

ОЦЕНКА МАТРИЦЫ КОРРЕСПОНДЕНЦИЙ В ВИДЕ ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ СО СМЕШАННЫМИ ОГРАНИЧЕНИЯМИ

А.Ю.Михайлов, И.М. Головных, Р.Ю. Лагерева
Иркутский государственный технический университет
e-mail:road@istu.irk.ru

В зарубежной теории и практике проектирования транспортных систем уже с 70-х годов уделялось большое внимание методам восстановления матриц корреспонденций в виде потоков транспортных средств (Estimation of Origin-Destination Matrix from Traffic Counts). Методы восстановления существующих матриц корреспонденций отнесены специалистами “Мировой дорожной ассоциации” (PIARC) к числу важнейших инструментов анализа УДС. В американском руководстве Highway Capacity Manual 2000 восстановлению матриц корреспонденций уделено большое внимание (Глава 29 “Corridor analysis”) и изложен метод такой оценки применительно к магистральным дорогам. По данной тематике имеется обширнейшая библиография, которую невозможно представить в рамках статьи.

В нашей стране в первую очередь получила развитие теория расчетов пассажиропотоков, в том числе, были предложены методы оценки существующих матриц корреспонденций на маршрутном пассажирском транспорте [1]. Собственно методам восстановления матриц корреспонденций в виде потоков транспортных было посвящено лишь несколько исследований [3,4], что объяснялось относительно меньшим интересом к проблемам проектирования и реконструкции УДС.

Продолжающийся рост автомобильного парка и объемов движения в российских городах требует совершенствования инструментария проектирования и оценки УДС. Поэтому представляется интересным рассмотреть возможности решения рассматриваемой задачи применительно к российским условиям, учитывая, в первую очередь применяющиеся методы обследований интенсивности движения. Как правило, такие обследования проводятся в пиковые часы с подсчетом транспортных средств разных типов на всех направлениях движения на каждом из обследуемых перекрестков. Поэтому формулируется задача регрессионной оценки матрицы корреспонденций для следующих условий:

- исходными данными являются замеры интенсивности движения на отдельных участках улично-дорожной сети, выполняемые в определенное время в будние дни (например, вечерний час пик 17.00—19.00);

- исходные данные об интенсивности движения содержат ошибки (см. рис. 3), вызванные проведением замеров в разные дни и ошибками самих подсчетов интенсивности движения.

Следует привести очень важную для понимания рассматриваемой задачи цитату из HCM 2000 [6]: “Количество элементов матрицы корреспонденций всегда превосходит количество сегментов сети. Цель оценивания состоит не в определении точной матрицы корреспонденций, а в нахождении такой, которая достаточно близка к ней и соответствует данным интенсивности движения.” (глава 29 “Corridor analysis” с. 29-30). Следует пояснить, что под сегментами в оригинальном тексте понимаются дуги графа, которым представлено описание сети.

В последние годы восстановление матриц корреспонденций по замерам на дугах сети применяется для информационных сетей (Y. Zhang и др. [7], A. Medina и др. [8]). Также отмечается (A. Medina и др. [8]), что в рассматриваемых задачах количество оцениваемых корреспонденций N (корреспонденции $x_i, i = 1, 2, \dots, N$) значительно превосходит количество дуг K , на которых известны нагрузки, т.е. $N \gg K$.

В отличие от HCM 2000 рассматривается не “транспортный коридор”, а участок УДС. С целью использования распределения потоков по принципу “все или ничего” (all or nothing) УДС разделяется на отдельные “маршруты”, или (как в HCM 2000) “направления движения” (рис 1 – 2). Таким образом, матрицы описывающая принадлежность корреспонденций дугам графа будет состоять из 1 и 0 (наличие или отсутствие данной корреспонденции).

