M2 internship : Signal processing with application to epidemiology

Title : Sparsity-based signal processing with application to epidemiology

Supervisors : Barbara Pascal, Sébastien Bourguignon, Gwenaël Samain

Contact : barbara.pascal@cnrs.fr, sebastien.bourguignon/gwenael.samain@ls2n.fr,

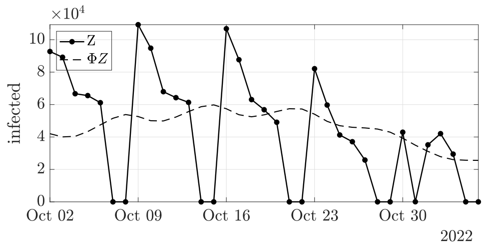
Location : Laboratoire des Sciences du Numérique de Nantes, Centrale Nantes

Date and duration : Spring 2023 ; 4 to 6 months

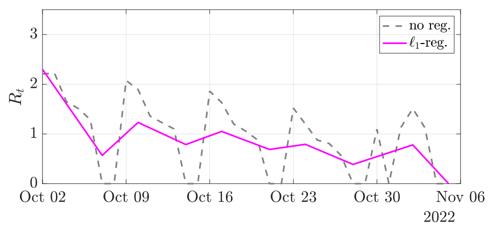
Salary : $\sim$ 600 euros per month (standard internship compensation)

Abstract : Since the beginning of the Covid-19 pandemic, daily new infections (solid line, left plot) have been monitored. Yet, it is not informative enough about the dynamics of the pandemic and decision makers rapidly got interested in the *reproduction number* R_t at day t, quantifying the number of people infected by one sick individual : $\mathsf{R}_t > 1$, corresponds to exponential spread of the virus ; while for $\mathsf{R}_t < 1$, the epidemic is vanishing. R_t is linked to daily new infections by Cori's model [1] stating that the number of new cases follows a nonstationary Poisson distribution whose instantaneous parameter at day t is proportional to R_t. Direct estimation of R_t provides very irregular estimates (dashed gray line, right plot). To obtain realistic estimates, it has been proposed to enforce temporal regularity using a ℓ_1 penalization on the second order derivative of R_t [2]. Yet, the ℓ_1-norm only enables to obtain *approximate* sparsity : the second-order derivative of the obtained estimates are not sparse (see the magenta curve on right plot). Recent approaches [3] have proposed numerical schemes enforcing sparsity by using a ℓ_0-norm based regularization term. The corresponding solution is then likely to provide more accurate estimates, despite their much higher numerical complexity. In the considered case, however, the relatively small dimension of the problems should make it possible to use such exact approaches. For the moment, these techniques have mostly be used coupled to quadratic data-fidelity terms, which does not cover the estimation of R_t proposed in [2].

The purpose of this project are : 1) to get familiar with variational tools for the estimation of its reproduction number ; 2) to understand the numerical challenges raised by the ℓ_0-norm ; 3) to identify the major deadlocks in using an ℓ_0-norm penalization in the estimation of the reproduction number R_t ; 4) to develop a propotype algorithm in Matlab that implements the branch-and-bound optimization of the ℓ_0 norm solution.



Daily new infections in France



Estimated reproduction number

Bibliography :

[1] Cori, A., Ferguson, N. M., Fraser, C., & Cauchemez, S. (2013). A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology*, vol. 178, no 9, p. 1505-1512.

[2] Abry, P., Pustelnik, N., Roux, S., Jensen, P., Flandrin, P., Gribonval, R., ... & Garnier, N. (2020). Spatial and temporal regularization to estimate COVID-19 reproduction number R (t) : Promoting piecewise smoothness via convex optimization. *Plos One*, vol. 15, no 8.

[3] Ben Mhenni, R., Bourguignon, S., & Ninin, J. (2021). Global optimization for sparse solution of least squares problems. *Optimization Methods and Software*, 1-30.