

СЕКЦИЯ 1

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

Д. О. БРЫКИН, С. А. ЛЕБЕДЕВ

Московский физико-технический институт, г. Москва

УДК 519.2; 519.86

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ ПРОГНОЗИРОВАНИЯ РЕДКИХ ПРОДАЖ В 1С: CROSTON, SBA, TSB И МОДИФИЦИРОВАННЫЙ TSB

Аннотация. Редкие продажи составляют значительную долю ассортимента многих предприятий. В статье сравниваются классические алгоритмы Croston, SBA и TSB с SQL-ориентированной модификацией TSB (Mod-TSB), реализуемой средствами платформы 1С. Эксперимент на 10000 сгенерированных временных рядов показал, что Mod-TSB снижает RMSE на $\approx 3\%$ и сокращает время расчета на $\sim 30\%$.

Ключевые слова: редкие продажи, прогнозирование спроса, 1С, Croston, SBA, TSB.

Введение. В условиях современного рынка и стремительной цифровизации вопрос точного прогнозирования спроса приобретает все большую значимость. Одной из наиболее сложных задач здесь является предсказание «редких продаж» — таких позиций, где количество продаж за период незначительно, а интервалы между покупками нерегулярны. Некорректный прогноз может привести либо к избытку неликвидных запасов, либо к нехватке востребованного товара в момент спроса, что негативно сказывается на финансовых показателях компании.

В работе рассматриваются данные алгоритмы в контексте внедрения в решения, работающие на платформе 1С. Программные продукты 1С занимают лидирующую долю рынка бизнес-приложений. При этом, с технической точки зрения, платформа обладает рядом ограничений, связанных с интерпретируемостью языка 1С. В стандартных инструментах 1С отсутствуют встроенные механизмы для специальных статистических расчетов, что затрудняет внедрение «тяжелых» алгоритмов. Поэтому при внедрении алгоритмов нужно искать баланс между точностью прогнозирования и вычислительной эффективностью.

В настоящем исследовании проводится сравнительный анализ четырех моделей для прогнозирования редких продаж, реализованных в среде 1С: классических Croston, SBA (Syntetos-Boylan Approximation), TSB (Teunter-Syntetos-Babai). Они широко применяются в ритейле и логистике благодаря своей простоте и приемлемой точности. Также, была предложена модификация TSB с заменой экспоненциального сглаживания на обычную среднюю, которую будем называть Mod-TSB. Такая модификация реализуется проще в запросах 1С и может быть производительнее на больших объемах данных за счет выноса логики на сторону СУБД. Для оценки точности указанных методов RMSE (Root Mean Square Error).

Границы редких/разреженных временных рядов расплывчаты. В данной работе взята классификация авторов [1], которые считают, что товары, у которых метрика ADI (Average

Demand Interval, среднее время между заказами) составляет 1.4 и более (то есть имеют от 30% периодов без продаж) являются редкими. Верхняя граница установлена в 90% периодов без продаж.

Ниже представлено краткое описание исследуемых моделей.

Модели Croston и SBA. Классический метод Croston предполагает отдельный расчет двух показателей: среднего объема спроса (non-zero demand) и средней частоты (интервала) между поступающими заказами [2]. Оценки обновляются с помощью экспоненциального сглаживания. Недостатком Croston считается завышение итогового прогноза. В ответ на это была предложена модификация SBA, которая вводит корректирующий коэффициент, призванный устранять систематическое смещение, тем самым улучшая точность прогноза в ряде случаев [3].

Классический TSB и Mod-TSB. TSB отличается от Croston тем, что использует экспоненциальное сглаживание сразу для вероятности ненулевого спроса и для средней величины спроса в ненулевые периоды. В отличие от Croston и SBA, TSB дает более сглаженный и стабильный результат [4], что особенно полезно при наличии «пульсирующих» паттернов спроса. TSB сглаживает вероятность ненулевой продажи p_t и средний размер заказа z_t с коэффициентами α , β , как показано на формуле 1.

$$\begin{aligned} p_t &= p_{t-1} + \alpha(o_t - p_{t-1}) \\ z_t &= z_{t-1} + \beta(y_t - z_{t-1}) \end{aligned} \quad (1)$$

где $o_t \in \{0,1\}$ — индикатор продажи, y_t — объем заказа. Прогноз $f_t = p_t * z_t$. Алгоритм быстро адаптируется к изменениям, но требует «долгой памяти» и нескольких рекурсивных вычислений в каждый момент времени.

