

PACS 87.15.Aa

© 2008 г. В.А. МОЛЧАНОВ, канд. физ-мат. наук  
(Национальный институт здравоохранения, Хельсинки, Финляндия),

А.И. МИХАЛЬСКИЙ, канд. техн. наук  
(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

## ОЦЕНКА ДИНАМИКИ ФАКТОРОВ РИСКА МЕТОДОМ ДИНАМИЧЕСКОЙ РЕГРЕССИИ

Рассматривается задача оценки динамики факторов риска и других индикаторов здоровья по данным ряда популяционных обследований, проводимых в различные годы на одной и той же возрастной категории населения. В случае нелинейной временной зависимости средних значений изучаемых величин простая интерполяция показателей по возрастным группам приводит к некорректным оценкам из-за неучета когортной динамики. Описывается метод построения и идентификации динамической регрессионной модели здоровья популяции, основанный на оценке когортной динамики показателей, что позволяет прогнозировать ожидаемые уровни факторов риска и выяснять причинные связи. Работоспособность модели продемонстрирована на примере обработки результатов обследований, проведенных в 1982, 1987 и 1992 гг. в Северной Карелии (Финляндия).

### 1. Введение

Оценка текущего состояния здоровья популяции основывается на результатах одномоментных обследований, проводимых среди населения выбранного возрастного диапазона в течение короткого промежутка времени (нескольких месяцев). Такое обследование называется кросс-секционным. Динамика здоровья популяции может быть оценена по данным нескольких независимых кросс-секционных обследований, разнесенных по времени, например на 5 лет. При этом в каждом обследовании регистрируются изучаемые параметры (систолическое и диастолическое кровяное давление, уровень холестерина, индекс массы тела и др.), однако в разных обследованиях обследуются, вообще говоря, разные лица. Эта схема отличается от так называемых долговременных когортных исследований, в которых в различные моменты времени обследуются одни и те же дожившие до обследования индивидуумы [1–5]. Теоретически многоразовые индивидуальные обследования дают более надежную информацию о процессах, происходящих в исследуемой когорте (которая, однако, не представляет всю популяцию), поскольку позволяют непосредственно наблюдать изменение показателей как с возрастом, так и во времени. Наиболее известны в настоящее время долговременные когортные исследования, направленные на изучение факторов риска возникновения хронических заболеваний, старения и инвалидизации населения. Среди них Фрэмлингэмское исследование (Framingham Heart Study) [1], Балтиморское исследование старения (Baltimore Longitudinal Study of Aging, BLSA) [2], Национальное исследование инвалидизации (National Long Term

Care Surveys, NLTCS) [3], Исследование стареющих датских близнецов (Longitudinal Study of Aging Danish Twins, LSADT) [4]. Организация подобных обследований преследует фундаментальные научные цели и требует значительных затрат.

В отличие от когортных кросс-секционные обследования более экономичны. Обычно такие обследования организуются на регулярной основе для выяснения существующих тенденций в здоровье населения, оценки эффектов профилактических мероприятий и построения прогнозов. При анализе кросс-секционных данных оценка временных трендов в зависимости от возраста основывается на сравнении средних значений или процентилей изучаемых параметров, представленных распределениями по возрастным категориям на разные моменты обследования. Аналогично исследуются и соответствующие эпидемиологические индикаторы (распространенность в населении повышенного давления крови, уровня холестерина, избытка массы тела и пр.) [6–12]. Для практического применения разработаны различные модификации методов оценки временных изменений по возрастным показателям. К ним относится “тренд по линейной регрессии”, который является основным методом в выполнении проекта всемирной организации здравоохранения MONICA по изучению факторов риска сердечно-сосудистых заболеваний [7–11]. При этом сначала путем построения линейной регрессии оценивается временной тренд для каждой возрастной группы. При необходимости проводится преобразование изучаемого параметра. На следующем этапе полученные тренды агрегируются с фиксированными весами, чтобы получить стандартизованный по возрасту показатель. В [12] описана модификация этого подхода, при которой по всему массиву данных строится логистическая регрессия, что соответствует агрегации трендов с весами, соответствующими изучаемой популяции. Полученные агрегированные тренды исследуются методами корреляционного анализа для проверки гипотез, сформулированных в рамках проекта MONICA.

При проведении регулярных кросс-секционных обследований (скринингов) часто возрастной диапазон обследуемого контингента достаточно широк (40 и более лет), при том что интервал между последовательными скринингами составляет 3–10 лет. В некоторых исследованиях в юношеских возрастах используют возрастные группы в 5–8 лет, при этом скрининги проводятся ежегодно либо через год. В результате в одной и той же возрастной группе в разные годы наблюдения оказываются представители одной и той же когорты рождения, что приводит к смешиванию возрастных и когортных эффектов и искажению оценок. Для разделения подобных эффектов используется метод APC (age, period, cohort) [13, 14], пригодный для учета высокой степени изменчивости временных трендов в зависимости от возраста, когорты и времени [10]. Недостатком метода APC является линейная параметризация по возрасту, когорте и времени, используемая при его реализации.

Другой популярный в демографии метод – метод Ли-Картера основан на разделении возрастного и временного эффектов [15, 16]. При этом зависимость от возраста описывается непараметрически, в то время как зависимость от времени задается линейной функцией. Этот метод хорошо зарекомендовал себя при прогнозировании смертности [14], однако его применение к анализу других показателей, связанных со здоровьем, представляется ограниченным.

В статье рассматривается подход – метод динамической регрессии, не требующий задания параметрической модели для изучаемой характеристики состояния популяции и позволяющий строить оценки по результатам нескольких последовательных кросс-секционных обследований. Это достигается за счет рассмотрения популяции как динамического объекта, состоящего из совокупности различных когорт рождения. Структура организации наблюдений за состоянием популяции рассматривается как схема измерений на плоскости с координатами “возраст” и “время”, а динамические свойства когорт описываются когортными трендами, которые восстанавливаются с учетом неполноты наблюдений. Программная реализация метода

предусматривает стабилизацию оценок либо с учетом их гладкости, либо через их агрегирование по времени и возрасту. В статье приводятся результаты оценок динамики изменения индекса массы тела по реальным данным для населения Северной Карелии (Финляндия).

## 2. Модель динамики популяции

Создание математической модели является первым шагом при изучении реально-го объекта, которым в рассматриваемом случае является население и его здоровье. Эта модель позволяет осуществлять анализ реальных наблюдений и может использоваться при моделировании реального объекта для тестирования различных методов оценивания и прогнозирования. Построенная математическая модель должна допускать эффективную численную идентификацию по имеющимся данным, результаты которой могут быть использованы в практических целях для прогнозирования, проверки эффективности профилактических мероприятий и разработки оптимальной стратегии управления здоровьем населения. Некоторые подходы к построению подобных моделей рассмотрены в [17].

Естественный способ описания сложного динамического объекта заключается в представлении его в виде совокупности элементов, которые можно проследить во времени как некие физические реальности. В популяции изменения во времени состоят из изменений во времени в когортах одного года рождения, составляющих популяцию.

