

Расчет показателей ведется с применением балльного метода оценки, при котором эффективность работы предприятия рассчитывается как отношение фактически полученных результатов к базисным. За нормативное значение эффективности берется оценка, равная 100 баллам.

Численные значения экономических и социальных показателей определяются из плановых документов, статистической и оперативной отчетности, например:

$$X_i = \frac{P_i^{\Phi}}{P_i^{\delta}} \times 100,$$

где X_i – процентное отношение i -го частного показателя эффективности, %;

P_i^{Φ} – фактическое значение i -го экономического и социального показателя конечного результата за отчетный год, нат.ед.;

P_i^{δ} – базисное значение i -го показателя конечного результата (план, норматив) за отчетный период, нат.ед.

Стимулирование реализуется путем введения в оценку эффективности работы скорректированных значений показателей по формуле:

$$Y_i = f(x_i),$$

где Y_i – численное значение скорректированного i -го показателя конечного результат, %;

$f(x_i)$ – математическая функция стимулирования i -го показателя.

Следовательно, при 100%-ном выполнении базисного показателя скорректированный показатель также должен быть равен 100%, а при отклонениях рассчитывается по конкретной функции стимулирования в зависимости от важности и экономического значения показателя.

Практическая ценность данных расчетов определяется возможностью оценки динамики функционирования персонала предприятия в различные периоды времени.

Литература:

Санталайнел Т. Управление по результатам / Пер. с фин. М.: Прогресс, 1993 – 318 с.

УДК 681

ОБ АЛГОРИТМЕ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ, ОСНОВАННОМ НА КРИТЕРИИ СОГЛАСИЯ

Е.Г.Чекал, А.А.Чичев

1. Задача классификации объектов, являющаяся одной из проблем статистического анализа данных, представляет собой важную составную часть разнообразных автоматизированных систем поддержки принятия решений в различных отраслях человеческой деятельности.

Ниже предлагается алгоритм классификационной процедуры, использующий в качестве решающего правила разбиения множества объектов критерий согласия для анализируемого расстояния и выборки расстояний между объектами.

Алгоритм строится исходя из следующих предположений.

Пусть задано множество объектов, характеризующихся количественными и качественными параметрами.

Параметры могут быть константами или являться функциями времени и некоторых параметров. Значения параметров задаются дискретно (в неравномерные промежутки времени). Аппроксимация функций допускает достаточно достоверную экстраполяцию на прогнозируемый интервал времени (определяется эмпирически в зависимости от области применения). Значения функций могут быть не определены в некоторые моменты времени. Начиная с любого момента времени, данные по одним объектам могут отсутствовать, и, наоборот, появиться данные по новым объектам.

Погрешности значений параметров не заданы.

Требуется создать процедуру выявления естественного расслоения объектов на четко выраженные сгустки объектов в условиях, когда: отсутствуют какие-либо оценки размеров и количества классов; нет образцов классов (обучающей выборки).

В случае временной зависимости:

результаты классификации должны быть известны в каждый момент времени, включая начальный (т.е. отсутствуют накопленные данные);

классы должны быть стабильны в течение некоторого (задаваемого эмпирически) периода времени; причем допускается включение в класс объектов с противоположными изменениями данных, не приводящих к разрушению класса.

По результатам классификации требуется определить характеристики классов.

Сам алгоритм работы процедуры кластеризации должен быть предсказуем и согласовываться с интуитивными представлениями человека о классах.

Реализация алгоритма на ЭВМ должна обеспечить работные времена, укладывающиеся во временные рамки реального процесса обработки данных с целью принятия решения.

2. Итак, пусть задано множество объектов $\{X_i, i=1..n\}$ с параметрами $\{X_j, i=1..n, j=1..m\}$.

2.1. Случай пространственно-статистической выборки, когда параметры $\{X_j, i=1..n, j=1..m\}$ не являются функциями времени и независимы.

Разбиение на классы осуществляется в несколько этапов. Сначала по первому параметру, затем по второму и т.д. Решающее правило разбиения рассмотрим на примере одного параметра.

Пусть $\{X_i^l, i=1..n\}$ значения первого параметра объектов.

Выбор метрики пространства объектов зависит в основном от главных целей исследования, физической и статистической природы объектов. Примеры метрик подробно рассмотрены в [1]. Авторами статьи метрика определяется следующим образом.

Множество $\{X_i^l, i=1..n\}$ упорядочивается по возрастанию и расстояния определяются по формуле (1)

$$d_i = d(X_i^l, X_{i+1}^l) = \frac{|X_{i+1}^l - X_i^l|}{\sum_{k=1}^n |X_i^k|} \quad (1)$$

Далее рассматривается выборка расстояний между соседними значениями в вариационном ряду $\{X_i^l, i=1..n\}$. Очевидно, что сравнительный анализ d_i может выявить сгустки объектов (по первому параметру) или, напротив, разрежение между скоплениями объектов в том случае, когда некоторые значения d_i значительно отличаются (аномальны) от остальных значений.

Строится вариационный ряд $\{d_i, i=1..n-1\}$ по возрастанию. Из полученного ряда исключаются малые значения, которые не отражают сущность цели или которыми можно пренебречь. Например, в радиолокации одни и те же объекты могут наблюдаться различными станциями. И потому в множестве исходных данных об объектах могут присутствовать очень близкие значения параметров и соответственно очень малые значения d_i . Но поскольку решается задача группирования, а не идентификации объектов, такие значения d_i нужно исключить.