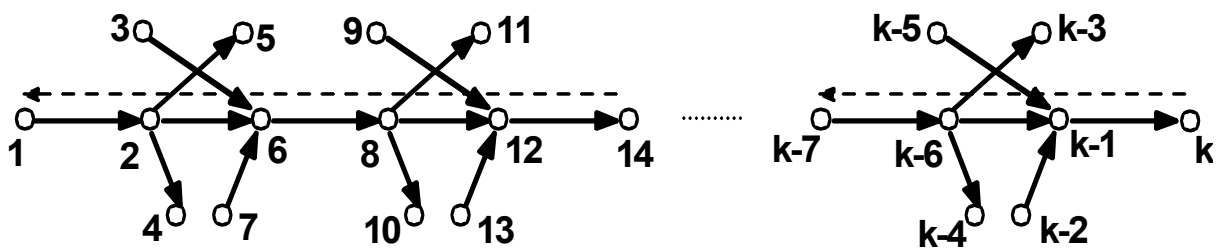
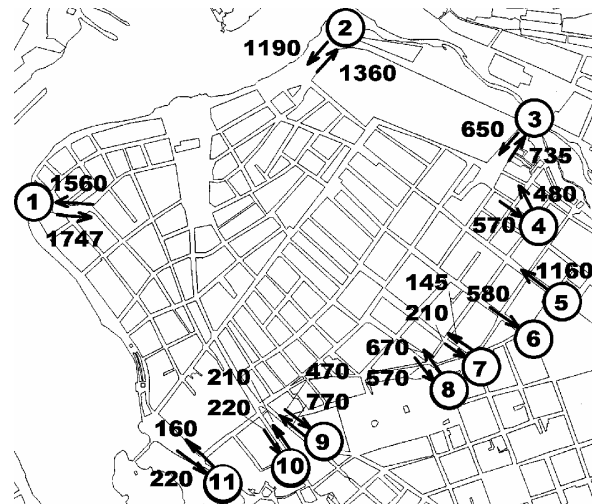
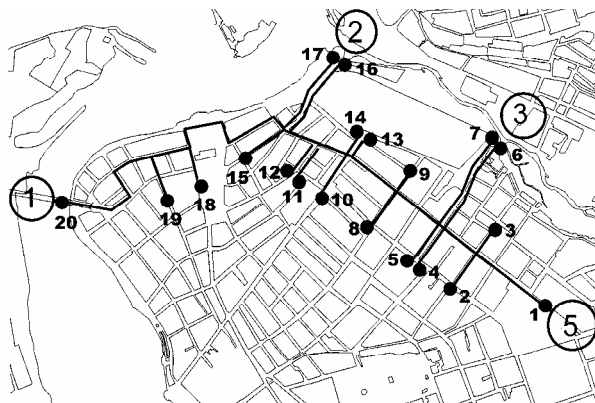


Рис. 1. Представление маршрута движения в виде ориентированного графа для определения матрицы корреспонденций: 1,3,4,5,7,9,10,11,13, ..., k-4, k-5, k-4, k-3, k-2, k - вершины графа, в которых возникают и заканчиваются корреспонденции (O-D pairs)

а) вершины матрицы корреспонденций транзитных потоков (“загрузочные” вершины) центра Иркутска и интенсивности движения в вечерний час пик, физ.ед./ч



б) граф для оценки матрицы корреспонденций из вершины 1 в вершины 1,2,3, ..., 20, используемой для оценки транзитных потоков из вершины 5 в вершины 1,2,3



в) транзитные потоки из вершины 5 в вершины 1,2,3, полученные при оценке матрицы корреспонденций для графа УДС, представленного вершинами 1,2,3, ..., 20,

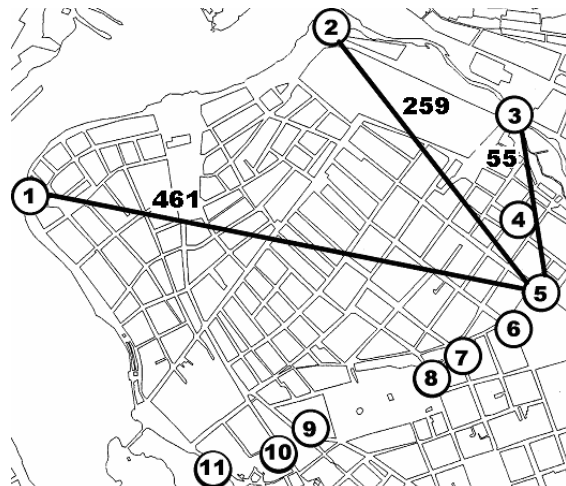


Рис. 2. Определение матрицы корреспонденций для транзитных потоков в центральной части Иркутска: а – исходные потоки на водах и выходах в центральную часть города, для которых рассчитывается матрица корреспонденций; б – один из маршрутов движения, на которые разделяется УДС, и для которых строятся графы и рассчитываются матрицы корреспонденций; в – результат оценки транзитных потоков, для одного из входов в центральную часть

Отдельно рассматриваемое направление движения (маршрут) представляется в виде графа (рис. 1 и 3), для которого определяется своя матрица корреспонденций. Соответственно начальными и конечными пунктам корреспонденций являются поворотные потоки на перекрестках, т.е. входящие потоки на улицу или выходящие с нее (рис. 1). Вместо центроидов расчетных транспортных районов используются “нагрузочные” вершины сети (рис. 2), образующиеся на границах рассматриваемого участка УДС, в которых начинаются и оканчиваются корреспонденции.

