



doi 10.26089/NumMet.v26r101

УДК 517.958:57;
517.988.8

Идентификация параметров SEIR-подобной эпидемиологической модели с помощью итеративно регуляризованного метода Гаусса–Ньютона

М. М. Кокурин

Марийский государственный университет, Институт цифровых технологий,
Йошкар-Ола, Российская Федерация

ORCID: 0000-0002-1878-3846, e-mail: kokurin@nextmail.ru

А. В. Гаврилова

Марийский государственный университет,
Йошкар-Ола, Российская Федерация

ORCID: 0000-0002-2499-6303, e-mail: anst-val@yandex.ru

Абду Рауф Чеха Мохамед

Российский университет дружбы народов имени Патриса Лумумбы,
Москва, Российская Федерация

ORCID: 0009-0002-1516-5492, e-mail: a8du.rauf@yandex.ru

Аннотация: В статье изучается вопрос о восстановлении коэффициентов SEIR-подобных эпидемиологических моделей на основе данных эпидемиологической статистики. Предлагается подход к решению этой задачи, основанный на итеративно регуляризованном методе Гаусса–Ньютона, а также его модификации, позволяющей находить квазирешения нелинейных операторных уравнений. Учитываются различные виды погрешностей в эпидемиологической статистике. Обсуждается возможность прогнозирования распространения эпидемий, а также определение показателей ее заразности и опасности.

Ключевые слова: эпидемиологическая модель, SEIR, нелинейное операторное уравнение, итеративно регуляризованный метод Гаусса–Ньютона.

Благодарности: Работа поддержана грантом Российского научного фонда № 22-71-10070, <https://rscf.ru/project/22-71-10070/>.

Для цитирования: Кокурин М.М., Гаврилова А.В., Абду Рауф Чеха Мохамед. Идентификация параметров SEIR-подобной эпидемиологической модели с помощью итеративно регуляризованного метода Гаусса–Ньютона // Вычислительные методы и программирование. 2025. **26**, № 1. 1–16. doi 10.26089/NumMet.v26r101.



Identification of Parameters for a SEIR-like Epidemiological Model Using Iteratively Regularized Gauss–Newton Method

Mikhail M. Kokurin

Mari State University, Institute of Digital Technologies, Yoshkar-Ola, Russia
 ORCID: 0000-0002-1878-3846, e-mail: kokurin@nextmail.ru

Anastasiya V. Gavrilova

Mari State University, Yoshkar-Ola, Russia
 ORCID: 0000-0002-2499-6303, e-mail: anst-val@yandex.ru

Abdou Raouf Cheha Mohamed

Peoples’ Friendship University of Russia named after Patrice Lumumba, Moscow, Russia
 ORCID: 0009-0002-1516-5492, e-mail: a8du.rauf@yandex.ru

Abstract: We study the problem of identifying the coefficients based on known statistical data for a SEIR-like epidemiological model. The proposed approach to solving this problem is based on the iteratively regularized Gauss–Newton method. We also use one modification of this method, which is able to find quasi-solutions of nonlinear operator equations. We take into account various types of errors in epidemiological statistics. The possibility of predicting the spread of epidemics, as well as determining indicators of its contagion and danger, is discussed.

Keywords: epidemiological model, SEIR, nonlinear operator equation, iteratively regularized Gauss–Newton method.

Acknowledgements: The work was supported by the grant of the Russian Science Foundation No. 22–71–10070, <https://rscf.ru/project/22-71-10070/>.

For citation: M. M. Kokurin, A. V. Gavrilova, and Abdou Raouf Cheha Mohamed, “Identification of Parameters for a SEIR-like Epidemiological Model Using Iteratively Regularized Gauss–Newton Method,” *Numerical Methods and Programming*. 26 (1), 1–16 (2025). doi 10.26089/NumMet.v26r101.

1. Постановка задачи. Изучается SEIR-подобная эпидемиологическая модель

$$\left\{ \begin{array}{l} \dot{e}(t) = (1 - e(t) - i(t) - r_e(t) - r_i(t) - p(t))(a_e e(t) + a_i i(t)) - (\varkappa + \rho)e(t) \\ \dot{i}(t) = \varkappa e(t) - (\beta + \mu)i(t) \\ \dot{r}_e(t) = \rho e(t) - \gamma_e r_e(t) \\ \dot{r}_i(t) = \beta i(t) - \gamma_i r_i(t) \\ \dot{r}_o(t) = \beta i(t) \\ \dot{p}(t) = \mu i(t) \end{array} \right. \quad (1)$$

и ставится задача восстановления ее неизвестных параметров по данным эпидемиологической статистики на примере распространения эпидемии COVID-19 в России и зарубежных странах в 2021 г. SEIR-подобным эпидемиологическим моделям и подходам к определению их коэффициентов посвящено множество публикаций, см., например, [1–8] и ссылки, приведенные в этих работах. Модель (1) основана на разделении населения рассматриваемого ареала на следующие группы: бессимптомные зараженные с относительной численностью $e(t)$, симптомные больные с относительной численностью $i(t)$, выздоровевшие после бессимптомного заражения и имеющие иммунитет (относительная численность $r_e(t)$), выздоровевшие после симптомной болезни и имеющие иммунитет (относительная численность $r_i(t)$), погибшие (относительная численность $p(t)$), восприимчивые к инфекции (относительная численность $s(t) = 1 - e(t) - i(t) - r_e(t) - r_i(t) - p(t)$). Здесь под относительной численностью понимается отношение количества людей, принадлежащих к той или иной группе, к общей численности популяции. Значения



функций $i(t)$ и $p(t)$ можно считать приближенно известными из эпидемиологической статистики заражений и смертности. В то же время статистика выздоровлений не соответствует ни одной из описанных выше функций. Дело в том, что модель (1) учитывает постепенную утрату иммунитета переболевшими и их переход из группы “выздоровевших” в группу “восприимчивых”, что, однако, не отражается на эпидемиологической статистике выздоровлений. Поэтому в модель (1) добавлена дополнительная функция $r_o(t)$, показывающая относительную численность “наблюдаемых выздоровевших”. Эти люди не образуют отдельную группу, и каждый из них входит в одну из перечисленных выше групп. Например, если человек переболел с симптомами и выздоровел, то он становится “наблюдаемым выздоровевшим” и переходит в группу r_i . В то же время бессимптомные переболевшие не включаются в группу r_i и не относятся к “наблюдаемым выздоровевшим”. Различие между членами группы r_i и “наблюдаемыми выздоровевшими” связано с фактором угасания иммунитета, приобретенного после перенесенной болезни. Когда переболевший человек утрачивает иммунитет к заболеванию, он переходит из группы r_i в группу s (восприимчивые) и затем может повторно заразиться и перейти в другие соответствующие группы, однако остается “наблюдаемым выздоровевшим”. Утрата иммунитета не отражается на статистике выздоровлений в первую очередь потому, что невозможно отследить и указать точный момент этого события. Отметим, что более простые эпидемиологические модели, не учитывающие фактор утраты иммунитета, не способны адекватно описывать волнообразный характер эпидемии.

Деление всех зараженных на бессимптомных и симптомных является условным. В реальности степень выраженности симптомов может быть разной у разных больных и этот фактор связан с заразностью таких больных. Поэтому в некоторых SEIR-подобных эпидемиологических моделях используется деление на большее количество классов, например вводятся дополнительные классы госпитализированных и критических больных (см., например, [2, 8]). Однако каждое подобное уточнение модели приводит к увеличению количества неизвестных коэффициентов и усложнению задачи их восстановления, что может нивелировать потенциально большую точность таких моделей. Поэтому в данной работе мы используем деление зараженных только на два класса. Отметим также, что границу между этими двумя классами можно провести по-разному, и это не обязательно должна быть граница между строгим отсутствием и наличием симптомов. В данной работе под “бессимптомными больными” мы понимаем в первую очередь тех, кто не обращается за медицинской помощью и поэтому не включается в статистику заражений.

Уравнения системы (1) основаны на предположении, что переходы людей из одной группы в другую осуществляются с некоторыми постоянными интенсивностями, обусловленными свойствами инфекционного агента (кроме переходов из группы s в группу e , т.е. заражений, интенсивность которых зависит от количества зараженных). Так, α есть интенсивность перехода из группы бессимптомных зараженных в группу больных с симптомами, ρ есть интенсивность выздоровления бессимптомных зараженных, β есть интенсивность выздоровления симптомных больных, μ есть интенсивность гибели симптомных больных, γ_e и γ_i — интенсивности утраты иммунитета бессимптомными зараженными и симптомными больными соответственно. Под интенсивностью перехода здесь понимается вероятность перехода в единицу времени. Точнее, значение α интенсивности перехода из одной группы в другую означает, что вероятность такого перехода за время Δt для каждого человека из первой группы равна $\alpha \Delta t + o(\Delta t)$ при $\Delta t \rightarrow 0$. Коэффициенты a_e и a_i описывают заразность бессимптомных зараженных и симптомных больных соответственно. В общем случае эти коэффициенты могут зависеть от времени. Мы рассматриваем простейший случай, когда исследуемый отрезок времени $[0, T]$ разбит на S равных частей и функции $a_e(t)$ и $a_i(t)$ постоянны на каждом из полученных отрезков. Через a_{ek} и a_{ik} мы обозначаем значения функций $a_e(t)$ и $a_i(t)$ на k -м из этих отрезков, $1 \leq k \leq S$. В примерах ниже выбирается значение $S = 2$.

