

ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ
МАТЕМАТИКИ
имени М.В. Келдыша
РОССИЙСКОЙ АКАДЕМИИ НАУК

А.Б. Гончаренко, В.Т. Жуков

АНАЛИЗ ДВУХ МОДЕЛЕЙ ДВИЖЕНИЯ НАСЕЛЕНИЯ

Москва, 2005

А.Б. Гончаренко, В.Т. Жуков. АНАЛИЗ ДВУХ МОДЕЛЕЙ ДВИЖЕНИЯ НАСЕЛЕНИЯ

Email: al-gonch@yandex.ru, zhukov@kiam.ru

АННОТАЦИЯ

Рассмотрены две математические модели мобильности групп населения. Для определения параметров моделей ставятся задачи минимизации нелинейных функционалов. Обсуждаются вычислительные трудности решения этих задач и строятся алгоритмы их решения. Приводится анализ работоспособности моделей на разных типах движения населения:

A.B. Goncharenko, V.T. Zhukov. Analysis of two model of population mobility.

ABSTRACT

Two mathematical models of population mobility are considered. For definition of model parameters the problems of functional minimization are set. Computational difficulties of numerical solution are discussed and the algorithms are constructed. Efficiency of these model are studied for a few types of group mobility.

СОДЕРЖАНИЕ

Введение	3
Параметризация моделей мобильности	3
1.1. Интегральный показатель качества жизни (различия групп)	3
1.2. Интегральный показатель сходства групп	5
1.3. Построение «pull-push» моделей движения населения.	6
Оценка параметров моделей движения населения	9
2.1. Метод оценки параметров.....	9
2.2. Выбор метода решения задачи минимизации	11
2.3. Алгоритм решения.....	14
2.4. Методика проверки качества моделей	19
2.5. Методика проверки гипотез относительно полученных оценок параметров	20
Полученные результаты	22
3.1. Используемые данные и общие результаты.	22
3.2. Относительное качество жизни в различных группах.	24
3.3. Сходство групп	26
3.4. Комплексный анализ, основанный на сходстве и различии групп.	29
Заключение	1
Литература	1

Введение

Существует достаточно большое количество моделей движения населения и трудовых ресурсов, опирающихся на самые различные предположения о природе межгрупповых потоков и предназначенных для получения большого спектра выводов о подвижности населения. Как правило, в них включают в качестве параметров различные характеристики групп:

- непосредственно относящиеся к населению (естественный прирост, доля городского и сельского населения, национальный состав, степень этнической однородности и т.п.);
- характеристики уровня экономического развития (отраслевая структура хозяйства, размеры капитальных вложений в производственную и непроизводственную сферы, темпы роста продукции, уровень развития транспортной сети и т.п.);
- характеристика уровня жизни населения (средний уровень текущей и ожидаемой заработной платы, различия в уровне заработной платы сравниваемых регионов, уровень занятости и уровень безработицы регионов входа и выхода, уровень развития культуры, здравоохранения, образования, торговли и обслуживания, жилищная обеспеченность населения и т.п.);
- природные характеристики (показатели климатических, почвенных и гидрологических условий и т.п.);
- характеристики миграционных связей в прошлом.

Однако при решении, каких-либо конкретных задач и проверке гипотез о степени влияния того или иного фактора (группы факторов) на процесс межгрупповых переходов возникают значительные трудности ввиду того, что далеко не все из вышеперечисленных характеристик поддаются формализации.

В предлагаемой работе рассматривается в некотором смысле «обратный» по отношению к вышеизложенному подход, а именно, предлагается по известной статистике о численности переходов оценить наиболее общие интегральные показатели, характеризующие подвижность населения, и уже по ним попытаться сделать выводы относительно природы межгрупповых потоков и свойств самих групп.

Параметризация моделей мобильности

1.1. Интегральный показатель качества жизни (различия групп)

Как уже отмечалось выше, многие характеристики групп не поддаются формализации, вследствие чего сравнение качества жизни в этих группах, выраженное через их характеристики, также крайне затруднено, так как нет подходящего способа его измерения. Но известно, что человек ощущает это качество и делает приемлемый для себя выбор. Одним из проявлений такого выбора является подвижность людей, их «голосование ногами» за лучшие условия. Поэтому, наблюдая и измеряя подвижность населения, можно сделать выводы об относительном качестве жизни в разных группах.

Построить интегральный показатель, позволяющий сравнивать уровни качества жизни в различных группах, можно на основе подхода используемого в теории поиска работы [16], где вводится предположение, что человек при принятии решения о совершении перехода сравнивает свое настоящее положение с возможным будущим, причем будущее обычно представляется случайным и характеризуется функцией распределения предложений другими группами. В теории поиска работы рассмотрен только доход и средняя выгода человека при переходе. Если предложение может быть описано многими характеристиками, то сравнение среднего улучшения всех характеристик затруднено, и лучше рассмотреть просто вероятность улучшения.

Далее каждый человек будет идентифицирован набором «условий жизни»^{*}, которыми он обладает, т.е. его доходом, уровнем жилья, квалификации и т.д. Пусть x_j - уровень j -го «условия жизни» ($j=1, \dots, m$) у случайно выбранного человека группы i ($i=1, \dots, k$). Любая группа i или совокупность всех ее членов может быть охарактеризована функцией распределения уровней всех m «условий» по членам группы $F_i(z_1, z_2, \dots, z_m)$. Вполне естественно предположить, что человек имеет возможность сравнивать свое положение (набор $\mathbf{x} = (x_1, x_2, \dots, x_m)$) с положением других людей. В теории для сравнения векторов часто используют функцию полезности $u(z_1, z_2, \dots, z_m)$. Таким образом, у любого случайно выбранного в группе человека, характеризуемого уровнями $\mathbf{x} = (x_1, x_2, \dots, x_m)$, полезность $U = u(\mathbf{x})$ будет случайной величиной с функцией распределения

$$G_i(U) = \int_{A(U)} dF_i(\mathbf{x}), \quad i=1, \dots, k, \quad (1)$$

где $A(U) = \{\mathbf{z} : u(\mathbf{z}) \leq U\}$.

Человек, обладающий уровнями «условий жизни» $\mathbf{x}$, (полезностью U), будет улучшать свои уровни $\mathbf{x}$ при переходе в группу j , если область притяжения (в терминах «условий жизни») будет равна множеству $A(U) = \{\mathbf{z} : u(\mathbf{z}) > U\}$. Поэтому вероятность улучшения его качества жизни (полезности U для фиксированного человека) при переходе в группу j будет равна $q_j(U) = \int_U^\infty dG_j(x) = 1 - G_j(U) = \bar{G}(U)$, так как качество жизни в новой группе j лучше, чем у человека, обладающего полезностью U , лишь у доли людей в группе j , равной $1 - G_j(U)$. Так как человек из группы i может быть выбран произвольно, то относительное качество жизни в группе j лучше, чем у людей из группы i , равно $Eq_j(U)$, т.е.

$$q_{ij} = \int \bar{G}_j(U) dG_i(U) \quad \text{для любых } i, j = 1, \dots, k. \quad (2)$$

^{*} под «условиями жизни» понимается вся совокупность условий, в которых протекает процесс перехода между группами, при этом предполагается, что человек способен сравнивать «условия жизни» и принимать на основе этого сравнения решения о переходе из группы в группу.

Полученный показатель можно трактовать как интегральный показатель, характеризующий привлекательность группы j для группы i .

q_{ij} обладает следующими полезными свойствами:

Свойство q1. $q_{ij} > 0$ и $q_{ij} + q_{ji} = 1$ $i, j = 1, \dots, k$, [15], [13].

Свойство q2. Величины q_{ij} инвариантны относительно любого фиксированного монотонного преобразования полезностей в группах i и j [13].

Таким образом, показатель q_{ij} позволяет однозначно определить соотношение между группами с точки зрения их взаимной привлекательности (относительного уровня качества жизни в них).

По типу действия q_{ij} можно условно считать показателем «различия» между группами: чем больше привлекательность (отличие) группы j по отношению к группе i (чем выше качество жизни в группе j по сравнению с i), тем больше поток из i в j , и, соответственно, меньше из j в i (в силу свойства q1).

1.2. Интегральный показатель схождения групп

Однако в рамках данного подхода (в терминах функции полезности) остаются неучтенными такие факторы («условия жизни») как, например расстояние между регионами, языковое сходство и т.д., которые обладают тем свойством, что невозможно определить предпочтение одной группы по отношению к другой на их основе (например, нельзя утверждать, что регион «А» более предпочтителен, чем регион «В» в силу того, что между ними расстояние «Х», так как с помощью того же аргумента можно показать обратное). Все эти факторы отвечают одному условию: они симметричны относительно групп входа и выхода. Как только становится возможным определить отношение предпочтения между двумя группами с помощью такого фактора (например, язык в группе «А» стал более предпочтителен, чем язык в группе «В» при принятии решения индивидуумом о переходе), то автоматически этот фактор учитывается в функции полезности и, соответственно, в интегральном показателе предпочтения q_{ij} .

Следовательно, необходимо ввести еще один обобщенный показатель, характеризующий симметричное влияние «условий жизни» на межгрупповые потоки - d_{ij} , обладающий:

Свойство d1. $d_{ij} = d_{ji}$, $i, j = 1, \dots, k$

Не ограничивая общности, можно принять, что $0 \leq d_{ij} \leq 1$ и поток из i в j , и из j в i тем больше, чем больше d_{ij} . Следовательно, по типу действия d_{ij} можно условно считать показателем «сходства» между группами: чем больше сходство – тем больше поток.

Например, подобным образом действует «расстояние» между группами (с точностью до наоборот, чем больше расстояние – тем меньше поток), также обладающее свойством симметрии и достаточно часто включаемое в различные модели движения населения.