Через  $C$  обозначим область на плоскости:

$$C = \{(y, a) : y \in [y_{\min}, y_{\max}], a \in [a_{\min}, a_{\max}]\},$$

где  $y$  – календарное время,  $a$  – возраст, выраженный в годах. В общем случае состояние популяции, определенное на этом компакте, описывается набором случайных величин  $X_i$ ,  $i = 1, \dots, L$ , соответствующих набору измеряемых индикаторов, определенных в каждой точке  $(y, a)$  области  $C$ . Для описания метода динамической регрессии ограничимся рассмотрением одного показателя, примером которого может служить индекс массы тела (Body Mass Index, BMI). Для простоты описания динамической модели будем считать, что справедливо соотношение

$$(1) \quad X(y, a) = v(y, a) + \varepsilon,$$

где  $\varepsilon$  – случайный процесс, определенный на  $C$ , такой что  $E(\varepsilon) = 0$ ,  $D(\varepsilon) = \sigma^2$ ,  $\forall (y, a) : (y, a) \in C$ . Отсюда следует, что

$$v(y, a) = E(X(y, a)).$$

Изменение среднего значения индикатора  $X$  для когорты рождения в точке  $(y, a)$  по интервалу времени  $dt$  запишем в виде:

$$(2) \quad v(y + dt, a + dt) = v(y, a) + u(y, a)dt + o(dt), \quad o(dt)/dt \rightarrow 0 \quad \text{при} \quad dt \rightarrow 0.$$

Функция  $u(y, a)$  представляет собой скорость изменения рассматриваемого параметра вдоль движения когорты в результате внешнего воздействия, вызванного окружающей средой. Это внешнее воздействие в точке  $(y, a)$  не зависит от свойств когорты, проходящей во время  $y$  возраст  $a$ . Более того, теоретически сам факт существования внешнего воздействия не зависит от факта существования когорты. В этом заключается суть концепции теории поля, выраженной в терминах описываемой модели и первоначально введенной Куртом Левиным [18]. Для удобства будем пользоваться

терминами “средние уровни” или “уровни” для значений функции  $v(y, a)$  и “когортные тренды” или “С-тренды” для значений функции  $u(y, a)$ . В более сложной модели результат воздействия представляет собой сумму результатов внешнего воздействия, обусловленных средой, и компонента, обусловленного текущим состоянием когорты. В этом случае в уравнении (2) член  $u(y, a)$  заменяется суммой  $u(y, a) + bv(y, a)$ , где  $b$  является дополнительным модельным параметром.

Пусть  $v_0(y, a)$  обозначает граничную величину  $v(y, a)$  на границе области  $C$  для когорты, проходящей через точку  $(y, a)$ :

$$v_0(y, a) = v(y - \delta, a - \delta), \quad \text{где } \delta = \min(y - y_{\min}, a - a_{\min}).$$

Тогда  $v(y, a)$  можно представить в виде

$$(3) \quad v(y, a) = v_0(y, a) + \int_0^{\delta} u(y - t, a - t) dt.$$

Таким образом, значения функции  $v(y, a)$  могут быть вычислены в любой точке области  $C$  по заданным граничным условиям  $v_0(y, a)$  и когортному тренду  $u(y, a)$ .

Обобщение модели (1)–(2) на многомерный случай получается, если функции  $v(y, a)$  и  $u(y, a)$  рассматриваются как вектор-функции, соотношение  $D(\varepsilon) = \sigma^2$  заменяется соотношением  $Cov(\varepsilon) = \Sigma$ ,  $\Sigma$  обозначает диагональную матрицу, а в выражении  $u(y, a) + bv(y, a)$  параметр  $b$  представляет собой матрицу.

### 3. Идентификация модели динамики популяции

#### 3.1. Общая формулировка задачи

Рассмотрим случай, когда каждое индивидуальное наблюдение представлено в виде тройки  $(x_k, y_k, a_k)$ ,  $k = 1, \dots, K$ , где  $x_k$  – значение индикатора, определенное в момент времени  $y_k$  для возраста  $a_k$ . Задача идентификации модели динамики популяции заключается в оценке функций  $v_0(y, a)$  и  $u(y, a)$  на  $C$  по имеющимся наблюдениям  $(x_k, y_k, a_k)$ ,  $k = 1, \dots, K$ .

Одна из возможностей для решения этой проблемы заключается в минимизации в функциональном пространстве функционала средне-квадратичного отклонения

$$(4) \quad I(u, v_0) = \sum_{k=1}^K (x_k - v(y_k, a_k))^2$$

при дополнительных ограничениях на функции  $u(y, a)$  и  $v_0(y, a)$  в виде, например, кусочной непрерывности или ограниченной вариации. Для практического применения удобно преобразовать сформулированную проблему к дискретному виду и воспользоваться преимуществом простоты анализа и адаптации численных методов, доступных в рамках стандартных статистических пакетов.

#### 3.2. Модель с дискретными временем и возрастом

Пусть  $i$  и  $j$  обозначают соответственно время и возраст в годах. Запишем дискретизированные представления для уравнений (1)–(4). Пусть  $P(i, j)$  является параллелограммом, заданным вершинами  $\{(i, j - 1), (i, j), (i + 1, j + 1), (i + 1, j)\}$  и не включающим свои левую и верхнюю границы

$$(5) \quad P(i, j) = \{(a, y) : y \in [i, i + 1), \quad a \in ((j - 1) + (y - i), j + (y - i))\}.$$

Будем предполагать, что в области  $P(i, j)$  когортные тренды  $u(\cdot, \cdot)$  принимают постоянное значение, а уровни индикатора  $v(\cdot, \cdot)$  при изменении возраста  $a$  принимают постоянное значение и линейно с постоянным наклоном  $u(\cdot, \cdot)$  изменяются при изменении времени  $y$ . Эти условия записываются в виде

$$(6) \quad \begin{aligned} u(y, a) &= u(i, j), \quad \forall i, j, y, a : (y, a) \in P(i, j) \\ v(y, a) &= u(i, j) \times (y - i) + v(i, j), \quad \forall i, j, y, a : (y, a) \in P(i, j). \end{aligned}$$

Минимальные и максимальные значения для величин  $i$  и  $j$  запишем пользуясь определением (5) в виде

$$\begin{aligned} (i_{\min}, j_{\min}) &: (y_{\min}, a_{\min}) \in P(i_{\min}, j_{\min}), \\ (i_{\max}, j_{\max}) &: (y_{\max}, a_{\max}) \in P(i_{\max}, j_{\max}). \end{aligned}$$

В дальнейшем будем пользоваться относительной шкалой для времени и возраста, считая за ноль их минимальные в области  $P(i, j)$  значения.

Рассмотрим две функции  $u(i, j)$  и  $v(i, j)$ . Функция  $u(i, j)$  определена на элементах целочисленной двумерной области

$$U = \{(i, j) : i \in [0, I], j \in [0, J]\},$$

а функция  $v(i, j)$  определена на элементах целочисленной двумерной области

$$V = \{(i, j) : i \in [0, I + 1], j \in [0, J + 1]\},$$

где  $I = i_{\max} - i_{\min}$ ,  $J = j_{\max} - j_{\min}$ . Уравнение (2) можно переписать в виде

$$(7) \quad v(i + 1, j + 1) = v(i, j) + u(i, j), \quad \forall (i, j) \in U.$$

Пусть  $v_0(i, j)$  обозначает значение функции  $v(\cdot, \cdot)$  на нижней или левой границе области  $V$  в точке, через которую проходит когорта рождения, проходящая также через точку  $(i, j)$ , т.е.

$$(8) \quad v_0(i, j) = v(i - \delta, j - \delta), \quad \delta = \min(i, j).$$

Комбинируя (7) и (8), перепишем уравнение (3) в виде

$$(9) \quad v(i, j) = v_0(i, j) + \sum_{m=1}^{\delta} u(i - m, j - m).$$

Из (9) следует, что если известны значения функции  $v(\cdot, \cdot)$  на границе области  $V$  и функция  $u(\cdot, \cdot)$  известна на всей области  $U$ , то значения функции  $v(\cdot, \cdot)$  могут быть рассчитаны во всех точках области  $V$ . С учетом (1), (2) и (6) для каждого доступного наблюдения  $(x_k, y_k, a_k)$ ,  $k = 1, \dots, K$  запишем

$$(10) \quad x_k = v_0 + \sum_{m=1}^{\delta} u(i - m, j - m) + (y_k - i) u(i, j) + \varepsilon_k,$$

где  $Var(\varepsilon_k) = \sigma^2$ ,  $Cov(\varepsilon_k, \varepsilon_l) = 0$ ,  $k \neq l$ .