Мы предполагаем, что эпидемиологическая статистика заражений, выздоровлений и смертности соответствует значениям функций $i(t)$, $r_o(t)$ и $p(t)$, называемых ниже наблюдаемыми величинами, в равноотстоящие моменты времени $t_i = i\tau$, $0 \leq i \leq K$, $\tau = T/K$, так что $t_0 = 0$, $t_K = K\tau = T$. Следует учитывать, что статистические значения $i_1, \dots, i_K, r_1, \dots, r_K, p_1, \dots, p_K$ являются приближенными и могут отличаться от соответствующих им значений функций $i(t_1), \dots, i(t_K), r(t_1), \dots, r(t_K), p(t_1), \dots, p(t_K)$. Именно, статистические значения могут содержать погрешности следующих видов.

1. Систематическое занижение статистики, обусловленное ограниченностью возможностей медицинской системы. Мы предполагаем, что все три статистики (заражений, выздоровлений и смертности) занижены в каждом ареале примерно в одинаковое количество λ раз, которое может быть установлено путем сравнения статистики смертности с данными об избыточной смертности в рассматриваемом ареале по сравнению с периодом времени до пандемии.

2. Временные сдвиги эпидемиологической статистики относительно реальных значений наблюдаемых величин, обусловленные возможными задержками между заболеванием, выздоровлением, смертью человека и включением этой информации в статистику.
3. Малые случайные погрешности, не относящиеся к двум предыдущим видам.

Рассмотрим подходы к исследованию системы (1), позволяющие учитывать эти три вида погрешностей.

2. Решение задачи восстановления коэффициентов. Изучение модели (1) имеет две основные цели: прогнозирование распространения эпидемии на будущие промежутки времени и вычисление эпидемиологических параметров, таких как индекс репродукции \mathcal{R}_0 (определяющий заразность инфекции), летальность, вероятность заболеть в единицу времени и др. В свою очередь, эта информация позволяет оценить степень опасности инфекции и целесообразность введения тех или иных мер борьбы с эпидемией. Для достижения обеих целей необходимо найти неизвестные коэффициенты $\alpha, \rho, \beta, \mu, \gamma_e, \gamma_i$, а также $a_{ek}, a_{ik}, 1 \leq k \leq S$, и начальные значения всех функций в системе (1), в том числе ненаблюдаемых величин $e(t), r_e(t), r_i(t)$.

Для нахождения коэффициентов β и μ предлагается использовать следующий алгоритм [9]. Рассмотрим векторы $\tilde{\mathbf{i}} = (i_1, \dots, i_K)$, $\tilde{\mathbf{r}} = (r_1, \dots, r_K)$, $\tilde{\mathbf{p}} = (p_1, \dots, p_K)$, содержащие статистические данные, относящиеся к моментам времени t_1, \dots, t_K . Рассмотрим также векторы $\tilde{\mathbf{d}}\mathbf{i} = (di_1, \dots, di_K)$, $\tilde{\mathbf{d}}\mathbf{r} = (dr_1, \dots, dr_K)$, $\tilde{\mathbf{d}}\mathbf{p} = (dp_1, \dots, dp_K)$, координаты которых получены из соответствующих статистических данных с помощью подходящей формулы численного дифференцирования. Мы используем формулу

$$df_i = \frac{1}{h} \left(\frac{1}{12} f_{i-2h} - \frac{2}{3} f_{i-h} + \frac{2}{3} f_{i+h} - \frac{1}{12} f_{i+2h} \right), \quad 1 \leq i \leq K,$$

где вместо буквы f могут стоять i, r и p . Эта формула использовалась также в работе [7]; она имеет пятый порядок точности и относится к классу, описанному в [10, гл. 2]. Шаг дискретизации h выбирается равным пяти суткам. Если бы статистические данные были известны точно, если бы формула численного дифференцирования также была точной (т.е. были бы справедливы равенства $di_i = i'(t_i)$, $dr_i = r'_o(t_i)$, $dp_i = p'(t_i)$, $1 \leq i \leq K$) и если бы используемая математическая модель в точности описывала распространение эпидемии, то из системы (1) следовали бы векторные равенства $\tilde{\mathbf{d}}\mathbf{r} = \beta \tilde{\mathbf{i}}$, $\tilde{\mathbf{d}}\mathbf{p} = \mu \tilde{\mathbf{i}}$. Значит, оценить степень адекватности используемой эпидемиологической модели в применении к имеющимся данным можно, вычислив отношения

$$\frac{(\tilde{\mathbf{d}}\mathbf{r}, \tilde{\mathbf{i}})}{\|\tilde{\mathbf{d}}\mathbf{r}\| \|\tilde{\mathbf{i}}\|}, \quad \frac{(\tilde{\mathbf{d}}\mathbf{p}, \tilde{\mathbf{i}})}{\|\tilde{\mathbf{d}}\mathbf{p}\| \|\tilde{\mathbf{i}}\|}, \quad (2)$$

равные косинусам углов между векторами $\tilde{\mathbf{d}}\mathbf{r}, \tilde{\mathbf{i}}$ и между векторами $\tilde{\mathbf{d}}\mathbf{p}, \tilde{\mathbf{i}}$ соответственно. Если величины (2) близки к единице, то сами углы между указанными векторами близки к нулю, так что векторы $\tilde{\mathbf{d}}\mathbf{r}, \tilde{\mathbf{d}}\mathbf{p}, \tilde{\mathbf{i}}$ приблизительно сонаправлены и модель можно использовать.

Если отношения (2) оказались далеки от единицы, причину этого можно искать во временных сдвигах статистических данных относительно реальных значений функций i, r_o, p . По-разному сдвигая три статистики друг относительно друга и для каждого их взаимного положения рассчитывая указанные отношения, можно определить искомые величины сдвигов из условия максимальной близости обоих косинусов к единице. Вместе с использованием информации об избыточной смертности этот подход позволяет устранить погрешности входных данных, относящиеся к первым двум видам.

Отношения $\|\tilde{\mathbf{d}}\mathbf{r}\|/\|\tilde{\mathbf{i}}\|, \|\tilde{\mathbf{d}}\mathbf{p}\|/\|\tilde{\mathbf{i}}\|$, найденные с учетом сдвигов статистики, выбираются в качестве начальных приближений к значениям искомых коэффициентов β и μ . Затем эти значения уточняются с помощью итеративно регуляризованного метода Гаусса–Ньютона, устойчивого относительно погрешностей третьего вида. Перейдем к описанию этого метода.

Мы представляем задачу восстановления неизвестных параметров модели (1) в виде операторного уравнения

$$\mathbf{F}(\mathbf{x}) = \mathbf{f}, \quad (3)$$

где \mathbf{x} — вектор, содержащий все неизвестные параметры модели, \mathbf{f} — вектор, содержащий значения наблюдаемых величин в моменты времени t_1, \dots, t_K , а вектор $\mathbf{F}(\mathbf{x})$ содержит рассчитанные из системы (1) с параметрами \mathbf{x} значения тех же наблюдаемых величин в те же моменты времени. Для решения урав-



нения (3) применяется итеративно регуляризованный метод Гаусса–Ньютона [11]

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathcal{P}_Q(\overline{\mathbf{x}}_{n+1}), \\ \overline{\mathbf{x}}_{n+1} &= \boldsymbol{\xi} - (\mathbf{F}'^*(\mathbf{x}_n)\mathbf{F}'(\mathbf{x}_n) + \alpha_n\mathbf{E})^{-1}\mathbf{F}'^*(\mathbf{x}_n)(\mathbf{F}(\mathbf{x}_n) - \mathbf{f}_\delta - \mathbf{F}'(\mathbf{x}_n)(\mathbf{x}_n - \boldsymbol{\xi})). \end{aligned} \quad (4)$$

Используется также модификация этого метода [7, 12], позволяющая находить не только решения, но и квазирешения операторных уравнений, т.е. точки минимума функционала невязки $\|\mathbf{F}(\mathbf{x}) - \mathbf{f}\|$:

