УДК 51.77

Т. С. Бабичева[1,2], *А. В. Гасников*[2,3], *А. А. Лагуновская*[1,2,3], *М. А. Мендель*[2]

[1]Институт прикладной математики им. М. В. Келдыша РАН
[2]Московский физико-технический институт (государственный университет)
[3]Центр исследований транспортной политики Института экономики транспорта и транспортной политики, Национальный исследовательский университет «Высшая школа экономики», Москва

Двухстадийная модель равновесного распределения транспортных потоков

В работе описываются некоторые непроработанные ранее детали в многостадийном подходе к транспортному моделированию. В основе подхода лежит теорема о потенциальности специальной популяционной игры, возникающей при объединении модели равновесного распределения потоков по путям и модели формирования корреспонденций.

Ключевые слова: равновесие Нэша, модель расчета матрицы корреспонденций, модель равновесного распределения потоков по путям, многостадийная модель.

T. S. Babicheva[1,2], *A. V. Gasnikov*[2,3], *A. A. Lagunovskaya*[1,2,3], *M. A. Mendel*[2]

[1]Keldysh Institute of Applied Mathematics
[2]Moscow Institute of Physics and Technology (State University)
[3] Research Centre for Transport Policy Studies, Institute for Transport Economics and Transport Policy Studies, National Research University «Higher School of Economics», Moscow

Two-stage model of equilibrium distributions of traffic flows

This paper describes some previously unexamined features in a multistage approach to transport modelling. The approach described is based on a theorem on the potentiality of the special population game that arises while the model of equilibrium flow distribution over paths and the model of correspondence formation are combined.

Key words: Nash equilibrium, model of correspondence matrix calculation, model of equilibrium flow distribution over paths, multistage model.kwe Nash equilibrium, model of correspondence matrix calculation, equilibrium model of distributed flows, multistage model.

1. Введение

В последние годы наблюдается резкое увеличение доступной транспортной информации, на которую, скажем, 15 лет назад нельзя было рассчитывать. Прежде всего речь идет о данных сотовых операторов и данных GPS / ГЛОНАСС треков автомобилей. Уже имеется положительный опыт использования этой информации в транспортных приложениях. К примеру, данные сотовых операторов активно использует сейчас Генплан г. Москвы для оценивания перемещений (мобильности) населения г. Москвы и области. По-видимому, в ближайшее время этого можно ожидать и от РЖД [1].

Данные GPS и ГЛОНАСС треков автомобилей и общественного транспорта уже достаточно давно использует сервис Яндекс.Пробки. Важной и актуальной сейчас задачей является разработка необходимого математического обеспечения, позволяющего максимально использовать эту (достаточно дешевую) информацию огромных объемов и вставлять её в модели.[1] Это порождает, в свою очередь, задачу разработки моделей, «соизмеримых» этим новым данным.

[1]Можно также ожидать в перспективе расчет исходя из трековой информации транспортных налогов за пользование теми или иными дорогами. Это представляется намного удобнее, чем турникеты на въезде (сейчас в большинстве случаев реализуется именно сценарий с турникетами). Можно показать, что существует (Сэндхольм, [2]) такой способ взимания плат за проезд по дорогам, чтобы равновесие в транспортной сети (Нэша–Вардропа) соответствовало социальному оптимуму.

Авторам не раз приходилось сталкиваться с философией, преобладающей до недавнего времени среди московских чиновников: «зачем разрабатывать что-то свое, ведь наверняка на Западе уже все давно придумали, надо просто найти, что подходит, купить и использовать». Сейчас уже однозначно можно сказать, что такая стратегия совершенно не работает.

Связано это с тем, что большинство существующих пакетов транспортного моделирования разрабатывалось еще в прошлом веке, т.е. тогда, когда приходилось ориентироваться в основном на данные достаточно дорогостоящих опросов и на данные датчиков, установленных на дорогах. Другими словами, эти пакеты не подстроены под данные, на которые сейчас переходит транспортное моделирование.

Имеется и много вопросов к тому, какие модели и алгоритмы используются в этих (и других) программных продуктах. Например, неочевидным элементом почти всех этих продуктов является использование в качестве одного из блоков модели равновесного распределения транспортных потоков по путям Бэкмана [3]. Эта во многом хорошая модель, тем не менее, имеет довольно много недостатков (Нестеров–де Пальма [4]). Например, калибровка такой модели требует знания функций затрат на ребрах графа транспортной сети (эти функции связывают время в пути по ребру с величиной транспортного потока по этому ребру), причем сама модель оказывается довольно чувствительной к выбору этих функций, которые в модели Бэкмана, как правило, предполагаются монотонно возрастающими. Достаточно сказать, что в случае наличия платных дорог для расчета оптимальных плат за проезд требуется вычислять, например, производные этих неизвестных функций. Существование таких функций в модели Бэкмана является одним из основных предположений и одновременно одним из самых слабых мест. Реальные данные показывают, что предположение о классе функций затрат не выполняется. Но даже если предположить, что такая зависимость все же существует, то по-прежнему остается другая проблема: как калибровать модель Бэкмана, то есть откуда брать эти зависимости. Не получится ли переобучения у создаваемой нами модели?

Другим неочевидным элементом этих программных продуктов являются используемые вычислительные алгоритмы: контроль их робастности к неточности (неполноте) данных, ошибкам округления (поскольку возникают задачи огромных размеров, то такие ошибки могут интенсивно накапливаться).