Введем вектор  $\mathbf{z}$ , компонентами которого являются значения  $v_0(i, j)$  и  $u(i, j)$ ,

$$(11) \quad \mathbf{z} = (\mathbf{v}_0^T \mid \mathbf{u}^T)^T,$$

где

$$\mathbf{v}_0 = (v(I+1, 0), \dots, v(0, 0), \dots, v(0, J+1))^T,$$

$$\mathbf{u} = (u(0, 0), \dots, u(0, J), \dots, u(I, 0), \dots, u(I, J))^T.$$

Обозначив вектор коэффициентов в формуле (10) через  $\mathbf{b}_k$ , перепишем (10) в виде

$$(12) \quad x_k = (\mathbf{b}_k, \mathbf{z}) + \varepsilon_k.$$

Пусть  $\mathbf{B}_1$  обозначает матрицу, составленную вектор-строками  $\mathbf{b}_k^T$  из (12),  $\mathbf{x}_1$  обозначает вектор-столбец, составленный из переменных  $x_k$ . Проблема минимизации функционала (4) эквивалентна проблеме минимизации скалярной функции  $S_1(\mathbf{z})$ , определенной как

$$S_1(\mathbf{z}) = (\mathbf{B}_1 \mathbf{z} - \mathbf{x}_1)^T (\mathbf{B}_1 \mathbf{z} - \mathbf{x}_1).$$

Отметим, что если  $\text{rank}(\mathbf{B}_1) = \text{dim}(\mathbf{z})$ , то оценки, полученные путем безусловной минимизации функции  $S_1(\mathbf{z})$ , единственны. Легко убедиться, что это условие выполняется, только если наблюдения произведены во всех элементах области  $P(i, j)$ , что на практике является скорее исключением, чем правилом.

Минимизация функции  $S_1(\mathbf{z})$  эквивалентна поиску решения уравнения (12) и приводит к неустойчивому решению в силу плохой обусловленности матрицы  $\mathbf{B}_1^T \mathbf{B}_1$  при ее большой размерности. Для стабилизации решения модифицируем проблему минимизации функции  $S_1(\mathbf{z})$ , накладывая дополнительные ограничения на допустимое решение двумя независимыми способами: с помощью сглаживания и с помощью агрегирования.

### 3.3. Сглаживание

Рассмотрим функционал гладкости

$$(13) \quad S_2(\mathbf{z}) = \sum_{i=0}^{I+1} \sum_{j=1}^J (2v(i, j) - v(i, j-1) - v(i, j+1))^2 +$$

$$+ \sum_{j=0}^{J+1} \sum_{i=1}^I (2v(i, j) - v(i-1, j) - v(i+1, j))^2.$$

Каждое слагаемое в (13) пропорционально квадрату второй разности функции  $v(\cdot, \cdot)$  по возрасту или по времени в точке  $(i, j)$  и является аналогом второй частной производной. Обозначив через  $\mathbf{B}_2$  матрицу с элементами, соответствующими коэффициентам в (13), и, воспользовавшись (9) и (11), представим  $S_2(\mathbf{z})$  в виде

$$S_2(\mathbf{z}) = (\mathbf{B}_2 \mathbf{z})^T (\mathbf{B}_2 \mathbf{z}).$$

Для стабилизации решения наложим ограничение  $S_2(\mathbf{z}) \leq \alpha$  при некотором  $\alpha \geq 0$ . Отметим, что  $S_1(\mathbf{z})$ ,  $S_2(\mathbf{z})$  являются квадратичными функционалами в векторном пространстве  $\mathbf{E}_n$  размерности  $n = \text{dim}(\mathbf{z})$ . Оптимизационную проблему для оценки решения (12) запишем в виде

$$(14) \quad \min_{\mathbf{z} \in \mathbf{E}_n} S_1(\mathbf{z}) \quad \text{при ограничении} \quad S_2(\mathbf{z}) \leq \alpha, \quad \alpha \geq 0.$$

Условие оптимальности для решения этой проблемы может быть записано на основании следующей теоремы, сформулированной для общего случая.

*Теорема 1. Пусть  $S_k(\mathbf{x})$  – квадратичные функции вида*

$$S_k(\mathbf{x}) = (\mathbf{B}_k \mathbf{x} - y_k)^T (\mathbf{B}_k \mathbf{x} - y_k),$$

*заданные в  $\mathbf{E}_n$ ,  $k = 0, \dots, K$ ,  $K > 0$ . Пусть  $P1$  будет оптимизационной проблемой  $\min_{\mathbf{x} \in \mathbf{E}_n} S_0(\mathbf{x})$  при ограничении  $S_k(\mathbf{x}) \leq \alpha_k$ ,  $\alpha_k > 0$ ,  $k = 1, \dots, K$ . Допустим, что система ограничений непротиворечива, то есть существует такое  $\mathbf{x}$ , что одновременно выполняются все ограничения  $S_k(\mathbf{x}) \leq \alpha_k$ ,  $k = 1, \dots, K$ .*

*I. Пусть  $\mathbf{x}^*$  является решением проблемы  $P1$ . Тогда существуют неотрицательные числа  $\lambda_k$ ,  $k = 1, \dots, K$ , удовлетворяющие условиям дополняющей нежесткости*

$$\lambda_k (S_k(\mathbf{x}^*) - \alpha_k) = 0$$

*и такие, что  $\mathbf{x}^*$  одновременно является решением проблемы  $P2$ :*

$$\min_{\mathbf{x} \in \mathbf{E}_n} \left( S_0(\mathbf{x}) + \sum_{k=1}^K \lambda_k S_k(\mathbf{x}) \right).$$

*II. Соответственно, если  $\mathbf{x}^*$  является оптимальным решением проблемы  $P2$  для заданного набора  $\lambda_k \geq 0$ ,  $k = 1, \dots, K$ , удовлетворяющего условию дополняющей нежесткости, то  $\mathbf{x}^*$  является и глобальным решением проблемы  $P1$ , если*

$$S_k(\mathbf{x}^*) \leq \alpha_k, \quad k = 1, \dots, K.$$

Доказательство теоремы приведено в Приложении.

Обозначим:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ 0 \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} \mathbf{I}_1 & 0 \\ 0 & \lambda \mathbf{I}_2 \end{pmatrix},$$

где  $\mathbf{I}_1$  и  $\mathbf{I}_2$  – единичные матрицы. Модель наблюдений запишем в виде

$$\mathbf{x} = \mathbf{Bz} + \varepsilon, \quad E(\varepsilon) = 0, \quad D(\varepsilon) = \sigma^2 \mathbf{W}^{-1}.$$

Из теоремы 1 следует, что решение проблемы (14) одновременно является и решением проблемы

$$(15) \quad \min_{\mathbf{z} \in \mathbf{E}_n} S(\mathbf{z}),$$

где  $S(\mathbf{z}) = (\mathbf{Bz} - \mathbf{x})^T \mathbf{W} (\mathbf{Bz} - \mathbf{x}) = S_1(\mathbf{z}) + \lambda S_2(\mathbf{z})$ . Отметим, что в обозначениях теоремы 1 это является формулировкой проблемы  $P2$  для проблемы  $P1$ , представленной в (14).