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathcal{P}_Q(\overline{\mathbf{x}}_{n+1}), \\ \overline{\mathbf{x}}_{n+1} &= \boldsymbol{\xi} - ([\mathbf{F}'^*(\mathbf{x}_n)\mathbf{F}'(\mathbf{x}_n)]^2 + \alpha_n\mathbf{E})^{-1}[\mathbf{F}'^*(\mathbf{x}_n)\mathbf{F}'(\mathbf{x}_n)]\mathbf{F}'^*(\mathbf{x}_n)[\mathbf{F}(\mathbf{x}_n) - \mathbf{f}_\delta - \mathbf{F}'(\mathbf{x}_n)(\mathbf{x}_n - \boldsymbol{\xi})]. \end{aligned} \quad (5)$$

Использование этих методов регуляризации позволяет получать решения, устойчивые относительно погрешностей третьего вида. Векторы \mathbf{x}_0 и $\boldsymbol{\xi}$ — априори выбираемые начальные приближения для искомого вектора неизвестных коэффициентов \mathbf{x} , \mathbf{E} — единичный оператор, вектор \mathbf{f}_δ содержит данные эпидемиологической статистики, обработанные описанным выше способом для устранения погрешностей первых двух видов. Последовательность параметров регуляризации $\{\alpha_n\}$ в методах (4), (5) должна удовлетворять условиям

$$0 < \alpha_{n+1} \leq \alpha_n, \quad n = 0, 1, \dots; \quad \sup_{n=0,1,\dots} \frac{\alpha_n}{\alpha_{n+1}} < \infty; \quad \lim_{n \rightarrow \infty} \alpha_n = 0.$$

В наших расчетах в качестве последовательности параметров регуляризации выбирается геометрическая прогрессия $\alpha_n = 10^{-3} \cdot 0.7^{n+1}$. Наконец, Q есть допустимое множество изменений параметров модели, предполагаемое выпуклым и замкнутым, а \mathcal{P}_Q — оператор метрического проектирования на Q . В наших расчетах мы относим к множеству Q наборы параметров, содержащие только неотрицательные значения и удовлетворяющие условиям $\varkappa, \rho, \beta, \mu, a_{ek}, a_{ik} \geq 10^{-4}$. Смысл поиска квазирешений вместо решений заключается в том, что в силу приближенности эпидемиологической модели уравнение (3) может не иметь решений. Другие возможные подходы к решению нелинейных операторных уравнений описаны, например, в [11, 13].

Отличительной особенностью итерационных методов решения некорректных задач, к которым относится задача (3), является необходимость выбора момента останова итераций. Если итерации не останавливать, то методы (4) и (5), как правило, расходятся. Известны критерии останова итераций, априорные и апостериорные, которые можно использовать применительно к данным методам [7, 11, 12]. В настоящей работе используется упрощенный подход: мы останавливаем итерации в тот момент, когда достигается минимума невязка $\|\mathbf{F}(\mathbf{x}_n) - \mathbf{f}_\delta\|$. Следует отметить, что описанный упрощенный способ выбора момента останова итераций теоретически может приводить к ситуации “переобучения”, когда искомые коэффициенты подбираются так, что решение системы (1) становится слишком близким к наблюдаемым данным с их случайными погрешностями, а не к реальным значениям соответствующих величин. В теории машинного обучения известен способ предотвращения “переобучения” (см., например, [14]), основанный на минимизации ошибки не на тренировочном множестве данных, на основе которого происходит обучение, а на отдельном валидационном множестве. Применительно к нашей задаче аналогом тренировочного множества данных является эпидемиологическая статистика на промежутке времени $[0, T]$, а аналогом валидационного множества может служить эпидемиологическая статистика, соответствующая будущим моментам времени $t > T$. Другими словами, момент останова итераций разумно определять не из условия наилучшего приближения текущих статистических данных, служивших входными данными используемого метода, а из условия наилучшего прогноза будущих статистических данных. Конечно, этот способ был бы неприменим, если будущие моменты времени еще не настали и главная задача заключается как раз в составлении прогноза. Однако наши численные эксперименты показывают, что номер итерации, на которой достигается наилучший прогноз, обычно не сильно отличается от номера итерации, на которой достигается наилучшее приближение текущих статистических данных. Это позволяет считать, что в рассматриваемой задаче “переобучения” не происходит — вероятно, из-за того, что модель (1) слишком “бедна” для того, чтобы ее решения могли слишком хорошо подстраиваться под случайные погрешности статистических данных.

Методы (4) и (5) применялись для восстановления коэффициентов модели (1), соответствующих шести парам месяцев в составе 2021 года. Таким образом, длительность отрезка времени $[0, T]$ в расчетах составляла два месяца, а величина промежутка τ равнялась одним суткам. При использовании методов (4) и (5) применительно к каждой следующей паре месяцев в качестве начальных приближений a_{ek}, a_{ik} ,

$1 \leq k \leq S$, \varkappa , ρ , γ_e , γ_i , входящих в состав вектора $\mathbf{x}_0 = \boldsymbol{\xi}$, выбирались значения, полученные при решении задачи восстановления коэффициентов для предыдущей пары месяцев. Выбор начальных приближений к значениям параметров β и μ указан выше. Опишем выбор начальных приближений к значениям $e(0)$, $r_e(0)$, $r_i(0)$ ненаблюдаемых величин в системе (1), которые также относятся к неизвестным параметрам модели и необходимы для решения соответствующей задачи Коши. Из второго уравнения системы (1) следует равенство

$$e(0) = \frac{\dot{i}(0) + (\beta + \mu)i(0)}{\varkappa},$$

правая часть которого может быть приближенно найдена из статистики заражений с помощью численного дифференцирования. Полученное значение и принималось в качестве начального приближения к $e(0)$. Начальное приближение к $r_i(0)$ было взято равным значению $r_o(0)$, приближенно известному из эпидемиологической статистики. Наконец, начальное приближение к $r_e(0)$ определялось из правдоподобного приближенного равенства $\frac{r_e(0)}{r_i(0)} \approx \frac{e(0)}{i(0)}$.

Описанный подход с использованием метода (4) ниже называется алгоритмом 1, а с использованием метода (5) — алгоритмом 2. Для сравнения использовался также упрощенный подход к восстановлению коэффициентов (алгоритм 3), не учитывающий погрешности входных данных первых двух типов. Параметр S в этом упрощенном подходе принимался равным 1, т.е. параметры a_e и a_i считались постоянными на протяжении каждой пары месяцев. Для восстановления коэффициентов модели при этом использовался метод (4). В качестве начальных приближений к $e(0)$, $r_e(0)$ и $r_i(0)$ в алгоритме 3 выбирались значения $e(T)$, $r_e(T)$ и $r_i(T)$, полученные при расчете, соответствующем предыдущему промежутку времени.

3. Задача восстановления коэффициентов эпидемиологической модели в общем виде.

Сформулируем задачу восстановления неизвестных параметров эпидемиологической модели в более общем виде, который позволит применять развиваемый подход к анализу не только модели (1), но и других подобных моделей. Итак, пусть дана система из m дифференциальных уравнений относительно функций $X_1(t), \dots, X_m(t)$, образующих вектор-функцию $\mathbf{X}(t)$, с набором неизвестных параметров $\mathbf{x} = (k_1, \dots, k_N)$:

$$\frac{d\mathbf{X}}{dt} = \mathbf{g}(t, \mathbf{X}, \mathbf{x}), \quad \mathbf{X}(0) = \boldsymbol{\alpha}(\mathbf{x}). \quad (6)$$

Пусть величины X_1, \dots, X_l доступны для наблюдения, а величины X_{l+1}, \dots, X_m не доступны. В частности, начальные данные $X_j^{(0)} = X_j(0)$, $l + 1 \leq j \leq m$, неизвестны и должны входить в состав вектора неизвестных параметров \mathbf{x} . В модели (1) к наблюдаемым относятся величины i , r_o и p , а к ненаблюдаемым — величины e , r_e и r_i , т.е. $l = 3$, $m = 6$. Система (6) рассматривается на временном промежутке $t \in [0, T]$, на котором выделены моменты наблюдений $0 = t_0 < t_1 < \dots < t_K = T$.

Задачу восстановления неизвестных параметров системы (6) можно представить в виде операторного уравнения (3), в котором $\mathbf{f} \in \mathbb{R}^{Kl}$ есть вектор, включающий значения наблюдаемых величин X_1, \dots, X_l в моменты времени (t_1, \dots, t_K) . Далее, оператор \mathbf{F} ставит в соответствие вектору неизвестных коэффициентов \mathbf{x} значения наблюдаемых величин X_1, \dots, X_l в моменты (t_1, \dots, t_K) , определяемые решением системы дифференциальных уравнений (6) с вектором параметров \mathbf{x} . Обозначим значение функции X_j , где $1 \leq j \leq m$, в момент t , полученное при решении системы (6) с параметрами \mathbf{x} , через $X_j(t; \mathbf{x})$. Таким образом, компонентами вектора $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^{kl}$ являются величины $X_j(t_i; \mathbf{x})$, $1 \leq i \leq K$, $1 \leq j \leq m$. Задача заключается в нахождении такого вектора неизвестных параметров \mathbf{x} , при котором значения наблюдаемых величин, предсказанные эпидемиологической моделью и составляющие вектор $\mathbf{F}(\mathbf{x})$, совпадают с реальными значениями этих величин, составляющими вектор \mathbf{f} , или минимально от них отличаются.

