

Анализ и прогнозирование COVID-19

А.Е. Суроегина, Е.Н. Пелиновский
nsuroegina@mail.ru, enpeli@mail.ru



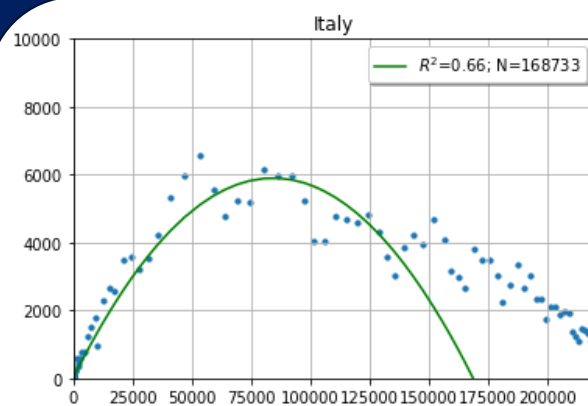
Аннотация

В данной работе проведены анализ и прогнозирование динамики развития COVID-19 с помощью логистического уравнения и модели Гомперца со сдвигом. Данные о заболеваемости коронавирусом в России взяты с сайта Всемирной организации здравоохранения. По скользящим данным за пять дней были построены аппроксимации данных моделей, которые позволяют сделать прогноз на несколько дней вперед. Обе модели с хорошей точностью приближаются к действительным показателям.

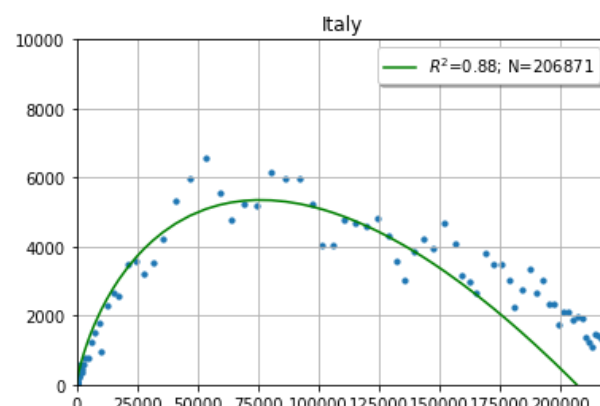
Анализ первой волны

Для описания скорости распространения и прогнозирования их последствий используется ряд нелинейных математических моделей различной сложности. Логистическое уравнение (уравнение Ферхульста) и модель Гомперца на примере распространения коронавирусной инфекции для первой волны рассмотрены в работах [1-2], однако, существуют и более сложные модели [3-4]. Но эпидемия не закончилась на первой волне, на данный момент в России на подходе уже новая.

Обе модели хорошо показали себя для первой волны. Однако, в работе [5] предлагается доступный работающий метод прогнозирования поведения эпидемии на различных этапах только для логистической модели со сдвигом.



Логистическая модель



Модель Гомперца

Основная идея автора состоит в том, чтобы взять данные за определенное количество дней (17 дней: известная скорость распространения на данный момент и на шестнадцать предшествующих), аппроксимировать по ним параболу

$$N_{n+1} = a(M - N_n)(N - N_n), \quad (1)$$

где N_n - количество всех больных, а N_{n+1} - скорость заболевания. Полученные коэффициенты a, M, N подставляются в решение уравнения (1):

$$N(t) = M + 0.5(N - M) \left(1 + \tanh(k(t - T)) \right), \quad (2)$$

где $k = 0.5a(N - M)$, а T вычисляется подстановкой в уравнение (2) числа заболевших на данный день. В результате чего получается прогноз заболеваемости на некоторое время вперед. Таким образом, получая новые данные и двигая кривую по ним, получается можно узнать дальнейшее поведение. Несмотря на простоту подхода, данная модель довольно хорошо работает.

Основные вопросы:

- Почему берутся данные за 17 дней, а не меньше, ведь для построения параболы достаточно трех точек?
- На сколько дней вперед можно сделать прогноз?
- Можно ли наложить этот алгоритм на другие модели?

