УДК 330.43

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОЙ РАСПРЕДЕЛЁННОСТИ ТОРГОВЫХ СЕТЕЙ

О. С. Балаш

Саратовский государственный университет
E-mail: olgabalash@mail.ru

В представленной статье рассматриваются проблемы регрессионного анализа пространственных данных на примере торговых сетей в городе Саратове с использованием закона розничной гравитации.

Ключевые слова: модели розничных продаж, пространственная регрессионная модель, розничная модель гравитации

Modeling of Spatial Distribution of Trading Networks

O. S. Balash

In this paper we consider the problem of regression analysis of spatial data on an example of trading networks in the city of Saratov are considered, using the law of retail gravitation.

Key words: retail sales models, spatial regression model, retail gravity model.

При открытии магазина или расширении розничной сети владелец должен не только иметь целостный взгляд на рыночные показатели, но и обладать информацией о территориальном распределении покупателей и конкурентов по городу.

В 1931 г. W. Reilly применил закон тяготения И. Ньютона к экономической географии[1]. Им сформулирован закон розничной гравитации, связывающий потребителей и продавцов с учетом их пространственной распределенности:

$$Z_{bc} = k m_b^{\beta_b} m_c^{\beta_c} m_{bc}^{\beta_{bc}},$$

где Z_{bc} – взаимодействие между розничным магазином b и потребителем c; m_b – характеристики магазина b; m_c – характеристики потребителя c; k – константа; β_b, β_c, β_{bc} – оцениваемые параметры; d_{bc} – расстояние между магазином b и потребителем c.

Гравитационная модель может включать не все параметры. Так, T. Lakshmanan и W. Hansen использовали только размер магазина и расстояние[2]. D. Gautschi, проведя тестирование по данным о торговых поездках, предположил, что при исключении переменных из модели завышается значение параметра расстояния от центра[3].

T. Stanley и M. Sewall обнаружили, что размер магазина не оказывает большого влияния на взаимодействие покупателей и продавца[4]. M. Eppli и J. Shilling построили модель по данным продаж методом наименьших квадратов (МНК). Они сделали вывод, что местоположение магазина не имеет большого значения, а переменная «расстояние» значительно завышается[5].

Более поздние исследования включают в модель параметры магазина, такие как тип, длительность работы, а также экономические и демографические характеристики покупателей и частоту посещения магазина. A. Okoruwa, J. Terza и H. Nourse тестировали уравнение регрессией Пуассона[6].

Многие авторы исследовали пространственную концентрацию магазинов. Они выяснили, что кластеризация разнородных магазинов способствует снижению общих затрат на поездку, а однородных магазинов – ценовой оптимизации для покупателя.

Основой гравитационной модели в пространстве является расстояние, но оно не может отображать всего спектра взаимозависимости потребителей и продавцов. При моделировании необходимо учитывать многоцелевое поведение клиентов. В работах M. Eppli и J. Benjamin исследовано влияние потребительского поведения на выбор магазина[7]. Тем не менее достаточно часто переменная «потребительское поведение клиента» в гравитационную модель не включается.

R. Pace и O. Gilley в 1997 г. предложили использовать SAR-модель *(structure-activity relationship model)* [8]:

$$(E - \alpha D)Y = (E - \alpha D)X\beta + \varepsilon,$$

где Y – вектор независимых переменных, размерности $n \times 1$; X – матрица, содержащая n наблюдений k независимых регрессоров; α – пространственный параметр; D – матрица пространственных весов размером $n \times n$; ε – вектор ошибок размером n; E – единичная матрица.

Эту модель можно представить как

$$Y = X\beta + \varepsilon,$$

где вектор ошибок $\varepsilon \sim N(0, \Omega)$ и обратная матрица пространственных весов

$$\Omega^{-1} = (I - \alpha D)^T (I - \alpha D).$$

Параметр $\alpha > 0$ показывает положительную пространственную зависимость. То есть ошибки, имеющие одинаковый знак, географически сгруппированы вместе. Если $\alpha < 0$, то имеет место отрицательная пространственная зависимость. Если $\alpha = 0$, то модель SAR сводится к модели МНК.

Было обнаружено, что потребители и их покупательная способность положительно пространственно зависят между собой. Совокупность отдельных магазинов и их способность к продажам также приводят к положительной пространственной зависимости среди магазинов розничной сети.

Матрица D имеет нули на диагонали. Для облегчения интерпретации сумма элементов строки D равна единице. Для устойчивости процесса будем предполагать, что пространственный параметр автокорреляции α находится в интервале (0, 1).

