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.

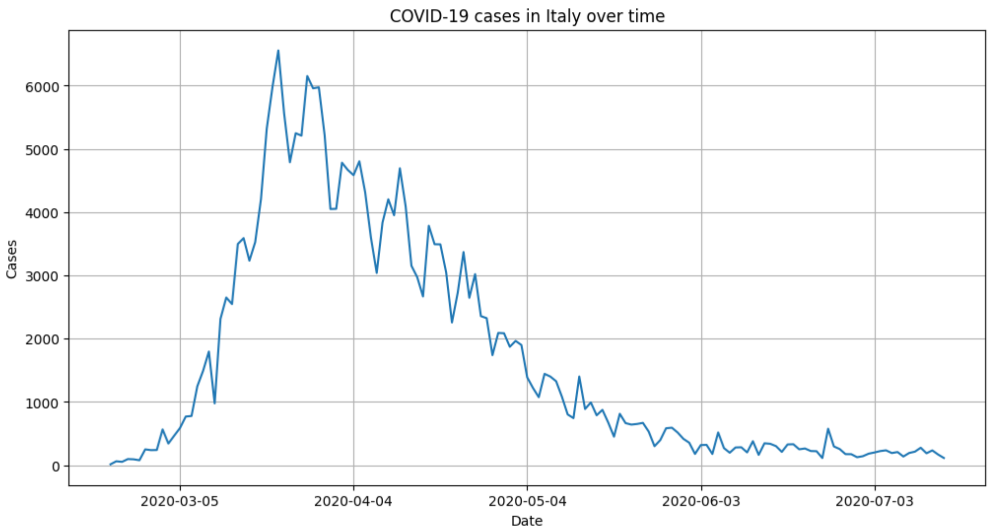


Figure 1: New infections per day in Italy

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:

$$N(t) = \frac{K}{1 + \left(\frac{K - N_0}{N_0}\right) e^{-rt}}$$

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

$$t^* = \frac{1}{r} \ln\left(\frac{K - N_0}{N_0}\right)$$

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:

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

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:

$$\ddot{N} = -\frac{V^{-\frac{1+\varepsilon}{\varepsilon}}}{\varepsilon}\left(\ddot{V} - \frac{1+\varepsilon}{\varepsilon}\frac{\dot{V}^2}{V}\right) \quad \rightarrow \quad t^* = \frac{1}{\varepsilon r}\ln\left(\varepsilon K N_0^{-\varepsilon} - \varepsilon\right)$$

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:

$$\beta_0 = -\alpha_2\varepsilon = \frac{r\varepsilon}{K}, \quad \beta_1 = \varepsilon(1-\alpha_1) = -r\varepsilon.$$

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:

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

Table 1

	$\varepsilon =: 1.0$	$\varepsilon =: 0.1$	$\varepsilon =: 0.01$	$\varepsilon =: 0.001$	$\varepsilon =: 0.0001$
const	0.0000*	0.0211***	0.0501***	0.0547***	0.0552***
	(0.0000)	(0.0006)	(0.0013)	(0.0014)	(0.0014)
x1	0.6348***	0.9278***	0.9433***	0.9446***	0.9447***
	(0.0071)	(0.0019)	(0.0014)	(0.0014)	(0.0014)
R-squared	0.9827	0.9994	0.9997	0.9997	0.9997
R-squared Adj.	0.9826	0.9994	0.9997	0.9997	0.9997
N	142	142	142	142	142

which can be useful for Ramsey testing

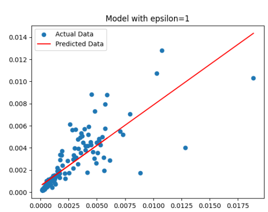
3 Estimation

3.1 Preliminary Regression

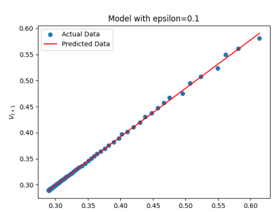
Let's regress the following model with OLS:

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

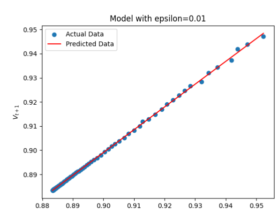
where $V_t = N_t^{-\varepsilon}$, for different values of ε.



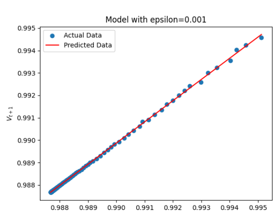
(a) Caption for plot 1



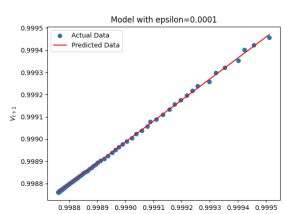
(b) Caption for plot 2



(c) Caption for plot 3



(d) Caption for plot 4



(e) Caption for plot 5

Figure 2: Regression Plot

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01

3.2 GMM estimation with Stata

Let's consider the model:

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

We assume that U_t is mean independent of all the N_s for $s \leq t-1$. For 3 parameters, at least 3 instruments are needed:

Table 2: GMM Regression Table

	N_1 N_2	N_1 N_2 N_3	N_1 N_2 N_3 N_4	N_1 N_2 N_3 $N \ln N$
α_1	4.280	7.049	7.457	7.503
	(6.681)	(.)	(.)	(.)
α_2	-2.528	-4.981***	-5.466***	-5.474***
	(6.514)	(0.0550)	(0.0453)	(0.0493)
ε	0.0210	0.0157***	0.0135***	0.0140***
	(0.0436)	(0.000896)	(0.000672)	(0.000730)
J		252496718.1	161236752.4	209017433.8
J_df	0	2	3	3
rank	3	2	2	2

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3 Ramsay Test

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