

Калибровка параметров модели расчета матрицы корреспонденций для г. Москвы

А. С. Иванова^{1,2,a}, С. С. Омельченко^{2,b}, Е. В. Котлярова^{2,c},
В. В. Матюхин^{2,d}

¹Национальный исследовательский университет «Высшая школа экономики»,
Россия, 101000, г. Москва, ул. Мясницкая, д. 20

²Национальный исследовательский университет «Московский физико-технический институт»,
Россия, 141701, г. Долгопрудный, Институтский пер., д. 9

E-mail: ^a anastasiya.s.ivanova@phystech.edu, ^b sergey.omelchenko@phystech.edu,
^c kotlyarova.ev@phystech.edu, ^d vladmatyukh@gmail.com

Получено 13.05.2020, после доработки — 30.08.2020.

Принято к публикации 03.09.2020.

В данной работе рассматривается задача восстановления матрицы корреспонденций для наблюдений реальных корреспонденций в г. Москве. Следуя общепринятому подходу [Гасников и др., 2013], транспортная сеть рассматривается как ориентированный граф, дуги которого соответствуют участкам дороги, а вершины графа — районы, из которых выезжают / в которые въезжают участники движения. Число жителей города считается постоянным. Задача восстановления матрицы корреспонденций состоит в расчете всех корреспонденций из района i в район j .

Для восстановления матрицы предлагается использовать один из наиболее популярных в урбанистике способов расчета матрицы корреспонденций — энтропийная модель. В работе, в соответствии с работой [Вильсон, 1978], приводится описание эволюционного обоснования энтропийной модели, описывается основная идея перехода к решению задачи энтропийно-линейного программирования (ЭЛП) при расчете матрицы корреспонденций. Для решения полученной задачи ЭЛП предлагается перейти к двойственной задаче и решать задачу относительно двойственных переменных. В работе описывается несколько численных методов оптимизации для решения данной задачи: алгоритм Синхорна и ускоренный алгоритм Синхорна. Далее приводятся численные эксперименты для следующих вариантов функций затрат: линейная функция затрат и сумма степенной и логарифмической функции затрат. В данных функциях затраты представляют из себя некоторую комбинацию среднего времени в пути и расстояния между районами, которая зависит от параметров. Для каждого набора параметров функции затрат рассчитывается матрица корреспонденций и далее оценивается качество восстановленной матрицы относительно известной матрицы корреспонденций. Мы предполагаем, что шум в восстановленной матрице корреспонденций является гауссовским, в результате в качестве метрики качества выступает среднеквадратичное отклонение. Данная задача представляет из себя задачу невыпуклой оптимизации. В статье приводится обзор безградиентных методов оптимизации для решения невыпуклых задач. Так как число параметров функции затрат небольшое, для определения оптимальных параметров функции затрат было выбрано использовать метод перебора по сетке значений. Таким образом, для каждого набора параметров рассчитывается матрица корреспонденций и далее оценивается качество восстановленной матрицы относительно известной матрицы корреспонденций. Далее по минимальному значению невязки для каждой функции затрат определяется, для какой функции затрат и при каких значениях параметров восстановленная матрица наилучшим образом описывает реальные корреспонденции.

Ключевые слова: модель расчета матрицы корреспонденций, энтропийно-линейное программирование, метод Синхорна, метод ускоренного Синхорна

Исследование выполнено при поддержке Российского фонда фундаментальных исследований (проект 18-29-03071 мк). Исследование В.В.Матюхина выполнено при поддержке Министерства науки и высшего образования Российской Федерации (госзадание) № 075-00337-20-03, номер проекта 0714-2020-0005.

© 2020 Анастасия Сергеевна Иванова, Сергей Сергеевич Омельченко, Екатерина Владимировна Котлярова,
Владислав Вячеславович Матюхин

Статья доступна по лицензии Creative Commons Attribution-NoDerivs 3.0 Unported License.
Чтобы получить текст лицензии, посетите веб-сайт <http://creativecommons.org/licenses/by-nd/3.0/>
или отправьте письмо в Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Calibration of model parameters for calculating correspondence matrix for Moscow

A. S. Ivanova^{1,2,a}, S. S. Omelchenko^{2,b}, E. V. Kotliarova^{2,c},
V. V. Matyukhin^{2,d}

¹Higher School of Economics – National Research University,
20 Myasnitskaya st., Moscow, 101000, Russia

²National Research University Moscow Institute of Physics and Technology,
9 Institute lane, Dolgoprudny, 141701, Russia

E-mail: ^a anastasiya.s.ivanova@phystech.edu, ^b sergey.omelchenko@phystech.edu,
^c kotlyarova.ev@phystech.edu, ^d vladmatyukh@gmail.com

Received 13.05.2020, after completion – 30.08.2020.

Accepted for publication 03.09.2020.

In this paper, we consider the problem of restoring the correspondence matrix based on the observations of real correspondences in Moscow. Following the conventional approach [Gasnikov et al., 2013], the transport network is considered as a directed graph whose edges correspond to road sections and the graph vertices correspond to areas that the traffic participants leave or enter. The number of city residents is considered constant. The problem of restoring the correspondence matrix is to calculate all the correspondence from the i area to the j area.

To restore the matrix, we propose to use one of the most popular methods of calculating the correspondence matrix in urban studies – the entropy model. In our work, which is based on the work [Wilson, 1978], we describe the evolutionary justification of the entropy model and the main idea of the transition to solving the problem of entropy-linear programming (ELP) in calculating the correspondence matrix. To solve the ELP problem, it is proposed to pass to the dual problem. In this paper, we describe several numerical optimization methods for solving this problem: the Sinkhorn method and the Accelerated Sinkhorn method. We provide numerical experiments for the following variants of cost functions: a linear cost function and a superposition of the power and logarithmic cost functions. In these functions, the cost is a combination of average time and distance between areas, which depends on the parameters. The correspondence matrix is calculated for multiple sets of parameters and then we calculate the quality of the restored matrix relative to the known correspondence matrix.

We assume that the noise in the restored correspondence matrix is Gaussian, as a result, we use the standard deviation as a quality metric. The article provides an overview of gradient-free optimization methods for solving non-convex problems. Since the number of parameters of the cost function is small, we use the grid search method to find the optimal parameters of the cost function. Thus, the correspondence matrix calculated for each set of parameters and then the quality of the restored matrix is evaluated relative to the known correspondence matrix. Further, according to the minimum residual value for each cost function, we determine for which cost function and at what parameter values the restored matrix best describes real correspondence.

Keywords: correspondence matrix calculation model, entropy linear programming, Sinkhorn method, accelerated Sinkhorn method

Citation: *Computer Research and Modeling*, 2020, vol. 12, no. 5, pp. 961–978 (Russian).

The research was supported by Russian Foundation for Basic Research (project 18-29-03071 mk). The research of V. V. Matyukhin was supported by the Ministry of Science and Higher Education of the Russian Federation (Goszadaniye) No. 075-00337-20-03, project No. 0714-2020-0005.

1. Введение

Методы моделирования матриц корреспонденций начали активно развиваться в 60-х годах прошлого века. Задача состояла в том, чтобы по $2n$ параметрам (численности активного населения в n районах и рабочих мест в этих районах) определить n^2 параметров — матрицу корреспонденций. Модели расчета матрицы корреспонденций базируются на характеристиках районов ($2n$ параметров) и матрицах затрат (n^2 известных чисел, характеризующих затраты на дорогу из одного района в другой: время, расстояние и т. д.). Самой первой моделью являлась гравитационная модель, в основу которой был положен аналог закона всемирного тяготения Ньютона–Гука. Однако в отличие от отмеченного закона физики параметры этого закона (показатель степени, с которым входят затраты между районами в оценку корреспонденций) неизвестны. В данной работе рассматривается один из наиболее популярных в урбанистике способов расчета матрицы корреспонденций, имеющий более чем сорокалетнюю историю, — энтропийная модель. В основе рассматриваемой модели лежит марковский процесс в пространстве огромной размерности, т. е. данный процесс порождает макросистему. Данный марковский процесс представляет собой ветвящийся процесс специального вида: модель стохастической химической кинетики. Вывод энтропийной модели приведен на базе бинарных реакций обменного типа, популярных в различного рода физических и социально-экономических приложениях моделей стохастической химической кинетики [Гардинер, 2009; Вайдлих, 2010]. В данной статье на базе энтропийных моделей расчета матрицы корреспонденций (см. книгу А. Дж. Вильсона [Вильсон, 1978], книгу Е. В. Гасниковой и А. В. Гасникова [Гасников и др., 2013]) предложен способ определения параметров таких моделей и выбор наилучшей. Стоит отметить, что ранее такого рода исследования сосредотачивались в основном только на гравитационной и энтропийной моделях, у которых имеется всего один параметр для калибровки. Данная работа сосредоточена на более общих и точных современных вариантах данных моделей, описанных в цикле работ Е. В. Гасниковой и А. В. Гасникова [Гасников и др., 2012; Гасников и др., 2013]. Калибровка таких моделей на реальных данных (были использованы данные по районам Москвы) ранее не осуществлялась. Отметим, что при описании процедуры калибровки для получения гарантированного (теоретически обоснованного) результата необходим аккуратный анализ возникающих задач оптимизации и используемых алгоритмов. Предложенный в статье анализ базируется на современных достижениях в двух направлениях: оптимальный транспорт (современный анализ оценок скорости сходимости алгоритма Синхорна [Dvurechensky et al., 2018]) и безградиентные методы невыпуклой оптимизации с неточным оракулом [Bogolubsky et al., 2016]. Отметим также метод балансировки [Шелейховский, 1936; Брэгман, 1967], который эквивалентен алгоритму Синхорна и может рассматриваться для решения задачи ЭЛП.