Наконец, сама философия, использующаяся в таких продуктах при построении равновесной модели города, также вызывает много вопросов. Пожалуй, самым слабым местом тут является хорошо известный подход многостадийного моделирования [5], который прогоняет отдельные блоки формирования загрузки транспортной сети и ищет неподвижную точку. В этом подходе, преобладающем сейчас повсеместно, никак не обосновывается существование неподвижной точки, не приводятся какие бы то ни было гарантии сходимости такого итерационного процесса (последовательной прогонки этих блоков с замыканием первого на последний / иногда предпоследний).

Таким образом, актуальной задачей, решению которой и посвящена настоящая работа, является разработка нового подхода (со строгим математическим обоснованием) к созданию комплексных интеллектуальных транспортных систем (КИТС) крупных городов. Ожидается, что КИТС позволит решать исходя из имеющихся данных нового типа, например, задачи долгосрочного планирования (развития) транспортной инфраструктуры города. В частности, помогать отвечать на вопросы: какой из проектов дорожного строительства оптимален, где пропуская способность дороги недостаточна, как изменится транспортная ситуация, если построить в этом месте торговый центр (жилой район, стадион), как оптимально взимать платы за проезд по платным дорогам и где стоит делать дороги платными, как правильно определять маршруты и расписание движения общественного транспорта, какой эффект даст выделение полос для общественного транспорта и т.п.

В дальнейшем изложении мы базируемся на работах [6–16]. В частности, приводим здесь доказательство теоремы 2, неявно использующейся в работах [7,8] и сформулированной (без доказательства) в работе [10]. Эта теорема представляет, на наш взгляд, основной элемент построения оригинальной многостадийной модели. Для наглядности изложения в данной работе мы ограничимся только двумя блоками — блоком расчета матрицы корреспонденций и блоком равновесного распределения потоков по путям.

2. Модель равновесного распределения потоков по путям

Далее мы опишем основную конструкцию (в немного упрощенном виде), на базе которой планируется строить многостадийную (в нашем случае двухстадийную) модель. Предлагаемый подход очень сильно отличается от всех существующих сейчас подходов к разработке многостадийной модели.

Далее рассматривается пример (более общий, чем транспортные потоки на сети) сетевого рынка одного товара (все, что далее написано, можно распространить на еще более общие постановки). Сердцевина подхода заключена в формулах (3), (8) и теореме 2.

Имеется ориентированный транспортный граф, каждое ребро которого характеризуется неубывающей функцией затрат $\tau_e(f_e)$ на прохождение этого ребра в зависимости от потока по этому ребру. Можно еще ввести затраты на прохождения вершин графа E, но это ничего не добавляет с точки зрения последующих математических выкладок. Часть вершин графа является источниками, часть – стоками (эти множества вершин могут пересекаться). В источниках O и стоках D имеются (соответственно) пункты производства и пункты потребления.

Задача поиска равновесия разбивается на две подзадачи разного уровня. На нижнем уровне, соответствующем быстрому времени, при заданных корреспонденциях $\{d_{ij}\}$ (сколько товара перевозится в единицу времени из источника i в сток j) идет равновесное формирование способов транспортировки товаров. В результате формируются функции затрат $T_{ij}(\{d_{ij}\})$. Исходя из этих затрат на верхнем уровне, соответствующем медленному времени, решается задача поиска конкурентного равновесия между производителями и потребителями с учетом затрат на транспортировку. Мы будем иметь дело с адиабатическим исключением быстрых переменных (принцип подчинения Г. Хакена) в случае стохастических динамик [8, 10].

Обозначим множество пар $w = (i, j)$ источник–сток OD, x_p — поток по пути p; P_w – множество путей, отвечающих корреспонденции w, $P = \bigcup\limits_{w\in OD} P_w$ — множество всех путей; $f_e(x) = \sum\limits_{p\in P} \delta_{ep}x_p$ – поток на ребре e (здесь и далее $x = \{x_p\}$, $f = \Theta x$), где $\delta_{ep} = \begin{cases} 1, & e \in p; \\ 0, & e \notin p; \end{cases}$ $\tau_e(f_e)$ — затраты на ребре e ($\tau_e'(f_e) \geqslant 0$); $G_p(x) = \sum\limits_{e\in E} \tau_e(f_e(x))\,\delta_{ep}$ — затраты на пути p; $X = \left\{x \geqslant 0 : \sum\limits_{p\in P_w} x_p = d_w,\ w \in OD\right\}$ — множество допустимых распределений потоков по путям, где d_w — корреспонденция, отвечающая паре w. Далее (теорема 1) мы приводим классические результаты, описанные в огромном числе литературных источников (см., например, [6] и цитированную там литературу).

