

First evidence of downturn

Santu Ghosh & Bhaswati Ganguli

Assistant Professor, Dept. Biostatistics, St John's Medical College, Bangalore
Professor, Dept. of Statistics, Calcutta University, Kolkata

Introduction

The following analysis explored possible downturn of COVID 19 infection rate in India. Daily active cases are taken for the analysis (Source: <https://www.covid19india.org/>). As daily number of active cases are count data, we converted that into log scale to make it continuous before apply semiparametric regression method to estimate the curve over time. The semiparametric regression with linear mixed model setup with piecewise cubic polynomial basis function was used (Durban et al 2005). Let assume Y_t be the log # of active case on day t, the regression model is defined as follows

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \sum_{k=1}^K u_k (t - K_k)_+^3 + \epsilon_t \quad (1)$$

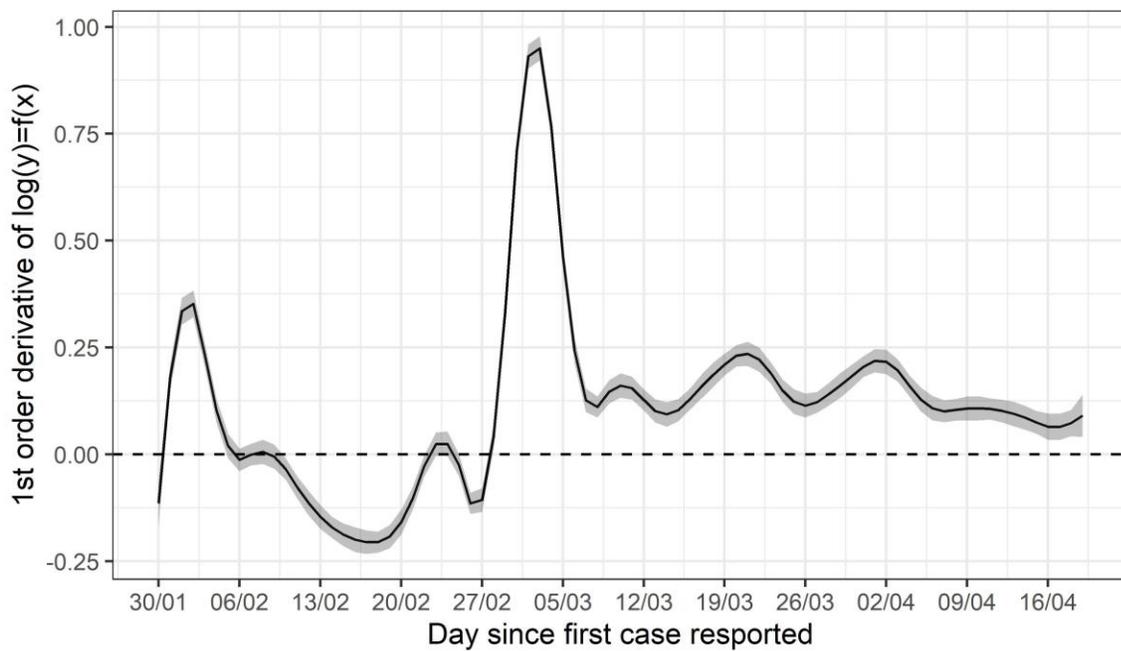
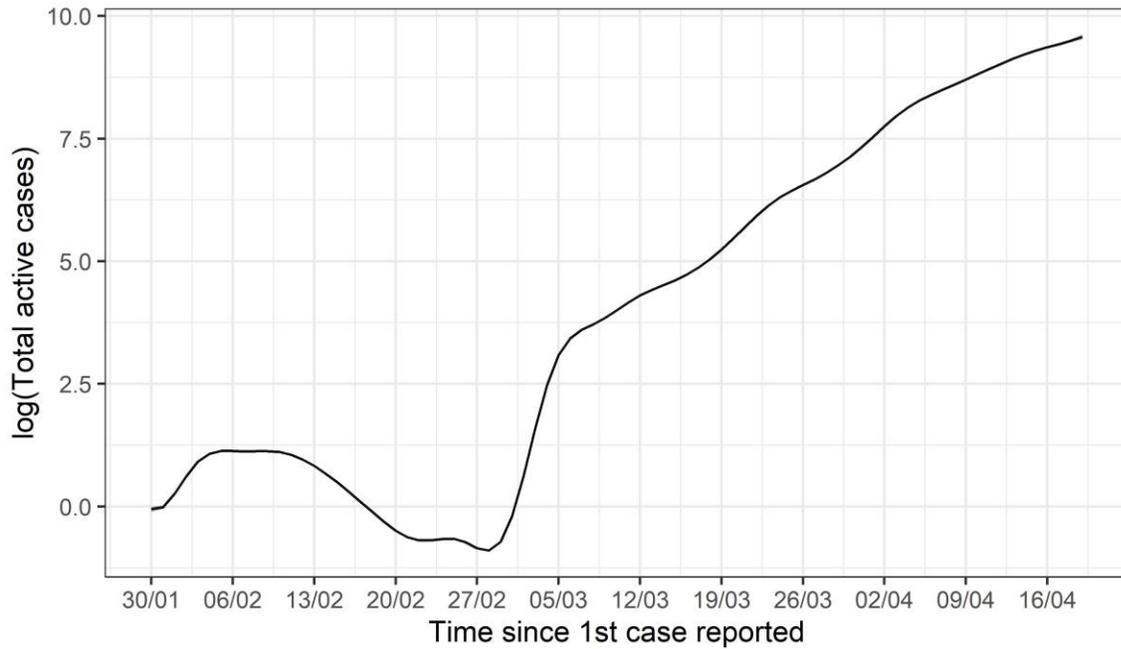
Where $u_k \sim N(0, \sigma_u^2)$ & $\epsilon_t \sim N(0, \sigma_e^2)$; $(t - K_k)_+^3 = \begin{cases} (t - K_k)^3, & t > K_k \\ 0, & \text{otherwise} \end{cases}$. The number of knots K is fixed and large enough (in this case K =40) to ensure the flexibility of the curve. The knots are chosen as quantiles of t with probabilities $1/(K+1), \dots, K/(K+1)$.

Understanding movement of the curve

First order derivative of a differentiable function always gives lot of information about the curve. The first derivative of a differentiable continuous function for given point is defined as the slope of the tangent line to a curve at the given point. Hence will indicate the movement of the curve, decrement of the slope indicates downward movement and increment indicates upward movement. The slope '0' may indicate a peak or dip or a plateau. One can easily assess mathematically estimating the slopes of tangents at all possible points the peak, dip or plateau observing unique characteristics of the features of them.

The objective of this exercise is to investigate the movement of curve over time and to estimate time to peak, evidence of downturn, time to dip or even time to plateau. Hence, the 1st order differentiation of the equation (1) was estimated by the method developed by Bandyopadhyay and Ganguli (2020). The time to peak, dip or plateau were explored based on simultaneous confidence interval of the derivative obtained by reverse regression technique mentioned by the above authors in their article.

The following figure (bottom panel) depicts the slope and its 95% CI of the semiparametric curve of log of the number of daily active cases of COVID 19 in India. The curve shows some downturn but doesn't reach at peak or dip yet.



Reference

1. M. Durban, J. Harezlak, M.P. Wand, and M.P. Carroll. Simple fitting of subject-specific curves for longitudinal data. *Statistics in Medicine*,24:1153–1167, 2005
2. Souvik Bandyopadhyay, Bhaswati Ganguli(2020). Estimating time to plateau(in communication)