Статья организована следующим образом.

В § 2 описан вывод энтропийной модели. В результате показано, что задачу расчета матрицы корреспонденций можно рассматривать как задачу энтропийно-линейного программирования.

В § 3 описано решение задачи энтропийно-линейного программирования из § 2. Предлагается перейти к двойственной задаче, которую будем решать с помощью алгоритма Синхорна, попеременно минимизируя двойственную целевую функцию по одному из двух блоков двойственных переменных. Также описан ускоренный алгоритм Синхорна, который позволяет быстрее решать рассматриваемую задачу.

В § 4 описывается задача оценки невязки между реальной матрицей и матрицей, восстановленной в результате решения задачи из § 3.

В § 5 проведены численные эксперименты по расчету матрицы корреспонденций по реальным данным для города Москвы для разных функций затрат. Проводится сравнительный анализ для определения функции затрат, которая наилучшим образом описывает реальные данные.

2. Постановка задачи.

Энтропийная модель расчета матрицы корреспонденций

В данном разделе, базируясь на работе [Вильсон, 1978], приведем эволюционное обоснование энтропийной модели. Дальнейшее изложение вывода идет согласно статье [Гасников и др., 2016].

Пусть в некотором городе имеется n районов. Общее число жителей города постоянно и равно N , при это выполняется $N \gg n^2$. Пусть $L_i \geq 0$ — это число жителей, выезжающих в типичный день за рассматриваемый промежуток времени из района i , а $W_j \geq 0$ — число жителей, приезжающих на работу в район j в типичный день за рассматриваемый промежуток времени. В рамках рассматриваемого подхода данные величины являются входными параметрами для модели, т. е. они не моделируются. При этом будут выполняться следующие соотношения:

$$\sum_{i=1}^n L_i = \sum_{j=1}^n W_j = N.$$

Обозначим через $d_{ij}(t) \geq 0$ число жителей, живущих в i -м районе и работающих в j -м в момент времени t . Мы предполагаем, что со временем жители могут меняться только квартирами, поэтому во все моменты времени $t \geq 0$ выполнено

$$d_{ij}(t) \geq 0, \quad \sum_{i,j=1}^{n,n} d_{ij}(t) = N, \quad \sum_{j=1}^n d_{ij}(t) = L_i, \quad \sum_{i=1}^n d_{ij}(t) = W_j, \quad i, j = 1, \dots, n.$$

Определим следующее множество:

$$Q = \left\{ d_{ij} \geq 0: \sum_{i,j=1}^{n,n} d_{ij} = N, \sum_{j=1}^n d_{ij} = L_i, \sum_{i=1}^n d_{ij} = W_j, i, j = 1, \dots, n \right\}. \quad (1)$$

Отметим, что основным стимулом к обмену места жительства для жителя города будут являться транспортные издержки, то есть для каждого жителя работать далеко от дома плохо из-за больших транспортных издержек. Будем считать, что эффективной функцией затрат [Гасников и др., 2013] является функция $R(T) = \frac{\gamma T}{2}$, где $T > 0$ — затраты на путь от дома до работы, которые определяются как временем так и расстоянием, а $\gamma > 0$ — настраиваемый параметр модели, который можно интерпретировать как цену единицы затрат на путь от работы до дома. Далее в работе будем подразумевать под функцией затрат только $T(\alpha, \beta, \gamma)$, где α, β, γ — настраиваемые параметры модели.

Динамику процесса можно описать следующим образом: пусть в момент времени $t \geq 0$ r -й житель живет в k -м районе и работает в m -м, а s -й житель живет в p -м районе и работает в q -м. Тогда $\lambda_{k,m;p,q}(t)\Delta t + o(\Delta t)$ есть вероятность того, что жители с номерами r и s ($1 \leq r < s \leq N$) поменяются квартирами в промежутке времени $(t, t + \Delta t)$. Вероятность обмена местами жительства зависит только от мест проживания и работы обменивающихся:

$$\lambda_{k,m;p,q}(t) = \lambda_{k,m;p,q} = \lambda N^{-1} \exp\left(\underbrace{R(T_{km}) + R(T_{pq})}_{\text{суммарные затраты до обмена}} - \underbrace{(R(T_{pm}) + R(T_{kq}))}_{\text{суммарные затраты после обмена}}\right) > 0,$$

где коэффициент $0 < \lambda = O(1)$ характеризует интенсивность обменов. Отметим, что совершенно аналогичным образом можно было рассматривать случай, когда жители могут обмениваться местами работы. То есть мы предполагаем некое равноправие агентов (жителей) внутри фиксированной корреспонденции и их независимость [Гасников, Гасникова, 2014; Ethier, Kurtz, 1986].

Согласно эргодической теореме для марковских цепей (в независимости от начальной конфигурации $\{d_{ij}(0)\}_{i,j=1,1}^{n,n}$) [Малышев, Пирогов, 2008; Гасников и др., 2013; Гардинер, 2009; Вайдлих, 2010; Sandholm, 2010; Боровков, 1999; Levin et al., 2009] предельное распределение совпадает со стационарным (инвариантным), которое можно посчитать (получается проекция прямого произведения распределений Пуассона на Q):

$$\lim_{t \rightarrow \infty} P(d_{ij}(t) = d_{ij}, i, j = 1, \dots, n) = Z^{-1} \prod_{i,j=1,1}^{n,n} \exp(-2R(T_{ij})d_{ij}) \cdot (d_{ij}!)^{-1} = \\ \stackrel{\text{def}}{=} p(\{d_{ij}\}_{i,j=1,1}^{n,n}),$$

где $\{d_{ij}\}_{i,j=1,1}^{n,n} \in Q$, а «статсумма» — Z . Отметим, что начальная конфигурация $\{d_{ij}(0)\}_{i,j=1,1}^{n,n}$ влияет на время выхода на стационарное состояние. При этом стационарное распределение $p(\{d_{ij}\}_{i,j=1,1}^{n,n})$ удовлетворяет условию детального равновесия [Гасников, Гасникова, 2014; Sandholm, 2010]:

$$(d_{km} + 1)(d_{pq} + 1)p(\{d_{11}, \dots, d_{km} + 1, \dots, d_{pq} + 1, \dots, d_{pm} - 1, \dots, d_{kq} - 1, \dots, d_{nn}\}) \lambda_{k,m;p,q} = \\ = d_{pm}d_{kq}p(\{d_{ij}\}_{i,j=1,1}^{n,n}) \lambda_{k,m;p,q}.$$

При $N \gg 1$ распределение $p(\{d_{ij}\}_{i,j=1,1}^{n,n})$ экспоненциально сконцентрировано на множестве Q в $O(\sqrt{N})$ -окрестности наиболее вероятного значения $d^* = \{d_{ij}^*\}_{i,j=1,1}^{n,n}$, которое определяется как решение задачи энтропийно-линейного программирования (ЭЛП) [Малышев, Пирогов, 2008; Гасников и др., 2013]:

$$\ln p(\{d_{ij}\}_{i,j=1,1}^{n,n}) = \max_{\{d_{ij}\}_{i,j=1,1}^{n,n} \in Q} -\gamma \sum_{i,j=1,1}^{n,n} d_{ij}T_{ij} - \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij}.$$

Это следует из теоремы Санова о больших отклонениях для мультиномиального распределения [Санов, 1957]. Отметим, что в данном параграфе описывается основная идея перехода к решению задачи ЭЛП при расчете матрицы корреспонденций. Более формально полученный результат представлен в [Гасников и др., 2016].