Определение 1. Распределение потоков по путям $x = \{x_p\} \in X$ называется равновесием (Нэша–Вардропа) в популяционной игре $\langle\{x_p\} \in X, \{G_p(x)\}\rangle$, если из $x_p > 0 (p \in P_w)$ следует $G_p(x) = \min\limits_{q\in P_w} G_q(x)$. Или, что то же самое:

$$\text{для любых } w \in OD,\ \ p \in P_w \text{ выполняется } x_p \cdot \left(G_p(x) - \min_{q\in P_w} G_q(x)\right) = 0.$$

Теорема 1. *Популяционная игра $\langle\{x_p\} \in X, \{G_p(x)\}\rangle$ является потенциальной. Равновесие x^* в этой игре всегда существует и находится из решения задачи выпуклой оптимизации*

$$x^* \in \operatorname{Arg}\min_{x\in X} \Psi(f(x)), \tag{1}$$

где

$$\Psi(f(x)) = \sum_{e \in E} \int\limits_0^{f_e(x)} \tau_e(z)\,dz = \sum_{e \in E} \sigma_e(f_e(x)).$$

Мы оставляем в стороне вопрос единственности равновесия. Отметим лишь, что при естественных условиях равновесное распределение потоков по ребрам f^* единственно. В частности, для этого достаточно, чтобы $\tau_e'(f_e) > 0$ для всех $e \in E$. Если дополнительно $f^* = \Theta x$ однозначно разрешимо относительно x (в реальных транспортных сетях часто случается, что число допустимых для перевозки путей меньше числа ребер, это как раз и приводит к однозначной разрешимости), то отсюда будет следовать, что равновесное распределение потоков по путям x^* единственно.

Удобно считать, что возрастающие функции затрат $\tau_e(f_e) := \tau_e^\mu(f_e)$ зависят от параметра $\mu > 0$, причем

$$\tau_e^\mu(f_e) \underset{\mu \to 0+}{\longrightarrow} \begin{cases} \bar{t}_e, & 0 \leqslant f_e < \bar{f}_e; \\ [\bar{t}_e, \infty), & f_e = \bar{f}_e. \end{cases}$$

В таком пределе (в пределе модели стабильной динамики [4, 7, 8, 10]) задачу выпуклой оптимизации 1 можно переписать как задачу ЛП:

$$\min_{\substack{f = \Theta x,\ x \in X \\ f \leqslant \bar{f}}} \sum_{e \in E} f_e \bar{t}_e.$$

Такого типа транспортные задачи достаточно хорошо изучены.

Для дальнейшего будет важно переписать задачу $\min\limits_{x \in X} \Psi(f(x))$ через двойственную:

$$\min_{x \in X} \Psi(f(x)) = \min_{f,x} \left\{ \sum_{e \in E} \sigma_e(f_e) : \quad f = \Theta x, \quad x \in X \right\} =$$

$$= \min_{f,x} \left\{ \sum_{e \in E} \max_{t_e \in \operatorname{dom} \sigma_e^*} [f_e t_e - \sigma_e^*(t_e)] : \quad f = \Theta x, \quad x \in X \right\} =$$

$$= \max_{t \in \operatorname{dom} \sigma^*} \left\{ \min_{f,x} \left[\sum_{e \in E} f_e t_e : \quad f = \Theta x, \quad x \in X \right] - \sum_{e \in E} \sigma_e^*(t_e) \right\} =$$

$$= \max_{t \geqslant \bar{t}} \left\{ \sum_{w \in OD} d_w T_w(t) - \left\langle \bar{f}, t - \bar{t} \right\rangle - \mu \sum_{e \in E} h\left(t_e - \bar{t}_e, \bar{t}_e, \bar{f}_e, \mu\right) \right\} =$$

$$\overset{\mu \to 0+}{=} \max_{t \geqslant \bar{t}} \left\{ \sum_{w \in OD} d_w T_w(t) - \left\langle \bar{f}, t - \bar{t} \right\rangle \right\}, \tag{2}$$

где $\sigma_e^*(t_e)$ — сопряженная функция к $\sigma_e(f_e)$, $T_w(t) = \min\limits_{p \in P_w} \sum\limits_{e \in E} \delta_{ep} t_e$ — длина кратчайшего пути из i в j ($w = (i,j)$) на графе, взвешенном согласно вектору t, $h\left(t_e - \bar{t}_e, \bar{t}_e, \bar{f}_e, \mu\right)$ — сильно выпуклая функция по первому аргументу. При этом

$$\tau_e^\mu(f_e(x(\mu))) \underset{\mu \to 0+}{\longrightarrow} t_e,$$

где $x(\mu)$ – равновесное распределение потоков, рассчитывающееся по формуле (1), а $t = \{t_e\}$ — решение задачи (2), при естественных условиях единственное. Описанный предельный переход позволяет переходить к задачам, в которых вместо функции затрат на ребрах $\tau_e(f_e)$ заданы ограничения на пропускные способности $0 \leqslant f_e \leqslant \bar{f}_e$ и затраты $\bar{t}_e$ на прохождения ребер, когда на ребрах нет «пробок» ($f_e < \bar{f}_e$).

Основным для дальнейшего выводом из этого всего является способ (основанный на применении теоремы Демьянова–Данскина–Рубинова, как правило, в градиентном варианте ввиду единственности t) потенциального описания набора $T(d) := \{T_w(t(d))\}$:

$$T(d) = \nabla_d \min_{x \in X(d)} \Psi(f(x)) = \nabla_d \max_{t \geqslant \bar{t}} \left\{ \sum_{w \in OD} d_w T_w(t) - \langle \bar{f}, t - \bar{t} \rangle - \mu \sum_{e \in E} h\left(t_e - \bar{t}_e, \bar{f}_e, \mu\right) \right\}. \tag{3}$$

В работах [8, 11, 17] приведены эволюционные динамики, приводящие к описанным здесь равновесиям. Отметим, однако, что если рассматривать Logit динамику (ограниченно рациональных агентов с параметром $\tilde{\gamma} > 0$ [17]), то задачу (1) необходимо будет переписать в виде

$$\min_{x \in X} \left\{ \Psi(f(x)) + \tilde{\gamma} \sum_{w \in OD} \sum_{p \in P_w} x_p \ln(x_p / d_w) \right\}. \tag{4}$$

В таком случае говорят, что вместо равновесия Нэша–Вардропа ищется стохастическое равновесие [16, 18].

