

## **Модель планирования сезонных продаж в современных корпоративных информационных системах**

*Работа посвящена проблеме автоматизации управления современным предприятием, а именно – использованию на предприятии современных информационных технологий планирования продаж. Предложена новая модель планирования продаж, ориентированная на сезонные товары и связанная с применением идей кластерного анализа. На реальных данных продемонстрирована конструктивность и эффективность модели.*

### **Введение**

Деятельность современных производственных и торговых предприятий характеризуется крайне высоким уровнем различного рода издержек, что приводит к постоянному балансированию на грани рентабельности. Цена управленческих ошибок в таких условиях становится очень значительной; это является основным побудительным мотивом внедрения на предприятиях средств информатизации – корпоративных информационных систем (КИС). Основной задачей таких корпоративных систем является создание информационной базы для принятия управленческих решений руководством предприятий, при этом содержание и форма представления информации должны обеспечивать принятие как можно меньшего числа ошибочных решений.

Номенклатура материалов и товаров крупной компании может исчисляться десятками и сотнями тысяч наименований. Известные технологии поддержки работы предприятия в этих условиях (ERP, MRP) ориентированы, прежде всего, на минимизацию складских запасов, что позволяет сократить издержки на хранение и привлечение оборотных средств. С другой стороны, в условиях конкуренции не менее актуальной становится задача удержания клиента, что предполагает гарантию наличия товара в нужный момент и постоянное расширение списка номенклатуры.

В этих условиях планирование продаж становится важнейшей функцией корпоративной информационной системы предприятия. Современные корпоративные системы (например, российские «Галактика», «1С», зарубежные «SAP», «Ахарта» и др.) реализуют различные технологии автоматизации планирования продаж: от простейших, до весьма продвинутых. Все они основаны, по сути, на ретроспективном анализе данных и экстраполяции на предстоящие периоды с использованием методов регрессии и статистики.

Планирование продаж сезонных товаров представляет особенно важную и сложную задачу, так как предполагает выборку номенклатурных позиций, продажи которых носят сезонный характер, оценку периодичности продаж и графика распределения продаж по периодам. Указанные параметры должны обеспечивать максимальную достоверность планирования, однако все современные КИС (в их стандартной конфигурации) предоставляют пользователю выбирать эти параметры исключительно интуитивно, «вручную»; при большом списке номенклатуры качественная оценка параметров сезонности таким образом может быть неосуществима.

В настоящей работе рассматривается модель задачи планирования сезонных продаж, ориентированная на реализацию в современных КИС и предлагается метод автоматизированной оценки параметров сезонности на основе технологий кластерного анализа.

### **Основные модельные представления**

В задачах экономического анализа минимальным периодом оценивания какого-либо показателя являются, как правило, сутки, соотнесённые с конкретной датой. В задачах планирования используются, главным образом, «укрупнённые» периоды: неделя, декада, месяц, квартал и т. д.

Пусть  $X_k$  - значение выбранного показателя (им может быть, например, прибыль от продаж, количество продаж, выручка от продаж и т.п.) в период с номером  $k$ . Модель изменения значения показателя  $X$  во времени может быть выражена формулой

$$X_k = G(k) + \eta(k), \quad (1)$$

где  $G(k)$  - функция, выражающая детерминированный закон эволюции величины  $X$  (тренд),  $\eta(k)$  - случайная величина, характеризующая отклонение фактического значения показателя от его тренда (здесь и далее будем считать, что  $\eta(k)$  - некоррелированная случайная величина с нулевым математическим ожиданием). При решении задачи планирования известная функция  $G(k)$  используется для экстраполяции значений показателя  $X$ , а свойства  $\eta(k)$  ложатся в основу оценки достоверности планирования (например, выраженной в форме доверительных интервалов) [1].

Примем, что тренд  $G(k)$  есть функция, представляемая как:

$$G(k) = A \cdot g(k), \quad (2)$$

где  $g(k)$  есть периодическая функция с периодом  $J$ , так что  $g(k) = g(k + J)$ ,  $\sum_{k=1}^J g(k) = 1$ ,

$A$  - величина, характеризующая суммированное значение показателя  $X$  за период.

Функция  $g(k)$ , таким образом, задает график распределения показателя  $X$  в течение периода. (Пояснение: авторы надеются, что употребление термина «период» как для обозначения номера интервала времени  $k$  в формуле (1), так и для обозначения свойства периодической функции  $g(k)$  в формуле (2) не затруднит читателю понимание текста).

Из практики хорошо известно, что период, свойственный продажам сезонных товаров, составляет 1 год, что однозначно идентифицирует соответствующую этому отрезку времени величину  $J$ . В обсуждаемом контексте основной проблемой планирования продаж становится, таким образом, идентификация тех номенклатурных позиций, для которых  $g(k)$  действительно является периодической функцией и определение вида функции  $g(k)$  в пределах годового периода.