3. Методы для решения задачи ЭЛП

В данном разделе приведем описание способов решения задачи ЭЛП, возникающей при расчете матрицы корреспонденций.

Как было показано в предыдущем разделе, задачу восстановления матрицы корреспонденций можно записать как следующую задачу оптимизации:

$$\min_{d_{ij} \in Q} f(d_{ij}) := \gamma \sum_{i,j=1,1}^{n,n} d_{ij}T_{ij} + \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij}, \quad (2)$$

где Q определяется как (1) и $T_{ij} := T_{ij}(\eta)$ — функция затрат на перемещение из района i в район j , которая зависит от вектора параметров η .

Введем следующую нормировку: $\sum_{i,j=1,1}^{n,n} d_{ij} = 1$; тогда ограничения можно переписать в следующем виде: $\sum_{j=1}^n d_{ij} = l_i$ и $\sum_{i=1}^n d_{ij} = w_j$, где $l_i = \frac{L_i}{N}$ $w_j = \frac{W_j}{N}$. И определим следующее множество:

$$\tilde{Q} = \left\{ d_{ij} \geq 0: \sum_{i,j=1}^{n,n} d_{ij} = 1, \sum_{j=1}^n d_{ij} = l_i, \sum_{i=1}^n d_{ij} = w_j, i, j = 1, \dots, n \right\}.$$

После введения нормировки получаем, что для задачи (2)

$$\gamma \sum_{i,j=1,1}^{n,n} N \cdot d_{ij} T_{ij} + \sum_{i,j=1,1}^{n,n} N \cdot d_{ij} \ln N \cdot d_{ij} = N \left\{ \gamma \sum_{i,j=1,1}^{n,n} d_{ij} T_{ij} + \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij} + \underbrace{\sum_{i,j=1,1}^{n,n} d_{ij} \ln N}_{=1} \right\}.$$

Тогда задача (2) переписывается в следующем эквивалентном виде:

$$\min_{d_{ij} \in Q} \gamma \sum_{i,j=1,1}^{n,n} d_{ij} T_{ij} + \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij}. \quad (3)$$

Далее введем два блока двойственных переменных: $\lambda^l \in \mathbb{R}^n$ и $\lambda^w \in \mathbb{R}^n$, где λ_i^l — множитель к ограничению $\sum_{j=1}^n d_{ij} = l_i$ и λ_j^w — множитель к ограничению $\sum_{i=1}^n d_{ij} = w_j$. Применим для решения задачи (3) метод множителей Лагранжа. Для этого запишем двойственную задачу:

$$\begin{aligned} \min_{d_{ij} \in Q} \gamma \sum_{i,j=1,1}^{n,n} d_{ij} T_{ij} + \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij} &= \\ &= \max_{\lambda^l, \lambda^w} \min_{\substack{\sum_{i,j=1,1}^{n,n} d_{ij} \geq 0 \\ d_{ij} \geq 0}} \left\{ \gamma \sum_{i,j=1,1}^{n,n} d_{ij} T_{ij} + \sum_{i,j=1,1}^{n,n} d_{ij} \ln d_{ij} + \sum_{i=1}^n \lambda_i^l \left(\sum_{j=1}^n d_{ij} - l_i \right) + \sum_{j=1}^n \lambda_j^w \left(\sum_{i=1}^n d_{ij} - w_j \right) \right\} = \\ &= \max_{\lambda^l, \lambda^w} \left\{ -\langle \lambda^l, l \rangle - \langle \lambda^w, w \rangle + \min_{d_{ij} \geq 0} \left\{ \sum_{i,j=1,1}^{n,n} d_{ij} (\gamma T_{ij} + \ln d_{ij} + \lambda_i^l + \lambda_j^w) + \nu \left(\sum_{i,j=1,1}^{n,n} d_{ij} - 1 \right) \right\} \right\} = \\ &= \max_{\lambda^l, \lambda^w} \left\{ -\langle \lambda^l, l \rangle - \langle \lambda^w, w \rangle + \left\{ \sum_{i,j=1,1}^{n,n} d_{ij} (\lambda^l, \lambda^w) (\gamma T_{ij} + \ln d_{ij} (\lambda^l, \lambda^w) + \lambda_i^l + \lambda_j^w + \nu) - \nu \right\} \right\}, \end{aligned}$$

где

$$d_{ij}(\lambda^l, \lambda^w) = \operatorname{argmin}_{d_{ij} \geq 0} \left\{ \sum_{i,j=1,1}^{n,n} d_{ij} (\gamma T_{ij} + \ln d_{ij} + \lambda_i^l + \lambda_j^w) + \nu \left(\sum_{i,j=1,1}^{n,n} d_{ij} - 1 \right) \right\}.$$

Используя условия оптимальности, получаем

$$\gamma T_{ij} + \ln d_{ij} + \lambda_i^l + \lambda_j^w + 1 + \nu = 0, \quad \sum_{i,j=1,1}^{n,n} d_{ij} = 1.$$

Решая данную систему уравнений и переопределяя $\lambda^l := -\lambda^l - \frac{1}{2}$ и $\lambda^w := -\lambda^w - \frac{1}{2}$, получаем, что

$$d_{ij}(\lambda^l, \lambda^w) = \frac{\exp(-\gamma T_{ij} + \lambda_i^l + \lambda_j^w)}{\sum_{i,j=1,1}^{n,n} \exp(-\gamma T_{ij} + \lambda_i^l + \lambda_j^w)} = \frac{B_{ij}(\lambda^l, \lambda^w)}{\mathbf{1}^T B(\lambda^l, \lambda^w) \mathbf{1}},$$

где $B_{ij}(\lambda^l, \lambda^w) = \exp(-\gamma T_{ij} + \lambda_i^l + \lambda_j^w)$. Подставляя это в двойственную задачу, получаем, что двойственная задача имеет вид

$$\max_{\lambda^l, \lambda^w} \tilde{\varphi}(\lambda^l, \lambda^w) := \langle \lambda^l, l \rangle + \langle \lambda^w, w \rangle - \ln(\mathbf{1}^T B(\lambda^l, \lambda^w) \mathbf{1}).$$

Перепишем задачу как задачу минимизации с точностью до знака:

$$\min_{\lambda^l, \lambda^w} \varphi(\lambda^l, \lambda^w) := \ln(\mathbf{1}^T B(\lambda^l, \lambda^w) \mathbf{1}) - \langle \lambda^l, l \rangle - \langle \lambda^w, w \rangle. \quad (4)$$

Для решения двойственной задачи рассмотрим метод альтернированной минимизации (алгоритм 1). Для удобства описания алгоритма введем следующее обозначение. Множество $\{1, \dots, n\}$ векторов $e_{i=1}^n$ ортонормированного базиса разделено на p непересекающихся блоков I_k , $k \in \{1, \dots, p\}$. Пусть $S_k(x) = x + \text{span}\{e_i: i \in I_k\}$ — подпространство, содержащее x , построенное на базисных векторах k -го блока.

Алгоритм 1. Метод альтернированной минимизации

- 1: **Input:** x^0 — starting point.
 - 2: **for** $k \geq 0$ **do**
 - 3: Choose $i_k \in 1, \dots, p$.
 - 4: Compute $x^{k+1} = \underset{x \in S_{i_k}(x^k)}{\text{argmin}} f(x)$.
 - 5: **end for**
 - 6: **Output:** x^k .
-

Отметим, что основной идеей данного алгоритма является минимизация по произвольно выбранному блоку переменных i_k на каждой итерации. Для задачи (4) мы будем рассматривать минимизацию по двум блокам: λ^l и λ^w . Согласно лемме 5 из [Guminov et al., 2020] шаг минимизации по блоку λ^l можно представить в следующем виде:

$$[\lambda^l]^{k+1} = [\lambda^l]^k + \ln(l) - \ln(B([\lambda^l]^k, [\lambda^w]^k) \mathbf{1}),$$

аналогичным образом можно представить шаг минимизации по блоку λ^w :

$$[\lambda^w]^{k+1} = [\lambda^w]^k + \ln(w) - \ln(B^T([\lambda^l]^k, [\lambda^w]^k) \mathbf{1}).$$

Учитывая это, для решения (4) получаем алгоритм 2.