Это замечание понадобится нам в дальнейшем.

3. Модель расчета матрицы корреспонденций

До сих пор матрица корреспонденций $\{d_w\}_{w \in OD}$ была задана по постановке задачи. Сейчас мы откажемся от этого условия, вводя в источники O производство, а в стоки D — потребление. Агенты «появляются» в тех пунктах производства, произведя товар в которых его можно с выгодой для себя реализовать в каком-нибудь из пунктов потребления. Это означает, что затраты на производство и затраты на транспортировку полностью окупаются последующей выручкой от реализации продукции в пункте потребления. Агенты, которых мы здесь считаем маленькими, будут «приходить» в систему до тех пор, пока существует цепочка (пункт производства–маршрут–пункт потребления), обеспечивающая им положительную прибыль. Важно отметить, что по ходу «наплыва» агентов транспортная сеть становится все более и более загруженной, что может сказываться на затратах на перевозку. В результате прибыль ранее пришедших агентов падает, что побуждает их перераспределяться, т.е. искать более выгодные цепочки. Возникает ряд вопросов. Например, сходится ли такая динамика (точнее, семейство динамик, отражающих рациональность агентов) к равновесию? Если сходится, то единственно ли это равновесие? Если равновесие единственно, то как его можно эффективно найти (описать)? Попытка ответить на эти вопросы (но, прежде всего, на последний вопрос) для достаточно широкого и важного в приложениях класса задач предпринята в последующей части статьи.

Рассмотрим сначала, для большей наглядности, отдельно потенциальный случай. А именно, тот случай, когда в источнике $i \in O$ производственная функция имеет вид $\sigma_i(f_i)$, где $f_i = \sum\limits_{k:(i,k)=e \in E} f_e = \sum\limits_{j:(i,j) \in OD} d_{ij}$, аналогично для стоков $j \in D$ определим функции полезности со знаком минус $\sigma_j(f_j)$, $f_j = \sum\limits_{k:(k,j)=e \in E} f_e = \sum\limits_{i:(i,j) \in OD} d_{ij}$. Все эти функции считаем выпуклыми. Мы обозначаем эти функции одинаковыми буквами, однако это не должно вызвать в дальнейшем путаницы ввиду характерных нижних индексов. Редуцируем рассматриваемую задачу к рассмотренной ранее задаче. Рассмотрим новый граф с множеством вершин $O \cup D$, соединенных теми же ребрами, что и в изначальном графе, добавив в него один дополнительный фиктивный источник и один дополнительный фиктивный сток. Этот фиктивный источник соединим со всеми источниками O, аналогично фиктивный сток соединим со всеми стоками D. Если существует путь из источника i в сток j в исходном графе, то в новом графе прочертим соответствующее ребро с функцией затрат $T_{ij}(d)$. Проведем дополнительное (фиктивное) ребро, соединяющее фиктивный источник с фиктивным стоком, затраты на прохождения которого отсутствуют. Получим в итоге ориентированный граф путей из источника в сток. Легко понять, что мы оказываемся «почти» в условиях предыдущего пункта (причем с графом более простого вида — с одним источником и стоком) с точностью до обозначений:

$$\{x_p\} \to \{d_{ij}\}, \quad \{\tau_e(f_e)\} \to \left\{\sigma_i'(f_i), T_{ij}(d), \sigma_j'(f_j)\right\}.$$

«Почти» – потому что, во-первых, затраты $T_{ij}(d)$ зависят от всего набора $\{d_{ij}\}$, а не только от d_{ij}, а во-вторых, не ясно, что в данном случае играет роль матрицы корреспонденций (в нашем случае это матрица 1×1, т.е. просто число). Начнем с ответа на второй вопрос. Мы считаем, что в источниках имеется потенциальная возможность производить неограниченное количество продукта, просто в какой-то момент перестает быть выгодным что-то производить и перевозить. Для этого, собственно, и было введено нулевое ребро, поток по которому обозначим d_0. То есть, другими словами, мы должны считать, что $\sum\limits_{(i,j)\in OD} d_{ij} + d_0 = \bar{d}$. Если $\bar{d}$ – достаточно большое, то равновесная конфигурация не зависит от того, чему именно равно $\bar{d}$, поскольку не требуется определять d_0. С первой проблемой можно разобраться, немного обобщив теорему 1. Предположим, что (см. формулу (3))

$$\exists \ \Phi(d) \text{ – выпуклая: } T(d) = \nabla\Phi(d). \tag{5}$$

Тогда имеет место ключевая

Теорема 2. *Популяционная игра*

$$\left\langle \{d_{ij}, d_0 \geqslant 0\}, \left\{G_{ij}(d) = \sigma_i'(f_i) + T_{ij}(d) + \sigma_j'(f_j), G_0(d) \equiv 0\right\}\right\rangle$$