С практической точки зрения важен вопрос о существовании и единственности решения проблемы (15) или, что эквивалентно, проверке условия

$$(16) \quad \text{rank}(\mathbf{B}) = \dim(\mathbf{z}).$$

*Теорема 2. Рассмотрим представление*

$$(17) \quad v(i, j) = \gamma_1 + \gamma_2 i + \gamma_3 j + \gamma_4 i j$$

*и такую матрицу  $\mathbf{B}_{1, \gamma}$ , что*

$$\mathbf{B}_{1, \gamma} \gamma = \mathbf{B}_1 \mathbf{z},$$

*где  $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)^T$ .*

Если

$$(18) \quad \text{rank}(\mathbf{B}_{1,\gamma}) \geq 4,$$

то справедливо (18).

Доказательство теоремы приведено в Приложении.

*Следствие.* Для существования единственного решения проблемы (16) достаточно иметь 4 такие наблюдения, что никакие 3 из них не лежат на одной прямой линии.

### 3.4. Агрегирование

Будем рассматривать агрегирование как преобразование данных с целью уменьшения числа переменных, входящих в модель, увеличения скорости вычисления и повышения устойчивости ответа к случайным возмущениям в данных. Когортные тренды  $u(\cdot, \cdot)$  представим в виде кусочно-постоянной функции внутри областей, заданных по возрасту точками разбиения  $a_k$ ,  $k = 1, \dots, K$  и по времени точками  $y_m$ ,  $m = 1, \dots, M$ :

$$u(i, j) = h(k, m) \quad \forall (i, j) : (a_{k-1} \leq i < a_k, \quad y_{m-1} \leq j < y_m), \\ a_0 = a_{\min}, \quad a_{N+1} = a_{\max}, \quad y_0 = y_{\min}, \quad y_{N+1} = y_{\max}.$$

Формально агрегирование задается такой матрицей  $\mathbf{G}$ , что

$$(19) \quad \mathbf{u} = \mathbf{G}\mathbf{h},$$

где  $\mathbf{h}$  – вектор с определенным образом упорядоченными компонентами  $h(k, m)$ ,  $\mathbf{u}$  – вектор, определенный в (11). Напомним, что вектор  $\mathbf{z}$  состоит из компонент  $v_0(i, j)$  и  $u(i, j)$ , упорядоченных в одномерный массив. Матрицу  $\mathbf{B}$  в (15) представим в виде  $\mathbf{B} = (\mathbf{B}_{v_0} \mathbf{B}_u)$  откуда

$$\mathbf{Bz} = \mathbf{B}_{v_0}\mathbf{v}_0 + \mathbf{B}_u\mathbf{u}.$$

Применив преобразование (19) к проблеме (15), получим представление для “агрегированной” проблемы

$$(20) \quad \mathbf{x} = \mathbf{B}_G\mathbf{z}_G + \varepsilon, \quad E(\varepsilon) = 0, \quad D(\varepsilon) = \sigma^2\mathbf{W}^{-1},$$

где  $\mathbf{B}_G = (\mathbf{B}_{v_0}\mathbf{B}_u\mathbf{G})$ ,  $\mathbf{z}_G = (\mathbf{v}_0^T \mathbf{h}^T)^T$ . Легко убедиться, что из единственности решения исходной проблемы (15) следует единственность решения и агрегированной проблемы (20).

## 4. Алгоритмическая реализация метода динамической регрессии

В этом разделе детально опишем алгоритм для решения агрегированной проблемы (20). Агрегирование  $S$ -трендов является одной из наиболее удобной в практической реализации возможностей. Как альтернативу агрегированию можно использовать сглаживание  $S$ -трендов. Будем пользоваться обозначением DRM-SA для этого случая.

Алгоритм реализован в виде программы для системы статистического анализа SAS [19] и состоит из следующих шагов.



**Шаг 1.** Ввод данных обследования и вычисление для каждого человека параметров:

$x$  – изучаемая переменная,

$i$  – календарный год даты измерения,

$j$  – возраст в полных годах на конец календарного года, в котором произведено измерение,

$yfr = (y - i)$  – часть текущего года, прошедшая от начала года до даты проведения измерения.

Определение минимальных и максимальных значений для  $i$  и  $j$  по множеству наблюдений, вычисление диапазона изменения по времени  $I$  как  $(i_{\max} - i_{\min})$  и по возрасту  $J$  как  $(j_{\max} - j_{\min})$ .

**Шаг 2.** Генерация матрицы  $\mathbf{B}_1$ . Для каждого объекта  $k$  генерируется рабочий массив  $r(0 : I + 1, 0 : J + 1)$  и производятся следующие присвоения:

$$r(i_k, j_k) = yfr_k,$$

$$r(i, j) = 1 \quad \text{для всех } i, j \text{ таких, что } i - j = i_k - j_k \text{ и } i < i_k,$$

$$r(i, j) = 0 \quad \text{иначе.}$$

Созданный таким образом массив преобразуется в одномерный вектор, состоящий из строк матрицы  $\mathbf{B}_1$ , упорядоченной согласно (11).

**Шаг 3.** Генерация матрицы  $\mathbf{B}_2$ . Для каждого слагаемого в (13) создается рабочий массив  $r(0 : I + 1, 0 : J + 1)$ , которому присваиваются нулевые начальные значения, диагональным элементам присваиваются значения 2, а элементам верхней и нижней субдиагоналей присваиваются значения  $-1$  в соответствии со структурой слагаемых в (13). Созданный таким образом массив преобразуется в одномерный вектор таким же образом, как и на предыдущем шаге, и определяет строку матрицы  $\mathbf{B}_2$ .

**Шаг 4.** Создание матрицы агрегирования  $\mathbf{G}$ . Сначала создается рабочий массив  $r(0 : I, 0 : J)$ , затем в цикле по точкам разбиения по возрасту и времени каждому элементу  $r(i, j)$  присваивается номер кластера, которому принадлежит точка  $(i, j)$ ,  $r(i, j) = k$ ,  $k = 1, \dots, K$ , величина  $K$  равна числу кластеров. После этого каждый из  $K$  столбцов матрицы  $\mathbf{G}$  создается сначала заменой  $r(i, j) = 1$  если  $r(i, j) = k$ ,  $r(i, j) = 0$  иначе, и затем преобразованием массива  $r$  в вектор, следуя правилу (11).