Алгоритм 2. Алгоритм Синхорна

- 1: **Input:** $x^0 = [[\lambda^l]^0, [\lambda^w]^0] = (0, \dots, 0) \in \mathbb{R}^{2n}$ — starting point.
 - 2: **for** $k \geq 0$ **do**
 - 3: **if** $k \bmod 2 = 0$ **then**
 - 4: Compute

$$\begin{aligned} [\lambda^l]^{k+1} &= [\lambda^l]^k + \ln(l) - \ln(B([\lambda^l]^k, [\lambda^w]^k) \mathbf{1}), \\ [\lambda^w]^{k+1} &= [\lambda^w]^k. \end{aligned}$$
 - 5: **else**
 - 6: Compute

$$\begin{aligned} [\lambda^l]^{k+1} &= [\lambda^l]^k, \\ [\lambda^w]^{k+1} &= [\lambda^w]^k + \ln(w) - \ln(B^T([\lambda^l]^k, [\lambda^w]^k) \mathbf{1}). \end{aligned}$$
 - 7: **end if**
 - 8: **end for**
 - 9: **Output:** $x^k = [[\lambda^l]^k, [\lambda^w]^k] \in \mathbb{R}^{2n}$.
-

Отметим, что алгоритмом альтернированной минимизации для задачи (4) является хорошо известный алгоритм Синхорна [Cuturi, 2013].

Также для оптимального решения задачи ЭЛП будем рассматривать ускоренный вариант метода альтернативной минимизации. Согласно [Guminov et al., 2020] в качестве основы ускоренного метода альтернативной минимизации используется традиционный адаптивный ускоренный

градиентный метод. Для данной задачи этот вариант метода оказался быстрее на практике, чем другие способы ускорения. Здесь мы не используем одномерную минимизацию, чтобы найти размер шага, а вместо этого мы адаптируемся к константе Липшица L . Анализ скорости сходимости этого алгоритма можно найти в [Guminov et al., 2020]. В нашем случае ускоренный метод альтернативной минимизации представлен в виде алгоритма 3.

Алгоритм 3. Ускоренный алгоритм Синхорна

1: **Input:** $x^0 := [[x^l]^0, [x^w]^0] = (0, \dots, 0) \in \mathbb{R}^{2n}$ — starting point, $L_0 = 1$, $a_0 = 0$.
 2: **repeat**
 3: Set $y^0 := [[y^l]^0, [y^w]^0] = x^0$.
 4: Set $v^0 := [[v^l]^0, [v^w]^0] = x^0$.
 5: $L_{k+1} = L_k/2$
 6: **while** True **do**
 7: Set $a_{k+1} = \frac{1}{2L_{k+1}} + \sqrt{\frac{1}{4L_{k+1}^2} + a_k^2 \frac{L_k}{L_{k+1}}}$
 8: Set $\tau_k = \frac{1}{a_{k+1}L_{k+1}}$
 9: Set $y^k = \tau_k v^k + (1 - \tau_k)x^k$
 10: Choose $i_k = \operatorname{argmax}_{i \in \{1,2\}} \|\nabla_i \varphi(y^k)\|^2$
 11: **if** $i_k = 1$ **then**
 12: Compute

$$\begin{aligned} [x^l]^{k+1} &= [y^l]^k + \ln(l) - \ln(B([y^l]^k, [y^w]^k)\mathbf{1}), \\ [x^w]^{k+1} &= [y^w]^k. \end{aligned}$$

 13: **else**
 14: Compute

$$\begin{aligned} [x^l]^{k+1} &= [y^l]^k, \\ [x^w]^{k+1} &= [y^w]^k + \ln(w) - \ln(B^T([y^l]^k, [y^w]^k)\mathbf{1}). \end{aligned}$$

 15: **end if**
 16: Set $v^{k+1} = v^k - a_{k+1} \nabla \varphi(y^k)$
 17: **if** $\varphi(x^{k+1}) \leq \varphi(y^k) - \frac{\|\nabla \varphi(y^k)\|^2}{2L_{k+1}}$ **then**
 18: Set $\widehat{d}^{k+1} = \frac{a_{k+1} d^k(y^k) + L_k a_k^2 \widehat{d}^k}{L_{k+1} a_{k+1}^2}$
 19: **break**
 20: **end if**
 21: Set $L_{k+1} = 2L_{k+1}$.
 22: **end while**
 23: **until** $|f(\widehat{d}^{k+1}) + \varphi(x^{k+1})| \leq \varepsilon_f$, $\|\widehat{d}^{k+1} \mathbf{1} - l\|_2 \leq \varepsilon_{eq}$, $\|(\widehat{d}^{k+1})^T \mathbf{1} - w\|_2 \leq \varepsilon_{eq}$
 24: **Выход:** \widehat{d}^{k+1} , x^{k+1} .

При этом вектор градиента функции (4) представляет из себя следующий вектор:

$$\nabla \varphi(\lambda^l, \lambda^w) = \left[\nabla_1 \varphi^T, \nabla_2 \varphi^T \right]^T,$$

где

$$\nabla_1 \varphi(\lambda^l, \lambda^w) = -l + \frac{B(\lambda^l, \lambda^w)\mathbf{1}}{\mathbf{1}^T B(\lambda^l, \lambda^w)\mathbf{1}}, \quad \nabla_2 \varphi(\lambda^l, \lambda^w) = -w + \frac{B^T(\lambda^l, \lambda^w)\mathbf{1}}{\mathbf{1}^T B(\lambda^l, \lambda^w)\mathbf{1}}.$$

4. Задача подсчета невязки для восстановленной матрицы затрат

В данном параграфе опишем задачу подсчета невязки между восстановленной матрицей по затратам и реальной матрицей корреспонденций. Подсчет невязки необходим, чтобы оценить, насколько хорошо выбранная функция затрат описывает реальные данные.

Для постановки задачи подсчета невязки между d_{ij} (исходной матрицей корреспонденций) и $\widehat{d}_{ij}(\alpha)$ (восстановленной матрицей корреспонденций) предположим, что в восстановленной матрице корреспонденций $\widehat{d}_{ij}(\alpha)$ шум является гауссовским (мы восстанавливаем матрицу неточно, с шумом). Тогда восстановленную матрицу можно рассматривать как нормально распределенную выборку $N(\theta, \sigma^2)$ и плотность вероятности нормального распределения можно рассчитать следующим образом:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\theta)^2}{2\sigma^2}},$$

где в качестве матожидания введенного гауссовского распределения будет выступать исходная матрица корреспонденций d_{ij} .

Следовательно, максимизируя правдоподобие, получим:

$$L(\widehat{d}_{ij}(\alpha)) = \prod_{i,j=1}^{n,n} (p(d_{ij}(\alpha))) = \prod_{i,j=1}^{n,n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(d_{ij}-\widehat{d}_{ij}(\alpha))^2/2\sigma^2} \rightarrow \max,$$

$$\log L(\widehat{d}_{ij}(\alpha)) = \sum_{i,j=1}^{n,n} \left\{ -\log \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} \cdot (d_{ij} - \widehat{d}_{ij}(\alpha))^2 \right\} \rightarrow \max,$$

$$\log L(\widehat{d}_{ij}(\alpha)) = -\log \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (d_{ij} - \widehat{d}_{ij}(\alpha))^2 \rightarrow \max.$$

Изменив знак перед выражением, получаем следующую задачу для подсчета невязки:

$$\min_{\alpha} \sum_{i,j=1,1}^{n,n} (d_{ij} - \widehat{d}_{ij}(\alpha))^2, \text{ или, в нормированном случае, } \frac{\min_{\alpha} \sum_{i,j=1,1}^{n,n} (d_{ij} - \widehat{d}_{ij}(\alpha))^2}{n^2}. \quad (5)$$

Отметим, что данная задача является задачей минимизации, которая зависит от параметра α (это может быть вектор параметров, в зависимости от количества параметров в рассматриваемой функции затрат).

Целевая функция полученной задачи является невыпуклой функцией. Для решения данной задачи предлагается использовать безградиентные методы. В частности, в рассматриваемой задаче для поиска оптимального параметра (параметров) α в возникающей при подсчете невязки задаче минимизации используется метод перебора, так как число параметров, в зависимости от функции затрат, — 1–3.