Сформулированная задача регрессионного оценивания корреспонденций предполагает использование статистических процедур устойчивых к выбросам (т.е. грубым ошибкам данных). Один из приемов робастной оценки параметров регрессии θ состоит в использовании вместо суммы квадратов отклонений регрессии

$$\sum_{i=1}^n e_i^2 \text{ минимизации целевой функции}$$

$$\sum_{i=1}^n |e_i|^v = \sum_{i=1}^n |y_i - \theta|^v, \quad (1)$$

где v – показатель степени, значения которого принимают в диапазоне $1 \leq v < 2$.

Оценки (1) предложено называть L_v оценками [3]. Устойчивость такой аппроксимации по сравнению с методом наименьших квадратов вызвана тем, что большие отклонения $(y_i - \theta)$ меньше влияют на целевую функцию. Поэтому показатель степени v интерпретируется некоторыми авторами как фильтр выбросов [3].

Частным случаем ($v = 1$) оценки (1) является минимизация суммы абсолютных модулей отклонений, т.е. метод наименьших модулей (МНМ)

$$\sum_{i=1}^n |e_i| = \sum_{i=1}^n \left| y_i - \sum_{j=1}^m a_j x_{ij} \right|, \quad (2)$$

Аппроксимацию (2) можно рассматривать как задачу линейного программирования с минимизацией целевой функции в виде взвешенной суммы модулей ошибок e_i [5]

$$\sum_{i=1}^m w_i (g_i + h_i), \quad (3)$$

где $g_i = \max |0, e_i|$, $h_i = \min |-e_i, 0|$; w_i – весовые коэффициенты, при ограничениях (в случае простого МНМ весовые коэффициенты w_i принимаются равными 1).

Ограничения задачи (3): $Ax + g - h = y$; $w \geq 0, g \geq 0, h \geq 0$.

При определении вектора корреспонденций x (вектор-столбец размерности $m \times 1$), решается задача линейного программирования для нового век-

тора переменных x_2 размерности $(m+2n) \times 1$. В целевую функцию (3), которой является сумма модулей ошибок, входит лишь часть вектора переменных x_2 . Матрица коэффициентов A размера $n \times m$ преобразуется в матрицу A_2 размером $n \times (m+2n)$, т.е. дополняется двумя диагональными матрицами d_1 и d_2 .

$$A_2 = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} & 1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2m} & 0 & 1 & \dots & 0 & 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nm} & 0 & 0 & 0 & 1 & 0 & 0 & \dots & -1 \end{bmatrix}$$

Вектор ошибок e^i определяется учетом результатов оценки вектора корреспонденций x^{i-1} на предыдущей итерации. В зависимости от знака ошибки e_i задаются элементы вектора x_2 , начиная с индекса $m+1$. Таким образом, задача (3) формулируется в следующем виде

$$\min \sum_{i=m+1}^{m+2n} w_i x_i, \quad (4)$$

при условии

$$A_2 x_2 = y; \quad x \geq 0.$$

При проведении первой итерации для определения весовых коэффициентов w_i необходим оцененный каким-либо методом вектор x^0 , при этом допускаются даже экспертные оценки [2]. Применительно к рассматриваемой задаче такой начальной оценкой x^0 могут быть старая матрица, матрица, полученная в результате предварительного расчета, в самом крайнем случае – экспертная оценка. Начальные значения корреспонденций x_{ij}^0 определялись автором с использованием простой гравитационной модели

$$x_{ij}^0 = a_i b_j A_i B_j, \quad (5)$$

где x_{ij} – количество транспортных средств, следующих из вершины i в вершину j ; a_i – интенсивность движения из i ; b_j – интенсивность движения в j ; A_i, B_j – коэффициенты балансировки по столбцам и строкам матрицы.