Пусть  $g_i(k)$  - функция  $g(k)$  для номенклатурной позиции с номером  $i$ . Введем величину  $d_{ij}$ , характеризующую метрику расстояния между графиками распределения показателей  $X$  в течение года для номенклатурных позиций с номерами  $i$  и  $j$  так что

$$d_{ij} = \sum_{k=1}^J |g_i(k) - g_j(k)|. \quad (3)$$

Соответствующую  $d_{ij}$  матрицу  $D$  назовем матрицей расстояний между годовыми графиками распределения.

Среди множества номенклатурных позиций, как правило, находятся такие, которые характеризуются сходным «типом сезонности». Так, продажи одних товаров приходятся, главным образом, на зимний период, других – на летний, третьих – на межсезонье. У товаров с одним типом сезонности функции  $g(k)$  будут схожими. С точки зрения модельных представлений (3), такие товары характеризуются близко расположенными в  $J$ -мерном пространстве точками с координатами  $g(k)$ . Матрица  $D$  является входными данными для определения тех областей (кластеров)  $J$ -мерного пространства, в которых группируется наибольшее число близко расположенных точек  $g(k)$ . Соответствующие таким кластерам номенклатурные позиции могут быть идентифицированы как товары, обладающие одинаковым типом сезонности, и, следовательно, характеризующиеся одинаковым графиком распределения  $g(k)$  для решения задачи планирования.

Таким образом, в рамках модельных представлений (1), (2) и (3) может быть поставлена задача о нахождении на заданной метрике (3) центров кластеров на множестве точек  $g_i(k)$ , выделении точек, принадлежащих тому или иному кластеру и экстраполяции трендов  $G(k)$  согласно (2).

## Метод решения задачи

Известно несколько подходов к решению задачи кластеризации по известной матрице  $D$ . Первая группа методов связана с кластеризацией при заданном числе кластеров (четкий и нечеткий алгоритм С-средних [2, 3], нейронные сети Кохонена [4] и др.). Вторая группа не требует задания количества кластеров, кластеры идентифицируются в процессе работы алгоритма (горный алгоритм [5, 6]).

Кластеризация при заданном числе кластеров в рассматриваемой задаче означает, по сути, задание количества возможных графиков распределения  $g(k)$ . Такой подход применим в том случае, когда номенклатура товаров характеризуется четко выделенными группами и каждая группа имеет свой, присущий только ей тип сезонности. На практике такая ситуация встречается весьма нечасто. Поэтому в настоящей работе рассматривается подход, где количество кластеров определяется параметрами самого алгоритма.

Сущность алгоритма горной кластеризации состоит в том, что вначале определяют точки, которые могут быть центрами кластеров. Затем для каждой такой точки рассчитывается значение потенциала, показывающего возможность формирования кластера в ее окрестности. Чем плотнее расположены объекты в окрестности потенциального центра кластера, тем выше значение его потенциала. После этого итерационно выбираются центры кластеров среди точек с максимальными потенциалами.

В настоящей работе в качестве центров возможных кластеров предлагается выбирать сами объекты кластеризации, то есть оценки графиков распределения  $g_i(k)$  для каждой номенклатурной позиции, так что

$$\hat{g}_i(k) = \frac{X_k^{(i)}}{\sum_{j=1}^J X_j^{(i)}}, \quad k = \overline{1, J},$$

где  $X_k^{(i)}$  - значение выбранного показателя в период с номером  $k$  для номенклатурной позиции с номером  $i$ .

Для расчета потенциала точки  $\hat{g}_i(k)$  выбрана следующая функция потенциала:

$$P(\hat{g}_i(k)) = \sum_{j=1}^M \exp(-\alpha \cdot \hat{d}_{ij}),$$

где  $\hat{d}_{ij} = \sum_{k=1}^J |\hat{g}_i(k) - \hat{g}_j(k)|$  - оценка соответствующего коэффициента матрицы расстояний,  $M$  - количество номенклатурных позиций, участвующих в кластерном анализе,  $\alpha$  - число, характеризующее масштаб расстояний  $d_{ij}$ ,  $\exp(\cdot)$  - оператор экспоненты.