В рамках настоящей работы предлагается использовать модификацию классического TSB, которую обозначим как Mod-TSB. Вместо экспоненциального сглаживания для оценки вероятности и/или величины спроса в ненулевые периоды используется обычная средняя, вычисляемая через SQL запрос, используя SQL диалект 1C. Такой подход эффективнее с точки зрения вычислительных ресурсов. Платформа 1C накладывает существенные ограничения на производительность при больших объемах данных, поскольку язык является интерпретируемым [5]. Преимущество Mod-TSB в более быстрой обработке временных рядов при сопоставимой точности прогноза. Реализация описывается формулой 2.

$$\begin{aligned} p_t &= AVG(d_{t-k+1..t}) \\ p_t &= AVG(d_{t-k+1..t}) \end{aligned} \quad (2)$$

Результаты. Для сравнения алгоритмов было сгенерировано 100 наборов по 100 SKU длиной 365 дней. Для каждого ряда задана слабая недельная сезонность и умеренный возрастающий тренд 0.001, шанс возникновения продажи 15%. Для Croston, SBA и TSB коэффициенты α , β подбирались в сетке 0.01–0.30 с шагом 0.05 для минимизации RMSE. Выборка делилась на обучающий и тренировочные наборы, на тренировочный набор выделялось последние 10 % данных. Для MOD-TSB окно k фиксировалось равным длине ряда. В качестве метрики ошибок использовалось RMSE. В таблице 1 показан RMSE одного из наборов, а также итоговые средние значения.

Сравнение моделей по RMSE

<i>SKU</i>	<i>SKU_1</i>	<i>SKU_10</i>	<i>SKU_11</i>	...	<i>SKU_97</i>	<i>SKU_98</i>	<i>SKU_99</i>	<i>AVG RMSE</i>
Croston	1.69	2.47	1.91	...	1.85	1.93	1.47	1.83
SBA	1.65	2.46	1.89	...	1.85	1.90	1.47	1.82
TSB	1.64	1.63	1.74	...	1.55	1.53	1.46	1.66
Mod_TSB	1.60	1.61	1.70	...	1.53	1.46	1.42	1.61

По итогу, метод, в среднем, оказывается даже лучше классической версии TSB на $\approx 3\%$ по RMSE. Причина выигрыша MOD-TSB объясняется тем, что на коротких рядах экспоненциальное сглаживание «не успевает разогнаться», тогда как простое среднее работает стабильнее. Использование окна k равного длине ряда избавляет от поиска α , β и минимизирует вычислительные издержки. Если длина ряда большая или выражена сильная сезонность, классический TSB начинает выигрывать благодаря «длинной памяти». Кроме того, MOD-TSB дольше «помнит» устаревшие тренды, что может привести к запаздыванию прогноза на быстрорастущих позициях.

Так же, помимо возросшей точности, данный метод имеет еще несколько преимуществ. Один SQL-запрос с двумя AVG-агрегациями полностью выполняется в СУБД, снижая нагрузку на RAM и CPU сервера 1C и ускоряя расчет до 30 %. Метод целиком реализуется штатными средствами 1C; он подходит для облачных тарифов, не требует внешних библиотек и дорогих процессоров.

Заключение. Предложенная модификация TSB демонстрирует, что простота не обязательно означает потерю качества. При коротких историях и ограниченных ресурсах она обеспечивает лучшую (RMSE ниже, в среднем, на 3%) или сопоставимую точность при ощутимой экономии времени. При длинных рядах и ярко выраженной сезонности классический TSB может быть точнее.

СПИСОК ЛИТЕРАТУРЫ

1. Ghobbar A. A. Sources of intermittent demand for aircraft spare parts within airline operations / A. A. Ghobbar, C. H. Friend // Journal of Air Transport Management. — 2002. — Т. 8, № 4. — С. 221-231.
2. Croston J. D. Forecasting and Stock Control for Intermittent Demands / J. D. Croston // Operational Research Quarterly (1970-1977). — 1972. — Т. 23, № 3. — С. 289.
3. Syntetos A. A. The accuracy of intermittent demand estimates / A.A. Syntetos, J.E. Boylan // International Journal of Forecasting. — 2005. — Vol. 21, № 2. — P. 303-314.
4. Teunter R. H. Intermittent demand: Linking forecasting to inventory obsolescence / R.H. Teunter, A.A. Syntetos, M. Zied Babai // European Journal of Operational Research. — 2011. — Т. 214, № 3. — С. 606-615.
5. Платформа «1C: Предприятие» — что под капотом? — URL: <https://habr.com/ru/companies/1c/articles/269611/> (дата обращения: 28.04.2025).