Однако в качестве обзора приведем описание еще нескольких безградиентных методов для задач невыпуклой оптимизации. Рассмотрим метод имитации отжига, для работы которого не требуется гладкость функции [Zhigljavsky, 2008]. Он является вариантом метода случайного поиска и известен как алгоритм Метрополиса. Для задач оптимизации имитация процесса может быть произведена следующим образом. Вводится параметр T , который имеет смысл температуры, и в начальный момент ему устанавливается значение T_0 . Набор переменных, по которым

происходит оптимизация, будет обозначаться как x . В качестве начального состояния системы выбирается произвольная точка. Далее запускается итерационный процесс: на каждом шаге из множества соседних состояний случайно выбирается новое \widehat{x} . Если значение функции в этой точке меньше, чем значение в текущей точке, то эта точка выбирается в качестве нового состояния системы. В ином случае (т. е. если $f(\widehat{x}) > f(x)$) такой переход происходит с вероятностью P , зависящей от температуры T , текущего состояния и кандидата на новое состояние \widehat{x} следующим образом:

$$P = e^{-\frac{f(\widehat{x})-f(x)}{T}}.$$

Также стоит упомянуть метод ломаных, который применим к классу функций одной переменной, удовлетворяющих условию Липшица [Васильев, 2002]. Говорят, что функция $f(x)$ удовлетворяет условию Липшица, если найдется такая константа $L > 0$, что

$$|f(x) - f(y)| \leq L \cdot |x - y| \quad \forall x, y \in [a, b].$$

Пусть функция $f(x)$ удовлетворяет условию Липшица на отрезке $[a, b]$. Зафиксируем какую-либо точку $y \in [a, b]$ и определим функцию $g(x, y) = f(y) - L \cdot |x - y|$ переменной $a \leq x \leq b$. Функция $g(x, y)$ кусочно-линейна на $[a, b]$, и график ее представляет ломаную линию, составленную из отрезков двух прямых, имеющих угловые коэффициенты L и $-L$ и пересекающихся в точке $(y, f(y))$. Также в силу липшицевого условия имеем

$$g(x, y) = f(y) - L \cdot |x - y| \leq f(x, y) \quad \forall x \in [a, b],$$

причем $g(y, y) = f(y)$. Из этого следует, что график функции $f(x)$ лежит выше ломаной $g(x, y)$ при всех $x \in [a, b]$ и имеет с ней общую точку $(y, f(y))$. Данное свойство ломаной $g(x, y)$ можно использовать для построения метода. Этот метод начинается с выбора произвольной точки $x_0 \in [a, b]$ и составления функции $g(x, x_0) = f(x_0) - L \cdot |x - x_0| = p_0(x)$. Следующая точка x_1 определяется из условий $p_0(x_1) = \min_{x \in [a, b]} (p_0(x))$ ($x_1 \in [a, b]$), причем $x_1 = a$ или $x_1 = b$. Далее берется новая функция $p_1(x) = \max(g(x, x_1), p_0(x))$ и очередная точка x_2 находится из условий $p_1(x_2) = \min_{x \in [a, b]} p_1(x)$ ($x_2 \in [a, b]$) и т. д. Пусть точки x_1, \dots, x_n ($n \geq 1$) уже известны. Тогда составляется функция

$$p_n(x) = \max(g(x, x_n), p_{n-1}(x)) = \max_{0 \leq i \leq n} g(x, x_i)$$

и следующая точка x_{n+1} определяется условиями

$$p_n(x_{n+1}) = \min_{x \in [a, b]} p_n(x), \quad x_{n+1} \in [a, b].$$

Если минимум $p_n(x)$ достигается в нескольких точках, то в качестве x_{n+1} можно взять любую из них. Таким образом, метод ломаных описан.

Также следует упомянуть алгоритм случайного мултистарта [Zhigljavsky, 2008]. Случайный мултистарт — это метод глобальной оптимизации, состоящий в многократном отыскании локальных минимумов из различных начальных точек. В своем первоначальном виде он неэффективен, однако некоторые из его модификаций могут быть полезны. Основная сложность при практической реализации метода состоит в следующем: для того чтобы с высокой надежностью отыскать точку глобального минимума, необходимо взять количество начальных точек для локальных алгоритмов существенно больше, чем число локальных минимумов функции, которое обычно неизвестно.

5. Восстановление матрицы корреспонденций для г. Москвы

В данном разделе приведены численные эксперименты для расчета матрицы корреспонденций для города Москвы с использованием алгоритмов, описанных в предыдущем разделе. Данные эксперименты были проведены на основе данных, собранных в результате опроса 2013 года по Москве и Московской области. Данные представлены csv файлом с пятью полями:

- 1) зона i ;
- 2) зона j ;
- 3) число жителей i , которые ездят на работу в j ;
- 4) среднее время, затраченное на поездку, в минутах;
- 5) среднее расстояние по прямой между домом и работой (в Москве точки отправления или прибытия определяются с точностью до ближайшего метро, в Зеленограде — до центра района, в области — до центра населенного пункта).

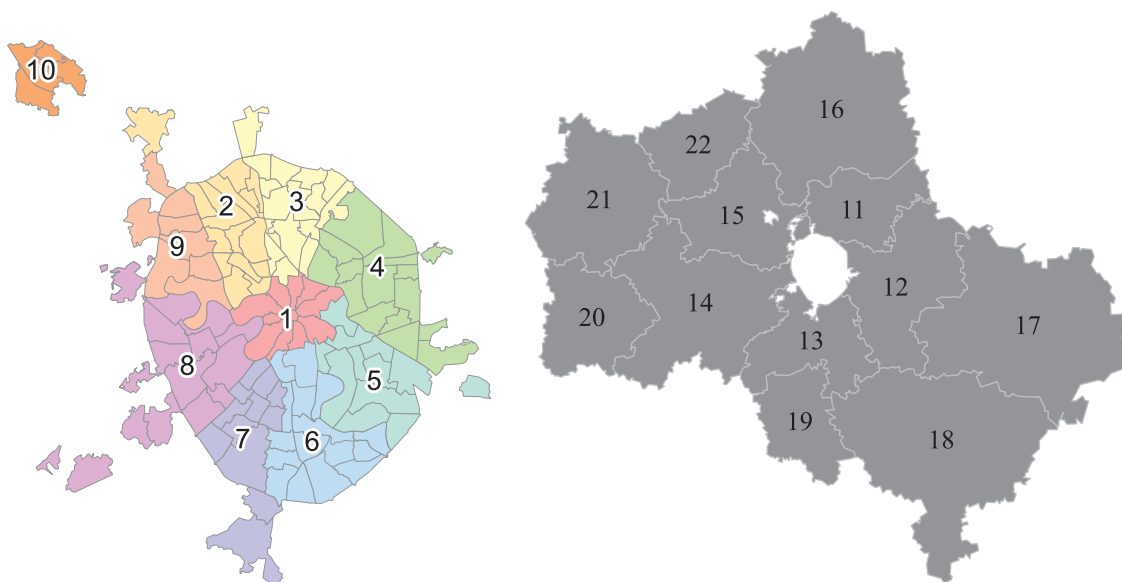


Рис. 1. Разбиение г. Москвы и МО по районам

Отметим, что зоны выбраны достаточно большими, чтобы уменьшить случайные ошибки. В Москве (в старых границах) зоны соответствуют округам, в области — нескольким средним районам. Всего есть 22 района, которые одновременно являются и источниками (место начала поездки), и стоками (место окончания поездки), однако не между всеми пунктами i – j есть корреспонденции. Суммарное число участников движения — 1965.

Далее приводится описание полученных результатов для разных функций затрат.

5.1. Линейная функция затрат

Для начала в качестве функции затрат рассмотрим линейную функцию от затрат, т. е.

$$T_{ij}(\alpha) = \alpha c_{ij},$$

где α — это калибруемый параметр.

Очевидно, что чем выше альфа, тем сильнее влияют затраты на проезд по пути между источником и стоком на соответствующее значение корреспонденций.

Рассмотрим следующие три варианта затрат.

- **Затраты — среднее время в пути.**

Рассмотрим простейшую модель, где c_{ij} — среднее время проезда от района i до района j . В рамках этой модели функция затрат имеет следующий вид:

$$T_{ij}(\alpha) = \alpha \cdot \text{time}_{ij}.$$

В данной функции в качестве калибруемого параметра выступает α . Подбор параметра происходил путем перебора по сетке $\alpha \in [0.01, 1]$ с шагом $1e-3$. График зависимости невязки (5) от параметра α представлен на рис. 2.