Проблема 1. Оптимальное количество значений

Был написан алгоритм, который позволил аппроксимировать кривые для различного количества точек: от 3 до 20. Чтобы оценить полученные результаты, было посчитано количество точек, наиболее приближенных к аппроксимированной кривой и к его решению. В результате чего было получено, что наиболее хорошие прогнозы получались при выборе данных от 4 до 7 дней. Однако, стоит отметить, что при фиксированном количестве точек, достаточно мало парабол было построено, чем хотелось бы, так как не любая парабола подойдет (ветви должны быть направлены вниз, точки пересечения с осью Ox должны быть положительны и т.п.). Поэтому алгоритм был изменен таким образом, что если «хорошая» парабола не строится, то мы добавляем точки из прошлого. Это позволило сделать более подробные прогнозы.

Анализ и прогнозирование COVID-19

А.Е. Суроегина, Е.Н. Пелиновский
nsuroegina@mail.ru, enpeli@mail.ru

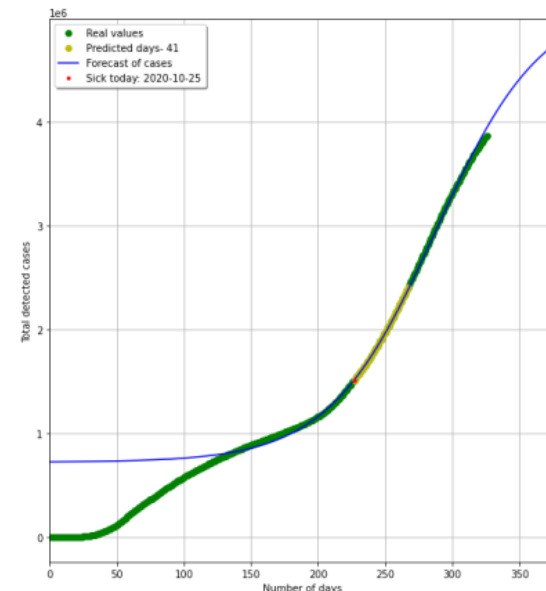
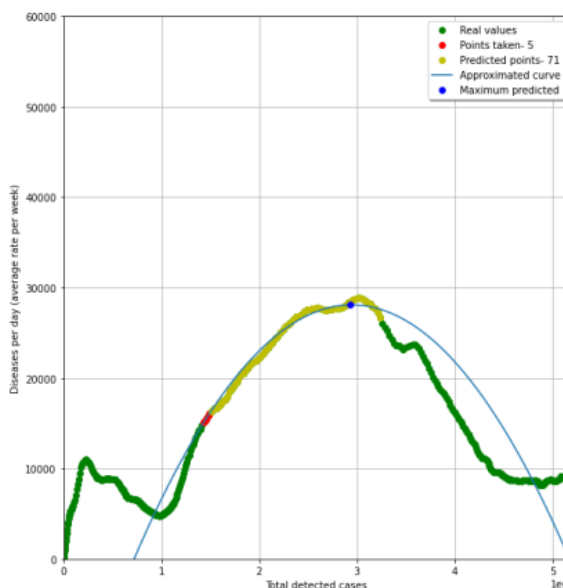


Логистическая модель со сдвигом

$$N_{n+1} = a(M - N_n)(N - N_n)$$

Рис 1. Зависимость средней скорости распространения, от общего количества для модели логистического уравнения со сдвигом.

Рис 2. Общее количество заболевших в день для модели логистического уравнения со сдвигом



Проблема 2. Оценка прогноза.