Минимизация целевой функции (4), рассматривается задача линейного программирования со смешанными ограничениями

$$\min w^T x, \quad (6)$$

при линейных ограничениях, в которых матрица A_2 получается на основе матрицы инцидентности A

$$A_2 x_2 = y; \quad (7)$$

и двухсторонних ограничениях

$$x^{lb} \leq x \leq x^{ub}. \quad (8)$$

Исходные данные включают:

A – матрицу инцидентности ранга $n \times m$, задает принадлежность корреспонденций дугам графа сети, ее элементы a_{ij} имеют значение 0 или 1. если корреспонденция j проходит по дуге i , то $a_{ij} = 1$, в противном случае $a_{ij} = 0$, где $i = 1, 2, \dots, n$, n – количество дуг графа УДС, на которых задана интенсивность движения, $j = 1, 2, \dots, m$, m – количество корреспонденций.

y – вектор-столбец значений интенсивности движения на дугах сети;

x^0 – начальное приближение значений корреспонденций – исходный вектор-столбец значений, который может быть задан: в виде старой матрицы корреспонденций, преобразованной в вектор; значениями корреспонденций, полученными в результате экспертной оценки; в виде оценки матрицы корреспонденций, полученной с использованием гравитационной модели;

lb – коэффициент, задающий нижние допустимые значения корреспонденций $x^{lb} = x lb$;

ub – коэффициент, задающий верхние допустимые значения корреспонденций $x^{ub} = x ub$;

div – коэффициент, задающий верхние допустимые значения искусственных переменных $x^{ub}_i = x_i / div$;

v – показатель степени весовых коэффициентов $w_i = |e_i|^{v-2}$.

В результате решения получаем векторы-столбцы:

x – восстановленные значения корреспонденций;

\hat{y} – теоретические значения интенсивности движения на дугах, рассчитанные на основе восстановленных корреспонденций $\hat{y} = A x$;

e – остатки (ошибки) регрессии – разность между наблюдаемыми и рассчитанными значениями интенсивности движения на дугах $e = y - A x$.

Оценка точности регрессии основана на парном сравнении замеренных y_i и оцененных \hat{y}_i значений потоков на дугах (т.е. анализ остатков регрессии $e = y - A * x$) и использовании статистик:

- критерий Стьюдента для парного сравнения исходных и оцененных значений потоков;
- средняя абсолютная разность пар замеренных и оцененных значений потоков $\bar{e}_{abs} = \sum_i |e_i| / n$;
- отношение средней абсолютной ошибки к среднему значению интенсивности движения $E = \bar{e}_{abs} / \bar{y}$.

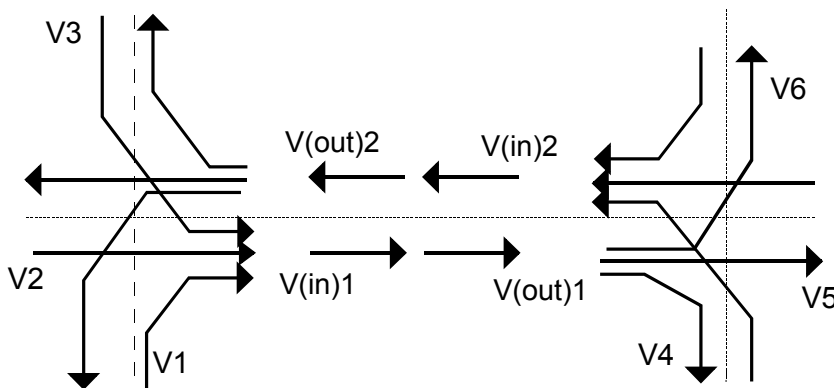
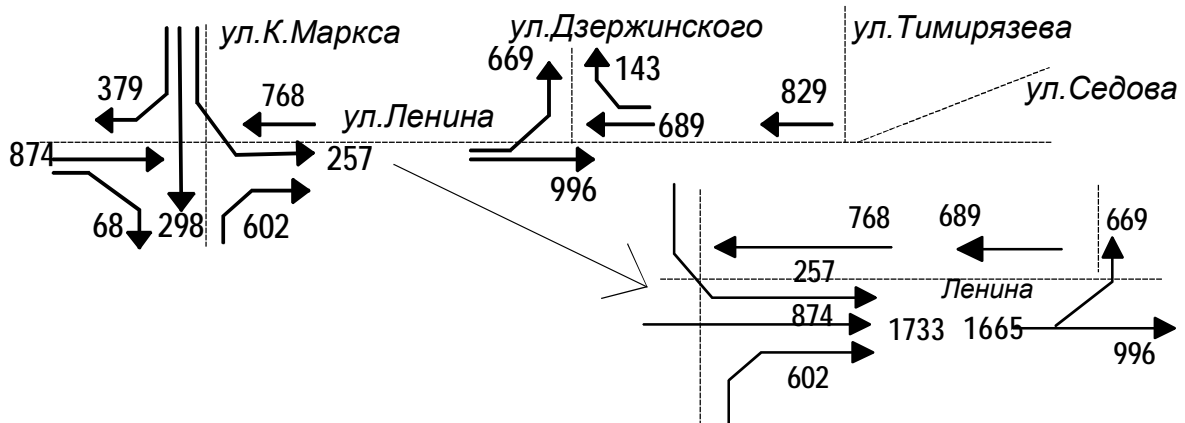
Для предварительной оценки качества исходных данных и наличия в них грубых ошибок (рис. 3а): предлагается рассматривать УДС в следующем виде (рис. 3б):

