Programming Assignment 1: GP Regression

Regression. The CSV data file new_cases_World-India contains four coumns: day number, date, number of new cases in world, number of new cases in India. The goal is to predict the number of new cases in India and the world for a day number (in future). Both the mean along with the uncertainty/variance value needs to be obtained. As a test set all the days in the month of September may be used. You may use separate GPs for the world and India. The data set is obtained from: https://ourworldindata.org/coronavirus-source-data

You may choose any GP prior, and corresponding hyperparameter values as you wish. However, choose a prior which captures the basic nature of the data. Assume a Gaussian noise with zero mean and variance of 0.1 denoting the error in reporting of cases.

Bonus: Perform hyperparameter optimization/model averaging. You may also choose non-statonary GP e.g., Wiener process

Submission: You need to submit two files - (i) a single program file (with your name, roll and dependencies mentioned in the header), (ii) a write-up (single pdf file) with your name/roll and the details of method used and visualization of results. For the program any language python/Julia/C/C++/java may be used. You may use basic data processing, linear algebra, random number generation libraries. But do not use GP functions directly.