Найденев

Разработан новый алгоритм кластеризации средств траекторных измерений, который позволяет устранить дублирование таких средств при решении задач программно-целевого планирования развития средств обеспечения испытаний. Проведена апробация этого алгоритма на примере обработки данных заявок на разработку ряда радиотехнических и оптико-электронных средств траекторных измерений.

Общие положения

Решение задач программно-целевого планирования развития средств траекторных измерений, как правило, начинается с обработки большого количества заявок на разработку и модернизацию средств такого класса, которые предлагаются научно-исследовательскими организациями, а также предприятиями промышленности – разработчиками средств обеспечения испытаний.

Как показал опыт разработки предложений по развитию средств обеспечения испытаний в соответствующие программные документы, количество заявок на разработку новых и модернизацию существующих средств траекторных измерений может достигать нескольких десятков. При этом имеют место случаи дублирования предложений по разработке таких средств траекторных измерений. Это обусловлено наличием большого количества предприятий, традиционно занимающихся разработкой новых средств траекторных измерений, а также нескольких заказчиков средств обеспечения испытаний, на которые замыкаются испытательные полигоны.

В связи с этим при программно-целевом планировании развития средств обеспечения испытаний первоочередной задачей является первичная обработка заявок на разработку новых средств траекторных измерений и модернизацию существующих.

До настоящего времени в теории и практике программно-целевого планирования развития средств траекторных измерений отсутствовали практически работоспособные алгоритмы автоматизированной обработки тематических карточек на разработку новых средств траекторных измерений с целью формирования отдельных классов средств, которые по области применения и своим тактико-техническим

характеристикам не дублировали бы друг друга и отличались от уже существующих на полигонах средств ТИ.

Для решения такой задачи был разработан и апробирован специальный алгоритм кластеризации средств траекторных измерений, который позволяет практически в автоматическом режиме формировать группы средств траекторных измерений, позволяющие в максимальной степени выполнять требования к зоне и точности измерений параметров, предусмотренных ТТХ, без дублирования разработок таких средств обеспечения испытаний.

Рассматриваемый алгоритм кластеризации средств траекторных измерений основан на теории однородности математически описываемых объектов.

В общем случае, понятие однородности объектов определяется заданием правила вычисления величины, характеризующей либо расстояние между двумя объектами из исследуемой совокупности, либо степень близости (сходства) тех же объектов.

Если обозначить через $\Xi_s (s = \overline{1, n})$ объекты (средства траекторных измерений), каждый из которых имеет $m (m = \overline{1, M})$ признаков (характеристик), то их совокупность $\bar{\Xi}^s$ может быть представлена в виде следующей матрицы:

$$\bar{\Xi}^s = \begin{bmatrix} \Xi_1 \\ \Xi_2 \\ \vdots \\ \Xi_s \\ \vdots \\ \Xi_n \end{bmatrix} = \begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1m} & \cdots & \xi_{1M} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2m} & \cdots & \xi_{2M} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi_{s1} & \xi_{s2} & \cdots & \xi_{sm} & \cdots & \xi_{sM} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi_{n1} & \xi_{n2} & \cdots & \xi_{nm} & \cdots & \xi_{nM} \end{bmatrix}. \quad (1)$$

Математические преобразования будем проводить в предположениях, что случайные величины $\xi_{sm} (m = 1, 2, \dots, M)$, образующие вектор-строку Ξ_s , являются независимыми и имеют одинаковую дисперсию.

Реализации многомерной случайной величины $\bar{\Xi}^s$ обозначим через матрицу \bar{X}^s , которая имеет следующую структуру:

$$\bar{X}^s = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_s \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2m} & \cdots & x_{2M} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{s1} & x_{s2} & \cdots & x_{sm} & \cdots & x_{sM} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \cdots & x_{nm} & \cdots & x_{nM} \end{bmatrix}.$$

Если одномерные случайные величины ξ_{sm} предполагаются нормально распределенными, то распределение случайной величины $\bar{\Xi}^s$ в M -мерном пространстве будет рассматриваться как многомерное нормальное. Тогда статистической моделью совокупности m -мерных наблюдений будет фиксированное множество m -мерных случайных величин $\bar{\Xi}^s$, распределенных нормально с плотностью:

$$f(X, M_s, K_s) = (2\pi)^{-\frac{M}{2}} \cdot (K)^{-\frac{M}{2}} \cdot \exp\left\{-\frac{1}{2}(X - M_s)^T \cdot K_s^{-1} \cdot (X - M_s)\right\}, \quad (2)$$

где M_s – вектор-строка математических ожиданий случайных величин ξ_{sm} ;

K_s – ковариационная матрица размерности $(M \times M)$.

Предположим, что множество $\bar{\Xi}^s$ можно представить в виде двух подмножеств $\bar{\Xi}^{R_1}$ и $\bar{\Xi}^{R_2}$. Так, множество $\bar{\Xi}^{R_1}$ образуется из множества $\bar{\Xi}^s$ путем отсечения n_1 строк, а множество $\bar{\Xi}^{R_2}$ образовано из оставшихся n_2 строк множества $\bar{\Xi}^s$. Таким образом, справедливо соотношение

$$\bar{\Xi}^{R_1} \cup \bar{\Xi}^{R_2} = \bar{\Xi}^s.$$

В этом случае критерием однородности множества $\bar{\Xi}^s$ для каждого элемента разбиения $\rho^2 \in P^2$ будет случайная величина [1, 2]:

$$\hat{\Xi}(\rho^2) = \frac{1}{n_1} \sum_{s \in R_1} \bar{\Xi}_s - \frac{1}{n_2} \sum_{s \in R_2} \bar{\Xi}_s. \quad (3)$$

Одномерные случайные величины, образующие M -мерный вектор $\hat{\Xi}(\rho^2)$, обозначим через $\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_m, \dots, \hat{\xi}_M$.