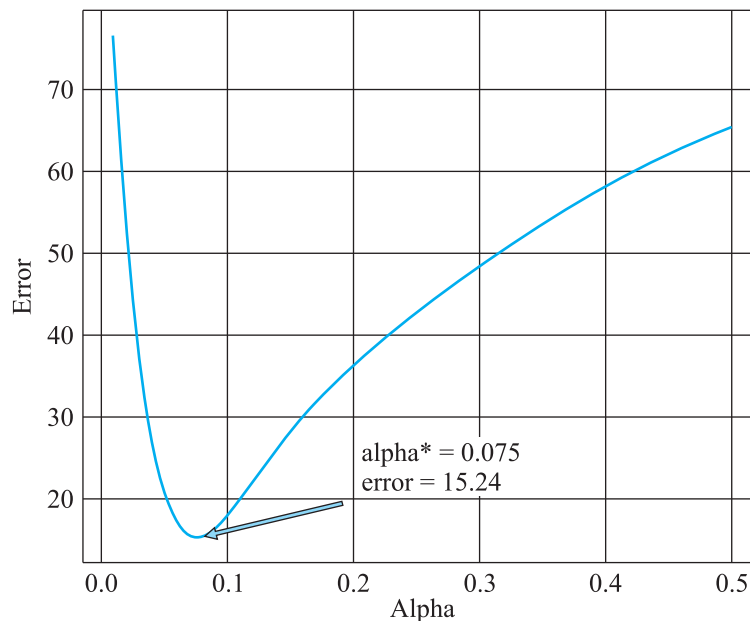


Рис. 2. Значение невязки (5) при функции затрат $T_{ij}(\alpha) = \alpha \cdot \text{time}_{ij}$ в зависимости от α

Также в результате перебора по сетке было найдено оптимальное значение $\alpha^* = 0.076$, невязка при данном значении параметра равна 15.24.

- **Затраты — среднее время в пути (степенная функция затрат).**

Рассмотрим усложненную двухпараметрическую модель, где c_{ij} — среднее время проезда от района i до района j в степени γ . В рамках этой модели функция затрат имеет следующий вид:

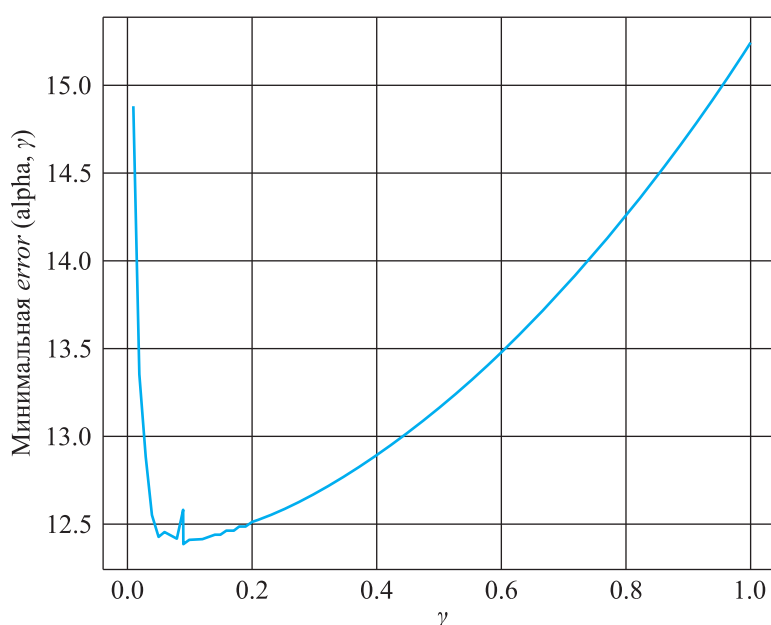
$$T_{ij}(\alpha, \gamma) = \alpha \cdot \text{time}_{ij}^{\gamma}. \quad (6)$$

В данной функции в качестве калибруемого параметра выступают два параметра: α и γ . Подбор параметров происходил путем перебора по сетке $\gamma \in [0.01, 1]$ с шагом $1e-2$ и при динамическом определении области перебора параметра α с целью определения окрестности, в которой достигается минимум невязки. Параметр γ определяет ширину углубления на рис. 2. В таблице 1 приведены лучшие комбинации параметров, а на рис. 3 — зависимость минимального значения невязки от γ .

Из графика хорошо видно, что введение степени γ матрицы временных затрат имеет смысл для минимизации невязки. Так, оптимальные значения $\gamma^* = 0.09$ и $\alpha^* = 26.760$ дают невязку, равную 12.385.

Таблица 1. Сравнение невязок для разных функций затрат

γ	α	Невязка
0.09	26.760	12.38466
0.1	23.770	12.40603
0.11	20.095	12.40931
0.12	18.290	12.41265
0.08	31.410	12.41749
0.05	56.250	12.42543
0.13	15.785	12.42546
0.07	37.230	12.43197
0.14	14.570	12.43495
0.15	12.750	12.43975

Рис. 3. Оптимальное значение невязки в зависимости от γ при функции затрат $T_{ij}(\alpha, \gamma) = \alpha \cdot \text{time}_{ij}^\gamma$

- **Затраты — комбинация времени и расстояния.**

Введем теперь также дополнительную зависимость от расстояния в затраты. Тогда функция затрат будет иметь следующий вид:

$$T_{ij}(\alpha, \beta, \gamma) = \alpha \cdot \text{time}_{ij}^\gamma \cdot \text{dist}_{ij}^\beta, \quad (7)$$

где dist_{ij} — расстояние между районами i и j , time_{ij} — среднее время в пути между районами i и j . В данной функции в качестве калибруемого параметра выступают три параметра: α , β и γ . Для проверки целесообразности добавления множителя dist_{ij}^β будем перебирать параметр $\beta \in [0, 0.5]$ с шагом 0.001. На рис. 4 показано изменение невязки в зависимости от β для оптимальных параметров при функции затрат (7).

Видим, что добавление множителя dist_{ij}^β в (6) имеет смысл, так как минимальная невязка достигается при $\beta = 0.005$. Сравним значения оптимальных невязок для (6) и (7) (рис. 5).

Оптимальное значение невязки достигается при $\alpha^* = 26.76$, $\gamma^* = 0.09$, $\beta^* = 0.005$ и равно 10.41226.

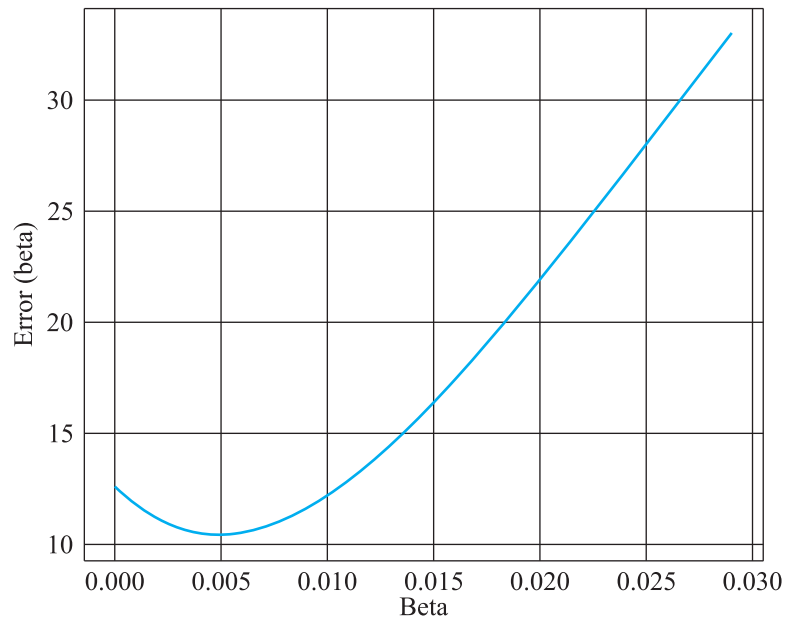


Рис. 4. Изменение невязки от β при функции затрат $T_{ij}(\alpha, \beta, \gamma) = \alpha \cdot \text{time}_{ij}^\gamma \cdot \text{dist}_{ij}^\beta$, где $\alpha^* = 26.76$ и $\gamma^* = 0.09$