является потенциальной. Равновесие d^* в этой игре всегда существует (*если $\sigma(\cdot)$ – сильно выпуклые функции, то равновесие гарантировано единственно*) *и находится из решения задачи выпуклой оптимизации*

$$d^* \in \arg\min_{d \geqslant 0} \tilde{\Psi}(d),$$

$$\tilde{\Psi}(d = \{d_{ij}\}) = \sum_{i\in O} \sigma_i \left(\sum_{j:(i,j)\in OD} d_{ij} \right) + \sum_{j\in D} \sigma_j \left(\sum_{i:(i,j)\in OD} d_{ij} \right) + \Phi(d). \tag{6}$$

Доказательство. Выпишем условие нелинейной комплиментарности (то есть, по сути, определение равновесия Нэша в популяционной игре, заданной в условии). Для этого занумеруем все индексы ij и 0 одним индексом k:

$$\text{для любых } k \text{ выполняется } d_k^* \cdot \left(G_k(d^*) - \min_{k'} \ G_{k'}(d^*) \right) = 0.$$

Действительно, допустим, что реализовалось какое-то другое равновесие $\tilde{d}$, которое не удовлетворяет этому условию. Покажем, что тогда найдется агент, которому выгодно поменять свой выбор. Действительно, тогда

$$\exists \ \tilde{k} : \tilde{d}_{\tilde{k}} \left(G_{\tilde{k}}\left(\tilde{d}\right) - \min_{k'} \ G_{k'}\left(\tilde{d}\right) \right) > 0.$$

Каждый агент, а множество таких агентов не пусто ($\tilde{d}_{\tilde{k}} > 0$), использующий стратегию $\tilde{k}$, действует не разумно, поскольку существует такая стратегия $\bar{k}$, $\bar{k} \neq \tilde{k}$, что $G_{\bar{k}}(d) = \min\limits_{k'} \ G_{k'}(d)$. Этот стратегия $\bar{k}$ более выгодна, чем $\tilde{k}$. Аналогично показывается, что при распределении d^* никому из агентов уже не выгодно отклоняться от своих стратегий.

Покажем, что рассматриваемая нами игра принадлежит к классу так называемых потенциальных игр. В нашем случае это означает, что существует такая функция

$\tilde{\Psi}(d = \{d_{ij}\})$, что $\partial\tilde{\Psi}(d) \big/ \partial d_k = G_k(d)$ для любого k. В нашем случае легко проверить, что этому условию удовлетворяет (по условию) функция

$$\tilde{\Psi}(d = \{d_{ij}\}) = \sum_{i\in O}\sigma_i\left(\sum_{j:(i,j)\in OD} d_{ij}\right) + \sum_{j\in D}\sigma_j\left(\sum_{i:(i,j)\in OD} d_{ij}\right) + \Phi(d).$$

Таким образом, мы имеем дело с потенциальной игрой. Оказывается, что d^* — равновесие Нэша–Вардропа тогда и только тогда, когда оно доставляет минимум $\tilde{\Psi}(d)$ на множестве $d \geqslant 0$. Действительно, предположим, что $d^* \geqslant 0$ — точка минимума. Тогда, в частности, для любых p, q $(d_p^* > 0)$ и достаточно маленького $\delta d_p > 0$ выполняется:

$$-\frac{\partial\tilde{\Psi}(d^*)}{\partial d_p}\delta d_p + \frac{\partial\tilde{\Psi}(d^*)}{\partial d_q}\delta d_p \geqslant 0.$$

Иначе, заменив d^* на

$$\breve{d}^* = d^* + \left(\underbrace{\overbrace{0,...,0,-\delta}^{p}\, d_p, 0,...,0,\delta}_{q}\, d_p, 0,...,0\right) \geqslant 0,$$

мы пришли бы к вектору $\breve{d}^*$, доставляющему меньшее значение $\tilde{\Psi}(d)$ на множестве $d \geqslant 0$:

$$\tilde{\Psi}\left(\breve{d}^*\right) \approx \tilde{\Psi}(d^*) - \frac{\partial\tilde{\Psi}(d^*)}{\partial d_p}\delta d_p + \frac{\partial\tilde{\Psi}(d^*)}{\partial d_q}\delta d_p < \tilde{\Psi}(d^*).$$

Вспоминая, что $\partial\tilde{\Psi}(d) \big/ \partial d_p = G_p(d)$, и учитывая, что q можно выбирать произвольно, получаем:

$$\text{для любого } p, \text{ если } d_p^* > 0, \text{ выполняется } \min_q G_q(d^*) \geqslant G_p(d^*).$$

Но это и есть по-другому записанное условие нелинейной комплементарности. Строго говоря, мы показали сейчас только то, что точка минимума $\tilde{\Psi}(d)$ на множестве $d \geqslant 0$ будет равновесием Нэша–Вардропа. Аналогично рассуждая, можно показать и обратное: равновесие Нэша–Вардропа доставляет минимум $\tilde{\Psi}(d)$ на множестве $d \geqslant 0$. □

Если искать стохастическое равновесие, то функционал в теореме 2 необходимо энтропийно регуляризовать[2]. Из теоремы 2 можно углядеть необходимость искусственного введения потенциалов (двойственных множителей) в сами функции σ. А именно, предположим, что все эти функции σ – линейные с неизвестными наклонами. Тем не менее считается, что при этом известно, чему должны равняться в равновесии следующие суммы:

$$\sum_{j:(i,j)\in OD} d_{ij} = L_i, \qquad \sum_{i:(i,j)\in OD} d_{ij} = W_j \left(\sum_{i\in O} L_i = \sum_{j\in D} W_j = N\right). \tag{7}$$