Для нахождения решений и квазирешений уравнения (3) можно использовать методы регуляризации (4) и (5). В этих методах \mathbf{F}' есть производная Фреше оператора \mathbf{F} . Известно, что производная Фреше нелинейного оператора представляет собой линейный непрерывный оператор, действующий в тех же самых пространствах. Поскольку эти пространства конечномерны, оператор $\mathbf{F}'(\mathbf{x})$ в любой точке \mathbf{x} описывается матрицей Якоби нелинейного оператора \mathbf{F} . Компонентами этой матрицы являются частные производные $u_{j\sigma}(t_i) = \frac{\partial X_j(t_i; \mathbf{x})}{\partial k_\sigma}$ наблюдаемых функций X_1, \dots, X_l по параметрам k_1, \dots, k_N в моменты времени t_1, \dots, t_K . Эти производные могут быть найдены из системы дифференциальных уравнений [15, §23]:

$$\frac{du_{j\sigma}}{dt} = \left(\sum_{i=1}^m \frac{\partial g_j}{\partial X_i} u_{i\sigma} \right) + \frac{\partial g_j}{\partial k_\sigma}, \quad u_{j\sigma}(0) = \frac{\partial \alpha_j(\mathbf{x})}{\partial k_\sigma}, \quad 1 \leq j \leq l, \quad 1 \leq \sigma \leq N. \quad (7)$$



Заметим, что коэффициенты $\frac{\partial g_j}{\partial X_i}$, $\frac{\partial g_j}{\partial k_\sigma}$ в этой системе, вообще говоря, зависят от функций $X_j(t)$, $1 \leq j \leq m$. Это значит, что для решения уравнений (7) их следует объединить в одну систему с уравнениями (6).

4. Вывод формулы для коэффициента заразности \mathcal{R}_0 . Коэффициент \mathcal{R}_0 является широко известной характеристикой эпидемического процесса. Он показывает, сколько человек в среднем заражает один зараженный, находясь в полностью восприимчивом окружении. Количество заражений от каждого конкретного человека складывается из количества людей, зараженных им в бессимптомном состоянии, и, если симптомы появились, из количества людей, зараженных после этого. Подходы к вычислению этого коэффициента на основе различных эпидемиологических моделей описаны в [8], см. также приведенные там ссылки.

Пусть в некоторый момент времени $t = 0$ заразились \tilde{e}_0 человек, вначале они находятся в бессимптомном состоянии. Пусть $\tilde{e}(t)$ — количество бессимптомных зараженных среди этих \tilde{e}_0 человек в момент t . В частности, $\tilde{e}(0) = \tilde{e}_0$. В соответствии с моделью (1) величина $\tilde{e}(t)$ удовлетворяет следующей задаче Коши:

$$\begin{cases} \dot{\tilde{e}}(t) = -(\varkappa + \rho)\tilde{e}(t), \\ \tilde{e}(0) = \tilde{e}_0. \end{cases}$$

Решая эту задачу, получаем $\tilde{e}(t) = \tilde{e}_0 e^{-(\varkappa + \rho)t}$. Таким образом, к моменту t остается $\tilde{e}_0 e^{-(\varkappa + \rho)t}$ бессимптомных из рассматриваемых \tilde{e}_0 человек. Значит, вероятность остаться бессимптомным к моменту t для каждого из них равна $e^{-(\varkappa + \rho)t}$, а вероятность перехода в другое состояние (т.е. вероятность выздоровления или появления симптомов) равна $1 - e^{-(\varkappa + \rho)t}$. Среднее время перехода в другое состояние, т.е.

математическое ожидание момента перехода, равно $\int_0^{+\infty} t \lambda(t) dt$, где

$$\lambda(t) = (1 - e^{-(\varkappa + \rho)t})' = (\varkappa + \rho)e^{-(\varkappa + \rho)t}$$

есть плотность вероятности перехода. Таким образом, среднее время перехода равно

$$\int_0^{+\infty} t(\varkappa + \rho)e^{-(\varkappa + \rho)t} dt = \frac{1}{\varkappa + \rho}.$$

Согласно модели (1), за любой малый промежуток времени $[t, t + \Delta t]$ заболевают $\varkappa \tilde{e}(t) \Delta t + o(\Delta t)$ людей, а $\rho \tilde{e}(t) \Delta t + o(\Delta t)$ людей выздоравливают. Поэтому вероятности заболевания и выздоровления для бессимптомного больного соотносятся как $\varkappa : \rho$. Так как в сумме эти вероятности дают единицу, вероятность заболевания равна $\frac{\varkappa}{\varkappa + \rho}$, а вероятность выздоровления равна $\frac{\rho}{\varkappa + \rho}$.

Запишем уравнение для относительного количества $e(t)$ бессимптомных зараженных из (1) в виде:

$$\frac{de(t)}{dt} = a_e s(t)e(t) + a_i s(t)i(t) - (\varkappa + \rho)e(t).$$

Пусть численность людей в рассматриваемом ареале равна P . Тогда абсолютное количество бессимптомных зараженных равно $Pe(t)$ и изменяется со скоростью

$$\frac{d(Pe(t))}{dt} = Pa_e s(t)e(t) + Pa_i s(t)i(t) - P(\varkappa + \rho)e(t).$$

Здесь первое слагаемое описывает заражения восприимчивых лиц от бессимптомных и показывает, что скорость таких заражений равна $Pa_e s(t)e(t)$. Эти заражения производятся $Pe(t)$ бессимптомными зараженными. Значит, один бессимптомный зараженный заражает восприимчивых в среднем со скоростью

$$\frac{Pa_e s(t)e(t)}{Pe(t)} = a_e s(t).$$

Поэтому за время $\frac{1}{\varkappa + \rho}$, которое человек в среднем проводит в бессимптомном состоянии, он успеет заразить в среднем $\frac{a_e s(t)}{\varkappa + \rho}$ людей. После этого с вероятностью $\frac{\varkappa}{\varkappa + \rho}$ он переходит в симптомное состояние,

а с вероятностью $\frac{\rho}{\varkappa + \rho}$ выздоравливает. В первом случае остается вопрос: сколько еще он успеет заразить людей? Для ответа на этот вопрос необходимо найти среднее время нахождения в симптомном состоянии и заразность в этом состоянии.

Среднее время нахождения в симптомном состоянии найдем по аналогии с предыдущими рассуждениями. Пусть в момент $t = 0$ заболело \tilde{i}_0 человек, пусть $\tilde{i}(t)$ — количество тех из них, кто продолжает болеть к моменту t . Задача Коши для величины $\tilde{i}(t)$ имеет вид

$$\begin{cases} \dot{\tilde{i}}(t) = -(\beta + \mu)\tilde{i}(t), \\ \tilde{i}(0) = \tilde{i}_0. \end{cases}$$

Отсюда $\tilde{i}(t) = \tilde{i}_0 e^{-(\beta + \mu)t}$. Тогда к моменту t остается $\tilde{i}_0 e^{-(\beta + \mu)t}$ человек в симптомном состоянии из общего количества \tilde{i}_0 . Значит, вероятность к моменту t остаться в симптомном состоянии для каждого из них равна $e^{-(\beta + \mu)t}$. Вероятность перехода в другое состояние, т.е. вероятность выздоровления или гибели, к моменту t равна $1 - e^{-(\beta + \mu)t}$. Плотность вероятности перехода поэтому равна

$$\lambda(t) = (1 - e^{-(\beta + \mu)t})' = (\beta + \mu)e^{-(\beta + \mu)t}$$

и среднее время перехода получается равным

$$\int_0^{+\infty} t\lambda(t)dt = \int_0^{+\infty} t(\beta + \mu)e^{-(\beta + \mu)t}dt = \frac{1}{\beta + \mu}.$$

За время Δt один симптомный больной заражает $a_i s(t)\Delta t$ восприимчивых. Поэтому за время нахождения в симптомном состоянии он успеет заразить в среднем $\frac{a_i s(t)}{\beta + \mu}$ людей. Мы получаем, что один зараженный в среднем суммарно заражает

$$\mathcal{R} = \frac{a_e s(t)}{\varkappa + \rho} + \frac{\varkappa}{\varkappa + \rho} \cdot \frac{a_i s(t)}{\beta + \mu}$$

человек. Заметим, что найденный показатель заразности допускает представление

$$\mathcal{R} = \mathcal{R}_0 s(t), \quad \mathcal{R}_0 = \frac{a_e}{\varkappa + \rho} + \frac{\varkappa}{\varkappa + \rho} \cdot \frac{a_i}{\beta + \mu}.$$

Коэффициент \mathcal{R}_0 показывает степень заразности эпидемии без учета количества восприимчивых и позволяет оценить влияние мер по борьбе с эпидемией на ее заразность. Он равен среднему количеству людей, которых заражал бы один зараженный, если бы ни у кого не было иммунитета и все люди были бы восприимчивыми к инфекции, т.е. при $s(t) = 1$.