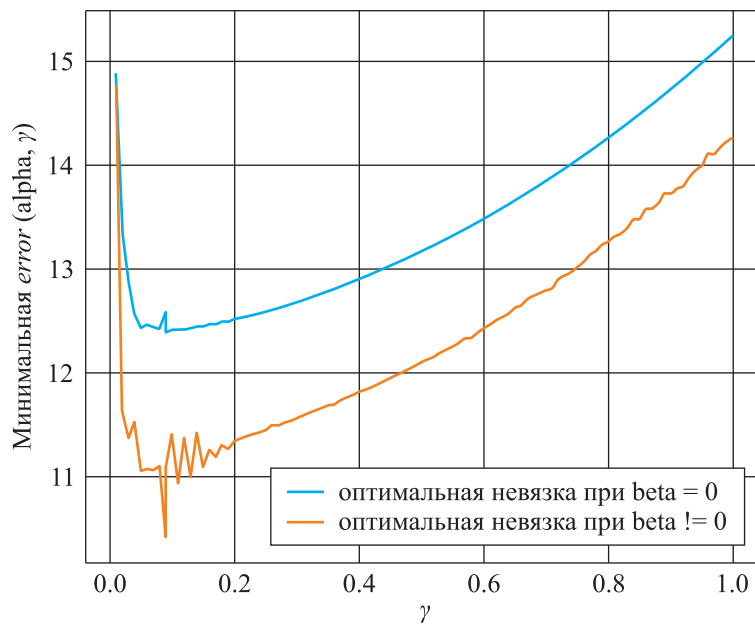


Рис. 5. Сравнение оптимальных невязок при функциях затрат $T_{ij}(\alpha, \beta, \gamma) = \alpha \cdot \text{time}_{ij}^\gamma \cdot \text{dist}_{ij}^\beta$ и $T_{ij}(\alpha, \gamma) = \alpha \cdot \text{time}_{ij}^\gamma$ в зависимости от γ

5.2. Сумма степенной и логарифмической функций затрат

Следуя подходу из [Гасников и др., 2013], рассмотрим модель, в которой функция затрат является суммой степенной и логарифмической, т. е.

$$T_{ij}(\alpha, \beta, \gamma) = \alpha c_{ij}^\gamma - \beta \ln c_{ij}. \quad (8)$$

В качестве обоснования использования данной функции можно привести следующее рассуждение: первое слагаемое отражает нежелательность больших затрат на дорогу, а второе от-

ражает возможность найти работу на расстоянии (среднем времени в пути) порядка c_{ij} от дома. Для данной функции в качестве затрат можно рассматривать затраты, равные среднему времени в пути, или затраты, равные расстоянию между районами. Данная модель зависит от трех параметров. В случае затрат, равных среднему времени в пути, функция затрат будет иметь следующий вид:

$$T_{ij}(\alpha, \beta, \gamma) = \alpha \text{time}_{ij}^{\gamma} - \beta \ln \text{time}_{ij}. \quad (9)$$

По результатам проведенных экспериментов можно заключить, что оптимальные значения невязки достигаются при $\beta = 0$, т.е. когда функция затрат (9) вырождается в (6). На рис. 6 показано, как ухудшается невязка при росте β при оптимальных (α, γ) .

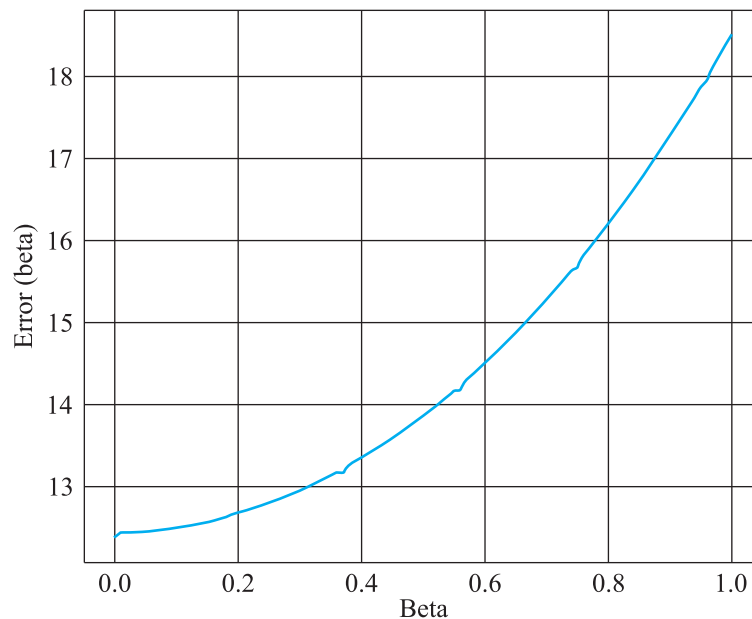


Рис. 6. Невязка в зависимости от β при функции затрат $T_{ij}(\alpha, \gamma, \beta) = \alpha \cdot \text{time}_{ij}^{\gamma} - \beta \ln \text{time}_{ij}$ при оптимальных (α, γ)

Если в (8) $c_{ij} = \text{dist}_{ij}$, то получим функцию затрат от среднего расстояния в пути:

$$T_{ij}(\alpha, \beta, \gamma) = \alpha \text{dist}_{ij}^{\gamma} - \beta \ln \text{dist}_{ij}. \quad (10)$$

Перебор по параметрам $\alpha \in [0.01, 10]$, $\beta \in [0, 0.5]$, $\gamma \in [0, 1]$ выдал оптимальную точку ($\alpha^* = 3.01$, $\beta^* = 0$, $\gamma^* = 0.25$) с невязкой 4.84729.

5.3. Сравнение функций затрат

В данном разделе приведем сравнение результатов, полученных выше. Обозначим t_{ij} как среднее время в пути между районами i и j , а d_{ij} — как среднее расстояние между районами i и j . В таблице 2 приведены минимальные значения невязок для различных функций затрат и соответствующие этим значениям невязок оптимальные параметры.

По итогам проведенных экспериментов заключаем, что оптимальной функцией затрат для восстановления матрицы корреспонденций по имеющимся данным является $T_{ij} = 3.01 \cdot d_{ij}^{0.25}$. Заметим, что функция затрат не зависит от времени. Это может говорить как о недостаточности объема временных данных, все-таки для измерения средних значений по времени требуется больше наблюдений, так и о необходимости продолжить исследования по подбору иных функций затрат.

Таблица 2. Сравнение невязок для разных функций затрат

Функция затрат $T_{ij}(\alpha, \beta, \gamma)$	Невязка	Оптимальные параметры
$\alpha \cdot t_{ij}$	15.24	$\alpha = 0.076$
$\alpha \cdot t_{ij}^\gamma$	12.38	$\alpha = 26.76, \gamma = 0.09$
$\alpha \cdot t_{ij}^\gamma \cdot d_{ij}^\beta$	10.41	$\alpha = 26.76, \beta = 0.005, \gamma = 0.09$
$\alpha t_{ij}^\gamma - \beta \ln t_{ij}$	12.38	$\alpha = 26.76, \beta = 0, \gamma = 0.09$
$\alpha d_{ij}^\gamma - \beta \ln d_{ij}$	4.84	$\alpha = 3.01, \beta = 0, \gamma = 0.25$

6. Заключение

В статье была предложена и научно обоснована технология калибровки моделей расчета матрицы корреспонденций. В отличие от известных ранее (однопараметрических) моделей расчета матрицы корреспонденций в статье рассматривались многопараметрические модели. Также отметим важное отличие предлагаемых исследований от известных ранее. Проводимые ранее исследования в данном направлении, как правило, не обосновывались строго. Более того, в подавляющем числе случаев это считалось простой задачей, и про это не особо подробно писали в соответствующих работах. Однако уже для случая двух параметров задача становится существенно более затратной, и от выбора способа ее численного решения существенно зависит эффективность подхода. Собственно, вычислительным аспектам в данной статье и было уделено наибольшее внимание. Отметим, что проводимые в данной работе исследования (точнее, их теоретическое обоснование) стали возможны благодаря современным достижениям в области математического моделирования транспортных потоков (появился новый цикл многопараметрических моделей расчета матрицы корреспонденций, обобщающих энтропийную модель) и тонкому анализу скорости сходимости алгоритма Синхорна, использующегося для расчета матрицы корреспонденций (напомним, что модель расчета матрицы корреспонденций сводится к системе $2n$ нелинейных уравнений, для решения которой и используется алгоритм Синхорна).