То есть имеются скрытые от того, кто моделирует, потенциалы (параметры) $\left\{\lambda_i^L, \lambda_j^W\right\}$ [11], которые могут быть рассчитаны исходя из дополнительной информации. Применительно к модели расчета матрицы корреспонденций выписанные дополнительные условия (7) однозначно определяют все неизвестные потенциалы. Однако при этом вместо задачи выпуклой оптимизации мы получаем минимаксную (седловую) задачу, выпуклую по $\{d_{ij}\} \geqslant 0$ и вогнутую, точнее даже линейную, по потенциалам $\left\{\lambda_i^L, \lambda_j^W\right\}$:

[2] К сожалению, строгое обоснование имеется только в случае известного (фиксированного) значения $\bar{d}$ (при этом можно считать $d_0 = 0$).

$$\min_{\substack{\{d_{ij}\} \geqslant 0 \\ \sum\limits_{(i,j)\in W} d_{ij} = N}} \max_{\{\lambda_i^L, \lambda_j^W\}} \left[\sum_{i\in O} \lambda_i^L \cdot \left(\sum_{j:(i,j)\in OD} d_{ij} - L_i \right) + \sum_{j\in D} \lambda_j^W \cdot \left(W_j - \sum_{i:(i,j)\in OD} d_{ij} \right) + \Phi(d) + \gamma \sum_{i,j\in OD} d_{ij} \ln\left(d_{ij}/N\right) \right]. \tag{8}$$

Эта задача всегда имеет решение, притом единственное. Именно эта задача аккумулирует в себе сразу несколько блоков (в нашем случае — два) многостадийной модели, о который мы ранее много раз упоминали.

4. Численный поиск равновесия в многостадийной модели

Задачу (8) также можно переписать следующим образом (не ограничивая общности, считаем $\sum\limits_{i=1}^{n} L_i = \sum\limits_{j=1}^{n} W_j = 1$ и используем другие обозначения переменных, соответсвующих работе [14])

$$\min_{\substack{\sum\limits_{j:(i,j)\in OD} d_{ij} = L_i,\ \sum\limits_{i:(i,j)\in OD} d_{ij} = W_j \\ \sum\limits_{(i,j)\in OD} d_{ij} = 1}} \max_{t \geqslant \bar{t}} \left\{ \sum_{(i,j)\in OD} d_{ij} \ln d_{ij} + \sum_{(i,j)\in OD} c_{ij}(t)\, d_{ij} + g(t) \right\} =$$

$$= \max_{t \geqslant \bar{t}} \max_{\lambda,\mu} \left\{ \langle \lambda, L \rangle + \langle \mu, W \rangle - \ln \left(\sum_{(i,j)\in OD} \exp\left(-c_{ij}(t) - 1 + \lambda_i + \mu_j\right) \right) + g(t) \right\} =$$
$$= -\min_{t \geqslant \bar{t}} f(t), \tag{9}$$

где

$$f(t) = \min_{\lambda,\mu} \left\{ \ln \left(\sum_{(i,j)\in OD} \exp\left(-c_{ij}(t) - 1 + \lambda_i + \mu_j\right) \right) - \langle \lambda, L \rangle - \langle \mu, W \rangle - g(t) \right\}.$$

Заметим, что расчет градиента $\nabla f(t)$ (в ряде транспортных приложений вогнутые функции $c_{ij}(t)$ — негладкие, тогда вместо градиентов стоит понимать суперградиенты $c_{ij}(t)$ и субградиент выпуклой функции $f(t)$) осуществляется по следующей формуле (Демьянова–Данскина–Рубинова):

$$\nabla f(y) = -\frac{\sum\limits_{(i,j)\in OD} \exp\left(-c_{ij}(t) + \lambda_i^* + \mu_j^*\right) \nabla c_{ij}(t)}{\sum\limits_{(i,j)\in OD} \exp\left(-c_{ij}(t) + \lambda_i^* + \mu_j^*\right)} - \nabla g(t) =$$
$$= -\sum_{(i,j)\in OD} d_{ij}\left(\lambda^*, \mu^*\right) \nabla c_{ij}(t) - \nabla g(t),$$

где (λ^*, μ^*) — одно из решений задачи (9).

Внутренняя задача максимизации по (λ, μ) в (9) может быть явно решена по μ при фиксированном λ и наоборот. Собственно, таким образом, получается метод балансировки расчета матрицы корреспонденций по энтропийной модели (тесно связанный с методом Синхорна) как метод простой итерации для явно выписываемых условий экстремума (принципа Ферма): $\lambda = \Lambda(\mu)$, $\mu = \mathrm{M}(\lambda)$.

Оператор $(\lambda, \mu) \to (\Lambda(\mu), \mathrm{M}(\lambda))$ является сжимающим в метрике Биркгофа–Гильберта ρ. Это означает, после $N \sim \ln\left(\sigma^{-1}\right)$ итераций метода балансировки можно получить такие (λ_N, μ_N), что $(\{(\lambda_*(t), \mu_*(t))\}$ — двумерное аффинное множество решений)

$$\rho\left((\lambda_N, \mu_N); \{(\lambda_*(t), \mu_*(t))\}\right) \leqslant \sigma.$$

Причем на практике наблюдается очень быстрая сходимость, т.е. коэффициент пропорциональности небольшой. Таким образом, мы можем приближенно (с хорошей точностью) решить внутреннюю задачу.