5. Показатели опасности эпидемии. В этом разделе выводятся формулы для двух показателей, характеризующих опасность эпидемии. Первый из них, который мы обозначим через Ω , равен вероятности погибнуть в случае заражения. Второй показатель Ψ равен вероятности серьезно заразиться за единицу времени, где под серьезным заражением понимается заражение с последующим смертельным исходом.

Выразим показатель Ω через коэффициенты модели (1). Рассуждая как в разделе 4, заметим, что вероятность выздоровления для симптомного больного равна $\frac{\beta}{\beta + \mu}$, а вероятность смерти равна $\frac{\mu}{\beta + \mu}$, так как эти вероятности соотносятся как $\beta : \mu$. Вероятность Ω погибнуть в случае заражения равна произведению вероятности заболеть в случае заражения, найденной выше и равной $\frac{\varkappa}{\varkappa + \rho}$, и вероятности $\frac{\mu}{\beta + \mu}$ погибнуть в случае заболевания. Таким образом,

$$\Omega = \frac{\varkappa}{\varkappa + \rho} \cdot \frac{\mu}{\beta + \mu}.$$

Перейдем к выводу формулы для показателя Ψ , равного вероятности серьезно заразиться за единицу времени. Для определенности будем искать значение этого показателя в момент времени $t = 0$. Заметим, что производная по времени относительного количества $s(t) = 1 - e(t) - i(t) - r_e(t) - r_i(t) - p(t)$ восприимчивых к инфекции людей равна

$$\dot{s}(t) = -\dot{e}(t) - \dot{i}(t) - \dot{r}_e(t) - \dot{r}_i(t) - \dot{p}(t).$$



Комбинируя это равенство с уравнениями модели (1), получаем:

$$\dot{s}(t) = -a_e s(t)e(t) - a_i s(t)i(t) + \gamma_e r_e(t) + \gamma_i r_i(t). \quad (8)$$

Первое и второе слагаемые в правой части (8) описывают заражение восприимчивых лиц от бессимптомных зараженных и от больных с симптомами. Третье и четвертое слагаемые описывают прибавку количества восприимчивых за счет утраты иммунитета выздоровевшими.

Рассмотрим в интересующий нас момент $t = 0$ достаточно большую группу, состоящую из \tilde{s}_0 человек, восприимчивых к инфекции. Обозначим через $\tilde{s}(t)$ количество восприимчивых среди этих \tilde{s}_0 человек к моменту t , тогда количество заразившихся к этому моменту равно $\tilde{s}_0 - \tilde{s}(t)$. Также пусть $e_0 = e(0)$, $i_0 = i(0)$. Так как нас интересует вывод показателя опасности Ψ в начальный момент времени, будем рассматривать только моменты t , достаточно близкие к нулю, так что можно считать $e(t) \approx e_0$, $i(t) \approx i_0$. Замечая, что третье и четвертое слагаемые в (8) не имеют отношения к рассматриваемой группе из \tilde{s}_0 человек, приходим к задаче Коши

$$\begin{cases} \dot{\tilde{s}}(t) = -\tilde{s}(t)(a_e e_0 + a_i i_0), \\ \tilde{s}(0) = \tilde{s}_0. \end{cases}$$

Отсюда $\tilde{s}(t) = \tilde{s}_0 e^{-(a_e e_0 + a_i i_0)t}$. Мы видим, что вероятность заразиться за время t у каждого члена исходной группы из s_0 восприимчивых равна $1 - e^{-(a_e e_0 + a_i i_0)t}$. Вероятность заражения за единицу времени равна производной от этой функции при $t = 0$, т.е. $a_e e_0 + a_i i_0$. Так как вероятность смертельного исхода в случае заражения равна Ω , искомая вероятность заразиться с последующим смертельным исходом за единицу времени равна

$$\Psi = (a_e e_0 + a_i i_0)\Omega = (a_e e_0 + a_i i_0) \frac{\chi}{\chi + \rho} \cdot \frac{\mu}{\beta + \mu}.$$

Величину Ψ мы будем рассматривать как основной показатель опасности эпидемии.

6. Результаты расчетов. Построен комплекс программ в среде Maple, осуществляющий решение задачи восстановления коэффициентов эпидемиологических моделей в соответствии с предлагаемым подходом и позволяющий вычислять показатели заразности и опасности эпидемии, а также строить прогноз распространения эпидемии.

В качестве входных статистических данных по России использовалась официальная эпидемиологическая статистика, публиковавшаяся на сайте “Новости о коронавирусе в России | стопкоронавирус”¹, по зарубежным странам — эпидемиологическая статистика, представленная на справочном сайте Worldometers — real time world statistics. Статистика от Worldometers была выбрана в связи с тем, что она включает данные по выздоровлениям за весь 2021 г., в отличие от, например, статистики от Центра системных исследований и инжиниринга при Университете Джона Хопкинса. Для расчета показателя занижения статистики λ использовалась информация об избыточной смертности в 2021 г., опубликованная в электронном издании Our World in Data (<https://ourworldindata.org/>). Значения избыточной смертности в каждом из изучаемых ареалов были разделены на приращение количества погибших от COVID-19 за 2021 г. согласно статистике от Worldometers. Соответствующие коэффициенты занижения статистики λ приведены в табл. 1.

Для определения сдвигов статистики путем анализа значений величин (2) использовался один большой промежуток времени с 1 января по 31 октября 2021 г., для того чтобы результаты расчета минимально зависели от случайных флуктуаций статистических данных. Мы не включили в этот промежуток два последних месяца, чтобы даже при максимальных возможных сдвигах на результаты не влияла волна эпидемии, вызванная штаммом “омикрон”, так как полноценный учет этой волны потребовал бы использования статистики 2022 г. Статистики заражений и выздоровлений сдвигались относительно статистики смертности на целое количество дней от -20 до 20 , и выбирались сдвиги, доставляющие максимум отношениям (2). Рассчитанные величины задержки τ_i статистики заражений и задержки τ_r статистики выздоровлений по сравнению со статистикой смертности приведены в табл. 1. То, что для разных стран эти задержки имеют различный знак, требует интерпретации. Положительные значения τ_i и τ_r легко объяснить тем, что между моментом заражения или выздоровления и занесением этого факта в эпидемиологическую статистику проходит значительное время, в то время как смерть человека от инфекции

¹ URL сайта <https://стопкоронавирус.рф>

Таблица 1. Показатель уровня вакцинации v , показатель занижения статистики λ , задержки статистики τ_i и τ_r
 Table 1. Vaccination rate v , understatement coefficient λ , statistics delays τ_i, τ_r

Ареал / Region	v	λ	τ_i	τ_r
Россия / Russia	1.01	2.8	9	20
Армения / Armenia	0.59	2.1	3	8
Казахстан / Kazakhstan	0.91	4.6	1	6
Аргентина / Argentina	1.72	1.3	-8	-3
Канада / Canada	1.79	1	-15	-10
Германия / Germany	1.84	1	-20	-13
Дания / Denmark	2.02	1.2	-20	-14

фиксируется быстрее. Отрицательные значения τ_i и τ_r могут быть связаны с внутренним недостатком модели (1), в которой не учитывается ненулевая минимальная длительность болезни перед гибелью больного. Из-за этого фактора, например, при наступлении новой волны эпидемии всплеск смертности начинается только через некоторое время после всплеска заболеваемости, но в рамках модели (1) это может интерпретироваться как задержка статистики смертности, т.е. как отрицательные значения τ_i и τ_r . Отметим, что в табл. 1 отчетливо видна положительная корреляция между задержками τ_i, τ_r , отражающими трудности в своевременной постановке диагноза и обработке медицинских данных, и коэффициентом занижения статистики λ .

Для нахождения неизвестных коэффициентов модели (1) применительно к каждой паре месяцев в составе 2021 г. применялись алгоритмы 1–3, описанные в разделе 2. Выбранные нами начальные приближения для указанных алгоритмов применительно к первому промежутку времени (январь и февраль 2021 г.) указаны в табл. 2. Эти значения были подобраны экспериментально. Впрочем, результаты расчетов в целом устойчивы по отношению к изменениям начальных приближений. Значения $i(0)$ и $r_o(0)$ в табл. 2 брались из эпидемиологической статистики, с учетом ее сдвигов и занижения (в алгоритмах 1 и 2) или без учета этих факторов (в алгоритме 3). Выбор начальных приближений к параметрам β и μ в алгоритмах 1 и 2 описан в разделе 2.