**Шаг 5.** Выбор параметра  $\lambda$ . При заданном начальном значении параметра  $\lambda$  вычисляются:

точечная оценка

$$\tilde{\mathbf{z}} = (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{x},$$

значения функционалов

$$\tilde{S}_1 = (\mathbf{B}_1 \tilde{\mathbf{z}} - \mathbf{x})^T (\mathbf{B}_1 \tilde{\mathbf{z}} - \mathbf{x}), \quad \tilde{S}_2 = (\mathbf{B}_2 \tilde{\mathbf{z}})^T (\mathbf{B}_2 \tilde{\mathbf{z}}), \quad \tilde{S} = \tilde{S}_1 + \lambda \tilde{S}_2,$$

оценка дисперсии случайной ошибки в модели

$$s^2 = \frac{\tilde{S}}{K - p},$$

$K$  – число строк в матрице  $\mathbf{B}_1$ ,  $p$  – число столбцов в матрице  $\mathbf{B}_1$ . Ковариационная матрица оценок

$$\mathbf{C} = (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} s^2.$$

Среднеквадратичная ошибка элементов вектора  $\mathbf{v}_0$  – вектора значений изучаемого индикатора на нижней или левой границе области анализа вычисляется по формуле

$$s_{v_0}^2 = \frac{1}{n_{v_0} - 2} \sum_{m=2}^{n_0-1} c_{m,m},$$

а индекс гладкости оценки – по формуле

$$f_{sm} = \frac{s_{v_0}^2}{\tilde{s}_2}, \quad \text{где} \quad \tilde{s}_2 = \frac{\tilde{S}_2}{n_2},$$

где  $n_2$  – число слагаемых в выражении для  $S_2$ , равное числу строк в матрице  $\mathbf{B}_2$ . Требуемая величина гладкости  $f_{sm}$  и ее допустимая относительная погрешность  $\delta$  задаются перед началом итераций. Условие достижения заданной гладкости записывается в виде

$$(21) \quad \left| \frac{f_{sm} s_{v_0}^2}{\tilde{s}_2} - 1 \right| \leq \delta.$$

Если это условие не выполняется, то вычисляется новое значение параметра  $\lambda$

$$\lambda_{next} = \frac{\lambda f_{sm} \tilde{s}_2}{s_{v_0}^2}$$

и вычисления шага 5 повторяются с этим значением  $\lambda = \lambda_{next}$ . Итерации прекращаются при выполнении условия (21).

**Шаг 6.** Вычисление заключительных статистик, проверка результатов и построение графиков. На этом шаге для каждой пары соседних кластеров проверяется гипотеза о неизменности когортных трендов. Для этого попарно сравниваются С-тренды в текущем и соседнем кластере, соответствующем более старой возрастной группе либо соответствующем следующему году (если таковые существуют). Для двух кластеров проверяется гипотеза  $\mathbf{H}_0 : z_i - z_j = 0$ , где  $z_i$  – компонента вектора  $\mathbf{z}$ , соответствующая С-тренду в  $i$ -м кластере. При проверке гипотезы используется  $F$ -статистика

$$F = \frac{(\tilde{z}_i - \tilde{z}_j)^2}{c_{i,i} - 2c_{i,j} + c_{j,j}}.$$

Соответствующие доверительные вероятности для  $F$ -статистики вычисляются с помощью стандартных функций системы SAS.

Оценка вектора  $\mathbf{u}$  вычисляется по агрегированному вектору с помощью формулы (19), функция  $v(i, j)$  вычисляется по формуле (9). В заключение с помощью графических процедур SAS [20] строится набор трехмерных графиков, показывающих функции  $v(i, j)$ ,  $u(i, j)$  с графическим изображением результатов проверки гипотезы о значимости разности между соседними кластерами. Исходные данные представляются усредненными по времени и возрасту значениями, прогнозируемые величины средних значений изучаемого индикатора, равные величинам  $v(i, j)$ , также представляются на графиках.

## 5. Пример применения метода динамической регрессии

Для демонстрации применения и иллюстрации качества работы метода динамической регрессии использовались данные трех кросс-секционных обследований,

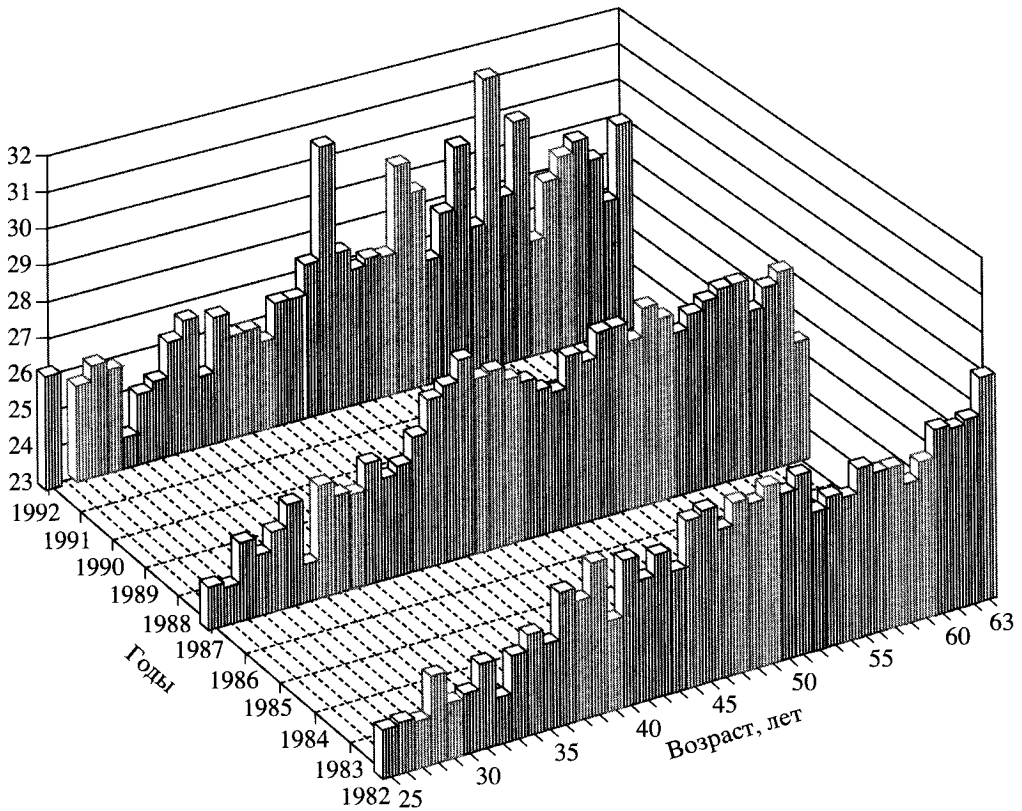


Рис. 1. Средние значения индекса массы тела (BMI) по однолетним возрастным группам, полученные при обследованиях в 1982, 1987 и 1992 годах.

проведенных в Северной Карелии, Финляндия, в 1982, 1987 и 1992 гг. В каждом обследовании формировалась случайная выборка людей, которые затем стратифицировались по 10-летним возрастным группам (25–34, 35–44, 45–54, 55–64) и по полу. В описываемом примере используются только данные обследования мужчин. Число обследованных мужчин в 1982, 1987 и 1992 гг. равнялось 1537, 1481 и 673 соответственно. Все обследования начинались в начале года, обследования в 1982 и 1987 гг. были закончены за четыре месяца, в 1992 – за три месяца. Регистрировались и измерялись следующие параметры: пол, дата рождения, дата обследования, вес и рост.

В модель для анализа включены следующие переменные:

BMI – индекс массы тела, определяемый как отношение  $\text{вес(кг)}/\text{рост}^2(\text{м})$ ,

AGE – возраст в полных годах, определенный как год обследования минус год рождения,

YEAR – год обследования.

При анализе использовалось агрегирование, при этом C-тренды задавались кусочно-постоянными функциями внутри 10-летних возрастных групп и на 5-летнем временном периоде, что описывается 8 параметрами. Величина фактора гладкости задавалась равной 6.

Результаты анализа представлены в виде трехмерных графиков на рис. 1–4.