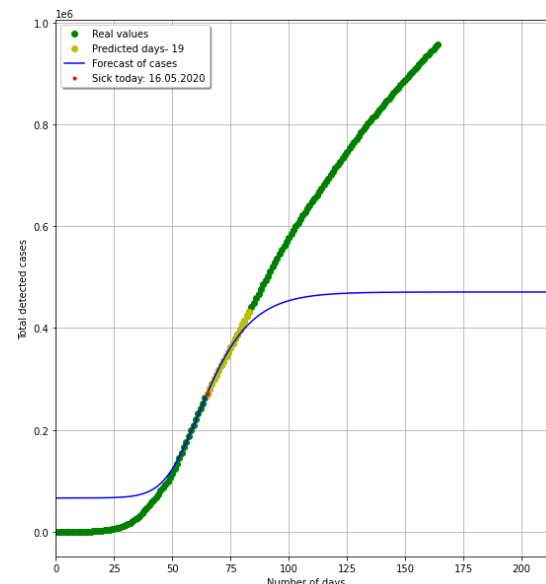
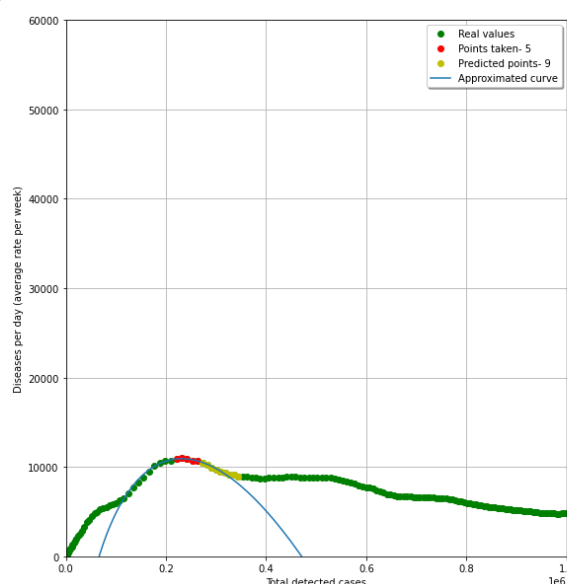
Второй этап оценки, был реализован при помощи подсчета наиболее приближенных значений к построенным графикам. Желтым цветом были отмечены точки, которые находились ближе к кривым. Несмотря на то, что правый график достаточно близок к фактическим значениям, расстояние будет в пределах 20 000, что практически не заметно в таких масштабах. Таким образом получили в среднем прогноз на 11 дней, что довольно хорошо, при условии, что взяты данные за пять дней.

Модель Гомперца со сдвигом

$$N_{n+1} = a(M - N_n)(N - \ln(N_n))$$

Рис 3. Зависимость средней скорости распространения, от общего количества для модели Гомперца со сдвигом.

Рис 4. Общее количество заболевших в день для модели Гомперца со сдвигом.



Проблема 3. Другие модели со сдвигом

Реализованный алгоритм был наложен на модель Гомперца со сдвигом: $N_{n+1} = a(M -$

Результаты

Каждая модель (логистическая и Гомперца со сдвигом) имеет свои плюсы и минусы, но в целом их поведение схоже, поэтому прогноз, построенный на основе этих моделей, довольно схож. Оптимальное количество точек для построения модели не больше 7, а средний прогноз на 11-12 дней.

Частично исследования поддержаны грантом РФФ 19-12-00253.

Литература

[1] E. Pelinovsky, A. Kurkin, O. Kurkina, M. Kokoulina, A. Epifanova - Logistic equation and COVID-19// Chaos, Solitons and Fractals. – 2020.- V.140.-P.110241

[2] E. Pelinovsky, M. Kokoulina, A. Epifanova, A. Kurkin, O. Kurkina, M. Tang, E. Macau, M. Kirillin - Gompertz model in COVID-19 spreading simulation // Chaos, Solitons and Fractals.- 2022.- V.-154. - P.111699

[3] Nicholas Tze Ping Pang, Assis Kamu, Mohd Amiruddin Mohd Kassim, Chong Mun Ho.- Monitoring the impact of Movement Control Order (MCO) in flattening the cumulative daily cases curve of Covid-19 in Malaysia: A generalized logistic growth modeling approach// Infectious Disease Modelling 2021. V. 6, P. 898-908

[4] Bruno Hebling Vieira, Nathalia Hanna Hiar, George C. Cardoso - Uncertainty Reduction in Logistic Growth Regression Using Surrogate Systems Carrying Capacities: a COVID-19 Case Study// Brazilian Journal of Physics.- 2022.- V.- 52. P.-15

[5] С.М. Абрамов, С.О. Травин - Об анализе, моделировании и прогнозе статистики эпидемии коронавируса в России // Цифровая экономика.- 2020.- Т. 2(10). С. 5-19