Построенный комплекс программ был применен к анализу эпидемиологической статистики, описывающей распространение эпидемии COVID-19 в России, Армении, Казахстане, Аргентине, Канаде, Германии и Дании. Перечисленные страны достаточно сильно отличаются друг от друга по размерам, численности населения, географическому местоположению, а также по объемам проведенной в 2021 г. вакцинации от COVID-19. Объем вакцинации будем измерять величиной v , равной отношению использованных к концу 2021 г. доз вакцины к численности населения. Отметим, что указанное отношение может быть больше единицы, так как каждый человек обычно получает несколько доз вакцины. Известно, что в России, Армении и Казахстане это отношение значительно ниже, чем в Аргентине, Канаде, Германии и Дании (табл.1). Данные о количестве доз вакцины были взяты из электронного издания Our World in Data. Отметим, что используемая нами эпидемиологическая модель (1) не учитывает вакцинацию в явном виде, но может учитывать ее косвенно. Именно, влияние вакцинации на распространение инфекции проявляется в изменении параметров модели при переходе от более ранних промежутков времени в составе 2021 г. к более поздним.

Таблица 2. Начальные приближения для методов (4) и (5) применительно к первому промежутку времени (01.21–02.21). Способ выбора параметров β и μ в алгоритмах 1 и 2 описан в разделе 2
 Table 2. The initial approximations for the methods (4) and (5) as applied to the first time interval (01.21–02.21). The method for identifying the parameters β and μ in Algorithms 1 and 2 is described in Section 2

Параметр Parameter	Алгоритмы 1 и 2 Algorithms 1 and 2	Алгоритм 3 Algorithm 3
\varkappa	0.01	0.06
ρ	0.06	0.07
β	—	0.03
μ	—	0.06
γ_e	0.01	0
γ_i	0.15	0
a_e	0.36	0.1
a_i	0.23	0.2
$e(0)$	$12i(0)$	$2i(0)$
$r_e(0)$	$6r_o(0)$	$2r_o(0)$



Разумно предположить, что вакцинация должна приводить к уменьшению описанных выше показателей \mathcal{R}_0 и Ψ . Для проверки этой гипотезы были найдены значения \mathcal{R}_0 и $\lg\Psi$ для каждой из перечисленных стран в каждую из шести пар месяцев в течение 2021 года. Эти значения были рассчитаны на основе коэффициентов модели (1), восстановленных по данным эпидемиологической статистики согласно алгоритмам 1–3. Расчеты показывают, что предложенные алгоритмы не позволяют устойчиво определять показатель \mathcal{R}_0 . Значения этого показателя, найденные с помощью различных алгоритмов, существенно отличаются друг от друга и от найденных значений этого же показателя в соседние промежутки времени. Для отдельных ареалов и промежутков времени получаются неправдоподобно высокие значения \mathcal{R}_0 , равные десяткам или сотням единиц, в сочетании с очень низкими значениями показателя \varkappa , что соответствует случаю чрезвычайно заразной инфекции, которая, тем не менее, очень редко проявляется в виде симптомной болезни. Таким образом, предложенные алгоритмы с трудом различают низкую заразность и низкую заболеваемость зараженных. По той же причине наши методы плохо подходят для определения показателя Ω , показывающего вероятность погибнуть при условии наличия заражения.

Вместе с тем предложенные методы определяют значения показателя опасности Ψ значительно более устойчиво. С помощью алгоритмов 1 и 2 были найдены значения $\lg\Psi$ для каждого ареала и каждого из двенадцати месяцев года (напомним, что каждый промежуток времени длительностью два месяца в рамках этих алгоритмов разбивается на $S = 2$ равных промежутка с различающимися значениями параметров a_e и a_i , что приводит к различным значениям показателя Ψ). Затем было найдено среднее из этих двенадцати значений для каждого ареала, а также их среднеквадратичное отклонение от среднего. С помощью упрощенного алгоритма 3 также были получены значения $\lg\Psi$ в каждой из шести пар месяцев и затем найдены их средние значения и среднеквадратичные отклонения для каждого ареала. Эти результаты приведены в табл. 3–5. В этих же таблицах для каждого расчета указана относительная невязка алгоритма, рассчитанная по формуле $\|\mathbf{F}(\mathbf{x}) - \mathbf{f}_\delta\| / \|\mathbf{f}_\delta\|$ и равная отношению среднеквадратичного отклонения рассчитанных значений наблюдаемых величин от соответствующих статистических данных к среднему квадратическому статистических данных.

Все три алгоритма показывают большие значения показателя опасности Ψ в странах с низким уровнем вакцинации, чем в странах с высоким уровнем вакцинации. Это в целом относится и к алгоритмам 1 и 2, учитывающим занижение статистики, и к алгоритму 3, в котором оно не учитывается. Более конкретно, страны с низким уровнем вакцинации (Армения, Казахстан, Россия) имеют средний показатель $\lg\Psi$ больше -5 согласно алгоритмам 1, 2 и больше -5.3 согласно алгоритму 3, в то время как для стран с высоким уровнем вакцинации (Аргентина, Канада, Германия, Дания) справедливо противоположное неравенство. Исключением являются лишь средние значения $\lg\Psi$ для Аргентины согласно алгоритму 1 и для Казахстана согласно алгоритму 3; при этом в обоих случаях два других алгоритма показывают средние значения $\lg\Psi$, соответствующие уровню вакцинации. Заметим, что упрощенный алгоритм 3 обычно показывает меньшее значение Ψ , чем алгоритмы 1 и 2; это можно объяснить тем, что в алгоритме 3 не учитывается систематическое занижение эпидемиологической статистики.

Отдельные примеры результатов работы программы показаны на рис. 1–4. Приведены графики, показывающие количество зараженных $I(t)$, количество выздоровевших $R(t)$ и количество погибших $P(t)$ согласно статистическим данным, умноженным на показатель занижения статистики λ (точечные линии), и согласно расчету из модели (1) с восстановленными коэффициентами на отрезке времени $[0, T]$ длительностью два месяца (красные линии). Отметим, что функции $I(t)$, $R(t)$ и $P(t)$ есть функции $i(t)$,

Таблица 3. Алгоритм 1, результаты расчетов

Table 3. Algorithm 1, calculation results

Ареал / Region	Относительная невязка Relative discrepancy	Среднее $\lg\Psi$ Average $\lg\Psi$	Среднеквадратичное отклонение $\lg\Psi$ Standard deviation of $\lg\Psi$
Россия / Russia	0.013	-4.59	0.24
Армения / Armenia	0.022	-3.01	2.17
Казахстан / Kazakhstan	0.023	-4.56	0.82
Аргентина / Argentina	0.014	-4.43	1.35
Канада / Canada	0.008	-5.17	0.36
Германия / Germany	0.049	-5.76	0.48
Дания / Denmark	0.023	-5.78	1.04

Таблица 4. Алгоритм 2, результаты расчетов

Table 4. Algorithm 2, calculation results

Ареал / Region	Относительная невязка Relative discrepancy	Среднее $\lg \Psi$ Average $\lg \Psi$	Среднеквадратичное отклонение $\lg \Psi$ Standard deviation of $\lg \Psi$
Россия / Russia	0.014	−3.62	0.83
Армения / Armenia	0.010	−3.71	1.10
Казахстан / Kazakhstan	0.027	−4.17	1.51
Аргентина / Argentina	0.015	−5.01	1.62
Канада / Canada	0.017	−5.47	0.34
Германия / Germany	0.036	−5.46	0.82
Дания / Denmark	0.048	−5.92	0.58

Таблица 5. Алгоритм 3, результаты расчетов

Table 5. Algorithm 3, calculation results

Ареал / Region	Относительная невязка Relative discrepancy	Среднее $\lg \Psi$ Average $\lg \Psi$	Среднеквадратичное отклонение $\lg \Psi$ Standard deviation of $\lg \Psi$
Россия / Russia	0.004	−5.26	0.24
Армения / Armenia	0.015	−4.05	1.16
Казахстан / Kazakhstan	0.009	−5.86	0.81
Аргентина / Argentina	0.008	−5.40	0.96
Канада / Canada	0.022	−5.58	0.49
Германия / Germany	0.015	−5.77	0.95
Дания / Denmark	0.096	−6.28	0.88

$r_o(t)$ и $p(t)$ из модели (1), умноженные на численность популяции рассматриваемого ареала. Также на графиках показан прогноз распространения эпидемии за пределами временного отрезка $[0, T]$, использованного для восстановления коэффициентов (синие линии). Для составления прогноза решалась задача Коши для системы (1) с найденными коэффициентами и значениями ненаблюдаемых величин $e(T)$, $r_e(T)$ и $r_i(T)$ в момент $t = T$; значения же наблюдаемых величин $i(T)$, $r_o(T)$ и $p(T)$ в этот момент брались из эпидемиологической статистики с учетом ее сдвигов и занижения. Видно, что в ряде случаев модель дает достаточно точный количественный прогноз на несколько недель, а иногда позволяет предугадывать качественное изменение указанных величин. Отметим, однако, что качественный прогноз наблюдается далеко не во всех расчетах. Вместе с тем эксперименты показывают, что подбор начальных приближений для итеративных методов (4) и (5) в большинстве случаев позволяет значительно улучшить получаемые результаты.