1.3. Построение «pull-push» моделей движения населения

Процедура построения показателей «сходства» и «различия» групп автоматически задает классификацию всех «условий жизни» (окружающей среды, в которой происходят переходы людей между группами). Согласно [2] классификация как процедура является частным случаем логической операции деления объёма понятий, которая должна отвечать следующим правилам: а) члены деления должны исключать друг друга (не должны пересекаться, быть частью другого); б) деление должно быть соразмерным (объём делимого понятия должен быть равен сумме объёмов членов деления - видов, классов); в) на каждом этапе деления должно использоваться одно и то же основание (некая совокупность существенных признаков). В нашем случае полученная классификация является простейшей дихотомией, у которой в качестве основания деления выбран принцип симметричности или асимметричности влияния «условий жизни» на потоки между группами. Проверка на предмет удовлетворения вышеприведенным трем правилам является тривиальной.

Таким образом, на всем множестве «условий жизни» задана классификация, разбивающая их на два класса. Для каждого класса построен интегральный показатель, обладающий вполне определенными аналитическими свойствами. Следовательно, поток из группы i в группу j может быть полностью описан с помощью двух показателей, учитывающих весь набор «условий жизни»: q_{ij} - характеризующего различие в относительном качестве жизни в группе j по сравнению с группой i и d_{ij} - отражающего степень сходства групп i и j . Причем поток из группы i в группу j тем больше, чем выше качество жизни в группе j по сравнению с i и чем больше сходство групп i и j . Простейшую модель, отражающую данные свойства, можно представить в виде:

$$r_{ij} = q_{ij}d_{ij}, \quad i, j = 1, \dots, k, \quad i \neq j \quad (3)$$

где, r_{ij} - интенсивность потока из группы i в j (всего групп k), q_{ij} и d_{ij} - введенные выше параметры, обладающие свойствами $q_{ij} + q_{ji} = 1$ и $d_{ij} = d_{ji}$.

Таким образом, мы имеем систему, состоящую из $k(k-1)$ уравнений, $k(k-1)/2$ неизвестных параметров q_{ij} и $k(k-1)/2$ неизвестных параметров d_{ij} . Если принять грубое предположение о детерминированности данной модели (в действительности исходные данные и параметры являются случайными величинами и необходимо рассматривать статистические модели и искать оценки параметров), то интенсивность переходов можно представить как отношение численности перешедших людей (за определенное время) из

группы i в j - v_{ij} к численности исходной группы i - n_i , тогда в явном виде можно получить выражения для:

$$q_{ij} = \frac{n_j v_{ij}}{n_i v_{ji} + n_j v_{ij}} \text{ и } d_{ij} = \frac{n_i v_{ji} + n_j v_{ij}}{n_i n_j}. \quad (4)$$

Как показано выше, все факторы, связанные со сравнением групп, уже учтены в интегральных показателях q_{ij} и d_{ij} . В том числе в них же учтены и тривиальные факторы, которые характеризуют только группу выхода (безотносительно к конечной группе) и только группу входа (безотносительно к исходной). Таким факторам можно дать следующую экономическую интерпретацию: человек для совершения перехода помимо потенциального желания перейти, основанного на сравнении качества жизни людей из исходной группы с качеством жизни в конечной, и оценки степени сходства групп, должен также иметь возможность покинуть исходную группу и возможность попасть в конечную, иначе, несмотря на высокую привлекательность и сходство переход может не состояться.

Поэтому, с содержательной точки зрения, при моделировании переходов людей между группами необходим учет в явном виде параметров: a_i - характеризующих способность человека покинуть группу и b_j - отражающих простоту попадания в группу (т.е. ее доступность, насыщенность населения и рынка трудовых ресурсов).

Таким образом, если удастся очистить наблюдаемые статистическими органами межгрупповые потоки от ограничений, связанных с доступностью групп и способностью их покинуть, то появляется возможность оценки истинных относительных предпочтений (различий) групп и выявления групп, более подходящих друг к другу (сходных групп), которые отражаются в переходах людей.

Для достижения поставленной цели была использована хорошо известная pull-push модель поведения человека при выборе им направления своего перехода, в которой вероятность перехода отдельного человека определяется тремя группами условий:

- 1) способен ли человек покинуть исходную группу (например, способны ли люди этой группы реализовать свою собственность);
- 2) доступна ли для человека выбранная им группа (т.е. обстановка в выбранной группе благоприятна - он найдёт там работу, жильё и т.п.);
- 3) улучшатся ли условия жизни человека в новой группе после перехода по сравнению с условиями исходной (т.е. относительное качество жизни исходной группы хуже выбранной) и/или насколько сходны между собой исходная и конечная группы.

Третье условие, как уже отмечалось выше, определяет тягу людей, входящую в виде некоторого параметра в вероятности переходов. В первых pull-push моделях, применяемых при исследовании только миграции между

районами проживания, эти параметры связывались только близостью районов-групп, т.е. с расстоянием при переезде. Затем при применении моделей для анализа трудовой мобильности со сходством условий жизни в группах, объединяющих людей по месту работы. В последнем случае для определения близости (сходства) вводились новые характеристики, например, заработки, определяющие не только географическое расстояние. Смысл расстояния при определении потоков оставался прежним: чем меньше расстояние (больше сходство условий), тем больше поток между группами. Однако несколько иначе изменяет потоки качество жизни - оно подобно доходу может быть близким по величине и разным по направлению действия, т.е. теперь это не сходство условий, а их различие. Как сходство условий жизни, так и различие в качестве жизни влияют на потоки людей, но если сходство отражает симметрию потоков (чем ближе условия двух групп - тем больше потоки между ними), то различие качества жизни скорее их асимметрию (чем больше различие - тем больше поток в одном направлении и меньше в противоположном).

Таким образом, может быть сделано следующее предположение: интенсивность переходов из группы i в группу j пропорциональна параметрам привлекательности q_{ij} , сходства d_{ij} , способности покинуть исходную группу - a_i и доступности конечной группы - b_j . Однако теперь мы должны оценить всего $k(k-1)+2k$ неизвестных параметров ($k(k-1)/2$ параметров q_{ij} , $k(k-1)/2$ - d_{ij} , k - a_i и k - b_j) по $k(k-1)$ известным численностям переходов между группами v_{ij} . Решение данной задачи с помощью предложенного в данной работе подхода является невозможным, так как число оцениваемых параметров превышает число уравнений. Поэтому предлагается рассмотреть отдельно модель, позволяющую оценить относительное качество жизни в группах и отдельно модель, отражающую сходство групп. Такое упрощение позволит достичь основной цели данной работы – выявить относительные предпочтения групп и определить их степень сходства. Но при таком подходе мы, к сожалению, получим трудноинтерпретируемые оценки параметров a_i и b_j , так как в первом случае они будут искажены тем, что мы не учитываем в модели сходство групп, а во втором - их различие.

Итак, получены две модели:

относительного качества жизни в группах (различия) – $r_{ij} = a_i b_j q_{ij}$,

(I)

и сходства групп - $r_{ij} = a_i b_j d_{ij}$,

(II)

где $0 \leq a_i \leq 1$, $0 \leq b_j \leq 1$, $0 \leq q_{ij} \leq 1$, $q_{ij} + q_{ji} = 1$, $0 \leq d_{ij} \leq 1$, $d_{ij} = d_{ji}$

Теперь остается определить параметры движения людей по наблюдаемым данным о межгрупповых переходах. Но эти данные не могут быть

детерминированными, поэтому об определении параметров говорить не следует, а параметры движения людей можно лишь оценить на основе статистических данных. Методу оценки параметров посвящена следующая часть.

Оценка параметров моделей движения населения

2.1. Метод оценки параметров

Одним из наиболее обоснованных методов оценки параметров является метод максимального правдоподобия. В его основе лежит известный вид функции распределения наблюдаемых случайных величин, которыми здесь являются численности переходов между группами. В соответствии с работами [5] и [13] можно утверждать, что потоки между группами i и j за фиксированный промежуток времени независимы между собой и распределены по закону Пуассона. Известно, что параметром закона Пуассона будет среднее значение $\nu_{ij} = n_i r_{ij}$ - численности перешедших (за время наблюдения) из группы i в j (всего групп k). Поэтому модели (I) и (II) можно записать в виде:

$$E \nu_{ij} = n_i r_{ij} = n_i (a_i b_j q_{ij} + c_i \delta_{ij}), \quad 0 \leq a_i \leq 1, \quad 0 \leq b_j \leq 1, \quad 0 \leq q_{ij} \leq 1, \quad q_{ij} + q_{ji} = 1, \quad (1.I)$$

$$E \nu_{ij} = n_i r_{ij} = n_i (a_i b_j d_{ij} + c_i \delta_{ij}), \quad 0 \leq a_i \leq 1, \quad 0 \leq b_j \leq 1, \quad 0 \leq d_{ij} \leq 1, \quad d_{ij} = d_{ji}, \quad (1.II)$$

где межгрупповые потоки $\nu_{ij} \sim Pu(n_i r_{ij})$ и независимы $\forall i, j$:

$$Pu(n_i r_{ij}) = P(\nu_{ij} = m_{ij}) = \frac{(n_i r_{ij})^{m_{ij}}}{m_{ij}!} e^{-n_i r_{ij}}, \quad m_{ij} = 0, 1, 2, \dots \quad - \text{ численность наблюдаемых}$$

переходов из группы i в j .