Далее предлагается воспользоваться универсальным методом Ю.Е. Нестерова [19] для решения внешней задачи оптимизации по t. Этот метод оптимально адаптивно настраивается на гладкость функционала $f(t)$ на текущем участке пребывания итерационного процесса. Однако нам потребуется использовать этот метод в варианте с неточным оракулом, выдающим градиент, поскольку внутренняя задача решается приближенно. Все это и многое другое можно сделать с точными константами в оценках скоростей сходимости [14].

Описанную схему можно уточнить за счёт учёта специфики функций $g(t)$ и $c_{ij}(t)$ и наличия возможности эффективного расчёта их градиентов. Здесь возникает большое число нюансов, которые необходимо прорабатывать.

Задачи (8), (9), хотя и имеют выпуклую структуру, что не имеет место для способа формирования многостадийной модели по всем известным нам ранее предлагавшимся подходам, все равно представляются достаточно сложными для численного решения. Связано это с их размерами. Число вершин в графе реальной транспортной сети (например, Москвы и МО) измеряется десятками тысяч вершин и корреспонденций, то есть существует сто тысяч ребер и намного больше возможных путей. Здесь уместно отметить, что простейшая транспортная задача ЛП (Л.В. Канторовича) имеет трудоемкость решения, пропорциональную кубу числа источников/стоков (это оценка не улучшаема в общем случае). А у нас намного более сложная задача. Если учесть, что современный компьютер за одну секунду может сделать не более миллиарда операций с плавающей запятой, то понятно, что для реальных транспортных сетей даже решение обычной транспортной задачи уже представляет большую сложность (часы машинного времени). Основное планируемое усовершенствование в ближайшей перспективе – это сделать возможным за разумное время эффективно решать (в том числе на одном PC) задачи поиска равновесий в многостадийных моделях транспортных потоков. Это очень важно уметь делать быстро, поскольку такие задачи нужно много раз перерешивать, сравнивая перебором, например, различные сценарии транспортной застройки. К сожалению, от перебора здесь нельзя никак уйти, поэтому единственный выход – увеличение скорости одного сравнения (расчета равновесия при заданном сценарии).

Авторы выражают благодарность А.А. Шананину, Ю.Е. Нестерову за плодотворное обсуждение различных частей данной заметки.

Работа выполнена при поддержке гранта РФФИ 13-01-12007-офи_м.

Литература

1. http://rbcdaily.ru/industry/562949995887243

2. *Sandholm W.H.* Evolutionary implementation and congestion pricing // Review of Economic Studies. 2002. V. 69. P. 81–108.

3. *Beckmann M., McGuire C.B., Winsten C.B.* Studies in the economics of transportation // RM-1488. Santa Monica: RAND Corporation, 1955.

4. *Nesterov Y., de Palma A.* Stationary Dynamic Solutions in Congested Transportation Networks: Summary and Perspectives // Networks Spatial Econ. 2003. N 3(3). P. 371–395.

5. *Ortúzar J.D., Willumsen L.G.* Modelling transport. JohnWilley & Sons, 2011.

6. *Гасников А.В., Кленов С.Л., Нурминский Е.А., Холодов Я.А., Шамрай Н.Б.* Введение в математическое моделирование транспортных потоков. 2-е изд. / 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, PTV VISSION; с предисловием руководителя департамента транспорта г. Москвы М.С. Ликсутова. М.: МЦНМО, 2013. 427 с. http://www.mou.mipt.ru/gasnikov1129.pdf

7. *Ващенко М.П., Гасников А.В., Молчанов Е.Г., Поспелова Л.Я., Шананин А.А.* Вычислимые модели и численные методы для анализа тарифной политики железнодорожных грузоперевозок. М.: ВЦ РАН, 2014. 52 с. arXiv:1501.02205

8. *Гасников А.В., Дорн Ю.В., Нестеров Ю.Е., Шпирко С.В.* О трехстадийной версии модели стационарной динамики транспортных потоков // Математическое моделирование. 2014. Т. 26:6. С. 34–70. arXiv:1405.7630

9. *Гасников А.В.* Марковские модели макросистем. e-print, 2014. arXiv:1412.2720

10. *Гасников А.В.* Об эффективной вычислимости конкурентных равновесий в транспортно-экономических моделях // Математическое моделирование. 2015. Т. 27. arXiv:1410.3123 (принята к печати).

11. *Гасников А.В., Гасникова Е.В., Мендель М.А., Чепурченко К.В.* Эволюционные выводы энтропийной модели расчета матрицы корреспонденций // Математическое моделирование. 2016. Т. 28. arXiv:1508.01077 (принята к печати).

12. *Гасников А.В., Гасникова Е.В., Нестеров Ю.Е., Чернов А.В.* Об эффективных численных методах решения задач энтропийно-линейного программирования // ЖВМ и МФ. 2016. Т. 56. № 2. arXiv:1410.7719 (принята к печати).