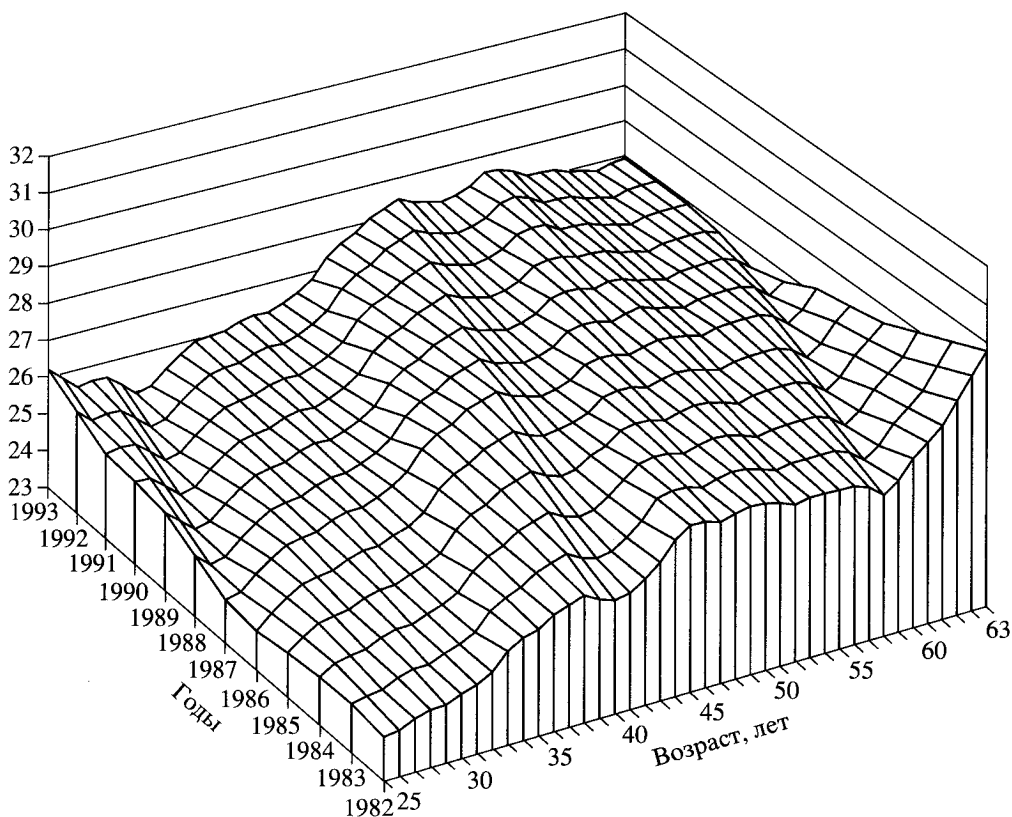


Рис. 2. Оценки среднего уровня индекса массы тела (ВМІ), рассчитанные в диапазоне возрастов 25–65 лет для периода 1982–1993 гг.

Рисунок 1 представляет значения изучаемой переменной ВМІ, полученные при обследованиях в 1982, 1987 и 1992 гг. Величины представляют средние значения ВМІ, вычисленные для каждого возраста и года, в которых имелись данные обследований. Усреднение проводилось не менее чем по 9 значениям, соответствующим одному возрасту. Чтобы графически выделить изменения, происходящие в когорте, столбцы, соответствующие одной когорте рождения, в различных обследованиях отмечены одинаковым тоном серого цвета.

Рисунок 2 изображает оценки среднего уровня ВМІ, рассчитанные методом динамической регрессии в диапазоне возрастов 25–65 лет для временного периода 1982–1993 гг.

На рис. 3 изображены С-тренды (когортные тренды) с 95% доверительными интервалами, нижняя часть которых показаны только на левой и передней границах рисунка. Проверка гипотез о равенстве С-трендов в сопредельных кластерах показала, что для всех возрастных групп, кроме группы 35–44 года, когортные тренды изменяются во времени статически недостоверно. Для возрастной группы 35–44 года произошло статистически значимое снижение ( $p < 0,05$ ) когортных трендов после 1986 г. При сравнении изменений когортных трендов по возрасту оказалось, что статистически значимые различия ( $p < 0,01$ ) наблюдаются только при переходе от возрастной группы 35–44 года к возрастной группе 45–54 года в период 1986–1992 гг. На рис. 4 представлены предсказанные с помощью модели оценки для наблюдаемых значений средних уровней ВМІ, показанных на рис. 1.

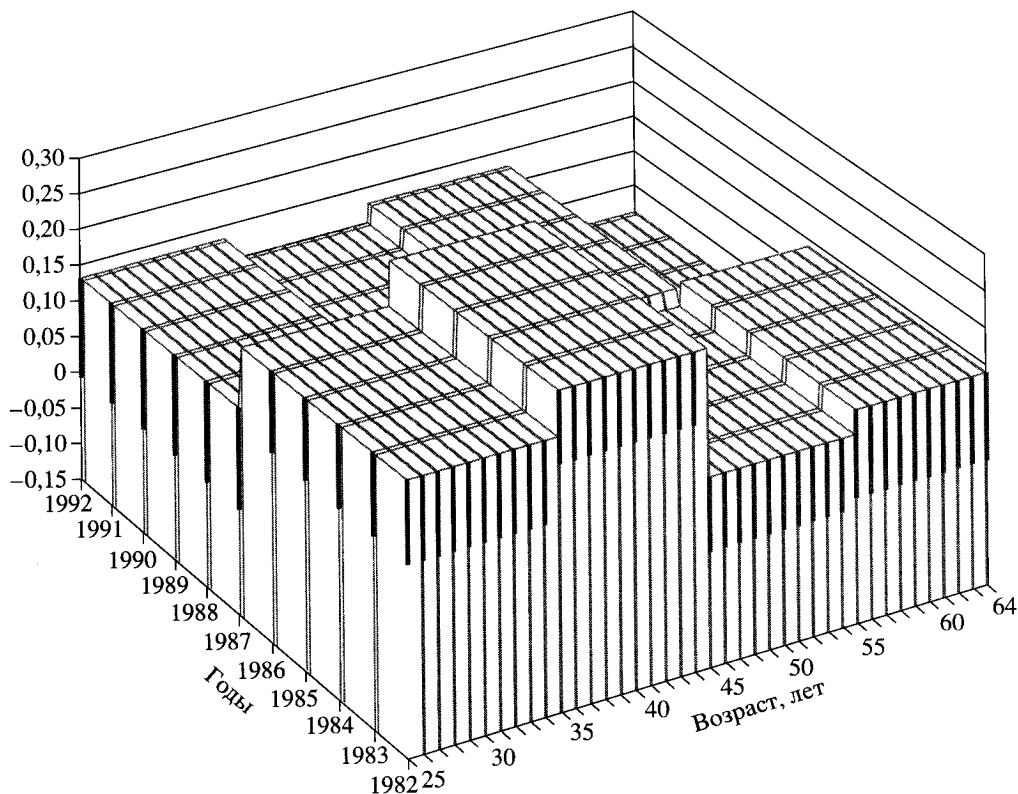


Рис. 3. Оценки С-трендов для среднего уровня индекса массы тела (ВМІ), рассчитанные в диапазоне возрастов 25–64 лет для периода 1982–1992 гг.

Из рис. 2 следует, что средний уровень ВМІ возрастает вдоль когорт на протяжении периода наблюдения, хотя он и различен для различных когорт рождения. Подъемы и падения уровней соответствуют переходам между когортами.