Идея метода максимального правдоподобия состоит в том, что находятся такие значения неизвестных параметров распределения, которые давали бы наибольшую вероятность наблюдаемых случайных величин. Для разных i и j выпишем вероятность наблюдения величин ν_{ij} - функцию правдоподобия. В силу независимости всех ν_{ij} вероятность наблюдения их совокупности равна произведению вероятностей наблюдения каждого отдельного ν_{ij} , т.е.

$$\text{для модели с «различием»} - L(\mathbf{a}, \mathbf{b}, \mathbf{q}) = \prod_{i=1}^k \prod_{\substack{j=1 \\ i \neq j}}^k \frac{[n_i a_i b_j q_{ij}]^{\nu_{ij}}}{\nu_{ij}!} e^{-n_i a_i b_j q_{ij}} \quad (2.I)$$

$$\text{для модели со «сходством»} - L(\mathbf{a}, \mathbf{b}, \mathbf{d}) = \prod_{i=1}^k \prod_{\substack{j=1 \\ i \neq j}}^k \frac{[n_i a_i b_j d_{ij}]^{\nu_{ij}}}{\nu_{ij}!} e^{-n_i a_i b_j d_{ij}} \quad (2.II)$$

Далее будет рассмотрена не точка максимума функций правдоподобия (2) при рассматриваемых ограничениях, а точка минимума логарифма функций $L(\mathbf{a}, \mathbf{b}, \mathbf{q})$ и $L(\mathbf{a}, \mathbf{b}, \mathbf{d})$, взятая со знаком минус:

$$l(\mathbf{a}, \mathbf{b}, \mathbf{q}) = - \sum_{i=1}^k \sum_{\substack{j=1 \\ i \neq j}}^k (\nu_{ij} \ln n_i + \nu_{ij} \ln a_i + \nu_{ij} \ln b_j + \nu_{ij} \ln q_{ij} - \ln \nu_{ij}! - n_i a_i b_j q_{ij}) \quad (3.I)$$

$$l(\mathbf{a}, \mathbf{b}, \mathbf{d}) = - \sum_{i=1}^k \sum_{\substack{j=1 \\ i \neq j}}^k (v_{ij} \ln n_i + v_{ij} \ln a_i + v_{ij} \ln b_j + v_{ij} \ln d_{ij} - \ln v_{ij}! - n_i a_i b_j d_{ij}) \quad (3.II)$$

Нахождение минимума функций (3.I) и (3.II) при сделанных ограничениях является задачей нелинейного программирования.

Решение данной задачи для (3.II) найти достаточно легко, несмотря на ее многоэкстремальность и невыпуклость, обусловленную наличием слагаемых $n_i a_i b_j d_{ij}$. Действительно, сделав замену переменных: $A_i = \ln a_i$, $B_j = \ln b_j$, $D_{ij} = \ln d_{ij}$, получаем выпуклую задачу с выпуклыми ограничениями:

$$l(\mathbf{A}, \mathbf{B}, \mathbf{D}) = \sum_{i=1}^k \sum_{\substack{j=1 \\ i \neq j}}^k (v_{ij} \ln n_i + v_{ij} A_i + v_{ij} B_j + v_{ij} D_{ij} - \ln v_{ij}! - n_i \exp(A_i + B_j + D_{ij})) \quad (4.II)$$

$A_i \leq 0$, $B_j \leq 0$, $D_{ij} \leq 0$, $D_{ij} = D_{ji}$, для которой существует и единственно решение.

Многоэкстремальность и невыпуклость функции (4.I) так легко обойти не удастся ввиду более сложного (асимметричного) вида ограничений на параметры q_{ij} . Попытка свести исходную задачу к выпуклой путем поиска соответствующего преобразования неизбежно приводит к значительному усложнению вида ограничивающих условий, что делает полученную задачу не менее сложной, чем исходная. Если же пытаться выписать двойственную задачу к исходной или уже преобразованной, то мы получим функционал, нахождение точной верхней грани которого подразумевает решение системы нелинейных уравнений аналогичных исходной задаче в явном виде, что тоже сделать не удастся.

Поэтому в данной работе предпринимается попытка найти решение задачи (3.I) с соответствующими ограничениями в исходном виде. К сожалению, в данном случае нам не удастся строго доказать существование, единственность и устойчивость решения, однако мы можем оценить качество рассматриваемой модели по фактическим данным и хотя бы таким образом апостериорно подтвердить или опровергнуть корректность полученного решения.

Существенным в данной ситуации является определение начального приближения, которое, учитывая многоэкстремальный характер минимизируемой функции, должно выбираться достаточно близко к точке глобального минимума (иначе большинство методов может в качестве решения дать значения локальных минимумов). Успех решения задачи (3.I) (и (4.II)) во многом определен достаточно обоснованным выбором начального приближения. Для параметров q_{ij} и d_{ij} , в качестве начального приближения предлагается брать рассчитанные по фактическим данным значения интегрального показателя относительного предпочтения групп и сходства групп из обобщенной модели (не учитывающей параметры доступности и способности выйти из группы)

$$q_{ij}^0 = \frac{n_j v_{ij}}{n_i v_{ji} + n_j v_{ij}}, \quad d_{ij}^0 = \frac{n_i v_{ji} + n_j v_{ij}}{n_i n_j}. \quad (5)$$

Для параметров a_i и b_j начальные приближения берутся как средние интенсивности выхода и входа:

$$a_i^0 = r_i = \frac{1}{k} \sum_{j=1}^k \frac{v_{ij}}{n_i}, \quad b_j^0 = r_j = \frac{1}{k} \sum_{i=1}^k \frac{v_{ij}}{n_i}. \quad (6)$$

Для решения задачи минимизации избавимся от ограничения типа равенства в условиях задачи (3.1), переписав функцию правдоподобия в виде:

$$l = - \sum_{i=1}^k \left(\sum_{j=i+1}^k (v_{ij} \ln a_i + v_{ij} \ln b_j + v_{ij} \ln q_{ij} - n_i a_i b_j q_{ij}) + \right. \\ \left. + \sum_{j=1}^{i-1} (v_{ij} \ln a_i + v_{ij} \ln b_j + v_{ij} \ln(1 - q_{ji}) - n_i a_i b_j (1 - q_{ji})) \right) + c \quad (7.1)$$

из ограничений останутся тогда только $0 \leq a_i \leq 1$, $0 \leq b_j \leq 1$, $0 \leq q_{ij} \leq 1$ (сумма членов, не зависящих от параметров a_i , b_j , q_{ij} обозначена через c).

2.2. Выбор метода решения задачи минимизации

Все методы многоэкстремальной оптимизации можно условно разделить на две группы. Для первых существуют какие-либо точные утверждения об их сходимости к глобальному минимуму, для вторых приходится ограничиваться некоторыми правдоподобными рассуждениями об их разумном поведении в многоэкстремальной ситуации, такие методы условно можно назвать эвристическими. Несмотря на свою обоснованность точные методы, как правило, имеют незначительную практическую ценность, поэтому далее мы будем рассматривать различные варианты эвристических методов.

В многоэкстремальных задачах приходится отличать истинные значения глобального минимума от различных стационарных точек (в которых градиент равен нулю): локальных минимумов, максимумов и седловых точек. Рассмотрим, как в этой ситуации работают основные методы минимизации.

Градиентный метод в задаче поиска минимума многоэкстремальной функции в окрестности невырожденного минимума (все собственные значения гессиана положительны [11]) при определенных условиях (выборе подходящей длины шага) сходится к этому минимуму независимо от того, является он локальным или глобальным. Учитывая наличие ограничений в нашей задаче, при решении ее с помощью градиентного метода нужно пользоваться его непосредственным обобщением – методом проекции градиента [14].

В нашем случае этот метод приобретает достаточно простой вид:

$$x^{s+1} = (x^s - \gamma \nabla l(x^s))_0^1, \quad (8)$$

где через x^s - обозначена s -я итерация вектора переменных: a_i , $i = 1, \dots, k$; b_j , $j = 1, \dots, k$; q_{ij} , $i < j = 1, \dots, k$;

градиент функции (7.1) - $\nabla l(x)$:

$$\begin{aligned}\frac{\partial l}{\partial a_i} &= -\sum_{j=i+1}^k \left(\frac{v_{ij}}{a_i} - n_i b_j q_{ij} \right) - \sum_{j=1}^{i-1} \left(\frac{v_{ij}}{a_i} - n_i b_j (1 - q_{ji}) \right), \quad i = 1, \dots, k, \\ \frac{\partial l}{\partial b_j} &= -\sum_{i=1}^{j-1} \left(\frac{v_{ij}}{b_j} - n_i a_i q_{ij} \right) - \sum_{i=j+1}^k \left(\frac{v_{ij}}{b_j} - n_i a_i (1 - q_{ji}) \right), \quad j = 1, \dots, k, \\ \frac{\partial l}{\partial q_{ij}} &= -\left(\frac{v_{ij}}{q_{ij}} - n_i a_i b_j \right) + \left(\frac{v_{ji}}{1 - q_{ij}} - n_j a_j b_i \right), \quad i < j = 1, \dots, k,\end{aligned}\tag{9}$$

выражение $(x)_0^1 = \begin{cases} x, & 0 \leq x \leq 1 \\ 0, & x < 0 \\ 1, & x > 1 \end{cases}$ обозначает проектор на множество заданных в

рассматриваемой задаче ограничений; параметр γ задает длину шага.

В наиболее простых вариантах градиентного метода параметр γ принимается постоянным и для обеспечения сходимости в общем случае должен, например, удовлетворять условию $0 < \gamma < 2/L$, где L - константа из условия Липшица для градиента функции: $\|\nabla l(x) - \nabla l(y)\| \leq L\|x - y\|$, оценка, которой является отдельной задачей и для многоэкстремальных функций выполнима не всегда. Понятно, что в данном случае конструктивный выбор γ крайне затруднен.

Применение этого метода в чистом виде для различных γ (в частности для вышеуказанного условия, в котором L выбиралась как оценка верхней границы спектра матрицы Гессе в различных точках), как и следовало ожидать, не принесло успеха. Однако следующий простейший прием выбора подходящей длины шага на каждой итерации в редких случаях приводит к решению задачи. А именно, задаются произвольные $0 < \varepsilon < 1$, $0 < \alpha < 1$ и некоторое γ . На каждой итерации вычисляется $l(x^s - \gamma \nabla l(x^s))$ и проверяется неравенство $l(x^s - \gamma \nabla l(x^s)) \leq l(x^s) - \varepsilon \gamma \|\nabla l(x^s)\|^2$. Если оно выполняется, то $x^{s+1} = (x^s - \gamma \nabla l(x^s))_0^1$, если же нет, то γ заменяется на $\gamma \alpha$ и проверка повторяется. Как правило, для нахождения подходящей длины шага требуется конечное число дроблений γ . Однако, главный недостаток градиентного метода – его медленную сходимость для плохо обусловленных задач обойти с помощью такого приема не удастся.