На первом шаге кластеризации рассчитывают потенциал каждой точки и выбирают точку  $\hat{g}_i^{(1)}(k)$  с наибольшим потенциалом  $P(\hat{g}_i^{(1)}(k))$ ; эта точка и будет центром первого кластера. На втором шаге пересчитывают значения потенциалов всех точек таким образом, чтобы исключить влияние потенциала уже найденного кластера; для этого от текущих значений потенциала вычитают вклад центра найденного кластера по формуле

$$P_2(\hat{g}_i(k)) = P(\hat{g}_i(k)) - P(\hat{g}_i^{(1)}(k)) \cdot \exp(-\beta \cdot \hat{d}_{iz}),$$

где  $\beta$  - число, характеризующее размер кластеров,  $z$  - индекс точки, являющейся центром первого кластера. Центр второго кластера - это точка  $\hat{g}_i^{(2)}(k)$  с максимальным значением потенциала  $P_2(\hat{g}_i^{(2)}(k))$ . Аналогично находятся центры всех следующих кластеров. Итерационная процедура пересчета потенциалов и выделения центров кластеров продолжается до тех пор, пока максимальное значение потенциала превышает некоторый заданный порог.

Принадлежность точки  $\hat{g}_i(k)$  к тому или иному кластеру определяется по расстоянию до центра этого кластера. Точка считается принадлежащей кластеру, если расстояние до него (по сравнению с расстоянием до центров других кластеров) минимально и не превышает некоторого заданного значения  $\beta^*$ . Если же расстояние от точки до центров всех найденных кластеров больше, чем  $\beta^*$ , то считается, что она не принадлежит ни одному из кластеров.

### Результаты решения задачи

Эксперимент по планированию продаж с использованием предлагаемой методики проводился на реальных данных о продажах крупной компании, занимающейся торговлей автозапчастями, номенклатура товаров которой включает более 13 тысяч наименований, при этом продажи по 2.5 тысячам наименований носят регулярный характер. С учетом специфики товаров и учетной политики предприятия были приняты следующие значения параметров задачи: показатель  $X_k$  равен количеству проданных товаров; интервал времени между  $X_k$  и  $X_{k+1}$  равен одному месяцу; период  $J$  равен 12 месяцам, причем  $k=1$  соответствует январю, а  $k=12$  - декабрю. Параметр  $\alpha$  был выбран равным 2 (максимальное возможное значение  $d_{ij}$ ), параметр  $\beta$  взят равным 8; максимальное расстояние до центра кластера  $\beta^*$  равно 0.2.

На рисунке 1а показаны значения потенциалов  $P(\hat{g}_i(k))$  на первом шаге алгоритма (сплошная линия), втором и третьем шаге (пунктир), а, также ещё трех шагах (точки).

Видно, что, начиная с четвертого кластера, влияние их на потенциалы точек  $\hat{g}(k)$  становится очень небольшим (очень небольшая разница между графиками из точек на рисунке). Это означает, что в окрестностях найденных центров четвертого и последующих кластеров точки, по сути, не группируются. Это подтверждается и анализом, результаты которого представлены на рисунке 1в. Здесь показано количество номенклатурных позиций, принадлежащих первому найденному кластеру (152 позиции), второму (31 позиция), третьему (45 позиций), четвертому и пятому (4 и 3 позиции соответственно). Из результатов следует, что из рассматриваемого множества точек  $\hat{g}(k)$  можно, по сути, выделить только три кластера с указанным числом точек.

На рисунке 1б показаны значения  $\hat{g}(k)$  точек, принадлежащих первому кластеру (пучок графиков из точек, левая часть рисунка) и значения  $\hat{g}(k)$  точки, являющейся центром первого кластера (жирная сплошная линия, левая часть рисунка). По сути дела, показана годовая динамика продаж товаров, отнесенных к первому типу сезонности, соотнесённая с эталонной динамикой первого типа сезонности (в качестве каковой взята годовая динамика точки – центра кластера). В правой части рисунка 1б показаны соотнесенные с найденной эталонной динамикой фактические значения  $\hat{g}(k)$ , относящиеся к началу следующего года. Видно, что левая и правая часть рисунка вполне согласуются между собой, то есть товары, идентифицированные по результатам анализа данных за год как сезонные первого типа, продолжают сохранять тот же тип сезонности и на следующий год.

Аналогичная картина имеет место и для второго, и для третьего кластера – рисунки 1г и 1е. Из сравнительного анализа рисунков 1б, 1г и 1е видно, что в рассмотренном массиве данных о количестве продаж может быть выделено три типа сезонности: низкий уровень продаж в январе-марте с резким ростом в апреле и стабильным уровнем до конца года (первый тип сезонности); стабильный уровень продаж с января по июль с ростом осенью и падением к концу года (второй тип сезонности); низкий уровень в зимний период и высокий с мая по октябрь (третий тип сезонности). Такие типы сезонности действительно характерны для некоторых групп автотоваров.

