1 Introduction

The aim of this work is to estimate the parameters of the SIR model. In particular, I will focus on the first wave of contagion of the Coronavirus pandemic in Italy.

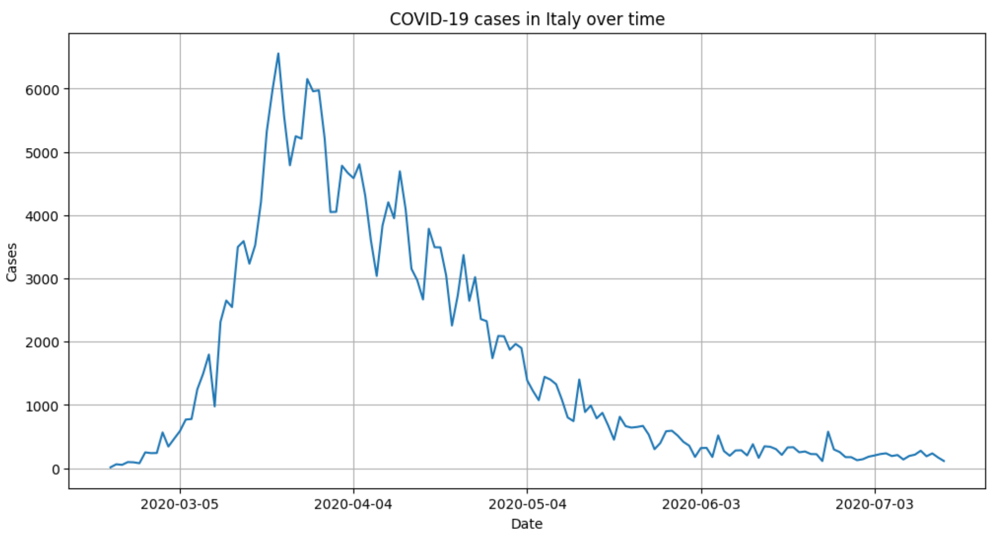


Figure 1: New infections per day in Italy

2 Model

2.1 Standard SIR model

My first try was with the following differential equation for the total infected population:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right), \quad N(0) = N_0$$

Which admits the following solution:

$$N(t) = \frac{K}{1 + \left(\frac{K-P_0}{P_0}\right)e^{-rt}}$$

That has a flex (i.e. the max of the new infections curve) in:

$$t^* = \frac{1}{r}\ln\left(\frac{K-P_0}{P_0}\right)$$

To use a linear regression, the differential equation can be discretized into:

$$N_{t+1} = \alpha_1 N_t + \alpha_2 N_t^2$$

So the parameter can be found using:

$$r = \alpha_1 - 1, \quad K = \frac{1-\alpha_1}{\alpha_2}, \quad P_0 = 1$$

2.2 Bernoulli equation model

Recently, I found that:

$$N_{t+1} = \alpha_1 N_t + \alpha_2 N_t^{1+\varepsilon} \tag{1}$$

With $\varepsilon \in (0,1)$ would work much better to reproduce the right-hand skewness of the contagion curve.
This equation too has a close form solution, being a Bernoulli differential equation:

$$V = N^{-\varepsilon}, \qquad \frac{\dot{V}}{V} = -r\varepsilon\frac{\dot{N}}{N}, \quad \Rightarrow \quad \dot{V} = -r\varepsilon V + \frac{r\varepsilon}{K}.$$

Which has close form solution:

$$V(t) = K^{-1} + \left(V_0 - K^{-1}\right) e^{-r\varepsilon t} \quad \Rightarrow \quad N(t) = \frac{1}{\left(K^{-1} + \left(N_0^{-\varepsilon} - K^{-1}\right) e^{-r\varepsilon t}\right)^{\varepsilon}}.$$

The maximum of its derivative is determined by solving:

$$\ddot{N} = -\frac{V^{-\frac{1+\varepsilon}{\varepsilon}}}{\varepsilon}\left(\ddot{V} - \frac{1+\varepsilon}{\varepsilon}\frac{\dot{V}^2}{V}\right)$$

Where:

$$\dot{V} = -r\varepsilon\left(V_0 - K^{-1}\right) e^{-r\varepsilon t}, \qquad \ddot{V} = (r\varepsilon)^2\left(V_0 - K^{-1}\right) e^{-r\varepsilon t}.$$

In terms of finite differences:

$$V_{t+1} = \beta_0 + \beta_1 V_t,$$

where:

$$\beta_0 = \frac{r\varepsilon}{K}, \quad \beta_1 = r\varepsilon.$$

2.3 Taylor expansion for small ε

Assuming $\varepsilon \simeq 0$, we can expand $N_t^{1+\varepsilon}$ as:

$$N_t^{1+\varepsilon} \simeq N_t\left(1 + \varepsilon \ln N_t + \frac{\varepsilon^2(\ln N_t)^2}{2} + \frac{\varepsilon^3(\ln N_t)^3}{6}\right). \tag{2}$$