Попытка несколько увеличить скорость сходимости за счет использования информации, полученной на предыдущих итерациях – так называемые многошаговые методы, также не принесла удовлетворительных результатов. В методе «тяжелого шарика» для рассматриваемой задачи крайне затруднен конструктивный выбор соответствующих параметров, обуславливающих его свойство «проскакивать» локальные минимумы и способность более быстро двигаться по дну «оврага» вблизи плохообусловленной точки минимума за счет «инерции» (см. [14])

Более обоснованный способ выбора параметров, определяющих направления спуска, реализован в методе сопряженных градиентов. В нем параметры находятся из решения вспомогательной оптимизационной задачи. Однако, решаемая нами задача обладает ограничениями и в том случае, если точка минимума лежит вне нашей допустимой области, скорость сходимости по-прежнему остается невысокой. Таким образом, данный метод может ускорить сходимость только в случае, если точка минимума лежит внутри области, задаваемой ограничениями, что априорно конечно тоже не известно.

Следующая группа методов основана на квадратичной аппроксимации. Однако, все эти методы, и в частности, наиболее простой - классический метод Ньютона, чувствительны к выбору начального приближения. Необходимость явного обращения матрицы Гессе на каждой итерации значительно увеличивает объем вычислений и время расчета.

Примерная структура матрицы Гессе для рассматриваемой задачи приведена ниже:

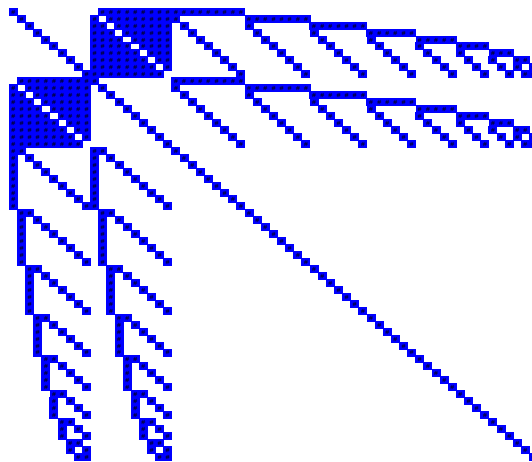


Рис. 1

Данная матрица является разреженной, но в практически важных случаях ее обращение, необходимое для реализации метода Ньютона, представляет собой непростую задачу. Обращать матрицу на каждой итерации с помощью стандартной техники, основанной на различных видах декомпозиции (LU, QR и т.д.), и представленной в пакете ОХ (с помощью которого производились расчеты), получалось далеко не всегда. Поэтому от методов, основанных на явном обращении гессиана, пришлось отказаться.

На практике достаточно большое распространение получили так называемые квазиньютоновские методы. В основе этого класса методов лежит идея восстановления квадратичной аппроксимации функции по значениям ее градиентов в ряде точек. Тем самым методы объединяют достоинства градиентного метода (не требуется вычисление матрицы вторых производных) и метода Ньютона (быстрая сходимость вследствие использования квадратичной аппроксимации).

Квазиньютоновские методы имеют общую структуру:

$$x^{s+1} = x^s - \gamma_s H_s \nabla l(x^s), \quad (10)$$

где матрица H_s пересчитывается рекуррентным способом на основе информации, полученной на s -ой итерации, так что $H_s - [\nabla^2 l(x^s)]^{-1} \rightarrow 0$. Таким образом, методы в пределе переходят в ньютоновский, что и объясняет их название. Отметим некоторые общие свойства методов такого типа. При любых равномерно положительно определенных H_s метод (15) обладает глобальной сходимостью (для многоэкстремальных функций застревает в любой стационарной точке), а при условии: $\|H_s - [\nabla^2 l(x^*)]^{-1}\| \rightarrow 0$, где x^* - невырожденная точка минимума $l(x)$ и $l(x)$ дважды непрерывно дифференцируема в окрестности x^* , в окрестности минимума метод сходится со сверхлинейной скоростью [11].

Наибольшее распространение получил метод Бroyдена-Флетчера-Гольдфаба-Шенно (БФГШ):

$$H_{s+1} = H_s + \frac{\rho_s p^s (p^s)^T - p^s (y^s)^T H_s - H_s y^s (p^s)^T}{(y^s, p^s)}, \quad (11)$$

где $\rho_s = \gamma_s + \frac{(H_s y^s, y^s)}{(y^s, p^s)}$, $H_0 > 0$; $p^s = -H_s \nabla l(x^s)$, $y^s = \nabla l(x^{s+1}) - \nabla l(x^s)$.

Однако, как уже отмечалось выше, этот метод чувствителен к выбору начального приближения, что ограничивает его применимость. Сходимость этого класса методов существенно локальная.

2.3. Алгоритм решения

Таким образом, поскольку наиболее универсальные методы решения задач многоэкстремальной оптимизации обладают, как мы видели, серьезными недостатками, то это вынуждает обращаться к более простым и грубым методам.

Заслуживает отдельного рассмотрения метод простых итераций. Хотя его скорость сходимости и невысока (не превышает линейную - скорость сходимости геометрической прогрессии) он обладает рядом достоинств. Во-первых, в нем не накапливается ошибка вычислений (неточность очередного приближения отражается только на числе итераций, а не на точности окончательного результата) и, во-вторых, он допускает более гибкую настройку для решения конкретных задач.

Рассмотрим необходимые условия минимума (7.1), т.е. приравняем нулю частные производные l по a_i , b_j и q_{ij} :

$$\begin{aligned} \frac{\partial l}{\partial a_i} &= -\sum_{\substack{j=1 \\ i \neq j}}^k \left(\frac{v_{ij}}{a_i} - n_i b_j q_{ij} \right) = 0, & \frac{\partial l}{\partial b_j} &= -\sum_{\substack{i=1 \\ i \neq j}}^k \left(\frac{v_{ij}}{b_j} - n_i a_i q_{ij} \right) = 0, \\ \frac{\partial l}{\partial q_{ij}} &= -\left(\frac{v_{ij}}{q_{ij}} - n_i a_i b_j \right) + \left(\frac{v_{ji}}{1 - q_{ij}} - n_j a_j b_i \right) = 0. \end{aligned} \quad (12)$$

Таким образом, необходимо решить полученную систему нелинейных уравнений относительно $k(k-1)/2$ неизвестных q_{ij} ($i < j$), k неизвестных a_i и k неизвестных b_j .

Прежде всего, сделаем замену переменных: $A_i = \ln a_i$, $B_j = \ln b_j$, $Q_{ij} = \ln q_{ij}$. Далее преобразуем систему (12) к следующему виду:

$$A_i = \ln \left[\frac{\sum_{j=1, j \neq i}^k v_{ij}}{\sum_{j=1, j \neq i}^k n_j e^{B_j + Q_{ij}}} \right], B_j = \ln \left[\frac{\sum_{i=1, i \neq j}^k v_{ij}}{\sum_{i=1, i \neq j}^k n_i e^{A_i + Q_{ij}}} \right], Q_{ij} = \ln \left[\frac{B_{ij} - \sqrt{B_{ij}^2 - 4A_{ij}v_{ij}}}{2A_{ij}} \right], i < j, \quad (13)$$

где $A_{ij} = n_i e^{A_i + B_j} - n_j e^{A_j + B_i}$, $B_{ij} = v_{ij} + v_{ji} + A_{ij}$, с ограничениями $A_i \leq 0$, $B_j \leq 0$, $Q_{ij} \leq 0$.

Выражение для Q_{ij} в (13) получено из условия

$$\frac{\partial l}{\partial q_{ij}} = - \left(\frac{v_{ij}}{q_{ij}} - n_i a_i b_j \right) + \left(\frac{v_{ji}}{1 - q_{ij}} - n_j a_j b_i \right) = 0 \quad \text{в системе (12), как решение}$$

соответствующего квадратного уравнения. Причем выбран его корень, всегда принимающий положительные значения (действительно, если $A_{ij} < 0$, то выражение под корнем в выражении для Q_{ij} в (13) положительно и само значения корня получается больше чем B_{ij} , соответственно, числитель дроби получается также отрицательным, для случая $A_{ij} > 0$ - рассуждения аналогичны).

Представим систему (13) в общем виде следующим образом:

$$x_l = \varphi_l(x_1, \dots, x_n), \quad l = 1, \dots, n, \quad (14)$$

где через x_l обозначены соответствующие переменные A_i , B_j и Q_{ij} и $n = k(k-1)/2 + 2k$.

Соответственно метод простых итераций будет выглядеть следующим образом:

$$x_l^{s+1} = \varphi_l(x_1^s, \dots, x_n^s), \quad l = 1, \dots, n. \quad (15)$$

Как оказалось, обеспечить сходимость данного метода можно с помощью элементарного тождественного преобразования, заключающегося в добавлении к правой и левой частям уравнения (14) переменной x_l и последующего деления обеих частей на 2

$$x_l = \frac{1}{2}(x_l + \varphi_l(x_1, \dots, x_n)), \quad l = 1, \dots, n. \quad (16)$$

В итоге получится эквивалентная система, для которой метод простых итераций:

$$x_l^{s+1} = \frac{1}{2}(x_l^s + \varphi_l(x_1^s, \dots, x_n^s)), \quad l = 1, \dots, n. \quad (17)$$

сходится (здесь опять же мы не можем отличить локальный минимум от глобального, попадание в глобальный минимум обеспечивается обоснованным

с содержательной точки зрения выбором начального приближением и возможностью последующей экономической интерпретации полученных значений параметров).