Задача была предложена Е. А. Нурминским и А. В. Гасниковым.

Список литературы (References)

- Боровков А. А. Эргодичность и устойчивость случайных процессов. — М.: УРСС, 1999.
Borovkov A. A. Ergodicity and Stability of Stochastic Processes. — Wiley series in Probability and Statistics, 1998. (Russ. ed.: *Borovkov A. A. Ergodichnost' i ustoichivost' sluchainykh protsessov.* — Moscow: URSS, 1999.)
- Брегман Л. Д. Доказательство сходимости метода Шелейховского для задачи с транспортными ограничениями // ЖВМ и МФ. — 1967. — Т. 7, № 1. — С. 147–156.
Bregman L. D. Dokazatel'stvo skhodimosti metoda Sheleikhovskogo dlya zadachi s transportnymi ogranicheniyami [Proof of the convergence of the Sheleikhovsky method for a problem with transport constraints] // *Computational Mathematics and Mathematical Physics.* — 1967. — Vol. 7, No. 1. — P. 147–156 (in Russian).
- Вайдлих В. Социодинамика: системный подход к математическому моделированию в социальных науках. — М.: УРСС, 2010.
Weidlich W. Sociodynamics: a System Approach to Mathematical Modelling in the Social Sciences. — Amsterdam: Harwood Academic Publishers, 2000. (Russ. ed.: *Weidlich W. Sotsiodinamika: sistemnyi podkhod k matematicheskomu modelirovaniyu v sotsial'nykh naukakh.* — Moscow: URSS, 2010.)
- Васильев Ф. П. Методы оптимизации. — М.: Факториал Пресс, 2002.
Vasiliev F. P. Metody optimizatsii [Methods of Optimization]. — Moscow: Faktorial Press, 2002 (in Russian).
- Вильсон А. Дж. Энтропийные методы моделирования сложных систем. — М.: Наука, 1978.
Wilson A. G. Entropy in urban and regional modeling. — Routledge, 2011. (Russ. ed.: *Wilson A. G. Entropiinye metody modelirovaniya slozhnykh sistem.* — Moscow: Nauka, 1978.)

- Гардинер К. В.* Стохастические методы в естественных науках. — М.: Мир, 1986.
Gardiner C. Stochastic methods. A Handbook for the Natural and Social Sciences. — Springer, 2009. (Russ. ed.: *Gardiner C.* Stokhasticheskie metody v estestvennykh naukah. — Moscow: Mir, 1986.)
- Гасников А. В., Гасникова Е. В.* Об энтропийно-подобных функционалах, возникающих в стохастической химической кинетике при концентрации инвариантной меры и в качестве функций Ляпунова динамики квазисредних // Математические заметки. — 2013. — Т. 94, № 6. — С. 819–827.
Gasnikov A. V., Gasnikova E. V. On Entropy-Type Functionals Arising in Stochastic Chemical Kinetics Related to the Concentration of the Invariant Measure and Playing the Role of Lyapunov Functions in the Dynamics of Quasiaverages // Mathematical Notes. — 2013. — Vol. 94, No. 2–14. — P. 819–827.
- Гасников А. В., Гасникова Е. В., Мендель М. А., Чепурченко К. В.* Эволюционные выводы энтропийной модели расчета матрицы корреспонденций // Математическое моделирование. — 2016. — Т. 28, № 4. — С. 111–124.
Gasnikov A. V., Gasnikova E. V., Mendel' M. A., Chepurchenko K. V. Evolyutsionnye vyvody entropiinoi modeli rascheta matritsy korrespondentsii [Evolutionary interpretations of entropy model for correspondence matrix calculation] // Mathematical Models and Computer Simulations. — 2016. — Vol. 28, No. 4. — P. 111–124 (in Russian).
- Гасников А. В., Гасникова Е. В., Федько О. С.* О возможной динамике в модели ранжирования web-страниц PageRank и модернизированной модели расчета матрицы корреспонденций // Труды МФТИ. — 2012. — Т. 4, № 2–14.
Gasnikov A. V., Gasnikova E. V., Fed'ko O. S. O vozmozhnoi dinamike v modeli ranzhirovaniya web-stranits PageRank i modernizirovannoi modeli rascheta matritsy korrespondentsii [Possible dynamics in the PageRank model of ranking web pages and the modernized model for calculating the correspondence matrix] // Trudy MFTI. — 2012. — Vol. 4, No. 2–14 (in Russian).
- Гасников А. В., Кленов С. Л., Нурминский Е. А., Холодов Я. А., Шамрай Н. Б.* Введение в математическое моделирование транспортных потоков / под ред. А. В. Гасникова, с приложениями М. Л. Бланка, К. В. Воронцова и Ю. В. Чеховича, Е. В. Гасниковой, А. А. Замятина и В. А. Малышева, А. В. Колесникова, Ю. Е. Нестерова и С. В. Шпирко, А. М. Райгородского, с предисловием руководителя Департамента транспорта г. Москвы М. С. Ликсутова. — М.: МЦНМО, 2013. — 427 с., 2-е изд.
Gasnikov A. V., Klenov S. L., Nurminsky E. A., Kholodov Ya. A., Shamray N. B. Vvedenie v matematicheskoe modelirovanie transportnykh potokov [Introduction to the mathematical modeling of traffic flows] // eds. A. V. Gasnikov with applications by M. L. Blank, K. V. Vorontsov and Yu. V. Chekhovich, E. V. Gasnikova, A. A. Zamyatin and V. A. Malyshev, A. V. Kolesnikov, Yu. E. Nesterov and S. V. Shpirko, A. M. Raigorodsky, with a foreword by the head of the Moscow Department of Transport M. S. Liksutov. — Moscow: MCCME, 2013. — 427 p., 2nd ed. (in Russian).
- Малышев В. А., Пирогов С. А.* Обратимость и необратимость в стохастической химической кинетике // Успехи математических наук. — 2008. — Т. 63, № 1 (379). — С. 3–36.
Malyshev V. A., Pirogov S. A. Reversibility and irreversibility in stochastic chemical kinetics // Russian Mathematical Surveys. — 2008. — Vol. 63, No. 1 (379). — P. 3–36.
- Санов И. Н.* О вероятности больших отклонений случайных величин // Математический сборник. — 1957. — Т. 42, № 1. — С. 11–44.
Sanov I. N. O veroyatnosti bol'shikh otklonenii sluchainykh velichin [On the probability of large deviations of random magnitudes] // Sbornik: Mathematics. — 1957. — Vol. 42, No. 1. — P. 11–44 (in Russian).
- Шелейховский Г. В.* Транспортные основания композиции городского плана. — Л.: Гипрогор, 1936.
Sheleikhovsky G. V. Transportnye osnovaniya kompozitsii gorodskogo plana [Transport bases of the composition of the city plan]. — Leningrad: Giprogor, 1936 (in Russian).
- Bogolubsky L., Dvurechensky P., Gasnikov A., Gusev G., Nesterov Y., Raigorodskii A. M., Tikhonov A., Zhukovskii M.* Learning supervised pagerank with gradient-based and gradient-free optimization methods // Advances in neural information processing systems. — 2016. — P. 4914–4922.
- Cuturi M.* Sinkhorn distances: Lightspeed computation of optimal transport // Advances in neural information processing systems. — 2013. — P. 2292–2300.
- Dvurechensky P. E., Gasnikov A. V., Kroshnin A. V.* Computational Optimal Transport: Complexity by Accelerated Gradient Descent Is Better Than by Sinkhorn's Algorithm // International Conference on Machine Learning. — 2018. — P. 1367–1376.

- Ethier N. S., Kurtz T. G.* Markov processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. — John Wiley and Sons Inc., New York, 1986.
- Guminov S., Dvurechensky P., Tupitsa N., Gasnikov. A.* Accelerated alternating minimization, accelerated Sinkhorn's algorithm and accelerated iterative Bregman projections // arXiv preprint arXiv:1906.03622. — 2020.
- Garber D., Hazan E.* A Linearly Convergent Conditional Gradient Algorithm with Applications to Online and Stochastic Optimization // SIAM Journal on Optimization. — 2013. — Vol. 26.
- Levin D. A., Peres Y., Wilmer E. L.* Markov chain and mixing times // American Mathematical Soc., Providence. — 2009.
- Sandholm W.* Population games and evolutionary dynamics // MIT press, 2010.
- Zhigljavsky A., Zilinskas A.* Stochastic global optimization // Springer Science & Business Media, 2008.