13. *Гасников А.В., Двуреченский П.Е., Дорн Ю.В., Максимов Ю.В.* Численные методы поиска равновесного распределения потоков в модели Бэкмана и модели стабильной динамики // Математическое моделирование. 2016. (подана) arXiv:1506.00293

14. *Гасников А.В., Двуреченский П.Е., Нестеров Ю.Е., Спокойный В.Г., Стецюк П.И., Суворикова А.Л., Чернов А.В.* Поиск равновесий в многостадийныйх транспортных моделях // ЖВМ и МФ. 2016. arXiv:1506.00292

15. *Гасников А.В., Гасникова Е.В.* Теория макросистем с точки зрения стохастической химической кинетики // Математическое моделирование. 2016. Т. 28. (подана).

16. *Бабичева Т.С., Гасников А.В., Гасникова Е.В., Двуреченский П.Е., Лагновская А.А.* Современные подходы к численному поиску стохастического равновесия в модели равновесного распределения потоков Бэкмана // ЖВМ и МФ. 2016. arXiv:1505.07492

17. *Sandholm W.H.* Evolutionary implementation and congestion pricing // Review of Economic Studies. 2002. V. 69. P. 81–108.

18. *Sheffi Y.* Urban transportation networks: Equilibrium analysis with mathematical programming methods. N.J.: Prentice–Hall Inc., Englewood Cliffs, 1985.

19. *Nesterov Yu.* Universal gradient methods for convex optimization problems // CORE Discussion Paper 2013/63. 2013.

References

1. http://rbcdaily.ru/industry/562949995887243

2. *Sandholm W.H.* Evolutionary implementation and congestion pricing. Review of Economic Studies. 2002. V. 69. P. 81–108.

3. *Beckmann M., McGuire C.B., Winsten C.B.* Studies in the economics of transportation. RM-1488. Santa Monica: RAND Corporation, 1955.

4. *Nesterov Y., de Palma A.* Stationary Dynamic Solutions in Congested Transportation Networks: Summary and Perspectives. Networks Spatial Econ. 2003. N 3(3). P. 371–395.

5. *Ortúzar J.D., Willumsen L.G.* Modelling transport. JohnWilley & Sons, 2011.

6. *Gasnikov A.V., Klenov S.L., Nurminskyi E.A., Holodov Y.A., Shamrai N.B.* Introduction to the mathematical modeling of traffic flows. Ed. Gasnikov A.V., with a foreword by hand of the Department of Transport in Moscow Liksutov M.S. M.: MCCME, 2013. (In Russian) http://www.mou.mipt.ru/gasnikov1129.pdf

7. *Vatschenko M.P., Gasnikov A.V., Molchanov E.G., Pospelova L.Y., Shananin A.A.* Analysis of tariff policy of a railway cargo transportation. M.: CC of RAS, 2014. 52 p. (In Russian) arXiv:1501.02205

8. *Gasnikov A.V., Dorn Y.V., Nesterov Y.E., Shpirko S.V.* On the three-stage version of stable dynamic model. Mathematical modelling. 2014. V. 26:6. P. 34–70. (In Russian) arXiv:1405.7630

9. *Gasnikov A.V.* Markov models of macrosystems. e-print, 2014. (In Russian) arXiv:1412.2720

10. *Gasnikov A.V.* About reduction of searching competetive equillibrium to the minimax problem in application to different network problems. Mathematical modelling. 2015. V. 27. (In Russian) arXiv:1410.3123 (accepted for publication).

11. *Gasnikov A.V., Gasnikova E.V., Mendel M.A., Chepurchenko K.V.* Evolutionary interpretations of entropy model for correspondence matrix calculation. Mathematical modelling. 2016. V. 28. (In Russian) arXiv:1508.01077 (accepted for publication).

12. *Gasnikov A.V., Gasnikova E.V., Nesterov Y.E., Chernov A.V.* Entropy linear programming. Computational Mathematics and Mathematical Physics. 2016. V. 56. N 2. (In Russian) arXiv:1410.7719 (accepted for publication).

13. *Gasnikov A.V., Dvurechensky P.E., Dorn Y.V., Maksimov Y.V.* Searching equillibriums in Beckmann's and Nesterov–de Palma's models. Mathematical modelling. 2016. (submitted) (In Russian) arXiv:1506.00293

14. *Gasnikov A.V., Dvurechensky P.E., Nesterov Y.E., Spokoyny V.G., Stetsyuk P.I., Suvorikova A.L., Chernov A.V.* Universal method with inexact oracle and its applications for searching equillibriums in multistage transport problems. Computational Mathematics and Mathematical Physics. 2016. (In Russian) arXiv:1506.00292

15. *Gasnikov A.V., Gasnikova E.V.* The theory of macrosystems in terms of stochastic chemical kinetics. Mathematical modelling. 2016. V. 28. (submitted) (In Russian).

16. *Babicheva T.S., Gasnikov A.V., Gasnikova E.V., Dvurechensky P.E., Lagunovskaya A.A.* Efficient calculation of stochastic equilibriums in the Beckmann's and stable dynamic models. Computational Mathematics and Mathematical Physics. 2016. (In Russian) arXiv:1505.07492

17. *Sandholm W.H.* Evolutionary implementation and congestion pricing. Review of Economic Studies. 2002. V. 69. P. 81–108.

18. *Sheffi Y.* Urban transportation networks: Equilibrium analysis with mathematical programming methods. N.J.: Prentice–Hall Inc., Englewood Cliffs, 1985.

19. *Nesterov Yu.* Universal gradient methods for convex optimization problems. CORE Discussion Paper 2013/63. 2013.

Поступила в редакцию 22.09.2015.