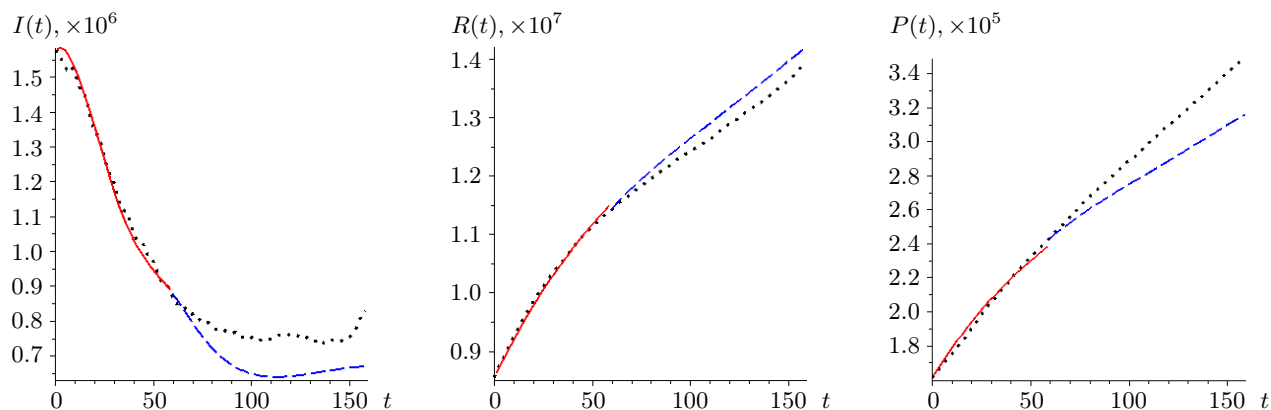


Рис. 1. Россия, алгоритм 1, январь–февраль 2021 г.

Fig. 1. Russia, Algorithm 1, January–February 2021

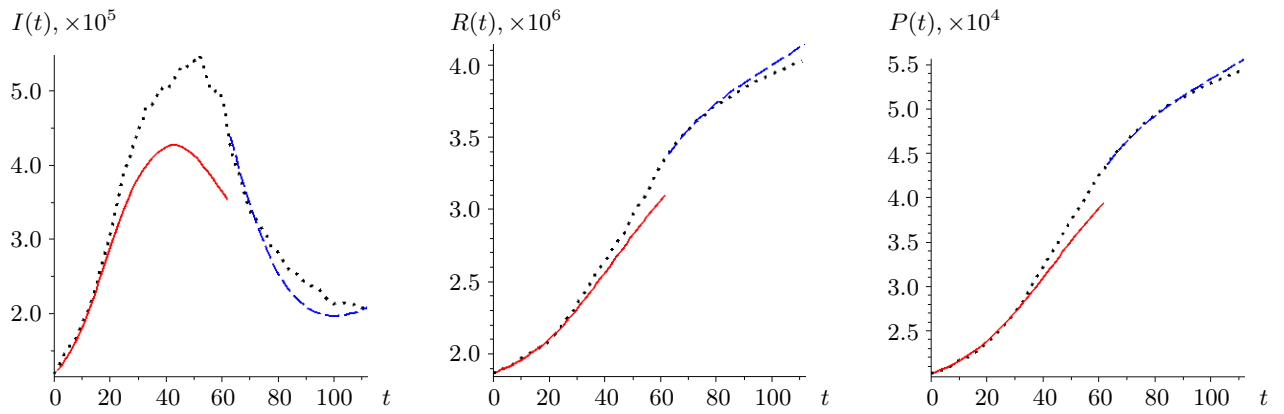


Рис. 2. Казахстан, алгоритм 1, июль–август 2021 г.

Fig. 2. Kazakhstan, Algorithm 1, July–August 2021

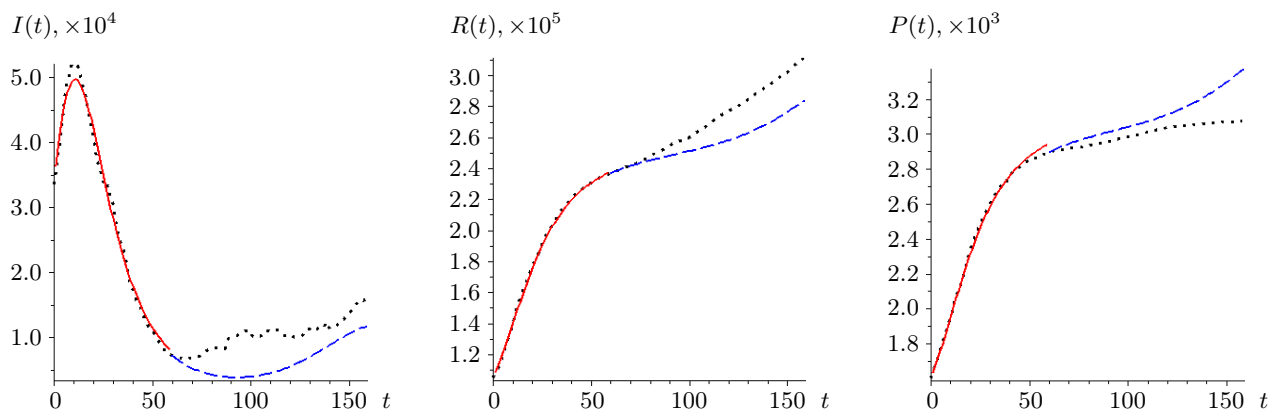


Рис. 3. Дания, алгоритм 2, январь–февраль 2021 г.

Fig. 3. Denmark, Algorithm 2, January–February 2021

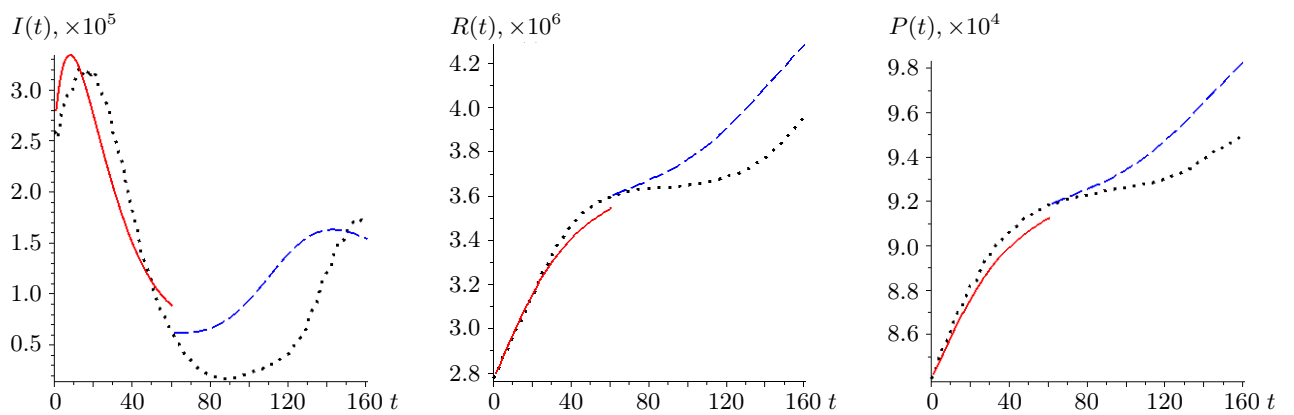


Рис. 4. Германия, алгоритм 2, май–июнь 2021 г.