Каждая составляющая вектора $\hat{\Xi}(\rho^2)$ определяется следующим образом:

$$\hat{\xi}_m = \frac{1}{n_1} \sum_{s \in R_1} \hat{\xi}_{sm} - \frac{1}{n_2} \sum_{s \in R_2} \hat{\xi}_{sm}. \quad (4)$$

Случайные величины ξ_m имеют выборочные значения, определяемые по формуле:

$$\hat{x}_m = \frac{1}{n_1} \sum_{s \in R_1} \hat{x}_{s_m} - \frac{1}{n_2} \sum_{s \in R_2} \hat{x}_{s_m} \quad (5)$$

Предполагая, что рассматриваемые случайные величины ξ_{s_m} распределены нормально с дисперсиями σ_m^2 , тогда случайные величины $\hat{\xi}_{m_1} = \frac{1}{n_1} \sum_{s \in R_1} \hat{\xi}_{s_m}$ и $\hat{\xi}_{m_2} = \frac{1}{n_2} \sum_{s \in R_2} \hat{\xi}_{s_m}$ будут иметь параметры нормального распределения – математическое ожидание соответственно M_{m_1} и M_{m_2} и дисперсию соответственно σ^2/n_1 и σ^2/n_2 .

В этом случае случайные величины $\hat{\xi}_m$, определяемые выражением (4), также будут распределены по нормальному закону, но со следующими параметрами:

$$\begin{aligned} M(\hat{\xi}_m) &= M_{m_1} - M_{m_2}; \\ D(\hat{\xi}_m) &= \sigma_m^2 (n_1 + n_2)(n_1 \cdot n_2)^{-1}. \end{aligned} \quad (6)$$

Исходя из проведенных рассуждений, и с учетом выражения (6) в качестве критерия для проверки гипотезы $H_0 : \hat{\Xi}(\rho^2) = \{0, 0, \dots, 0\}$ для всех $\rho^2 \in P^2$ об однородности двух выборок из множества Ξ^s можно принять критерий Пирсона (χ^2) [3] с M степенями свободы вида:

$$U(\rho^2) = \frac{(\hat{x}_{m_1} - \hat{x}_{m_2})^2 \cdot n_1 \cdot n_2}{\sigma_{B_m}^2 \cdot (n_1 + n_2)}, \quad (7)$$

где $\sigma_{B_m}^2$ выборочное значение дисперсии σ_m^2 , определяемое по формуле

$$\sigma_{B_m}^2 = \frac{1}{n_1 + n_2 - 1} \cdot \left[\sum_{s \in R_1} x_{s_m}^2 + \sum_{s \in R_2} x_{s_m}^2 - \frac{1}{n_1 + n_2} \cdot \left(\sum_{s \in R_1} x_{s_m} + \sum_{s \in R_2} x_{s_m} \right)^2 \right]. \quad (8)$$

Оценки \hat{x}_{m_1} и \hat{x}_{m_2} в выражении (7) определяются соответствующими суммами правой части равенства (5).

Если теперь в формулу (7) подставить значения $\sigma_{B_m}^2$, \hat{x}_{m_1} и \hat{x}_{m_2} , определяемые выражениями (6) и (8), то получим окончательную запись критерия однородности двух выборок в следующем виде:

$$U(\rho^2) = \frac{n_1 + n_2 - 1}{(n_1 + n_2) \cdot n_1 \cdot n_2} \cdot \sum_{m=1}^M \frac{\left(n_2 \sum_{s \in R_1} x_{s_m} - n_1 \sum_{s \in R_2} x_{s_m} \right)^2}{\sum_{s \in S} x_{s_m}^2 - \frac{1}{n_1 + n_2} \left(\sum_{s \in S} x_{s_m} \right)^2}. \quad (9)$$

Тогда гипотеза об однородности анализируемых выборок из множества $\bar{\Xi}^s$ может быть принята, если

$$U(\rho^2) \leq \chi_{\alpha, M}^2 \quad \text{для всех } \rho^2 \in P^2, \quad (10)$$

где α – уровень значимости, а M – число степеней свободы.

Если же неравенство (10) не выполняется, то гипотеза об однородности выборок отклоняется.

Необходимо отметить [3], что критерий (9) проверки однородности нескольких совокупностей (подмножеств) $\bar{\Xi}^{R_1}$ и $\bar{\Xi}^{R_2}$ множества $\bar{\Xi}^s$ не всегда является достоверным, однако этот недостаток можно устранить, если наложить следующие ограничения:

– должна иметься возможность ранжирования случайных величин $\Xi_s (s = \overline{1, n})$, входящих в многомерное множество $\bar{\Xi}^s$, в порядке возрастания величины одного из существенных для анализа признаков;

– не допускается менять заданное ранжированием расположение элементов совокупности по $m (m = \overline{1, M})$ признакам.

С учетом отмеченных ограничений данные о совокупности характеристик средств траекторных измерений объемом n могут быть представлены в виде матрицы (1), в которой $s = 1, 2, \dots, n$ будет означать положение элемента статистической совокупности в ранжированном ряду. Обозначим подмножество значений $s = \overline{1, t}$ через R_t , а подмножество значений $s = \overline{t+1, n}$ через \bar{R}_t .

