A data-driven approach to evaluate the impact of restrictive measures on the incidence of COVID-19 and economic indicators

Komorebi AI Technologies SL

During the past year several measures have been enforced by regional and central governments in order to mitigate the spread of SARS-Cov2 virus. These measures, known as Non-Pharmaceutical Interventions (NPIs) include limitations on social gatherings, closing or restricting commercial activities, and bounds on travel, among others. Despite the great effort invested in trying to understand the impact and efficiency of NPIs, there is still no clear consensus among the experts as to which are the most effective NPIs.
The enforcement of NPIs has a considerable political cost and many collateral effects on society, most particularly in slowing down the economical activity of a given region. Thus, there is also much interest in assessing the impact of NPIs on the economy.

Once these effects are reasonably understood, an optimal strategy can be reached for this multiobjective optimization problem: **find the combination of NPIs that reduce the incidence to reasonable values while having the smallest impact on the economy.**

In order to tackle this problem we propose a data-driven approach which aims to build a statistical model using data from the 50 spanish provinces during the past year. Incidence of COVID-19 has struck different provinces in a different manner at different times, and many of them have enforced a different set of NPIs. We plan to use the time series for these 50

provinces as natural experiments to derive a statistical model that estimates the impact of NPIs.

Similar approaches to estimate the impact of NPIs have been recently addressed in the literature [1-7]. The novelty of our approach is that this would be the **first study done entirely on data gathered in Spain**, where the main challenge is keeping track of the NPIs adopted at each province across the whole period. An ongoing project in collaboration with the Ministry of Health and the Ministry of Economy has built a set of NPI stringency indices that we can use for this study.

Work plan

The work for predicting the result of a match passes through 3 phases:

1. Data exploration and combination. Required data sources have already been scrapped by Komorebi for the sake of ease, although there is the possibility of adding new sources. The different sources must be understood and combined to be able to train the ML models.

2. Train machine learning models. The goal is to train time series forecasting models with a probabilistic output, in order to assess the uncertainty of the predictions. The different NPIs will be part of the input to the models.

3. Counterfactual analysis. Ultimately, we can evaluate the outcome of applying different restrictions on the economy and the spread of COVID-19. The goal is to build the Pareto frontier of this multi-target problem (minimize the incidence and the economic impact). The models will also take into account not only a single point prediction but the whole predictive distribution, in order to better assess the uncertainty and risk.

We expect that at the end of the modelling week, the students will have a thorough understanding of the problem and the methods used to solve it, as well as means to continue digging deeper into the topic if they so wish.

Pre-requisites

No specific prior knowledge is assumed beyond standard undergraduate courses on linear algebra, basic probability and statistics, optimization, and modelling. Also basic knowledge of a programming language (i.e. R, Python) is assumed. Some experience with handling structured data in Python or R will allow faster iteration and reaching better final results.

Datasets

The proposed problem involves merging data from different sources and creating descriptive analytics and statistical models to tackle relevant questions related to the COVID-19 pandemics in Spain. Data sources at our disposal include:

- **Demographic data:** population structured by age and geographic division.
- **Epidemiological data**: number of cases, ICU patients, deaths, etc. structured by age and geographic division
- **Non-pharmaceutical interventions**: restrictive measures as a function of time enforced by each regional administration, structured by sector and expressed as a stringency index.
- **Economic activity**: data from credit card transactions from several spanish banks, structured by commercial sector and geographic division.
- **Vaccination:** number of vaccinated individuals, structured by vaccine type, age and geographic division.
- **Mobility**: aggregated data from mobile phones provided by either Google (Google Mobility reports) or the Spanish Ministry of Transportation. In the first case, they are structured by province and sector, and expressed as a relative variation with respect to a reference period.

References

[1] Cowling, B. J., Ali, S. T., Ng, T. W., Tsang, T. K., Li, J. C., Fong, M. W., ... & Wu, J. T. (2020). Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *The Lancet Public Health*.

[2] Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., ... & Monod, M. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257-261.

[3] Karnakov, P., Arampatzis, G., Kičić, I., Wermelinger, F., Wälchli, D., Papadimitriou, C., & Koumoutsakos, P. (2020). Data driven inference of the reproduction number (R0) for COVID-19 before and after interventions for 51 European countries. *medRxiv*.

[4] Kucharski, A. J., Klepac, P., Conlan, A., Kissler, S. M., Tang, M., Fry, H., ... & CMMID COVID-19 Working Group. (2020). Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of SARS-CoV-2 in different settings. *medRxiv*.

[5] Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*.

[6] Vokó, Z., & Pitter, J. G. (2020). The effect of social distance measures on COVID-19 epidemics in Europe: an interrupted time series analysis. *GeroScience*, 1-8.

[7] Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., ... & Klimek, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature human behaviour*, 1-10.