В динамической модели С-тренды отражают совместный эффект действия контролируемых и неконтролируемых факторов, вызывающих изменение в когорте изучаемого индикатора по времени. Поэтому изменения С-трендов по годам при фиксированном возрасте может указывать на эффект профилактических мероприятий, в то время как изменение по возрастам может указывать на зависимые от возраста неконтролируемые изменения или различную восприимчивость к профилактике. С-тренды ясно демонстрируют снижение для периода 1987–1992 гг. по сравнению с периодом 1982–1986 гг. в возрастной группе 35–44 лет ( $p < 0,05$ ). Уровень С-трендов в этой возрастной группе выше по сравнению с более молодыми и с более старыми возрастными группами. Разница по сравнению с более старшей возрастной группой более значима ( $p < 0,01$ ). Никакие другие существенные различия между оценками в прилегающих кластерах не наблюдаются. Более углубленный анализ, возможно, потребует изучения хронологии событий, влияющих на социально-экономические характеристики и состояние здоровья населения за исследуемый период времени. Так, обнаруженное уменьшение С-тренда в возрастах 35–44 года можно связать с созданием новых рабочих мест в Северной Карелии за период 1987–1992 гг., уменьшившее отток населения из изучаемого региона, который в этой возрастной группе имел место в период с 1982 г. по 1986 г. и привел к снижению показателя среднего

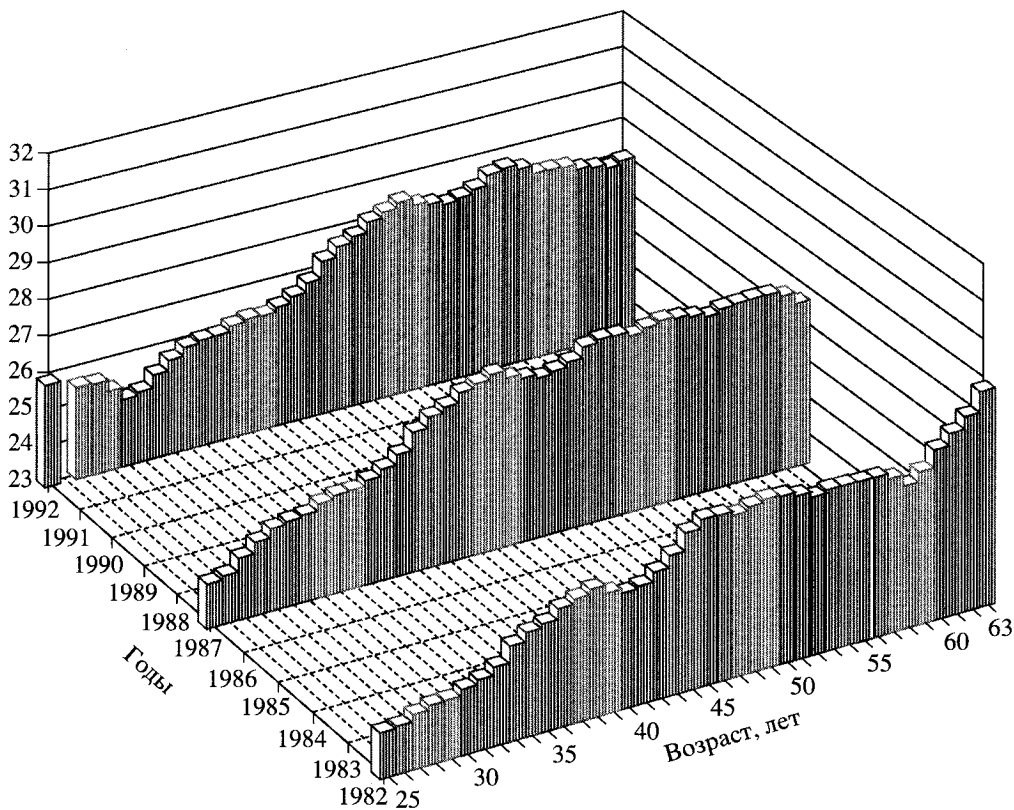


Рис. 4. Оценки среднего уровня индекса массы тела (ВМТ), рассчитанные для возрастов и лет, в которых проводились обследования.

уровня здоровья. Это снижение было обусловлено тем, что более здоровые люди уезжали, а оставались люди с худшим здоровьем и более высокими значениями факторов риска, к которым относится и индекс массы тела.

## 6. Заключение и обсуждение

Описанный метод разработан как альтернатива временным трендам, широко используемым до сих пор при анализе динамики факторов риска. Описанная модель характеризуется локальными когортными трендами, которые имеют ясную интерпретацию как изменение переменной состояния в единицу времени для элемента популяции – когорты. При формальном вычислении временных трендов в заданной возрастной группе получают изменение между двумя различными физическими элементами популяции – двумя различными когортами. Подобные оценки могут демонстрировать произвольное поведение, и с точки зрения динамических систем такие временные тренды не существуют в природе. Тем не менее в специальном случае, когда все повозрастные профили переменной состояния не меняются во времени, формально вычисленные временные тренды окажутся нулевыми для каждого возраста из изучаемого диапазона возрастов. Лишь в этом тривиальном случае временные тренды формально пригодны для анализа и прогноза.

В статье рассмотрено применение философии системного подхода к процессу построения модели реального сложного объекта – населения – в части, касающейся его

здоровья. При этом сначала на верхнем уровне агрегирования были выделены два взаимодействующих объекта: население и его «среда обитания». Далее системный анализ привел к декомпозиции изучаемого населения на когорты рождения как составляющие элементы, движущиеся со временем в фазовом пространстве (возраст – календарное время) и несущие во времени свой профиль состояния здоровья, выраженный в некоторых индикаторах – переменных состояния.

Предложенная модель использует концепцию *поля здоровья*, которая предполагает существование некоего потенциального воздействия, создаваемого средой обитания и действующего на население. Это воздействие, вообще говоря, зависит от календарного времени, возраста и вызывает в когортах рождения изменения изучаемых индикаторов здоровья со скоростью, определяемой силой этого воздействия.

На заключительной стадии конструирования модели был рассмотрен случай, в котором состояние здоровья характеризуется одним количественным индикатором и причем «сила» внешнего воздействия численно равна скорости изменения индикатора. Сохраняя модель достаточно реалистичной, эти упрощения позволили выделить основные свойства динамической модели здоровья и разработать метод ее идентификации – метод динамической регрессии.

На примере анализа данных индекса массы тела, полученных в трех последовательных независимых скринингах, было показано, что метод динамической регрессии позволяет получать реалистичные представления о динамике изучаемого индикатора. Метод динамической регрессии выявляет четкое различие между уровнями индикаторов и их *S*-трендами. С практической точки зрения именно *S*-тренды, а не уровни индикатора непосредственно модифицируются в результате спланированных профилактических мероприятий либо неспланированных воздействий среды на население. Результаты анализа, проведенного методом динамической регрессии служат основой и данными для следующего уровня анализа, в котором привлекается дополнительная информация и который направлен на поиск реалистичной интерпретации наблюдаемой динамики.

Одной из важных составляющих популяционного анализа является учет динамики изменения численности населения. При наличии существенной миграции в изучаемом населении наблюдаемые эффекты могут быть полностью либо частично обусловлены нестабильностью численности населения. Метод динамической регрессии позволяет эффективно учитывать это обстоятельство. Результаты анализа методом динамической регрессии могут использоваться напрямую для прогнозирования повозрастного профиля изучаемой переменной состояния в будущем. Для этого по оценкам уровней переменной состояния и по оценкам *S*-трендов в последнем году обследования строится линейный прогноз будущих значений переменной состояния в разных когортах рождения. Естественно, что подобный прогноз не применим для когорт рождения, не входящих в изучаемый диапазон возрастов в последнем году исследования.