Тогда множество объектов средств ТИ $\bar{\Xi}^s$ разделится на два следующих подмножества:

$$\bar{\Xi}^s = \Xi^{R_t} \cup \Xi^{\bar{R}_t}.$$

В случае, если $t=1$, то первая строка матрицы (1) представляет собой множество Ξ^{R_t} , а оставшиеся $n-1$ строк этой матрицы представляют множество $\Xi^{\bar{R}_t}$.

С учетом отмеченных ограничений и при нормальном законе распределения случайных величин $\Xi_s (s = \overline{1, n})$ гипотеза об однородности статистической совокупности может быть представлена как совместная проверка $(n-1)$ гипотез. При этом элемент ρ^2 задан на множестве разбиений P^2 пространства $\bar{\Xi}^s$ на взаимно дополняющие множества Ξ^{R_t} и $\Xi^{\bar{R}_t}$. Поэтому $\hat{\Xi}(\rho^2)$ определяется здесь в отличие от выражения (3) по формуле, учитывающей операцию дополнения:

$$\hat{\Xi}(\rho^2) = \frac{1}{t} \sum_{s \in R_t} \Xi_s - \frac{1}{n-t} \sum_{s \in \bar{R}_t} \Xi_s. \quad (11)$$

Тогда для проверки гипотезы об однородности последующих двух совокупностей средств ТИ может быть задана следующая функция:

$$U(\rho^2) = \frac{n-1}{n \cdot (n-t) \cdot t} \cdot \sum_{m=1}^M \frac{\left((n-t) \sum_{s=1}^t x_{s_m} - t \cdot \sum_{s=t+1}^n x_{s_m} \right)^2}{\sum_{s=1}^n x_{s_m}^2 - \frac{1}{n} \left(\sum_{s=1}^n x_{s_m} \right)^2}. \quad (12)$$

Функция (12) задана на множестве разбиений P^2 и представляет собой значения случайной величины, распределенной по закону χ^2 с M степенями свободы.

Результаты расчетов по формуле (12) признаются не противоречащими принятию гипотезы об однородности, если соблюдается неравенство:

$$U(\rho^2) \leq \chi_{\alpha, M}^2 \text{ для всех } \rho^2 \in P^2. \quad (13)$$

В случае, если хотя бы для одного $\rho^2 \in P^2$ имеет место, что $U(\rho^2) > \chi_{\alpha, M}^2$, то рассматриваемая совокупность не может быть признана однородной в силу отклонения нулевой гипотезы H_0 . В этом случае принимается конкурирующая гипотеза

за H_1 и возникает задача расчленения неоднородной совокупности средств ТИ на однородные группы.

Алгоритм автоматической кластеризации совокупности средств траекторных измерений

Алгоритм автоматической кластеризации средств траекторных измерений в задаче программно-целевого планирования развития таких средств основан на использовании предложенной гипотезы однородности многомерных исследуемых объектов. Сам алгоритм показан на рисунке 1.

Работа алгоритма начинается с ввода исходных данных об информационных признаках средств ТИ, т.е. матрицы \bar{X}^s , которая является реализацией случайной величины $\bar{\Xi}^s$.

В блоке 2 алгоритма проводятся ранжирование средств ТИ по наиболее существенному признаку (m) и формирование матрицы \bar{X}_p^s .

Дальнейшая часть алгоритма связана с решением итерационных задач проверки совокупности исследуемой генеральной совокупности средств ТИ на однородность и разбиением неоднородных совокупностей на пары групп.

Так, на первой итерации работы алгоритма проводятся вычисление значений величины критерия однородности $U(\rho^2)_s (s = \overline{1, k})$ согласно выражению (12). Эта операция выполняется с использованием блоков (4...8) алгоритма.

Вычисление значений $U(\rho^2)_s (s = \overline{1, k})$ проводится в блоке 9 и формируется вектор $\bar{U}(Jx1)$.

Далее в блоке 10 проводится проверка выполнения гипотезы H_0 о признании исследуемой генеральной совокупности средств ТИ однородной.