В модели пространственного распределения магазинов и покупателей мы специфицировали пространственную матрицу весов следующим образом:

$$D = wC + (1 - w)S,$$

где C и S – весовые матрицы для потребителей и магазинов соответственно и

$$0 \leq w \leq 1.$$

Если $w = 1$, то матрица D превращается в матрицу C, при $w = 0$ – в S.

Построим матрицу C, используя метод ближайших соседей с геометрически убывающими весами. В этом методе более близкому соседу присваивается больший вес, дальнему – меньший, убывающий по геометрической прогрессии.

Будем рассматривать только m ближайших клиентов для каждого потребителя.

Пусть $N^{(h)}$ представляет собой матрицу размером $n \times n$, для которой $N_y^{(h)} = 1$, если наблюдение j есть h ближайший сосед для i наблюдения ($h = 1, 2,\ldots, m$; $j = 1,2,\ldots, n$; $i \neq j$) и $N_y^{(h)} = 0$ и в противном случае.

Матрицы $N^{(1)}, N^{(2)}, \ldots$, представляют последовательность матриц соседей.

Обозначим ρ – знаменатель прогрессии весовых коэффициентов, для которых вес h соседа равен ρ^h при условии, что $0 \leq \rho \leq 1$.

Определим матрицу C следующим образом:

$$C = \frac{\sum_{h=1}^{m} \rho^h N^{(h)}}{\sum_{h=1}^{m} \rho^h},$$

имеющей в каждой строке сумму элементов, равную единице, и нули по диагонали.

Моделируя пространственную зависимость между магазинами, рассмотрим пять шагов построения эконометрической модели.

Первый шаг представляет собой триангуляцию Делоне среди магазинов. Триангуляция Делоне является аналогом двойной диаграммы Вороного, которая геометрически изображает связи между смежными магазинами[9]. Диаграмма Вороного конечного множества точек на плоскости представляет такое разбиение плоскости, при котором каждая область этого разбиения образует множество точек, более близких к одному из элементов этого множества, чем к любому другому элементу.

Каждый магазин в треугольнике Делоне обслуживает покупателей, которые, как правило, проживают в одном районе.

Матрица V размерности $n_s \times n_s$ представляет весовую матрицу триангуляции Делоне для n_s магазинов. В этом случае $V_{ij} = 1$, если наблюдения i и j принадлежат смежным треугольникам, и $V_{ij} = 0$ в противном случае.

Вторым шагом приводят строки к стандартизированному виду:

$$G_i = \frac{V_{ij}}{\sum_{\substack{j=1 \\ i \neq j}}^{n_s} V_{ij}} \cdot$$

Следовательно, G представляет собой вектор размером n_s, состоящий из единиц $G[1n_s] = [1n_s]$.

Третий этап состоит в агрегировании потребителей каждого магазина. Пусть $A_{ij} = 1$, если потребитель i делает покупки в магазине j, и $A_{ij} = 0$ в противном случае. Матрица A состоит из нулей и единиц размером $n_s \times n_s$.

Для четвертого шага стандартизируем транспонируемую матрицу A. Пусть R представляет собой диагональную матрицу размером $n_s \times n_s$ с элементами, равными обратной сумме столбцов матрицы A. В этом случае будет случайным RA^T и $RA^T[1] = [1]$.

Рассмотрим произведение матриц

$$S = AG\,(RA^T),$$

которое представляет собой количественную характеристику среднего пространственного уровня покупателя.

Пусть вектор переменной v размером $n \times 1$ – стоимость некой покупки. Вектор RA^Tv размерности $n_s{\times}1$ рассчитывается как средняя потребительская цена этой покупки для каждого магазина (агрегированного для n потребителей магазинов). Вектор GRA^Tv размерности $n_s{\times}1$ представляет пространственную среднюю потребительскую цену для конкурентных магазинов. Вектор $AGRA^Tv$ есть средняя цена приобретения для потребителей, которые покупали в соседних магазинах. Если обозначить как v ошибку, то $AGRA^Tv$ будет измерять средние ошибки близлежащих магазинов для каждого клиента. Если независимые переменные недостаточно специфицируемы для близлежащих магазинов, то $AGRA^Tv$ будет положительной для покупателей, которые покупали в данном магазине, и эта информация может улучшить прогноз. Учитывая это замечание, мы можем выполнять операции с S без необходимости сохранения $n \times n$ матрицы.

Оценка SAR-модели проводится методом максимизации функции правдоподобия (Pace, Barry, Slawson, Sirmans, 2002):

$$L(\alpha) = \kappa + \ln|I - \alpha D| - \frac{1}{2}\ln(SSE(\alpha)),$$

где $SSE(\alpha) = (Y - X\beta)^T (I - \alpha D)^T (I - \alpha D)(Y - X\beta)$;

κ – константа[10].