входящие на перегон $F(in)_i$ и выходящие $F(out)_i$ с перегона потоки образуют пары;

для оценки статистической однородности выборок значений интенсивности движения и выявления выбросов используются значения разностей пар

$$d_i = V(out)_i - V(in)_i. \quad (9)$$

а)



б)

Рис. 3. Возникновение ошибки при проведении обследований интенсивности движения (а) и представление перегона улицы в модели оценки ошибок подсчетов интенсивности движения (б)

Оценка качества данных интенсивности движения сводится к проверке гипотезы, что связанные выборки $V(out)_i$ и $V(in)_i$ принадлежат к одной генеральной совокупности, для их сравнения используется критерий Стьюдента для разностей пар. Кроме проверки близости выборок используются следующие критерии:

средняя абсолютная ошибка

$$\bar{d}_{abs} = \sum_i |d_i| / n. \quad (10)$$

отношение средней абсолютной ошибки к среднему значению интенсивности движения на перегоне в одном направлении

$$E = \bar{d}_{abs} / \bar{V}. \quad (11)$$

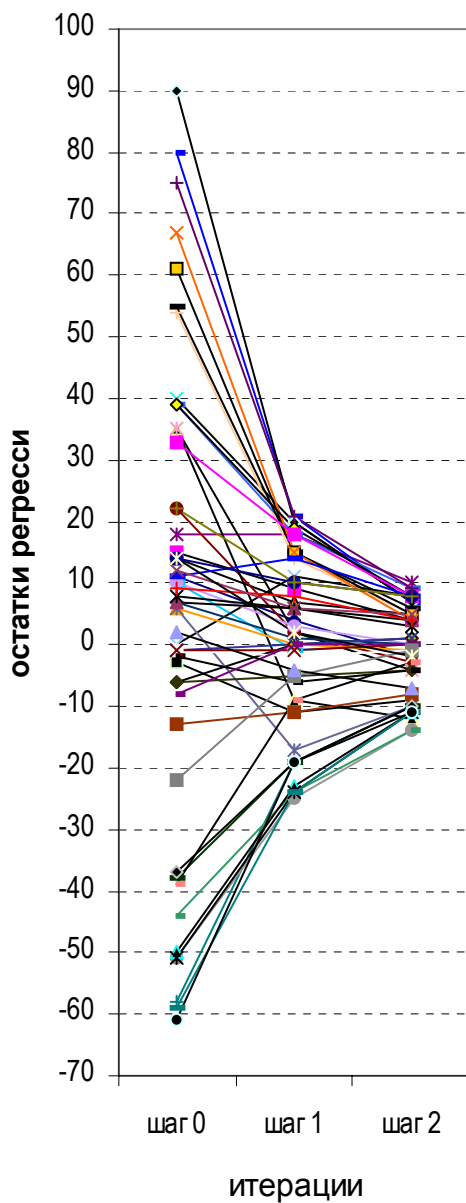
Для выявления грубых ошибок исходных данных следует использовать нормированное отклонение $\hat{z} = (d_i - \bar{d}) / S_d$, рассчитываемое для разностей пар входящих и выходящих потоков $d_i = V(out)_i - V(in)_i$.