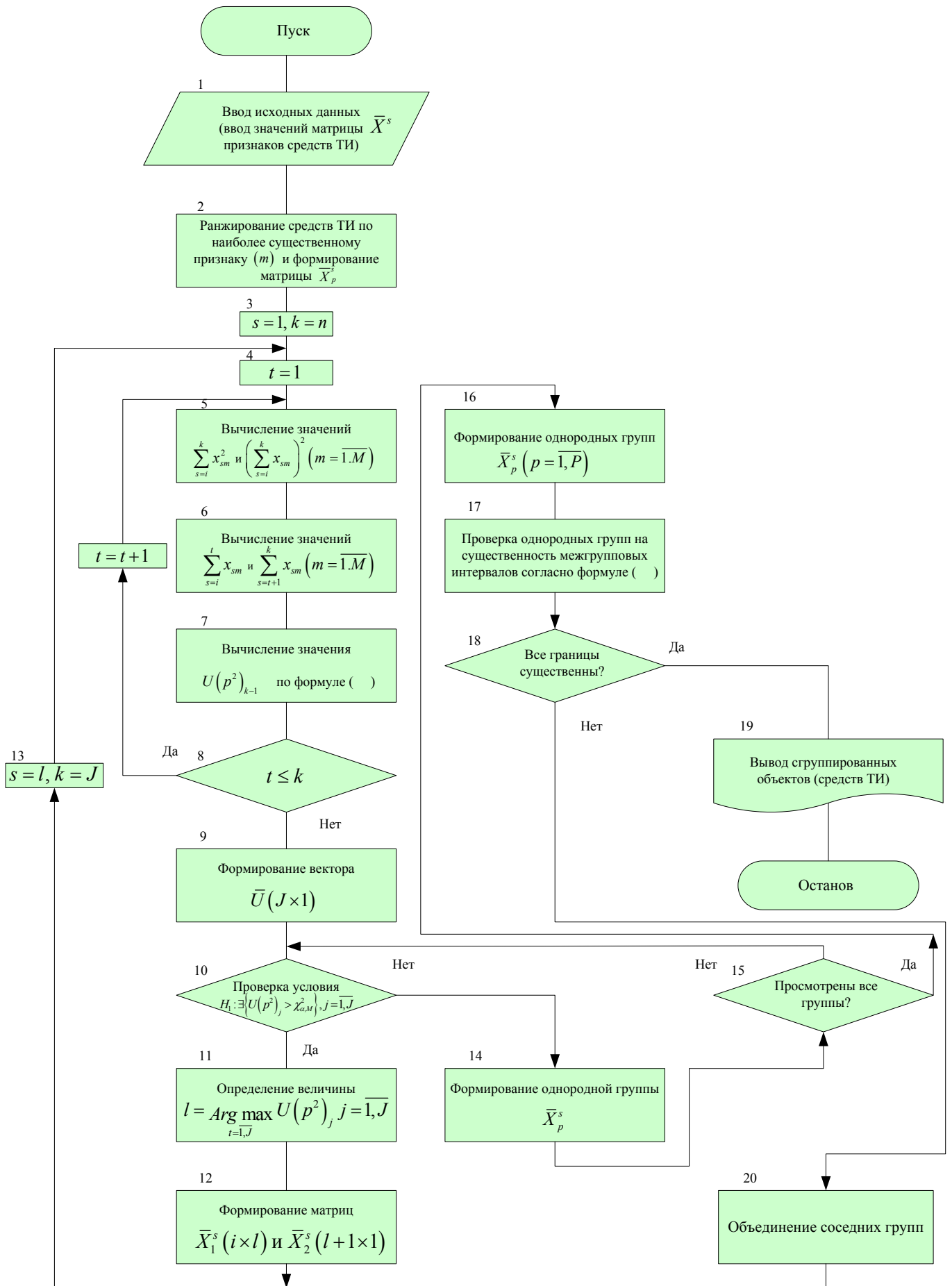


Рисунок 1 – Алгоритм автоматической кластеризации средств траекторных измерений

Проверка выполнения гипотезы H_0 проводится путем последовательной проверки выполнения конкурирующей гипотезы H_1 для всех значений $j = \overline{1, J}$, т.е. проверки выполнения следующего условия:

$$H_1 : \exists \left\{ U(\rho^2)_j > \chi_{\alpha, M}^2 \right\}, j = \overline{1, J}$$

В случае положительного решения в блоке 10 проводится разбиение генеральной совокупности средств ТИ на две группы, т.е. матрица \bar{X}^s разбивается на две следующие матрицы $\bar{X}_1^s (1 \times l)$ и $\bar{X}_2^s ((l+1) \times n)$ (блок 12 алгоритма).

Разбиение неоднородного множества средств траекторных измерений на однородные отличающиеся друг от друга группы должно базироваться на том варианте разбиения M -мерной совокупности, при которой достигается максимальное различие в комплексе признаков. Поэтому, если критерий однородности группы средств ТИ базировался на справедливости гипотезы H_0 , то при разбиении неоднородной совокупности этих средств необходимо опираться на принятии конкурирующей гипотезы $H_1 : \hat{\Xi}(\rho^2) \neq \{0, 0, \dots, 0\}$ для всех $\rho^2 \in P^2$.

Как показано в работе [2], критерий разбиения неоднородной совокупности на отличающиеся друг от друга однородные группы формулируется следующим образом.

Неоднородное множество ранжированных по существенному признаку многомерных объектов $\bar{\Xi}_p^s$ может быть разбито на два однородных подмножества $\bar{\Xi}_{1,l}^s$ и $\bar{\Xi}_{l,n-1}^s$, если значение l равно порядковому номеру строки матрицы \bar{U} , для которой максимальна величина $U(\rho^2)$, описываемая следующим выражением:

$$U(\rho^2) = \frac{n-1}{n \cdot (n-l) \cdot l} \cdot \sum_{m=1}^M \frac{\left((n-l) \sum_{s=1}^l x_{s_m} - l \cdot \sum_{s=l+1}^n x_{s_m} \right)^2}{\sum_{s=1}^n x_{s_m}^2 - \frac{1}{n} \left(\sum_{s=1}^n x_{s_m} \right)^2} \quad (14)$$

С использованием критерия (14) алгоритм группировки будет заключаться в следующем.

В векторе $\bar{U}(J \times 1)$ результатов вычисления критерия однородности $U(\rho^2)$ отыскивается максимальное значение критерия, не удовлетворяющего условию (13), и по соответствующему номеру l строки этого вектора разбивается матрица ранжированных рядов $\bar{\Xi}_p^s$ на две группы $\bar{\Xi}_{1,l}^s$ и $\bar{\Xi}_{l,n-1}^s$.

Каждая из полученных таким образом групп M -мерных данных проверяется на однородность. Если эти две группы окажутся однородными, то интервал группировки (граница разбиения множества $\bar{\Xi}_p^s$) будет только один и процесс группировки на этом заканчивается.

В случае, если гипотеза об однородности отвергается хотя бы для одной из групп, то ее нужно разделить на две части по максимальному значению критерия (14) и каждую из частей подвергнуть проверке на однородность. Процедура этого деления неоднородной статистической совокупности продолжается до тех пор, пока не будут образованы только однородные группы.