Идея данного преобразования достаточно проста – попытаться уменьшить норму матрицы производных. Для того чтобы исследовать сходимость данных итераций, нужно выяснить, какие значения принимает норма матрицы производных вблизи решения системы (14). Если норма матрицы производных меньше единицы, то норма вектора погрешностей убывает по геометрической прогрессии, что означает линейную сходимость метода. В данном случае показать в явном виде уменьшение нормы матрицы производных метода (16) по сравнению с (14) в окрестности точки минимума является достаточно сложной задачей, так как нет возможности выписать в явном виде само решение. Однако данное преобразование позволило решить задачу на всех рассмотренных реальных данных, в то время как процедура (14) – практически всегда расходилась.

Таким образом, простейшее преобразование привело к сходимости рассматриваемой задачи, но сходимость по-прежнему осталась достаточно медленной. Несколько ускорить этот процесс можно с помощью поправки Эйткена по трем последним итерациям:

$$x_l^{s+1} = \frac{1}{2}(x_l^s + \varphi_l(x_1^s, \dots, x_n^s)) - \frac{(x_l^s - x_l^{s-1})^2}{2x_l^{s-1} - x_l^s - x_l^{s-2}}, \quad l = 1, \dots, n \quad (18)$$

(такой метод называется ускорением Стефенсона (см. [10])). Также сходимость метода можно улучшить, используя найденное приближение x_l^{s+1} не только для вычисления следующей итерации, а используя его уже на данной итерации для вычисления следующих компонент [7].

Учитывая ограничения, накладываемые на задачу, в окончательном варианте метод приобретает вид:

$$x_l^{s+1} = \left(\frac{1}{2}(x_l^s + \varphi_l(x_1^{s+1}, \dots, x_{l-1}^{s+1}, x_l^s, \dots, x_n^s)) - \frac{(x_l^s - (x_l^{s-1})^2)}{2x_l^{s-1} - x_l^s - x_l^{s-2}} \right)_{-\infty}^0. \quad (19)$$

Так как метод (19) все же является модифицированным методом простых итераций, то оканчивать итерации можно по критерию сходимости [7] (проверяется одновременное выполнение для каждой компоненты):

$$\frac{(x_l^s - x_l^{s-1})^2}{|2x_{li}^{s-1} - x_l^s - x_l^{s-2}|} < \varepsilon. \quad (20)$$

Можно дополнительно попытаться несколько ускорить сходимость метода в случае, если траектория попала в окрестность решения, где уже будет работать какой-либо более быстрый алгоритм. А именно, предлагается через каждые n итераций (конструктивный выбор значения n в данной ситуации затруднен, поэтому для разных реальных данных этот параметр подбирался исходя из эвристических соображений, основанных на характере сходимости итераций), делать обновление - несколько пробных шагов по какому-либо более быстрому

методу, основанному на квадратичной аппроксимации, если этот метод не находит улучшения по какому-либо направлению, то возвращаемся обратно к методу простых итераций. Наиболее универсальными и хорошо себя зарекомендовавшими являются квазиньютоновские методы. Здесь мы в частности предлагаем использовать метод Бroyдена-Флетчера-Гольдфаба-Шенно (BFGS) [36]. Для ряда задач, в которых решение лежит внутри допустимого множества или хотя бы достаточно близко к его границе, данный метод значительно ускорил сходимость (как правило, это были задачи минимизации модели со сходством). Если же точка минимума лежит за границами ограничений - все методы сходятся очень медленно.

Необходимо отметить, что получаемые в процессе оценки системы нелинейных уравнений имеют достаточно большую размерность - $k(k-1)/2 + 2k$, где k - количество групп. Практически все получаемые на фактических данных задачи являются плохо обусловленными, что существенным образом отражается на количестве итераций. Ниже приведена таблица с параметрами сходимости для рассчитываемых в данной работе данных:

Таблица 1

Вид движения	k - кол-во групп	Кол-во нормальных уравнений	Модель с различием		Модель со сходством	
			Кол-во итераций	Число обусловленности матрицы Гессе*	Кол-во итераций	Число обусловленности матрицы Гессе
Межотраслевое, Латвия	12	90	1373	9.94E+03	509	1.14E+08
Миграция, СССР 67	15	135	186410	1.98E+07	304	2.18E+11
Миграция, СССР 73	15	135	35314	1.82E+08	337	3.64E+10
Миграция, СССР 89	15	135	8091	1.29E+10	426	4.34E+10
Миграция, Россия 96	12	90	53137	4.48E+06	160	6.17E+10
Миграция, Россия 97	12	90	85030	5.64E+08	183	1.76E+10

* Под числом обусловленности матрицы Гессе ($\nabla^2 l$) подразумевается отношение - $\mu = S/s$ (где S и s - верхняя и нижняя границы спектра матрицы Гессе соответственно) в найденной точке минимума, характеризующее обусловленность (устойчивость) данной задачи минимизации [11]. Большие значения данного показателя свидетельствуют об «овражном» характере исследуемой функции в некоторой окрестности точки минимума и, соответственно, ее плохой обусловленности.

Миграция, Россия 98	12	90	49989	7.44E+06	169	7.17E+10
Миграция, Россия 99	12	90	25015	4.10E+06	139	2.94E+11

Миграция, Нидерланды 75	40	860	3027	4.04E+10	198	1.28E+11
Миграция, Нидерланды 80	40	860	2641	1.06E+08	188	1.02E+12
Миграция, Нидерланды 85	40	860	1211	5.40E+10	175	2.70E+10
Миграция, Нидерланды 90	40	860	1230	2.02E+08	187	8.04E+11
Миграция, Нидерланды 95	40	860	1238	1.87E+11	196	8.26E+11

Все расчеты проводились с помощью консольной версии (свободно распространяемой для академических исследований и учебных целей) пакета ОХ (объектно-ориентированного матричного языка, разработанного в Nuffield College, University of Oxford для эконометрических расчетов, более подробная информация доступна по адресу - <http://www.nuff.ox.ac.uk/Users/Doornik/>).

2.4. Методика проверки качества моделей

Для проверки качества рассматриваемых моделей можно использовать традиционные для этих целей критерии: коэффициент множественной корреляции $\hat{R}^2$, скорректированный $\hat{R}_{adj}^2$ и F - критерий Фишера. Все эти тесты основываются на определенных предположениях регулярности для рассматриваемых случайных величин – нормальности и независимости распределения, а также независимости дисперсии от параметра распределения. Однако дисперсия случайной величины, имеющей пуассоновское распределение, равна значению параметра (см., например, [12]), т.е. не постоянна. Соответственно, несмотря на то, что в пределе пуассоновски распределенная случайная величина имеет нормальное распределение, мы должны скорректировать случайные величины ν_{ij} , то есть найти преобразование, стабилизирующее дисперсию, что автоматически будет означать ее независимость от параметра. Для пуассоновского распределения таким преобразованием является $1/2\sqrt{\nu_{ij}}$ [12]. Следовательно, если $\hat{a}_i, \hat{b}_j$ и $\hat{q}_{ij}$ ($\hat{d}_{ij}$) - состоятельные оценки и случайная величина ν_{ij} - имеет пуассоновское распределение, то при $n \rightarrow \infty$ случайные величины $\frac{1}{2}(\sqrt{\nu_{ij}} - \sqrt{n_i \hat{a}_i \hat{b}_j \hat{q}_{ij}})$ для модели с различием ($\frac{1}{2}(\sqrt{\nu_{ij}} - \sqrt{n_i \hat{a}_i \hat{b}_j \hat{d}_{ij}})$ - для модели со сходством) асимптотически нормальны с математическим ожиданием равным 0, дисперсией 1 и независимы.

С учетом сделанных замечаний получаем: $\hat{R}^2 = 1 - \frac{s_1}{s_0}$, где

$$s_0^2 = \sum_{i=1}^k \sum_{j \neq i}^k (\frac{1}{2} \sqrt{v_{ij}} - \bar{v})^2, \quad \bar{v} = \frac{1}{k(k-1)} \sum_{i=1}^k \sum_{j \neq i}^k \frac{1}{2} \sqrt{v_{ij}}, \quad s_1^2 = \sum_{i=1}^k \sum_{j \neq i}^k (\frac{1}{4} (\sqrt{v_{ij}} - \sqrt{n_i \hat{a}_i \hat{b}_j \hat{q}_{ij}})^2 \quad \text{для}$$

модели с различием и $s_1^2 = \sum_{i=1}^k \sum_{j \neq i}^k \frac{1}{4} (\sqrt{v_{ij}} - \sqrt{n_i \hat{a}_i \hat{b}_j \hat{d}_{ij}})^2$ для модели со сходством.

$\hat{R}^2$ обладает тем недостатком, что он зависит от количества оцениваемых параметров, в некоторой степени этого удастся избежать с помощью нормирования числом параметров знаменателя и числителя дроби: $\hat{R}_{adj}^2 = 1 - \frac{s_1^2}{s_0^2} \frac{t-1}{l}$, где $t = k(k-1)$ - общее количество наблюдений в случае, когда внутригрупповых переходов нет; $l = k(k+3)/2$ - количество неизвестных параметров,

F - критерий Фишера: $F = \frac{\hat{R}^2}{1 - \hat{R}^2} \frac{t-l-1}{l}$ - при больших n_i будет иметь

распределение, близкое к F -распределению (Фишера) с l и $l-t-1$ степенью свободы.

Часто потоки между группами округляются, т.е. эти показатели приводятся с различным числом значащих цифр. Для различных систем групп (стран, моментов исследования и т.д.) данные не одинаковы, т.е. они бывают приведены с разными масштабными множителями. Следовательно, критерий сравнения качества полученных оценок параметров модели движения населения, должен быть инвариантным к масштабу статистических данных. Выполнение этого свойства позволит сравнивать движение людей в различных системах групп (регионах стран, межотраслевое, профессиональное, социальное движение и т.д.) как между собой, так и в разные периоды времени. Все три рассмотренных критерия удовлетворяют данному требованию.