Fig. 4. Germany, Algorithm 2, May–June 2021

7. Заключение. В статье на основе итеративно регуляризованных методов типа Гаусса–Ньютона был разработан подход к решению задачи восстановления коэффициентов SEIR-подобной эпидемиологической модели по данным эпидемиологической статистики. Подход учитывает наличие в используемой статистике погрешностей разного вида, включая систематическое занижение, временные сдвиги и ма-

лые случайные погрешности. На основе предлагаемого подхода были разработаны и реализованы в виде программ в системе Maple три алгоритма решения изучаемой задачи. Эти алгоритмы были численно протестированы на реальных данных по распространению эпидемии COVID-19 в России и шести зарубежных странах. Продемонстрирована применимость построенных программ для прогнозирования распространения эпидемии, а также для оценки ее показателя опасности, что необходимо для принятия решений о тех или иных мерах борьбы с эпидемией. Приведенные графики показывают, что во многих случаях реальное распространение эпидемии может адекватно описываться изучаемой моделью с параметрами, идентифицированными с помощью разработанных алгоритмов.

Возможными дальнейшими направлениями исследований являются усовершенствование алгоритма определения временных сдвигов статистики, а также автоматизация подбора начального приближения для итеративно регуляризованных методов с целью получения стабильно качественных результатов расчетов. Кроме того, в дальнейших исследованиях планируется использовать математический аппарат дифференциальных уравнений с запаздыванием, дополняющий идею временных сдвигов эпидемиологической статистики.

Список литературы

1. *Kabanikhin S.I., Krivorot'ko O.I., Ermolenko D.V., et al.* Inverse problems in immunology and epidemiology // Eurasian J. Math. Comp. Appl. 2017. **5**, N 2. 14–35. <https://ejmca.enu.kz/assets/files/5-2-2.pdf>. Cited January 10, 2025.
2. *Криворотько О.И., Кабанихин С.И., Зятков Н.Ю. и др.* Математическое моделирование и прогнозирование COVID-19 в Москве и Новосибирской области // Сиб. журн. вычисл. матем. 2020. **23**, № 4. 395–414. doi 10.15372/SJNM20200404.
3. *Schiesser W.E.* Computational modeling of the COVID-19 disease: numerical ODE analysis with R programming. Singapore: World Scientific, 2020.
4. *Sameni R.* Mathematical modeling of epidemic diseases. A case study of the COVID-19 coronavirus // <https://arxiv.org/abs/2003.11371>. Cited January 10, 2025.
5. *Smirnova A., Bakushinsky A.* On iteratively regularized predictor–corrector algorithm for parameter identification // Inverse Problems. 2020. **36**, Article Number 125015. doi 10.1088/1361-6420/abc530.
6. *Smirnova A., deCamp L., Chowell G.* Forecasting epidemics through nonparametric estimation of time-dependent transmission rates using the SEIR Model // Bull. Math. Biol. 2019. **81**. 4343–4365. doi 10.1007/s11538-017-0284-3.
7. *Kokurin M.M., Kokurin M.Yu., Semenova A.V.* Iteratively regularized Gauss–Newton type methods for approximating quasi-solutions of irregular nonlinear operator equations in Hilbert space with an application to COVID-19 epidemic dynamics // Appl. Math. Comput. 2022. **431**, Article Number 127312. doi 10.1016/j.amc.2022.127312.
8. *Криворотько О.И., Кабанихин С.И.* О математическом моделировании COVID-19 // Сибирские электронные математические известия. 2023. **20**, № 2. 1211–1268.
9. *Кокурин М.М.* Введение в функциональный анализ и его приложения. Йошкар-Ола: Марийский гос. ун-т, 2022.
10. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. М.: БИНОМ, 2007.
11. *Бакушинский А.Б., Кокурин М.Ю.* Алгоритмический анализ нерегулярных операторных уравнений. М.: LENAND, 2012.
12. *Кокурин М.М.* Апостериорный останов в итеративно регуляризованных методах типа Гаусса–Ньютона для нахождения квазирешений нерегулярных операторных уравнений // Известия вузов. Математика. 2022. № 2. 29–42. doi 10.26907/0021-3446-2022-2-29-42.
13. *Тихонов А.Н., Леонов А.С., Ягола А.Г.* Нелинейные некорректные задачи. М.: Наука, 1995.
14. *Николенко С., Кадурич А., Архангельская Е.* Глубокое обучение. СПб.: Питер, 2022.
15. *Филлипов А.Ф.* Введение в теорию дифференциальных уравнений. М.: КомКнига, 2007.



Информация об авторах

Михаил Михайлович Кокурин — к.ф.-м.н., заведующий научно-исследовательской лабораторией; Марийский государственный университет, Институт цифровых технологий, пл. Ленина, 1, 424000, Йошкар-Ола, Российская Федерация.

Анастасия Валерьевна Гаврилова — младший научный сотрудник; Марийский государственный университет, пл. Ленина, 1, 424000, Йошкар-Ола, Российская Федерация.

Абду Рауф Чеха Мохамед — аспирант; Российский университет дружбы народов имени Патриса Лумумбы, ул. Миклухо-Маклая, 6, 117198, Москва, Российская Федерация.

References

1. S. I. Kabanikhin, O. I. Krivorot'ko, D. V. Ermolenko, et al., "Inverse Problems in Immunology and Epidemiology," *Eurasian J. Math. Comp. Appl.* **5** (2), 14–35 (2017). <https://ejmca.enu.kz/assets/files/5-2-2.pdf>. Cited January 10, 2025.
2. O. I. Krivorot'ko, S. I. Kabanikhin, N. Yu. Zyat'kov, et al., "Mathematical Modeling and Forecasting of COVID-19 in Moscow and Novosibirsk Region," *Sib. Zh. Vych. Mat.* **23** (4), 395–414 (2020) [*Numer. Anal. Appl.* **13** (4), 332–348 (2020)]. doi 10.1134/S1995423920040047.
3. W. E. Schiesser, *Computational Modeling of the COVID-19 Disease: Numerical ODE Analysis with R Programming* (World Scientific, Singapore, 2020).
4. R. Sameni, "Mathematical Modeling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus." <https://arxiv.org/abs/2003.11371>. Cited January 10, 2025.
5. A. Smirnova and A. Bakushinsky, "On Iteratively Regularized Predictor–Corrector Algorithm for Parameter Identification," *Inverse Probl.* **36**, Article Number 125015 (2020). doi 10.1088/1361-6420/abc530.
6. A. Smirnova, L. deCamp, and G. Chowell, "Forecasting Epidemics through Nonparametric Estimation of Time-Dependent Transmission Rates Using the SEIR Model," *Bull. Math. Biol.* **81**, 4343–4365 (2019). doi 10.1007/s11538-017-0284-3.
7. M. M. Kokurin, M. Yu. Kokurin, and A. V. Semenova, "Iteratively Regularized Gauss–Newton Type Methods for Approximating Quasi-Solutions of Irregular Nonlinear Operator Equations in Hilbert Space with an Application to COVID-19 Epidemic Dynamics," *Appl. Math. Comput.* **431**, Article Number 127312 (2022). doi 10.1016/j.amc.2022.127312.
8. O. I. Krivorot'ko and S. I. Kabanikhin, "On Mathematical Modeling of COVID-19," *Sib. Electron. Mat. Izv.* **20** (2), 1211–1268 (2023).
9. M. M. Kokurin, *Introduction to Functional Analysis and Its Applications* (Mari State University, Yoshkar-Ola, 2022).
10. N. S. Bakhvalov, N. P. Zhidkov, and G. M. Kobel'kov, *Numerical Methods* (BINOM, Moscow, 2007) [in Russian].
11. A. B. Bakushinsky and M. Yu. Kokurin, *Algorithmic Analysis of Irregular Operator Equations* (LENAND, Moscow, 2012) [in Russian].
12. M. M. Kokurin, "A Study of a Posteriori Stopping in Iteratively Regularized Gauss–Newton-Type Methods for Approximating Quasi-Solutions of Irregular Operator Equations," *Izv. Vyssh. Uchebn. Zaved., Mat., No. 2*, 29–42 (2022) [*Russ. Math.* **66** (2), 24–35 (2022)]. doi 10.3103/S1066369X22020062.
13. A. N. Tikhonov, A. S. Leonov, and A. G. Yagola, *Nonlinear Ill-Posed Problems* (Nauka, Moscow, 1995; CRC Press, London, 1998).
14. S. Nikolenko, A. Kadurin, and E. Arhangel'skaya, *Deep Learning* (Piter, Saint-Petersburg, 2022) [in Russian].
15. A. F. Filippov, *Introduction to the Theory of Differential Equations* (KomKniga, Moscow, 2007) [in Russian].

**Information about the authors**

Mikhail M. Kokurin — Ph. D., Head of Research Laboratory; Mari State University, Institute of Digital Technologies, Lenina ploshchad', 1, 424000, Yoshkar-Ola, Russia.

Anastasiya V. Gavrilova — Junior Research Assistant; Mari State University, Lenina ploshchad', 1, 424000, Yoshkar-Ola, Russia.

Abdou Raouf Cheha Mohamed — Postgraduate Student; Peoples' Friendship University of Russia named after Patrice Lumumba, Miklukho-Maklaya ulitsa, 6, 117198, Moscow, Russia.