Таким образом, в рассматриваемом алгоритме граница разбиения (l) исследуемой совокупности средств на две группы определяются в блоке 11 по следующему критерию:

$$l = \underset{l=1, \bar{J}}{\text{Arg max}} U(\rho^2)_j, j = \overline{1, \bar{J}}$$

Процесс нахождения неоднородных групп проводится итерационно с использованием блоков (4...13) рассматриваемого алгоритма.

В блоке 14 проводится формирование текущей однородной группы \bar{X}_p^s , прошедшей проверку на выполнение гипотезы H_0 в блоке 10 алгоритма.

В блоке 15 алгоритма проверяется условие: проверены ли все группы на однородность? В случае положительного решения в блоке 16 проводится формирование однородных групп средств ТИ $\bar{X}_p^s (p = \overline{1, \bar{P}})$.

После того, как найдены однородные группы, необходимо сделать проверку существенности межгрупповых границ.

Это вызвано тем, что критерий (14) более приспособлен для решения задач разбиения неоднородной совокупности на две части. Поэтому, когда число разбиений совокупности, диктуемое ее неоднородностью, более одного, процесс последовательного дробления совокупности может привести к появлению статистически неустойчивых границ между однородными группами. Выявление таких межгрупповых границ и их устранение из первоначально полученной группировки приводит к определению ее действительных границ.

Предположим, что в процессе первоначальной группировки совокупность разделена на q однородных групп. Обозначим через \hat{M}_k многомерное среднее k -ой группы ($k = 1, 2, \dots, q$). В принятых обозначениях нулевая гипотеза о том, что граница между группами $\bar{\Xi}_k^s$ и $\bar{\Xi}_{k+1}^s$ является статистически неустойчивой запишется в виде соотношения $H_0 : \hat{M}_k - \hat{M}_{k+1} = \{0, 0, \dots, 0\}$, а конкурирующая гипотеза примет вид $H_1 : \hat{M}_k - \hat{M}_{k+1} \neq \{0, 0, \dots, 0\}$.

Принятие нулевой гипотезы требует признания статистически неустойчивой границы между группами $\bar{\Xi}_k^s$ и $\bar{\Xi}_{k+1}^s$. Отклонение гипотезы H_1 способствует принятию альтернативной гипотезы H_1 .

Поэтому алгоритм проверки существенности межгрупповых границ основан на последовательном сравнении многомерных средних соседних групп: первой и второй, второй и третьей, третьей и четвертой и т.д.

Если граница признается существенной (устойчивой), то группы сохраняются. В противном случае, если граница окажется неустойчивой, то такие две группы объединяются в одну. При этом проверке подвергается граница между вновь образованной группой и последующей по порядку.

Установить существенность межгрупповых границ можно с помощью аналогичного (14) следующего критерия:

$$U(\bar{X}_k^s, \bar{X}_{k+1}^s) = \frac{n_k + n_{k+1} - 1}{(n_k + n_{k+1}) \cdot n_k \cdot n_{k+1}} \cdot \sum_{m=1}^M \frac{\left(n_{k+1} \sum_{s \in S_k} x_{s_m} - n_k \sum_{s \in S_{k+1}} x_{s_m} \right)^2}{\sum_{s \in S \cup S_{k+1}} x_{s_m}^2 - \frac{1}{n_k + n_{k+1}} \left(\sum_{s \in S \cup S_{k+1}} x_{s_m} \right)^2}. \quad (15)$$

Вычисленное значение критерия $U(\bar{X}_k^s, \bar{X}_{k+1}^s)$ сравнивается со значением $\chi_{\alpha, M}^2$. Если окажется, что значение критерия меньше соответствующего значения критерия $\chi_{\alpha, M}^2$, то межгрупповая граница считается статистически неустойчивой. В этом случае требуется провести объединение групп и последующее их совместное сравнение с группой $k+2$ и т.д.

В рассматриваемом алгоритме проверка однородных групп на существенность межгрупповых интервалов согласно критерию (15) проводится в блоке 17.

В случае, если все границы существенны (блок 18), то проводится вывод сгруппированных объектов (блок 20) и завершение работы алгоритма.

Если условие блока 18 не выполняется, то необходимо провести объединение таких соседних групп и провести повторную проверку этих групп на однородность в блоках (3...10) рассматриваемого алгоритма.

Кроме того, необходимо отметить ряд особенностей работы рассматриваемого алгоритма при обработке ограниченного числа характерных признаков средств траекторных измерений.

Такой случай может возникнуть, когда на одной из итераций работы алгоритма может возникнуть группа средств ТИ, у которых имеются одинаковые значения характерных признаков (тактико-технических характеристик). В этом случае для значений m , где имеются одинаковые значения признаков, величина знаменателя критерия (14) равна нулю и значение критерия $U(\rho^2)$ будет отсутствовать. В подобной ситуации работа алгоритма предусматривает исключение из рассмотрения тех характерных признаков $(X_m, m = \overline{1, M})$, которые имеют одинаковые значения (нет смысла проверять на близость случайные величины, которые имеют все одинаковые реализации), и оставление в рассмотрении оставшихся признаков (X'_m) , имеющих

различные значения. При этом изменяется допустимое значение распределения $\chi^2_{\alpha, M}$, поскольку уменьшается величина степеней свободы M .

Таким образом, разработанный алгоритм позволяет проводить первичную обработку данных о показателях качества средств ТИ, которые могут использоваться в качестве исходных данных в задачах программно-целевого планирования развития таких средств.