Для вычисления выражения $\ln|I - \alpha D|$ его можно разложить в степенной ряд следующим образом:

$$\ln\left|I - \alpha D\right| = \sum_{r=1}^{\infty} \frac{-\alpha^r tr(D^T)}{r} .$$

Используя данные, полученные по городу Саратову, была построена гравитационная модель:

$$1n\ Y = X\beta + \varepsilon ,$$

где Y – средний объем продаж магазинов.

Матрица X включает три блока переменных: X_{bc}– логарифм расстояния от магазина b до покупателя c; X_b – переменные, характеризующие магазин b; X_c – переменные, относящиеся к району, в котором покупает потребитель c; ε – случайная ошибка.

Матрица X_b – это размер торговой площади. Матрица X_c состоит из пространственных характеристик «среднее время в пути до магазина», «количество проживающих в районе», «численность женщин».

Для построения модели использовали метод наименьших квадратов. В результате получили значения коэффициентов (табл. 1).

Таблица 1

Гравитационная модель

Независимые переменные	Значения коэффициентов	Стандартная ошибка
Ln(distance)	– 1,9***	4,7
Ln(store sales)	– 3,105***	0,15
Ln(house)	– 0,5***	0,1
Ln(female)	0,3	2,1
Ln(pop)	5,4**	4,1
Ln(travel time)	– 0,3	3,1
Константа	12,3	9,02

Коэффициенты логарифма расстояния, размер магазина, численность проживающих в районе и количество квартир в районе значимы на 5%-ном уровне. Остальные – численность женщин, время в пути до магазина – незначимы. Все коэффициенты представляют собой эластичность продаж.

Далее с использованием тех же данных была построена SAR-модель, которая дала практически аналогичные результаты (табл. 2).

Таблица 2

SAR-модель

Независимые переменные	Значения коэффициентов	Стандартная ошибка
Ln(distance)	– 1,81***	5,7
Ln(store sales)	– 2,9***	0,2
Ln(house)	– 0,35***	0,05
Ln(female)	0,5	1,1
Ln(pop)	6,1**	3,1
Ln(travel time)	– 0,01	0,1
Константа	14,3	3,02

Модели достаточно адекватны данным (коэффициент детерминации в обоих случаях равен 0,45).

Анализ двух моделей подтверждает гипотезу, что население в основном делает покупки в близлежащих магазинах. Размер магазина не влияет на выбор покупателя, то есть покупатели чаще обращаются в магазины шаговой доступности. Время, затраченное на поход в магазин, незначимо, и соответствующий коэффициент имеет отрицательный знак, что говорит в пользу того, что покупатели все же чаще покупают в магазинах, расположенные рядом с домом.

Примечания

1. *Reilly W. J.* The Law of Retail Gravitation / W. J. Reilly Co. N.Y. 1931.
2. *Lakshmanan T. R., Hansen W.* G. A Retail Market Potential Model // Journal of the American Institute of Planner. 1965. № 13. P. 134–43.
3. *Gautschi D. A.* Specification of Patronage Models for Retail Center Choice // Journal of Marketing Research. 1981. № 18. P. 162–174.
4. *Stanley T. J., Sewall M. A.* Image Inputs to a Probabilistic Model: Predicting Retail Potential // Journal of Marketing. 1976. № 40. P. 48–53.
5. *Eppli M. J., Shilling J. D.* How Critical Is a Good Location to a Regional Shopping Center? // Journal of Real Estate Research. 1996. № 12. P. 459–468.
6. *Okoruwa A. A., Nourse H. O., Terza J. V.* Estimating Sales for Retail Centers: An Application of the Poisson Gravity Model // Journal of Real Estate Research. 1994. № 9. P. 85–97.
7. *Eppli M. J., Benjamin J. D.* The Evolution of Shopping Center Research: A Review and Analysis // Journal of Real Estate Research. 1994. № 9. P. 5–32.
8. *Pace R. K., Gilley O. W.* Using the Spatial Configuration of the Data to Improve Estimation // Journal of Real Estate Finance and Economics. 1997. № 14. P. 333–340.
9. *Calciu M., Salerno F.* A New Approach to Spatial Management of Retail Networks, Based on German School's Central Place Theory: Application to Bank Location / H. Mbhlbacher /J-P Flipo (eds.), Advances in Services Marketing. Gabler ; Wiesbaden. 1997.
10. *Pace R. K., Barry R., Slawson V. C. Jr., Sirmans C. F.* Simultaneous Spatial and Functional Form Transformations /(ed by L. Anselin, R. Florax // Advances in Spatial Econometrics, forthcoming, Springer-Verlag. Heidelberg, 2002.
11. *Barry R. P., Pace R. K.* Monte Carlo Estimates of the Log Determinant of Large Sparse Matrices // Linear Algebra and Applications. 1999. W289. P. 41–54.