В соответствии с имеющимися данными (материалы обследований: С.Петербург 1977—1978 гг., 1985—1988 г.г.; Иркутск 1995 г., центральная часть Иркутска 1998 г.) максимальные абсолютные значения ошибок достигают 450 – 470 авт./ч. (в одном направлении). Диапазон значений средней абсолютной ошибки составил $\bar{d}_{abs} = 53.8 - 190.6$, а значений отношения \bar{d}_{abs} / \bar{V} имеет разброс от 0.082 до 0.391. Средняя абсолютная ошибка составила 115.3 авт./ч., при этом средняя абсолютная ошибка составляет в среднем 19% интенсивности движения. Анализ ошибок обследований интенсивности движения позволило сформулировать требования к точности восстановления матриц корреспонденций. Представляется, что приемлемая точность восстановления корреспонденций $\pm 10 \dots \pm 20$ авт/ч, которую можно использовать в качестве допустимой невязки итераций.

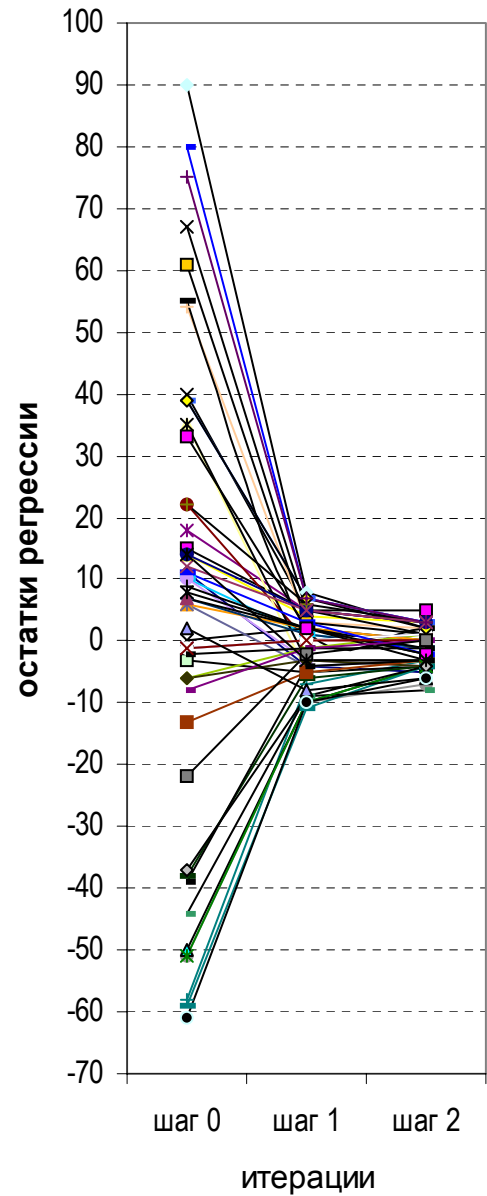
По результатам тестирования с применением искусственных данных без ошибок, искусственных загрязненных данных (максимальные ошибки ± 350 авт/ч) и реальных данных простой МНМ (все весовые коэффициенты $w_i = 1$) имеет более высокую скорость схождения по сравнению с взвешенным МНМ (рис. 4).

Анализ ошибок оценки корреспонденций \bar{d}_{abs} , значений потоков \bar{e}_{abs} и невязок оценок корреспонденций $\Delta d_i = |x_i^n - x_i^{n-1}| \leq 20$ (x_i^n и x_i^{n-1} – соответственно оценки корреспонденции x_i на итерациях $n-1$ и n) показывает, что необходимая сходимость достигается простым МНМ на 2-ой итерации (см. рис. 4). В этой связи для практического использования рекомендуется процедура восстановления матрицы с использованием простого МНМ или комбинированного МНМ (вторая итерация – взвешенный МНМ).

Установлено, что наибольший эффект на качество восстановления корреспонденций и значений потоков оказывают двухсторонние ограничения (8) корреспонденций $x_i \quad i= 1, 2, \dots, m$ и искусственных переменных $x_i, \quad i=m+1, m+2, \dots, m+2n$.



а)



б)

Рис. 4 Ошибки значений потоков $e_i = y_i - \hat{y}_i$ (остатки регрессии) на 0-й – 2-й итерациях: а – взвешенный МНМ; б – сочетание итераций: 1-я – простой МНМ, 2-я – взвешенный МНМ

В соответствии с результатами тестирования рекомендуются следующие значения ограничений:

нижние ограничения оцениваемых переменных $x_i^{lb} = kx_i, k=0$;

верхние ограничения оцениваемых переменных $x_i^{ub} = kx_i, k \geq 2$;

нижние ограничения искусственных переменных $x_i^{lb} = 0, i=m+1, m+2, \dots, m+2n$;

верхние ограничения искусственных переменных $x_{m+j}^{ub} = (1/div) x_j$ и $x_{m+n+j}^{ub} = (x_j/div)$ рассчитывать для $div \geq 30$, где x_j – модуль остатка рег-

рессии e_j , полученный на предыдущей итерации, n – количество дуг графа УДС, на которых заданы значения интенсивности движения; $j = 1, 2, \dots, n$.