Апробация алгоритма кластеризации средств траекторных измерений

Проверку разработанного алгоритма автоматической кластеризации средств траекторных измерений проведем на примере обработки значений ТТХ пятнадцати средств траекторных измерений, которые имеют возможность измерения дальности до летательного объекта, азимута и угла места на исследуемый объект. При этом рассматривается пять следующих показателей качества средств ТИ:

– среднеквадратическая погрешность измерения угловых координат летательного объекта $(\sigma_{\beta, \varepsilon})$;

– среднеквадратическая погрешности измерения наклонной дальности до летательного объекта (σ_D) ;

– максимальная дальность обнаружения и сопровождения летательного объекта (D_{\max}) ;

– максимальная угловая скорость сопровождения летательного объекта $(\dot{V}_{\beta, \varepsilon})$;

– максимальное угловое ускорение сопровождения летательного объекта $(\ddot{V}_{\beta, \varepsilon})$.

Значения показателей качества рассматриваемых средств ТИ приведены в таблице 1.

Рассматриваемые показатели качества средств ТИ будем рассматривать соответственно в виде пяти факторных признаков: X_1, X_2, X_3, X_4, X_5 .

Т а б л и ц а 1 – Значения показателей качества средств траекторных измерений

№№ средств	Тип средства траекторных измерений	Погрешности измерения угловых координат ($\sigma_{\beta,\varepsilon}$), угл. сек	Погрешности измерения дальности до объекта (σ_d), м	Максимальная дальность обнаружения и сопровождения объекта (D_{\max}), м	Максимальная угловая скорость сопровождения объекта ($\dot{V}_{\beta,\varepsilon}$), град/с	Максимальное угловое ускорение сопровождения объекта ($\ddot{V}_{\beta,\varepsilon}$), град/с ²
1	РЛС «Кама-Н»	200	10	2500	18	3
2	ОЭС «Янтарь»	5	1	50	30	20
3	РЛС «Леопард»	60	3	2500	20	4,5
4	ОЭС «Жасмин»	5	1	25	100	30
5	ОЭС «Жасмин -СС»	5	1	35	100	30
6	ОЭС «Жасмин – Р»	15	1	70	60	30
7	ОЭС «Визир-1»	15	0,3	20	45	25
8	ОЭС «Визир-2»	15	0,3	70	30	20
9	РЛС «Глория-2»	60	9	2500	60	60
10	ОЭС «Ока»	20	0,5	20	30	30
11	ОЭС «Виола»	10	2	50	30	30
12	ОЭС «Верста-ТМ»	5	1	1000	30	5
13	ОЭС «Верста-Т»	5	1	36000	30	5
14	РЛС «Глория-П»	360	0,4	1070	60	60
15	РЛС «Кама-М»	90	5	2500	20	4,5

Необходимо отметить, что каждый из пяти факторных признаков является важным.

Предварительно по каждому факторному признаку вычислим размах вариаций в соответствии со следующей формулой:

$$R_i = \frac{X_{(\max)_i}}{X_{(\min)_i}},$$

где $X_{(\max)_i}$ и $X_{(\min)_i}$ – соответственно максимальное и минимальное значение i -го фактора.

В нашем примере будем иметь: $R_1=72$; $R_2=33,3$; $R_3=1440$; $R_4=5,5$; $R_5=20$.

Столь высокие отклонения максимальных и минимальных значений признаков позволяют предположить, что совокупность, состоящая из 15 средств ТИ, неоднородна.

Далее в соответствии с разработанным алгоритмом необходимо ранжировать совокупность по одному из признаков, который по отношению к другим считается наиболее существенным. Примем в качестве такого признака среднеквадратическую погрешность измерения угловых координат (признак X_1). Ранжированный ряд представим в виде таблицы 2. В этой таблице в скобках записаны начальные номера средств траекторных измерений. В ходе всех дальнейших расчетов эти номера сохраняются, хотя участие в расчетах они не принимают. Сохранение номеров необходимо для того, чтобы определить в какую группу войдет конкретное средство ТИ.

Для проверки гипотезы однородности рассматриваемой совокупности средств ТИ используем критерий

$$U(\rho^2) = \frac{n-1}{n \cdot (n-t) \cdot t} \cdot \sum_{m=1}^M \frac{\left((n-t) \sum_{s=1}^t x_{s_m} - t \cdot \sum_{s=t+1}^n x_{s_m} \right)^2}{\sum_{s=1}^n x_{s_m}^2 - \frac{1}{n} \left(\sum_{s=1}^n x_{s_m} \right)^2}$$

Т а б л и ц а 2 – Ранжированный ряд средств траекторных измерений

№ п/п (номер средства ТИ)	X_1	X_2	X_3	X_4	X_5
1 (4)	5	1	25	100	30
2 (5)	5	1	35	100	30
3 (2)	5	1	50	30	20
4 (12)	5	1	1000	30	5
5 (13)	5	1	36000	30	5
6 (11)	10	2	50	30	20
7 (7)	15	0,3	20	45	25
8 (8)	15	0,3	70	30	20
9 (6)	15	1	70	60	30
10 (10)	20	0,5	20	30	30

№ п/п (номер средства ТИ)	X ₁	X ₂	X ₃	X ₄	X ₅
11 (3)	60	3	2500	20	4,5
12 (9)	60	9	2500	60	60
13 (15)	90	5	2500	20	4,5
14 (1)	200	10	2500	18	3
15 (14)	360	0,4	1070	60	60

Данный критерий исчисляется для множества разбиений $n=15$ и числа факторных признаков $M=5$. Число t изменяется от единицы, до числа, равного $n-l$.