Очевидно, что аномальные значения расстояний будут наибольшими в ряду $\{d_i, i=1..n-1\}$ и потому располагаться в конце вариационного ряда.

Просмотр d_i начинается с первого значения. Для него эмпирически определяется оценка. Далее рассматривается второе значение. Если d_2 согласуется (на основе некоторого критерия) с выборкой $d = \{d_1\}$, то d_2 включается в выборку d . Затем проверяется согласие d_3 и выборки $d = \{d_1, d_2\}$ и т.д. Процесс останавливается, если найдено значение d_s , не согласующееся с выборкой $d = \{d_1, d_2, \dots, d_{s-1}\}$. Все остальные значения расстояний, большие d_s , тем более будут аномальны.

Общая задача проверки согласия выборки d из вариационного ряда и некоторого значения d_s из этого ряда рассмотрена в [2].

Значение d_s считается аномальным, если

$$I_\lambda(s, 1) < p, \quad (2)$$

где p – пороговая вероятность, означающая вероятность ошибочного принятия аномальности значения d_{s+1} .

$$\lambda = P(d_{s+1})/P(d_s),$$

$P(d_s)$ - соответственно функция распределения d_s .

Для равномерно-распределенных или нормально-распределенных d_s критерий аномальности сформулирован [2] в виде, пригодном для реализации на ЭВМ.

Величина d_s считается аномальной, если

$$\left\{ \frac{n-1}{s+1} R\left(\frac{d_s}{d_{s+1}} \gamma_s \right) \right\}^s < p \quad (3)$$

Здесь

$$\gamma_s = \begin{cases} 0.5 b_s (1 + 1/6 b_s^2 + b_s^4), & \text{при } b_s = \frac{S+1}{n-1} < 5/6 \\ 1 - 3\sqrt{0.75(1 - b_s)}, & \text{при } b_s \leq 5/6 \end{cases}$$

функция $R(x)$ задается таблицей

x	$x \leq 0$	$0 < x \leq 1/2$	$1/2 < x \leq 1$	$x > 1$
$R(x)$	0	$2x - 4x^3/3$	$1 - 4(1-x)^3/3$	1

Пороговая вероятность задается равной 0,05 (0,01).

2.2. Случай пространственно-временной выборки, когда параметры $\{X_j^i(t), i=1..n, j=1..m\}$ являются функциями времени и взаимно независимы.

В каждый момент времени разбиение на классы производится по методике, изложенной для первого случая. Полученные классы текущего момента идентифицируются с классами предыдущего момента времени по принципу "ближайшего соседа". Для сокращения количества переборov, множество классов предварительно разделяется на скопления по критерию аномальности и уже в них применяется правило. Возможны ситуации, когда класс предыдущего момента времени не имеет своего продолжения в множестве классов текущего момента времени. Тогда он считается распавшимся. И, напротив, возможно появление нового класса.

2.3. Случай пространственно-временной выборки с зависимыми параметрами.

Предположим, что $X_i^1(t), X_i^2(t), \dots, X_i^k(t), i=1..n$ независимы, а $X_i^{k+1}, X_i^{k+2}, \dots, X_i^m, i=1..n$ зависят от $t, X_i^1(t), X_i^2(t), \dots, X_i^k(t)$, но независимы между собой.

Методика разбиения на классы аналогична второму случаю с той разницей, что в формуле расстояний (1) участвуют лишь $X_i^{k+1}, X_i^{k+2}, \dots, X_i^m$. Влияние параметров $X_i^1(t), X_i^2(t), \dots, X_i^k$, таким образом, учитывается косвенно.

3. При временной зависимости параметров объектов в ситуации, когда отсутствуют какие-либо значения – применяется экстраполяция значений соответствующих параметров предыдущих моментов времени. Экстраполяция осуществляется по закону, являющимся достаточным приближением реального. Закон подбирается по известным методикам. Например, можно использовать сглаженную предысторию значений параметров, накапливаемую в каждый момент времени.

Если значения какого-либо параметра не были известны ранее, то решение о принадлежности классу принимается по известным параметрам.

4. Когда параметры объектов являются функцией времени, авторами статьи применялась методика формирования предыстории объектов и использование накопленных данных для формирования обобщенных характеристик классов, которые использовались при идентификации объектов по данным текущего и предыдущего моментов времени.

5. Для улучшения стабильности классов во временном промежутке авторами применялась дополнительная проверка состава классов. Объекты классов экстраполировались на задаваемый эмпирически период стабильности. Далее применялась процедура разбиения на классы. Окончательно в классах оставались лишь те объекты, которые собирались в классы в течение всего периода стабильности.

6. По результатам классификации для каждого выявленного класса определяются выборочные моменты [3]: объем класса, среднее значение, выборочная дисперсия. Дополнительно определяются специфичные для физической природы объектов характеристики.

7. Идея, заложенная в алгоритме, представляет, в конечном счете, поиск пустот между объектами при просмотре одновременно всех объектов (так называемая параллельная процедура классификации [4]), что очень близко к человеческому интуитивному процессу выявления скопленений.

8. Ввиду того, что анализируются не все взаимные расстояния, а лишь специальная выборка, программная реализация алгоритма на ЭВМ показала возможность использования его в автоматизированных системах поддержки принятия решений, включая системы реального времени.

УДК 633.1

УСТОЙЧИВОСТЬ ЗЕРНОВОГО ПРОИЗВОДСТВА И КРИТЕРИИ ЕЕ ОЦЕНКИ

С.В. Басенкова, аспирантка

Зерновое производство по объему создаваемой продукции и социально-экономическому значению занимает ключевое место в обес-