2.5. Методика проверки гипотез относительно полученных оценок параметров

Для проверки ряда гипотез относительно входящих в рассматриваемые модели параметров нам понадобятся дисперсии полученных оценок. В [12] отмечалось, что при выполненных в данном случае условиях регулярности оценки максимального правдоподобия имеют асимптотически нормальное распределение и являются состоятельными и асимптотически эффективными оценками. В частности из этого следует, что для определения ковариационной матрицы оценок можно использовать неравенство Рао-Крамера, в котором связывается информационная матрица с ковариационной матрицей оценок:

$$D\hat{\theta} \geq - \left(E \frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right)^{-1}, \quad (21)$$

где θ_i - неизвестные параметры - a_i , b_j и q_{ij} или d_{ij} , а l - функция правдоподобия.

Данное неравенство задает нижнюю границу для ковариационной матрицы полученных оценок, с другой стороны оценки, полученные с помощью метода максимального правдоподобия, обладают свойством минимальности дисперсии, что позволяет использовать правую часть (21) – матрицу обратную информационной, в качестве оценки для дисперсий оценок максимального правдоподобия.

Согласно [12] мы можем считать, что оценки максимального правдоподобия для данной задачи асимптотически независимы. Следовательно, можем принять в качестве оценки дисперсий параметров моделей значения обратные диагональным элементам информационной матрицы:

$$\sigma^2_{\hat{a}_i} = \left[\sum_{\substack{j=1 \\ i \neq j}}^k \frac{n_i b_j q_{ij}}{a_i} \right]^{-1}, \quad \sigma^2_{\hat{b}_j} = \left[\sum_{\substack{i=1 \\ i \neq j}}^k \frac{n_i a_i q_{ij}}{b_j} \right]^{-1},$$

$$\sigma^2_{\hat{q}_{ij}} = \left[\frac{n_i a_i b_j}{q_{ij}} + \frac{n_j a_j b_i}{1 - q_{ij}} \right]^{-1}, \quad \sigma^2_{\hat{d}_{ij}} = \left[\frac{n_i a_i b_j}{d_{ij}} + \frac{n_j a_j b_i}{d_{ij}} \right]^{-1}. \quad (22)$$

Теперь мы можем проверить гипотезу о значимости различия между оценками среднего параметров a_i , b_j и q_{ij} , d_{ij} .

Так, например, для проверки основной гипотезы $H: q_{ij} = q_{ji}$ против альтернативы $H: q_{ij} \neq q_{ji}$, нужно рассмотреть разность оценок $\hat{q}_{ij} - \hat{q}_{ji}$. Чтобы эта разность была распределена по стандартному нормальному распределению, пронормируем ее среднеквадратичным отклонением, тогда случайная величина

$\frac{\hat{q}_{ij} - \hat{q}_{ji}}{\sqrt{\sigma^2_{\hat{q}_{ij}} + \sigma^2_{\hat{q}_{ji}}}}$ будет иметь стандартное нормальное распределение (в силу

свойства асимптотической нормальности оценок максимального правдоподобия [12]) и гипотеза $H: q_{ij} = q_{ji}$ отвергается при уровне значимости

α , если $\frac{|\hat{q}_{ij} - \hat{q}_{ji}|}{\sqrt{\sigma^2_{\hat{q}_{ij}} + \sigma^2_{\hat{q}_{ji}}}} > u_\alpha$ (где u_α - критическая граница стандартного нормального распределения).

Для оценки значимости мнения мигрантов из i об относительном качестве жизни в s по сравнению с l , можно проверять гипотезу $H: q_{il} = q_{is}$. Ее

следует отвергать при уровне значимости α , если $\frac{|\hat{q}_{il} - \hat{q}_{is}|}{\sqrt{\sigma^2_{\hat{q}_{il}} + \sigma^2_{\hat{q}_{is}}}} > u_\alpha$. В частности для оценки значимости мнения мигрантов из i об относительном качестве

жизни в j по сравнению с i и наоборот достаточно сравнить q_{ij} с $1/2$, соответственно гипотеза $H: q_{ij} = 1/2$ отвергается, если $\frac{|\hat{q}_{ij} - 1/2|}{\sigma_{\hat{q}_{ij}}} > u_{\alpha}$.

Для оценки значимости мнения мигрантов о сходстве групп i и j можно сравнивать аналогичным образом d_{ij} с 1 - гипотеза $H: d_{ij} = 1$ отвергается, если $\frac{|\hat{d}_{ij} - 1|}{\sigma_{\hat{d}_{ij}}} > u_{\alpha}$.

Полученные результаты

3.1. Используемые данные и общие результаты

Расчеты были сделаны для межреспубликанского движения в СССР в 1967[8], 1973[9] и 1989[6] годах, межрайонной миграции в России в 1996 - 1999 г., межрегиональной миграции в Нидерландах в 1975 - 1995 г., и межотраслевого движения трудовых ресурсов в Латвии в 1973 г. Выбор такого набора данных был продиктован стремлением продемонстрировать работоспособность моделей на разных типах движения населения: в России миграционные процессы характеризуются большими расстояниями между регионами и большим количеством населения в них с высоким уровнем расслоения; в Нидерландах - небольшие по численности близко расположенные друг от друга регионы; в Латвии также небольшие по численности группы, но уже принципиально другой тип движения населения - межотраслевая миграция. В виду громоздкости получаемых таблиц с параметрами приведем лишь значения статистик критериев - $\hat{R}^2$, $\hat{R}_{adj}^2$, F для рассматриваемых моделей.

Таблица 2

Вид Движения	Модель с различием			Модель со сходством		
	R^2	R_{adj}^2	F^* Статистика	R^2	R_{adj}^2	F^* Статистика
Межотраслево е, Латвия	0,96	0,87	10,20	0,95	0,83	7,86
Миграция, СССР 67	0,99	0,97	43,81	0,99	0,97	45,75
Миграция, СССР 73	0,99	0,96	40,24	0,99	0,98	75,47
Миграция, СССР 89	0,98	0,94	25,46	≈ 1	0,99	201,52

Миграция, Россия 96	0, 94	0,8 0	6,75	≈ 1	0,9 9	171,60
Миграция, Россия 97	0, 94	0,8 1	6,92	≈ 1	0,9 9	121,58
Миграция, Россия 98	0, 94	0,8 1	6,84	≈ 1	0,9 9	166,90
Миграция, Россия 99	0, 94	0,8 0	6,68	≈ 1	0,9 9	165,75
Миграция, Нидерланды 75	0, 79	0,5 3	3,07	0, 99	0,9 8	80,26
Миграция, Нидерланды 80	0, 78	0,5 1	2,84	0, 99	0,9 8	80,97
Миграция, Нидерланды 85	0, 77	0,4 9	2,75	0, 99	0,9 8	86,42
Миграция, Нидерланды 90	0, 79	0,5 2	2,99	0, 99	0,9 8	85,30
Миграция, Нидерланды 95	0, 76	0,4 7	2,60	0, 99	0,9 7	69,26

* - Число степеней свободы F статистики $l = k(k+3)/2$ и $t = k(k-1)$ для межотраслевого движения трудовых ресурсов в Латвии равно - $l = 90$ и $t = 132$, для межреспубликанского движения в СССР - $l = 135$ и $t = 210$, для межрайонной миграции в России - $l = 90$ и $t = 132$ и для межрегиональной миграции в Нидерландах - $l = 860$ и $t = 1560$. Значения p -уровней для F статистики (значимости отличия полученных с помощью модели оценок от фактических данных) для всех моделей не превышают 0.00001 (т.е. отличие модели от фактических данных незначимо), поэтому, чтобы не перегружать таблицу здесь они не приводятся.

Во всех случаях значения статистик $\hat{R}^2$, $\hat{R}_{adj}^2$, F свидетельствуют о достаточно высоком качестве обеих моделей, что косвенно подтверждает правильность найденного решения. Модель с различием наиболее хорошо описывает межреспубликанское движения в СССР (процент объясненной моделью дисперсии превышает 94%), для остальных видов движения этот показатель несколько ниже. Модель со сходством наиболее адекватна для межрайонной миграции в РФ (более 99% объясненной дисперсии). Модель с различием более хорошо относительно модели со сходством объясняет межотраслевое движение в Латвии ($F_{\text{(модель с различием)}} = 10,20$ против $F_{\text{(модель со}}$

сходством) = 7,86) и примерно на одном уровне обе модели описывают миграцию в СССР 1967 г., в остальных случаях предпочтительней оказалась модель со сходством.

Также можно заметить, что для межреспубликанского движения в СССР качество модели с различием убывает с годами, в то время как модель со сходством становится лучше (для миграции в РФ это различие стало еще более выраженным). Данный эффект свидетельствует о всевозрастающем влиянии стоимости переезда и расстояния, что влечет за собой сужение рынка труда.

3.2. Относительное качество жизни в различных группах.

Способ построения параметра q и предыдущие гипотезы позволяют ранжировать рассматриваемые группы по принципу привлекательности, что в данном случае можно трактовать как более высокий уровень качества жизни в одних группах по сравнению с другими. Мы можем последовательно выделять Парето-оптимальные группы, не доминируемые в смысле предпочтения заданного с помощью параметров q_{ij} никакими другими группами (т.е. выделить группу, наиболее привлекательную для всех остальных групп, затем группу, наиболее привлекательную для всех остальных, кроме первой и так далее). Как отмечалось выше, для сравнения двух групп i и j нам достаточно проверить гипотезу о значимости отличия q_{ij} от $1/2$ - т.е. мы можем считать, что группа j более предпочтительна, чем i в случае если выполняется

неравенство $\frac{\hat{q}_{ij}-1/2}{\sigma_{\hat{q}_{ij}}} > u_{\alpha}$, где u_{α} - критическая граница стандартного нормального распределения для некоторого уровня значимости α , т.е.