В результате проведенных последовательных вычислений заполняется таблица 3.

Т а б л и ц а 3 – Значение критерия однородности совокупности средств ТИ на первой итерации работы алгоритма

№ п/п	$t; n-t$	$U(\rho^2)$	№ п/п	$t; n-t$	$U(\rho^2)$
1	1;14	5,1	8	8;7	5,66
2	2;13	10,2*	9	9;6	5,98
3	3;12	6,4	10	10;5	5,98
4	4;11	5,26	11	11;4	4,67
5	5;10	6,15	12	12;3	2,82
6	6;9	5,38	13	13;2	1,98
7	7;8	5,48	14	14;1	1,04

Анализ таблицы 3 показывает, что значения $U(\rho^2)$ превышают значения критерия $\chi_{\alpha, M}^2 = 1,145$ при выбранном высоком уровне значимости α и пятью степенями свободы ($M=5$). Поэтому рассматриваемая совокупность средств ТИ не является однородной и требуется ее разбиение на две группы.

В соответствии с разработанным алгоритмом разбиения совокупности средств ТИ на однородные группы определяется число l , равное номеру строки таблицы, где значение критерия $U(\rho^2)$ максимально. Как видно из таблицы, в первой итерации процесса кластеризации $l=2$. Тогда совокупность средств ТИ разбивается на две группы следующим образом. В первую группу входят средства ТИ под порядковыми номерами 1 и 2, а во вторую – под номерами от 3 до 15 (см. таблицу 4).

Т а б л и ц а 4 – Совокупность средств ТИ после первой итерации работы алгоритма

Первая группа						Вторая группа					
№ п/п	X ₁	X ₂	X ₃	X ₄	X ₅	№ п/п	X ₁	X ₂	X ₃	X ₄	X ₅
1(4)	5	1	25	100	30	1(2)	5	1	50	30	20
2(5)	5	1	35	100	30	2(12)	5	1	1000	30	5
$\sum_{s=1}^2 X_{S_m}$	10	2	60	200	60	3(13)	5	1	36000	30	5
$\sum_{s=1}^2 X_{s,1}^2 = 50$; $\sum_{s=1}^2 X_{s,2}^2 = 2$; $\sum_{s=1}^2 X_{s,3}^2 = 1850$; $\sum_{s=1}^2 X_{s,4}^2 = 20000$; $\sum_{s=1}^2 X_{s,5}^2 = 1800$.						4(11)	10	2	50	30	20
						5(7)	15	0,3	20	45	25
						6(8)	15	0,3	70	30	20
						7(6)	15	1	70	60	30
						8(10)	20	0,5	20	30	30
						9(3)	60	3	2500	20	4,5
						10(9)	60	9	2500	60	60
						11(15)	90	5	2500	20	4,5
						12(1)	200	10	2500	18	3
						13(14)	350	0,4	1070	60	60

Далее подвергается анализу на однородность первая группа средств ТИ (итерация № 2).

Подсчет значения критерия $U(\rho^2)$ показывает, что его значение равно единице и группа 1 признается однородной.

На третьей итерации проводится проверка на однородность второй группы средств ТИ и подсчитывается значения критерия $U(\rho^2)$, которые представлены в таблице 5.

Т а б л и ц а 5 – Значения критерия однородности средств ТИ на третьей итерации работы алгоритма

№ п/п	t; n-t	$U(\rho^2)$	№ п/п	t; n-t	$U(\rho^2)$
1	1;12	1,09	7	7;6	6,02
2	2;13	3,25	8	8;5	6,17
3	3;10	7,94*	9	9;4	4,82
4	4;9	7,03	10	10;3	2,74
5	5;8	5,63	11	11;2	2,01
6	6;7	6,19	12	12;1	1,29

Анализ этой таблицы показывает, что рассматриваемая совокупность является неоднородной, поскольку

$$\exists U(\rho^2)_i \geq \chi_{\alpha, M}^2 (i = \overline{1, 12})$$

Поэтому рассматриваемая совокупность разбивается на следующие две группы, представленные в таблице 6.

Т а б л и ц а 6 – Совокупность средств ТИ после третьей итерации работы алгоритма

Первая группа						Вторая группа					
№ п/п	X ₁	X ₂	X ₃	X ₄	X ₅	№ п/п	X ₁	X ₂	X ₃	X ₄	X ₅
1(2)	5	1	50	30	20	1(11)	10	2	50	30	20
2(12)	5	1	1000	30	5	2(7)	15	0,3	20	45	25
3(13)	5	1	36000	30	5	3(8)	15	0,3	70	30	20
$\sum_{s=1}^3 X_{s,m}$	15	3	37050	90	30	4(6)	15	1	70	60	30
$\sum_{s=1}^3 X_{s,1}^2 = 75;$			$\sum_{s=1}^3 X_{s,4}^2 = 2700;$			5(10)	20	0,5	20	30	30
$\sum_{s=1}^3 X_{s,2}^2 = 3;$			$\sum_{s=1}^3 X_{s,5}^2 = 450$			6(3)	60	3	2500	20	4,5
$\sum_{s=1}^3 X_{s,3}^2 = 129700250;$						7(9)	60	9	2500	60	60
						8(15)	90	5	2500	20	4,5
						9(1)	200	10	2500	18	3
						10(14)	360	0,4	1070	60	60

Далее проводится проверка на однородность групп №1 и №2, представленных в таблице 6.

Значения критериев однородности $U(\rho^2)$ для этих двух групп представлены в таблице 7.