So:

$$N_{t+1} = (\alpha_1 + \alpha_2)N_t + \varepsilon\alpha_2 N_t \ln N_t + \varepsilon^2\alpha_2 N_t\frac{(\ln N_t)^2}{2} + \varepsilon^3\alpha_2 N_t\frac{(\ln N_t)^3}{6}. \tag{3}$$

3 Estimation

3.1 Preliminary Regression

Let's regress the following model with OLS:

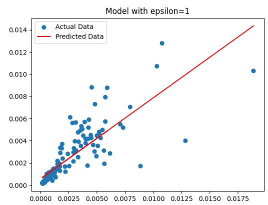
$$V_{t+1} = \beta_0 + \beta_1 V_t \tag{4}$$

where $V_t = N_t^{-\varepsilon}$, for different values of ε.

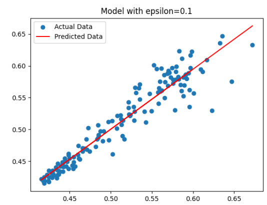
Table 1

	$\varepsilon =: 1$	$\varepsilon =: 0.1$	$\varepsilon =: 0.01$	$\varepsilon =: 0.001$	$\varepsilon =: 0.0001$
const	0.0006***	0.0293**	0.0454**	0.0474**	0.0476**
	(0.0002)	(0.0131)	(0.0221)	(0.0233)	(0.0234)
x1	0.7310***	0.9422***	0.9514***	0.9522***	0.9523***
	(0.0427)	(0.0253)	(0.0236)	(0.0234)	(0.0234)
R-squared	0.6771	0.9080	0.9206	0.9218	0.9219
R-squared Adj.	0.6748	0.9073	0.9200	0.9212	0.9213
N	142	142	142	142	142

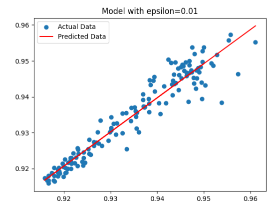
Standard errors in parentheses. $*p < .1$, $**p < .05$, $***p < .01$.



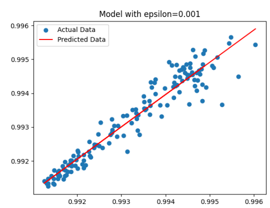
(a) Caption for plot 1



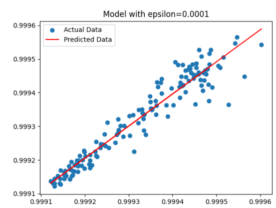
(b) Caption for plot 2



(c) Caption for plot 3



(d) Caption for plot 4



(e) Caption for plot 5

Figure 2: Regression Plot

3.2 GMM Regression of Taylor model

3.3 GMM Regression of full model

Let's consider the model:

$$N_t = \alpha_1 N_{t-1} + \alpha_2 N_{t-1}^{1+\varepsilon} + U_t \tag{5}$$

For a dataset of size T. We use GMM estimation, and since we have 3 parameters, we need to use at least 3 moment conditions. We consider different UCMR:

- Assuming U_t is orthogonal to N_{t-1} and N_{t-2} every period.
- Assuming U_t is orthogonal to N_{t-1} and N_{t-1}^{ε} every period.

3.4 Case 1

Let's assume that U_t is exogenous to N_{t-1} and N_{t-2} every period:

$$E[U_t|N_{t-2}, N_{t-1}] = 0 \tag{6}$$

Which implies the following unconditional moment restriction:

$$E\begin{bmatrix} 1 \\ N_{t-1} \\ N_{t-2} \end{bmatrix} U_t = 0 \tag{7}$$

In canonical form:

$$E\, g(N_{t-2}, N_{t-1}, N_t, \theta) = 0, \tag{8}$$

Where:

$$g(N_{t-2}, N_{t-1}, N_t, \theta) = \begin{bmatrix} 1 \\ N_{t-1} \\ N_{t-2} \end{bmatrix} \left(N_t - \alpha_1 N_{t-1} - \alpha_2 N_{t-1}^{1+\varepsilon} \right), \qquad \theta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \varepsilon \end{bmatrix} \tag{9}$$

Bookkeeping:

- The estimator is just identified:

$$\dim \theta = \dim g = 3 \tag{10}$$

- Function g is non-linear in the parameters.
- Its Jacobian is:

$$D(\theta) = \frac{\partial g}{\partial \theta} = \begin{bmatrix} \frac{\partial g}{\partial \alpha_1} & \frac{\partial g}{\partial \alpha_2} & \frac{\partial g}{\partial \varepsilon} \end{bmatrix} = - \begin{bmatrix} N_{t-1} & N_{t-1}^{1+\varepsilon} & \alpha_2 N_{t-1}^{1+\varepsilon} \ln N_{t-1} \\ N_{t-1}^2 & N_{t-1}^{2+\varepsilon} & \alpha_2 N_{t-1}^{2+\varepsilon} \ln N_{t-1} \\ N_{t-1} N_{t-2} & N_{t-1}^{1+\varepsilon} N_{t-2} & \alpha_2 N_{t-1}^{1+\varepsilon} N_{t-2} \ln N_{t-1} \end{bmatrix} \tag{11}$$

The estimator $\hat{\theta}$ is the minimizer:

$$\hat{\theta} = \arg\min_{\theta} \; E g^T(N_{t-2}, N_{t-1}, N_t, \theta) \, S \, E g(N_{t-2}, N_{t-1}, N_t, \theta) \tag{12}$$

for some weighting matrix S. Result:

$$\theta = \begin{bmatrix} 8.00 & -6.95 & 8.69\,e-04 \end{bmatrix} \tag{13}$$

3.5 Case 2

Let's assume that U_t is exogenous to N_{t-1} every period:

$$E[U_t | N_{t-1}] = 0 \tag{14}$$

Which implies the following unconditional moment restriction:

$$E \begin{bmatrix} 1 \\ N_{t-1} \\ N_{t-1}^{\varepsilon} \end{bmatrix} U_t = 0 \tag{15}$$

In canonical form:

$$E \, g(N_{t-1}, N_{t-1}^{\varepsilon}, N_t, \theta) = 0, \tag{16}$$

This time we find:

$$\theta = \begin{bmatrix} 8.00 & -6.95 & 8.69\,e-04 \end{bmatrix} \tag{17}$$

3.6 Numerical procedure

3.6.1 Try 1

1. Start by generating a random weighting matrix S_0.
2. Make a 3×3 grid of values for the parameters $(\alpha_1, \alpha_2, \varepsilon)$.
3. Compute the sample equivalent of $m(\theta) = Eg(N_t, N_{t+1}, \theta)$ for all points in the grid.
4. Take $\hat{\theta}^{(1)}$ for which the target function $Eg(N_t, N_{t+1}, \theta)\, S\, Eg^T(N_t, N_{t+1}, \theta)$ is minimum.
5. Estimate the variance as the sample equivalent of $Eg(N_t, N_{t+1}, \theta)\, g^T(N_t, N_{t+1}, \theta)$ for $\theta = \hat{\theta}(1)$:
$$\hat{V}^{(1)} = \frac{1}{T-1} \sum_{t=1}^{T-1} g(N_t, N_{t+1}, \hat{\theta}^{(1)}) g^T(N_t, N_{t+1}, \hat{\theta}^{(1)}) \tag{18}$$
6. Define a new weighting matrix $S^{(1)} = \left(\hat{V}^{(1)}\right)^{-1}$.
7. Iterate until convergence, i.e., until $\left|\hat{\theta}^{(t+1)} - \hat{\theta}^{(t)}\right|_i < \Delta$ for each component $i = 1, 2, 3$.
8. Try with different starting matrices S_0 and different grids $(\alpha_1, \alpha_2, \varepsilon)$.

3.6.2 Try 2

1. Start by generating a random weighting matrix S_0.
2. Numerically solve the FOCs for θ:
$$\hat{D}(\theta)\, S_0\, \hat{m}(\theta) = 0 \tag{19}$$
the resulting $\hat{\theta}^{(1)}$ minimize the target function $Em\theta)\, S\, m(\theta)$.
3. Estimate the variance as the sample equivalent of $Eg(N_t, N_{t+1}, \theta)\, g^T(N_t, N_{t+1}, \theta)$ for $\theta = \hat{\theta}(1)$:
$$\hat{V}^{(1)} = \frac{1}{T-1} \sum_{t=1}^{T-1} g(N_t, N_{t+1}, \hat{\theta}^{(1)}) g^T(N_t, N_{t+1}, \hat{\theta}^{(1)}) \tag{20}$$
4. Define a new weighting matrix $S^{(1)} = \left(\hat{V}^{(1)}\right)^{-1}$.
5. Iterate until convergence, i.e., until $\left|\hat{\theta}^{(t+1)} - \hat{\theta}^{(t)}\right|_i < \Delta$ for each component $i = 1, 2, 3$.
6. Try with different starting matrices S_0 and different grids $(\alpha_1, \alpha_2, \varepsilon)$.

For P-0 consider the stationary distribution as $t \to \infty$ and estimate t* as an average with that distribution