$P\left(\frac{\hat{q}_{ij}-1/2}{\sigma_{\hat{q}_{ij}}} > u_{\alpha}\right) = \alpha$ и $\frac{\hat{q}_{ij}-1/2}{\sigma_{\hat{q}_{ij}}} \in N(0,1)$. Однако вполне может возникнуть ситуация

когда однозначно упорядочить группы не удастся. Так, например, найдутся группы i , j и l такие, что группа i более привлекательна для группы j , а группа j более привлекательна для группы l , которая в свою очередь более привлекательна для группы i . Подобные случаи можно интерпретировать как циклы, поэтому отношение привлекательности между группами удобнее будет представлять в виде графа.

Ниже приведены соответствующие результаты для рассматриваемых потоков. Направление стрелки определяет предпочтение одних групп по отношению к другим (так, например, в 1967 г. Молдавия была значимо более привлекательна для Литвы, Латвии и Эстонии, которые в свою очередь были эквивалентны - значения параметров предпочтения между этими республиками не значимо отличались от $1/2$ - эту ситуацию отражает двунаправленная стрелка). Следует отметить, что мы рассматриваем истинные предпочтения, очищенные от ограничений, связанных с доступностью групп и способностью

их покинуть, поэтому направление стрелки не означает наличие неограниченного потока из менее привлекательной группы в группу с большей привлекательностью. На межгрупповые потоки оказывает также влияние возможность покинуть исходную группу и доступность конечной группы, которые отражаются в соответствующих показателях a_i и b_j .

Межреспубликанское движение в СССР, 1967г.

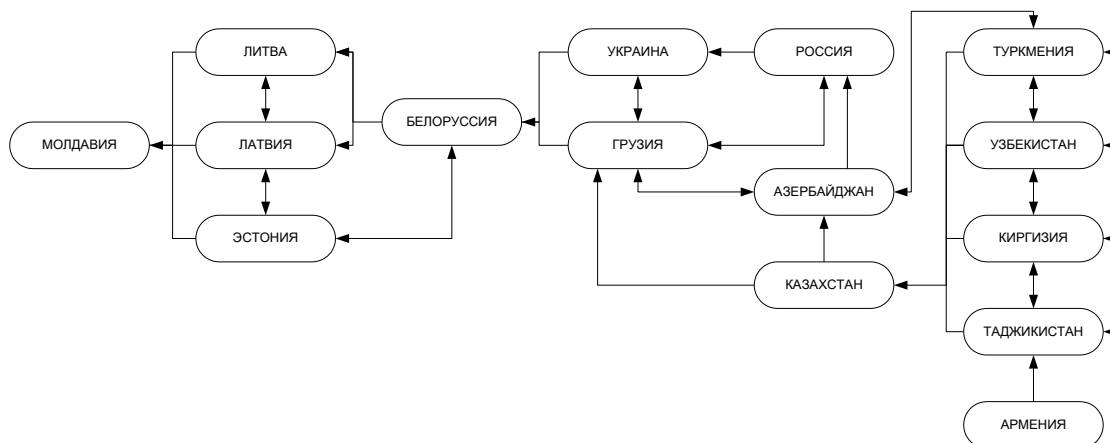


Рис. 2

Межреспубликанское движение в СССР, 1989г.

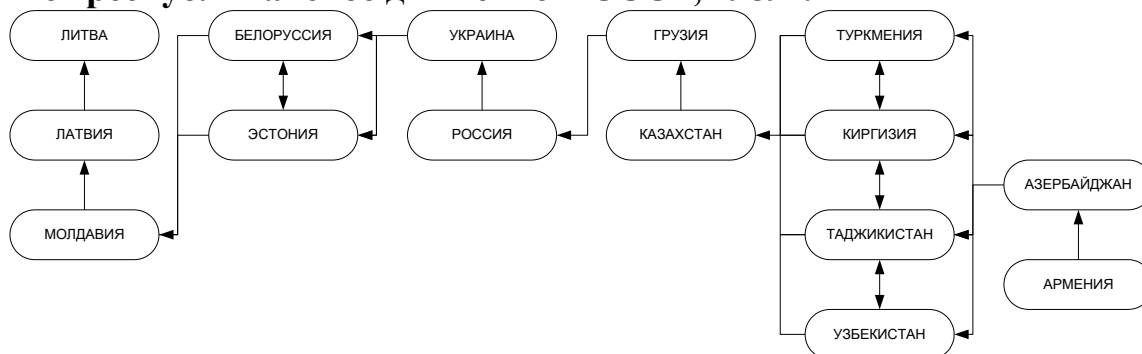


Рис. 3

На приведенных диаграммах (рис. 2, 3) видно, что на протяжении всех лет достаточно устойчивым был цикл (эквивалентность), который образовывали среднеазиатские республики. В 1967 г. был замечен цикл из России, Украины и Грузии, который к 1989 г. распался. Наиболее трудно интерпретируемые результаты касаются Молдавии и Армении. Оценка Молдавии в 1967г., возможно, завышена, однако в [4] отмечено, что на 1960-1969 г. приходилось наибольшее положительное сальдо межреспубликанской миграции для этой республики. Приток специалистов для народного хозяйства был обусловлен благоприятными климатическими условиями и относительно высоким уровнем социально-экономического развития. Однако поток в Молдавию был ограничен, и судя по относительно низкому значению показателя доступности $b_j=0.004$, можно предположить, что насыщенность региона была достаточно высока. С другой стороны, возможность выхода из республики была практически не ограничена - $a_i=1$. Поэтому, несмотря на

абсолютную привлекательность по сравнению с остальными республиками, прирост численности населения в Молдавии не был лавинообразным, так как не каждый мог туда попасть (требовались только специалисты народного хозяйства) и каждый практически «беспрепятственно» мог из республики уехать (т.е. помимо отсутствия каких-либо ограничений на выезд, также имелись средства на переезд и возможность найти место вне Молдавии, где мигранта готовы были принять).

С другой стороны, оценка Армении явно занижена. Опять же в [4] указывается, что на миграционные процессы в Армении оказывали влияние высокие показатели урбанизации и уровня образования титульного этноса, географическая удаленность от основного «миграционного поля» СССР (межреспубликанские миграционные связи Армении традиционно ограничивались в основном соседними республиками – Грузией и Азербайджаном), трудоизбыточность (так же как и в других республиках Закавказья и Средней Азии), богатые отходнические традиции, наличие обширной зарубежной диаспоры. Все это в совокупности видимо и определило наблюдаемый эффект. Удаленность от «миграционного поля» ограничивала как приток мигрантов, так и выход из республики, а высокий уровень развития в сочетании с трудоизбыточностью делали республику менее привлекательной ввиду наличия высокой конкуренции и повышенных требований к квалификации потенциального мигранта.

Также во всех рассматриваемых периодах у Армении были очень низкие значения склонности к выходу из группы и высокая «ненасыщенность» (практически полная доступность - $b_j=1$). Однако в 1989 г. склонность к выходу значительно возросла, что видимо, явилось следствием Спитакского землетрясения и событий в НКАО. В этом же году резко снизилась привлекательность Азербайджана, также принимавшем участие в Карабахском конфликте.

3.3. Сходство групп

Как уже отмечалось, параметр d_{ij} в модели со сходством можно интерпретировать как величину обратнопропорциональную обобщенному «расстоянию» или стоимости перехода из группы i в j (более точно, так как в рассматриваемой модели межгрупповые потоки принимаются пропорциональными параметру d_{ij} , то в качестве обобщенного «расстояния» нужно брать величину $\ln(1/d_{ij})$). Однако следует отметить, что для этой модели не всегда выполняется неравенство треугольника. Это затрудняет содержательную интерпретацию полученных результатов как расстояния между группами. Так, например, в евклидовой метрике получается, что в 1967 г. «расстояние» между Казахстаном и Литвой было значимо больше суммы «расстояний» между Казахстаном и Грузией и между Грузией и Литвой, подобные ситуации можно найти для всех приведенных выше данных. В

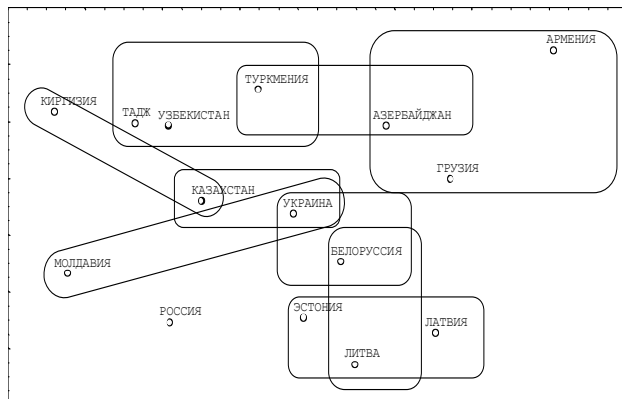
частности из этого можно сделать вывод, что при моделировании миграции с помощью различных параметров, отражающих конкретные характеристики групп (см. например [37], [38], [15]), введение в модель одного лишь симметричного параметра расстояния является явно не достаточным. Необходим также учет дополнительных характеристик, обладающих свойством симметрии.