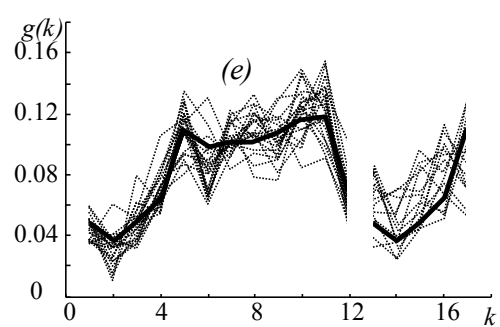
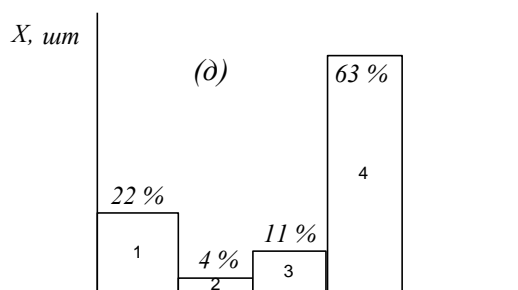
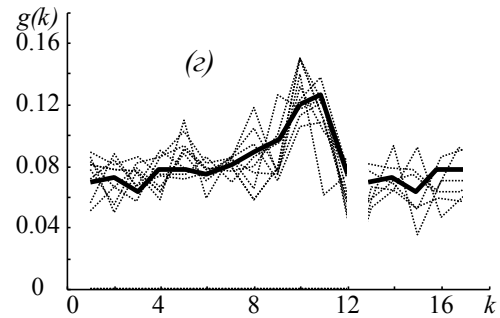
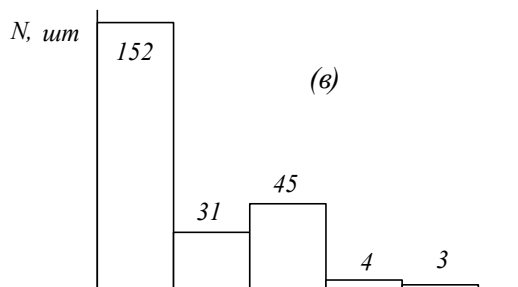
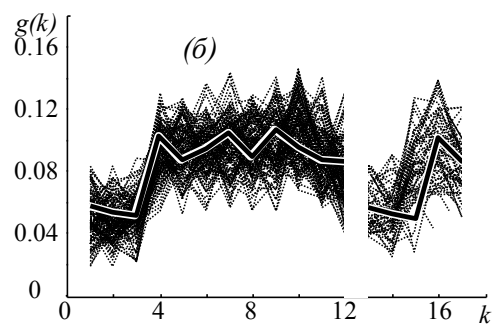
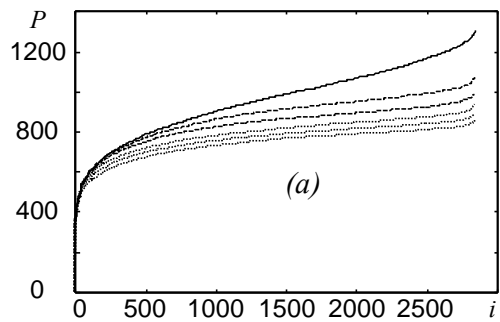
На рисунке 1д показано количественное распределение по кластерам тех товаров, продажи которых носят регулярный характер. На товары с сезонностью первого типа приходится 22% продаж, товары с сезонностью 2 и 3 типа занимают 4% и 11% соответственно. Остальные товары – 63% - не являются сезонными. Такое соотношение (более трети продаваемых товаров характеризуются ярко выраженной сезонностью) подтверждает актуальность применения соответствующих методов планирования.

Предлагаемая в работе модель планирования сезонных продаж адаптирована к данным корпоративной информационной системы «1С: Управление торговлей 8» и «1С: Управление производственным предприятием 8» и реализована в виде обработки платформы «1С: 8», основное назначение которой – предоставить менеджеру выборку сезонных номенклатурных позиций и рекомендации по планированию их закупок и продаж. Апробация программы на действующем предприятии подтвердила конструктивность и непротиворечивость предлагаемой методики.

Работа ориентирована на расширение функциональности современных корпоративных информационных систем.

## **Литература**

1. Гриняк В.М., Семенов С.М. Модель планирования продаж в современных корпоративных информационных системах // Естественные и технические науки. - 2009. - №1.
2. Bezdek J.C. Pattern Recognition with Fuzzy Objective Function. - New York: Plenum Press, 1981.
3. Babuska R. Fuzzy Modeling for Control. - Boston: Kluwer Academic Publishers, 1998.
4. Хайкин С. Нейронные сети. Полный курс. – М.: Вильямс, 2006.
5. Yager R., Filev D. Essentials of Fuzzy Modeling and Control. - John Wiley & Sons, 1984.
6. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику [Электронный ресурс] - Режим доступа: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php>





## **Подписи к рисункам**

**Рис. 1.** Результаты решения задачи планирования сезонных продаж