Из анализа данных таблицы 7 можно сделать следующие выводы:

– по результатам работы алгоритма на четвертой итерации делаем вывод о том, что первая группа средств ТИ распадается на две группы, первая из которых представляет самостоятельное средство под номером 1, а вторая – два оставшихся средства ТИ;

Т а б л и ц а 7 – Значения критерия однородности средств ТИ на четвертой и пятой итерации работы алгоритма

Первая группа			Вторая группа		
№ п/п	$t; n-t$	$U(\rho^2)$	№ п/п	$t; n-t$	$U(\rho^2)$
1	1;2	2,54*	1	1;9	2,03
2	2;1	1,25	2	2;8	3,54
Группа 1: $\max U(\rho^2) = 2,54$. Группа 2: $\max U(\rho^2) = 9,06$.			3	3;7	4,47
			4	4;6	7,54
			5	5;5	9,06*
			6	6;4	5,95
			7	7;3	3,48
			8	8;2	1,99
			9	9;1	1,24

– по результатам работы алгоритма на пятой итерации принимаем решение о разбиении второй группы на следующие две подгруппы, первая из которых включает первых пять средств ТИ, а вторая – оставшиеся пять средств.

Аналогичные вычисления значений критериев однородности и разбиения групп средств ТИ на подгруппы проводятся до тех пор, пока не будут получены однородные группы или группы не распадутся на отдельные средства траекторных измерений.

В таблице 8 показаны все этапы группирования средств ТИ на всех итерациях работы алгоритма для рассматриваемого варианта исходных данных. Номера средств ТИ, входящие в группы, соответствуют порядковым номерам в ранжированном ряду средств (см. таблицу 2).

Т а б л и ц а 8 – Характеристики итерационного процесса кластеризации средств ТИ

Номер итерации	Исходная группа средств ТИ	Первая группа средств ТИ	Вторая группа средств ТИ
1.	1;2;3;4;5;6;7;8;9; 10;11;12;13;14;15	1;2 (неоднородна)	3;4;5;6;7;8;9; 10;11;12;13;14;15 (неоднородна)
2.	1;2	Группа признана однородной	
3.	3;4;5;6;7;8;9;10; 11;12;13;14;15	3;4;5 (неоднородна)	6;7;8;9;10; 11;12;13;14;15 (неоднородна)
4.	3;4;5	3	4;5

Номер итерации	Исходная группа средств ТИ	Первая группа средств ТИ	Вторая группа средств ТИ
		(отдельное средство ТИ)	(неоднородна)
5.	6;7;8;9;10;11; 12;13;14;15	6;7;8;9;10 (неоднородна)	11;12;13;14;15 (неоднородна)
6.	4,5	Группа признана однородной	
7.	6;7;8;9;10	6	7;8;9;10
8.	11;12;13;14;15	11 (отдельное средство ТИ)	12;13;14;15 (неоднородна)
9.	12;13;14;15	12 (отдельное средство ТИ)	13;14;15 (неоднородна)
10.	13;14;15	13;14 (неоднородна)	15 (отдельное средство ТИ)
11.	13;14	Группа признана однородной	
12.	7;8;9;10	7;8 (неоднородна)	9;10 (неоднородна)
13.	7;8	7 (отдельное средство ТИ)	8 (отдельное средство ТИ)
14.	9;10	9 (отдельное средство ТИ)	10 (отдельное средство ТИ)

Таким образом, в результате работы алгоритма кластеризации средств ТИ были получены группы средств, представленные в таблице 9.

Т а б л и ц а 9 – Однородные группы средств траекторных измерений

№ группы	Средства траекторных измерений	№ группы	Средства траекторных измерений
1	ОЭС «Жасмин» ОЭС «Жасмин-СС»	8	ОЭС «Верста -ТМ» ОЭС «Верста-Т»
2	ОЭС «Янтарь»	9	РЛС «Кама-М» РЛС «Кама-Н»
3	ОЭС «Виола»	10	ОЭС «Жасмин -Р»
4	РЛС «Леопард»	11	ОЭС «Визир-1»
5	РЛС «Глория-2»	12	ОЭС «Визир-2»
6	РЛС «Глория»		
7	ОЭС «Ока»		

Анализ таблицы 9 показывает, что с помощью программно реализованного алгоритма проведено разбиение совокупности из 15 средств ТИ на 12 однородных групп.

Проверка существенности межгрупповых границ показала, что границы существенны и объединение соседних групп не требуется.

Анализ результатов апробации разработанного алгоритма автоматической кластеризации средств ТИ позволяет сделать ряд следующих выводов:

1. Разработанный алгоритм позволил выявить по рассматриваемым пяти признакам (техническим характеристикам) однородные оптико-электронные средства траекторных измерений («Жасмин» и «Жасмин – СС»), радиолокационные станции ТИ («Кама-Н» и «Кама-М») и ОЭС большой дальности («Верста – ТМ» и «Верста – Т»);

2. На пятой итерации алгоритм четко разграничил оптико-электронные и радиолокационные средства траекторных измерений;

3. Алгоритм является работоспособным и может использоваться для решения задачи построения первоначального номенклатурного ряда средств ТИ, который будет рассматриваться в качестве исходного в оптимизационных задачах программно-целевого планирования развития средств траекторных измерений.

Список использованных источников:

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрии.- М.: ЮНИТИ, 1998.
2. Общая теория статистики. Статистическая методология в изучении коммерческой деятельности. Под редакцией Башиной О.Э. – М.: Финансы и статистика, 1999.
3. Кильдишев Г.С. Основы анализа статистических данных. – М.: МЭСИ, 1978.