Для анализа групп с помощью показателей d_{ij} наиболее естественным было бы попытаться расположить их на плоскости, чтобы можно было оценить степень сходства-различия между ними визуально. Наиболее подходящим инструментом для данной задачи является метод многомерного шкалирования [1], который как раз позволяет располагать в пространствах небольшой размерности (в данном случае двумерном) исследуемые объекты, максимально сохраняя «реальные расстояния» между ними (этот метод воспроизводит не количественные меры сходств объектов, а лишь их относительный порядок, что вполне отвечает требованиям поставленной задачи). Ниже приведены результаты его применения на данных по межреспубликанскому движению в СССР 1967 и 1989 гг. Замкнутой линией обведены не значимо различные регионы (регионы i и j можно считать значимо различными, если

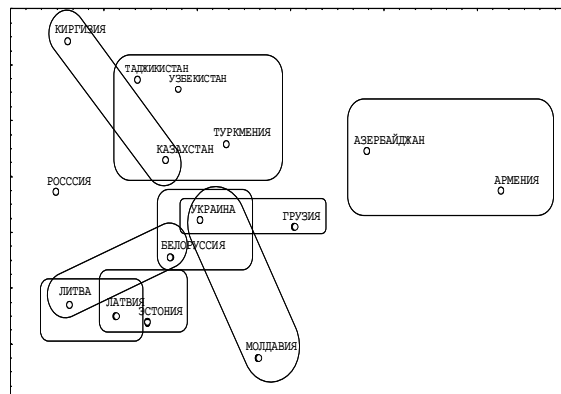
выполняется неравенство $\frac{1-\hat{d}_{ij}}{\sigma_{\hat{d}_{ij}}} > u_{\alpha}$, где u_{α} - критическая граница стандартного нормального распределения для некоторого уровня значимости α , т.е.

$$P\left(\frac{1-\hat{d}_{ij}}{\sigma_{\hat{d}_{ij}}} > u_{\alpha}\right) = \alpha \text{ и } \frac{1-\hat{d}_{ij}}{\sigma_{\hat{d}_{ij}}} \in N(0,1).$$

Межреспубликанское движение в СССР 1967г.



Межреспубликанское движение в СССР 1989г.

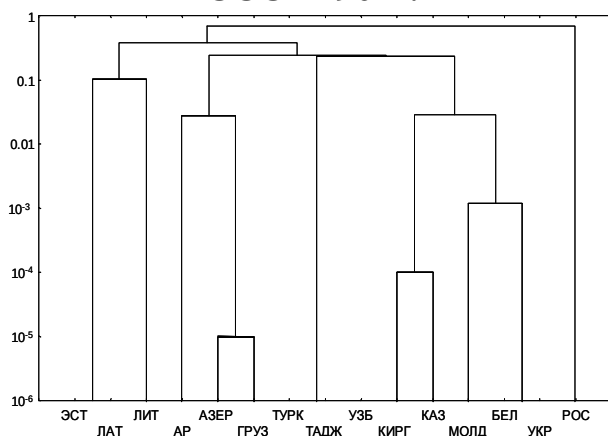


Как мы видим, сходные регионы (объединенные замкнутой линией) достаточно далеко отстоят друг от друга на плоскости, что является следствием невыполнения неравенства треугольника для рассматриваемого «обобщенного расстояния». Однако общее представление о структуре сходства различных регионов из этой диаграммы получить все-таки можно. Например, отчетливо проявляется сходство в группах среднеазиатских и прибалтийских республик. Достаточно наглядное подтверждение нашел высказанный в [4] тезис об удаленности Армении от основного «миграционного поля» СССР. Также вполне хорошо данная модель показала кризис 1989 г. в отношениях между Арменией и Азербайджаном. На диаграмме видно, как достаточно сильно отделились (в смысле «обобщенного расстояния») эти республики от остальных и стали сходно-неприемлемыми для большинства мигрантов.

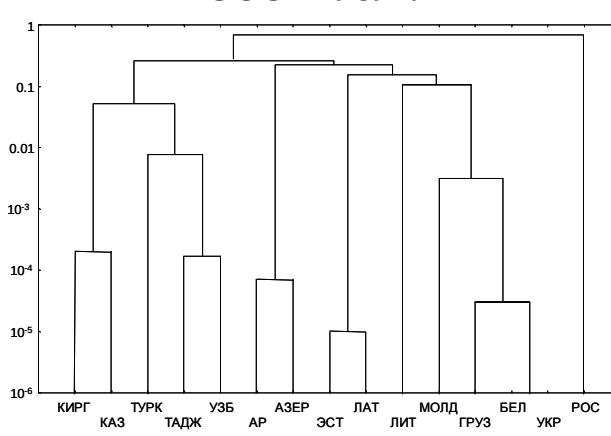
Помимо представленных диаграмм, для визуального анализа сходства групп также удобно использовать кластерный анализ [1]. Ниже представлен ряд диаграмм, полученных с помощью иерархического (древовидного) кластерного анализа. Назначение этого метода состоит в объединении объектов (в нашем случае групп) в кластеры, используя некоторую меру сходства или «расстояние» между объектами. Типичным результатом такой кластеризации является иерархическое дерево. О степени сходства групп можно судить по тому, на каком уровне «расстояния» («расстояние» между группами откладывается на вертикальной оси) они объединяются в кластеры.

Для определения сходства групп предлагается использовать метод «ближайшего соседа». В этом методе расстояние между двумя кластерами определяется расстоянием между двумя наиболее близкими группами (ближайшими соседями) в различных кластерах. Ниже приведены соответствующие примеры для межреспубликанского движения в СССР 1967 и 1989 гг.

Межреспубликанское движение в СССР 1967г.



Межреспубликанское движение в СССР 1989г.



На диаграмме, соответствующей 1967г. отчетливо видны кластеры среднеазиатских и прибалтийских республик, а также республик Закавказья. В 1989г. структура сходства несколько изменилась: Грузия выпала из кластера закавказских республик, а Литва из кластера прибалтийских, уменьшилось взаимное сходство среднеазиатских республик, хотя они и остались в одном кластере, также к ним добавились Киргизия и Казахстан, которые в 1967г. были более сходны с Молдавией, Украиной и Белоруссией, нежели с остальными среднеазиатскими республиками.

3.4. Комплексный анализ, основанный на сходстве и различии групп.

В заключении еще раз отметим необходимость комплексного подхода к анализу межгрупповых потоков, основанного как на сходстве, так и на различии групп. В частности, предложенные в данной работе модели позволяют провести в некоторой степени подобный анализ. А именно, наибольшего эффекта можно добиться, если анализировать регионы сразу по двум направлениям. Так, например, на диаграммах многомерного шкалирования видно, что в 1989г. Азербайджан с Арменией значительно отделились от остальных республик (т.е. их сходство с остальными республиками уменьшилось). Но на основании одного только расстояния нельзя сказать, в какую сторону произошло это движение, в лучшую или худшую. На этот вопрос можно ответить с помощью графа, на котором отражены различия (предпочтения) регионов. В 1989г. Армения с Азербайджаном имели наименьшую привлекательность среди всех остальных республик, в то время как в предыдущие года уровень предпочтения данных республик не был столь низок (Азербайджан был более привлекателен по сравнению со среднеазиатскими республиками, ситуация с Арменией описана выше). Т.е. привлекательность этих республик для мигрантов из остальных групп упала. Таким образом, с помощью показателя привлекательности мы

можем определить предпочтения в выборе мигрантами регионов (направление движения), а при помощи показателя сходства – оценить расстояние (стоимость перехода) между этими регионами.

Заключение

Предлагаемый подход позволяет выявлять сходство и различие в качестве жизни для любой пары групп в том виде, в котором оно представляется человеку. Таким образом, получен результат сопоставления условий жизни в различных группах основанный только лишь на знании численностей межгрупповых переходов и не требующий анализа большого числа взаимозависимых и зачастую трудноформализуемых факторов.

Переходы людей позволяют получать информацию даже о политической обстановке, что в полной мере проявилось в движении между республиками СССР в 1989 году. В этот период отношения между Арменией и Азербайджаном были очень далеки от приемлемых для людей, что тут же сказалось на их предпочтениях. Эти республики оказались на последних местах по качеству жизни.

Практические выводы о движении населения в полной мере сделать пока невозможно, поскольку в данной работе изучалась проблема получения метода, который давал бы возможность выделения сходства и различия в качестве жизни в группах по реакции на них самих людей, выражающейся в выборе лучших условий при переходах. В силу этого расчёты проводились только для того, чтобы понять, какую информацию можно извлечь из регистрируемых статистическими органами переходов людей между группами и, в частности, для подтверждения корректности найденного решения рассматриваемой в данной работе оптимизационной задачи.

Литература

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. - М.: ЮНИТИ, 1998.
2. Бутс Б., Дробышевский С., Кочеткова О., Мальгинов Г., Петров В., Федоров Г., Хехт А., Шеховцов А., Юдин А. Типология российских регионов. - М.: Институт экономики переходного периода – CEPRA, 2002.
3. Ван дер-Варден Б.Л. Математическая статистика. - М.: ИЛ, 1960.
4. Воробьева О.Д. Миграция населения. Выпуск 1. Приложение к журналу «Миграция в России». М.: 2001.
5. Д.Дж. Бартоломью. Стохастические модели социальных процессов. - М.: Финансы и статистика, 1985.

6. Демографический Ежегодник. - М.: Финансы и статистика, 1990.
7. Калиткин Н.Н. Численные Методы. - М.: Наука, 1978.
8. Механическое движение в СССР за 1967 г. Вестник статистики. 1968г. N10.
9. Миграция населения СССР за 1973 г. Вестник статистики. 1975г. N2.
10. Мэтьюз Д.Г., Финк К.Д. Численные методы. Использование Matlab. - М.: Издательский дом «Вильямс», 2001.
11. Поляк Б.Т. Введение в оптимизацию. - М.: Наука, 1983.
12. Рао С.Р. Линейные статистические методы и их применение. – М.: Наука, 1968
13. Староверов О.В. Комплексные факторы в моделях движения населения. Прикладной многомерный статистический анализ. 1978. Том 33. Вып. 4.
14. Федоренко Р. П. Приближенное решение задач оптимального управления. - М.: Наука, 1978
15. Dagum C. Inequality Measure between Income Distribution with Applications. Econometrica, 1980, v 43, N7.
16. Holmlund B. Labour Mobility. IUI, Stokholm, 1984.