В описанной модели используется сглаживание для уровней и агрегация для *S*-трендов. Подобный подход удовлетворяет целям использования модели для анализа изменений в популяции. Для прогнозирования, однако, лучше иметь более детальное представление *S*-трендов. В этом случае для *S*-трендов следует применять не агрегирование, а сглаживание. В заключение следует отметить, что упрощенное динамическое уравнение, использованное в настоящей модели, должно быть скорректировано если темп изменения может дополнительно зависеть от текущего уровня переменной состояния.

## ПРИЛОЖЕНИЕ

*Доказательство теоремы 1.* Функции  $f_0(x) = S_0(x)$  и  $f_k(x) = S_k(x) - a_k$ ,  $k = 1, \dots, K$  выпуклы и дифференцируемы в  $E_n$ . Оптимизационная проблема *P1*

может быть записана в виде  $\min_{x \in \mathbf{E}_n} f_0(x)$  при ограничениях  $f_k(x) \leq 0, k = 1, \dots, K$ .

Утверждения теоремы 1 следуют непосредственно из условия оптимальности теоремы Куна-Таккера [21] заменой  $f_0(x)$  на  $S_0(x)$  и  $f_k(x)$  на  $S_k(x) - \alpha_k, k = 1, \dots, K$ . Запись проблемы P2 в теореме 1 можно получить из условий Куна-Таккера для оптимального решения  $x^*$  проблемы

$$L(x^*, \lambda) = \min_{x \in \mathbf{E}_n} L(x, \lambda)$$

следующим образом. Если выполняются условия регулярности, то можно положить  $\lambda_0 = 1$  и записать функцию Лагранжа в виде

$$L(x, y) = \left( S_0(x) + \sum_{k=1}^K \lambda_k (S_k(x) - \alpha_k) \right),$$

а условие оптимальности Куна-Таккера в виде

$$\left( S_0(x^*) + \sum_{k=1}^K \lambda_k (S_k(x^*) - \alpha_k) \right) = \min_{x \in \mathbf{E}_n} \left( S_0(x) + \sum_{k=1}^K \lambda_k (S_k(x) - \alpha_k) \right),$$

или

$$\left( S_0(x^*) + \sum_{k=1}^K \lambda_k (S_k(x^*)) \right) = \min_{x \in \mathbf{E}_n} \left( S_0(x) + \sum_{k=1}^K \lambda_k (S_k(x)) \right).$$

Правая часть в последнем выражении представляет проблему P2. Теорема доказана.

*Доказательство теоремы 2.* Из (13) следует, что решение уравнения  $\mathbf{B}_2 \mathbf{z} = 0$  можно представить в виде (17), зависящим от 4-х параметров. Если справедливо (18), то среди уравнений системы  $\mathbf{B}_{1,\gamma} \gamma = x_1$  можно выделить систему из 4-х уравнений, имеющих единственно решение  $\gamma^*$ , которому соответствует  $\mathbf{z}^*$ . Другими словами можем сказать, что среди уравнений  $\mathbf{Bz} = \mathbf{x}$  существует подсистема, имеющая единственное решение  $\mathbf{z}^*$ , что означает, что матрица  $\mathbf{B}$  имеет ранг, равный  $\dim(\mathbf{z})$ , что и доказывает теорему.

## СПИСОК ЛИТЕРАТУРЫ

1. Dawber T.R., Meadors G.F., Moore F.E.J. Epidemiological approaches to heart disease: The Framingham Study // Am. J. Public Health. 1951. V. 41. P. 279-286.
2. Pearson J.D., Morrell C.H., Brant L.J. et al. Age-associated changes in blood pressure in a longitudinal study of healthy men and women // J. Gerontol. A Biol. Sci. Med. Sci. 1997. V. 52. P. M177-M183.
3. Manton K.G., Corder L., Stallard E. Chronic disability trends in elderly United States populations: 1982-1994 // Proc. Natl. Acad. Sci. USA. 1997. V. 94. P. 2593-2598.
4. Christensen K., McGue M., Yashin A. et al. Genetic and Environmental Influences on Functional Abilities in Danish Twins Aged 75 Years and Older // J. Gerontol. A Biol. Sci. Med. Sci. 2000. V. 55. P. M446-M452.
5. Михальский А.И., Яшин А.И. Управление старением и продолжительностью жизни // Проблемы управления. 2004. № 4. С. 46-53.
6. Mann C.J. Observational research methods. research design ii: cohort, cross sectional, and case-control studies // Emerg. Med. 2003. V. 20. P. 54-60.
7. MONICA Monograph and Multimedia Sourcebook / Ed. H Tunstall-Pedoe. Prepared by H. Tunstall-Pedoe, K. Kuulasmaa, H. Tolonen, M. Davidson, and S. Mendis with 64 other contributors for The WHO MONICA Project. Geneva: World Health Organization, 2003.



8. *Dobson J., Evans A., Ferrario M. et al.* Changes in estimated coronary risk in the 1980s: data from 38 populations in the WHO MONICA project // *Ann. Med.* 1998. V. 30. P. 199–205.
9. *Dobson A.J., Kuulasmaa K., Moltchanov V. et al.* Changes in cigarette smoking among adults in 35 populations in the mid-1980s // *Tobacco Control.* 1998. V. 7. P. 14–21.
10. *Kuulasmaa K., Tunstall-Pedoe H., Dobson A. et al.* Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations // *Lancet.* 2000. V. 355. P. 675–687.
11. *Gregg E.W., Cheng Y.J., Cadwell B.L. et al.* Secular trends in cardiovascular disease risk factors according to body mass index in us adults // *JAMA.* 2005. V. 293. P. 1868–1874.
12. *Kautiainen S., Rimpela A.H., Vikat A. et al.* Secular trends in overweight and obesity among Finnish adolescents in 1977–1999 // *Int. J. Obes. Relat. Metab. Disord.* 2002. V. 26. P. 544–552.
13. *Chen X., Li G., Unger J.B., Liu X. et al.* Secular trends in adolescent never smoking from 1990 to 1999 in California: and age-period-cohort analysis // *Am. J. Public Health.* 2003. V. 93. P. 2099–2104.
14. *Holford T.R.* Understanding the effects of age, period, and cohort on incidence and mortality rates // *Ann. Rev. Public Health.* 1991. V. 12. P. 425–457.
15. *Lee R.D., Carter L.R.* Modeling and Forecasting U.S. Mortality // *JASA.* 1992. V. 87. P. 659–671.
16. *Wilmoth J.R.* Computational methods for fitting and extrapolating the Lee-Carter model of mortality change // Technical report, Department of Demography, University of California, Berkeley. 1993.
17. *Moltchanov V.* The projection of the theory and methodology of the dynamic systems into epidemiological research // *Can. J. Cardiol.* 1993. V. 9. P. 88–89.
18. *Lewin K., Cartwright D.* Field theory in social science: selected theoretical papers. Harper, New York, 1951.
19. *SAS Institute Inc.* SAS Language Reference: Concepts, Version 8. Cary, NC: SAS Institute Inc., 1999.
20. *SAS Institute Inc.* SAS/GRAPH Software: Reference, Version 8. Cary, NC: SAS Institute Inc., 1999.
21. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука, 1974.

*Статья представлена к публикации членом редколлегии В.Н. Новосельцевым.*

Поступила в редакцию 30.11.2006