Восстановление матриц корреспонденций можно выполнять с использованием функций:

LP библиотеки Optimization Toolbox 2.0 версий пакета MATLAB 5.1 и 5.2;

LINPROG библиотеки Optimization Toolbox 2.2 версии пакета MATLAB 6.1;

Выбор пакета MATLAB для решения задачи восстановления матриц корреспонденций обусловлен тем, что пакет:

ориентирован на выполнение операций с векторами и матрицами, в том числе разряженными матрицами;

применяется для решения сетевых задач [7], в том числе на его основе осуществляется разработка приложений для транспортных задач, например, программа DelftOD Version 2.0 (автор Nanne J. van der Zijpp, Delft University of Technology - <http://www.delftod.tudelft.nl>);

достаточно известен и распространен в нашей стране, имеет сопровождение.

Предложенный метод полностью основан на применении стандартных математических библиотек и не требует специальных профессиональных знаний в области программирования, поэтому прост и доступен. Поскольку в составе пакета MATLAB есть средства разработки графического интерфейса и обмена данными с Excel имеется возможность разработки специального приложения для среды MATLAB. В дальнейшем авторы планируют рассмотреть возможности применения метода робастного восстановления корреспонденций пассажирских (межостановочных матриц корреспонденций) пешеходных потоков, а также разработать графический интерфейс на примере DelftOD Version 2.0 с помощью приложения «GUI» пакета MatLab (Graphical User Interfaces), при этом создать возможность синхронизации данных между табличным процессором Excel через приложение «Excel Link», для удобства сбора и подготовки исходных данных.

В завершение следует отметить, что опубликованное в августе 2003 решение задачи восстановления матриц корреспонденций в информационных сетях [8] также использует в качестве начальной итерации гравитационную модель, но следующей итерации рассматривается задача томографии, решаемая методами квадратичного программирования – т.е. применяется взвешенный метод наименьших квадратов. В нашем случае маршруты (см. рис. 1,2) описываются матрицами, имеющими плохую обусловленность, поэтому тесты по использованию взвешенного метода наименьших квадратов (1998 -2001 гг.) показали его непригодность.

ЛИТЕРАТУРА

1. *Артынов А.П., Скалецкий И.И.* Автоматизация процессов планирования и управления транспортными системам. - М.: Транспорт, 1981. - 280 с
2. *Демиденко У.З.* Линейная и нелинейная регрессия. – М.: Финансы и статистика, 1981. – 302 с.
3. *Киселева О.Н., Сена С.Л., Федоров В.П.* Определение матрицы существующих грузовых корреспонденций на основе обследований на магистральной сети. //Социально-экономические проблемы развития транспортных систем городов. Тез.докл. второй обл. эконом. конф. - Свердловск, 1988. с. 95 - 98.
4. *Мягков В.Н., Пальчиков Н.С., Федоров В.П.* Математическое обеспечение градостроительного проектирования. – Л.: Наука, 1989. – 144 с.
5. Петрович М.Л. Регрессионный анализ и его математическое обеспечение на ЕС ЭВМ: Практическое руководство. – М.: Финансы и статистика, 1982. – 199 с.
6. *Highway Capacity Manual.* // TRB, Washington, DC, 2000. – 1134 p.
7. *Medina A., Taft N., Salamatian K, Bhattacharyya S., Diot C.* Traffic Matrix Estimation: Existing Techniques and New Directions. SIGCOMM'02, August 19-23, 2002. – 15 p. Available: http://gaia.cs.umass.edu/measurement/traffic_matrix_estimation.pdf
8. *Yin Zhang, Roughan M., Duffield N., Greenberg A.* Fast Accurate Computation of LargeScale IP Traffic Matrices from Link Loads. SIGMETRICS'03, 2003, June 10–14, 12 p. Available: <http://www.research.att.com>

Английское название статьи: Robust Estimation of Origin-Destination Matrix: Using Optimization Toolbox (MATLAB)