

COVID-19 疫情数据的简要统计分析与预测

《MATLAB 程序设计》大作业汇报

王逸扬 (19300180016)

杨耕智 (19300180112)

2022 年 5 月 30 日

目录

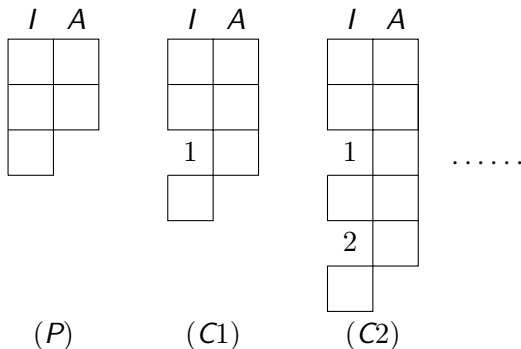
- ▶ 数据导入、缺失值处理的类 `StatisticsAnalysis`;
- ▶ 长期性分析（自回归分析、周期性）
- ▶ 短期性分析（聚类分析、基于 LSTM 模型的预测分析）

数据导入、缺失值处理的类 StatisticsAnalysis

参数:

- ▶ `TablePath, Table`;
- ▶ `ImportOptions` 修改 `detectImportOptions` 中的参数;
- ▶ `SelectTableOptions` 检索表格;
- ▶ `MissingValuesOptions` 缺失值探测、处理;
- ▶ `TagsGenerateOptions` 计算并添加统计指标到
`Table.Properties.CustomProperties`.

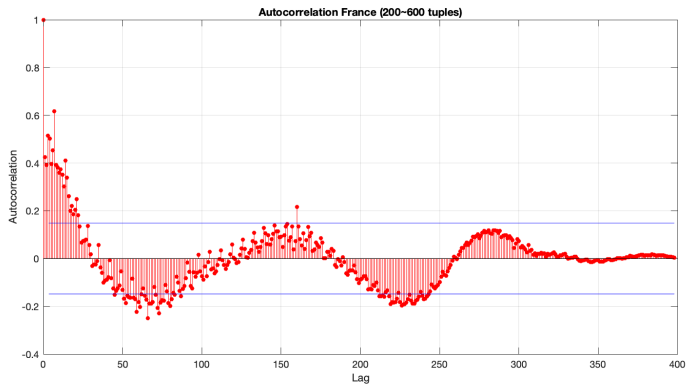
缺失值处理——「新增-累计」数据



- ▶ **需要插值的 *P*、*C* 两种情况的处理**：默认线性插值，也可以传入自定义的插值函数句柄，例如 `spline`；
- ▶ **增量数据不单调的处理**：不处理/衰减地分配到前几日/传入函数句柄处理。

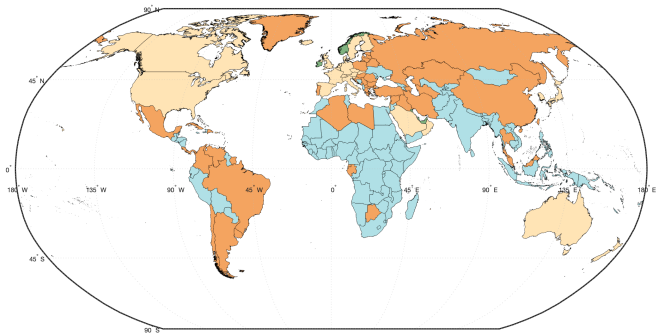
长期性分析

自回归函数 autocorr、偏自回归函数 parcorr、周期性.



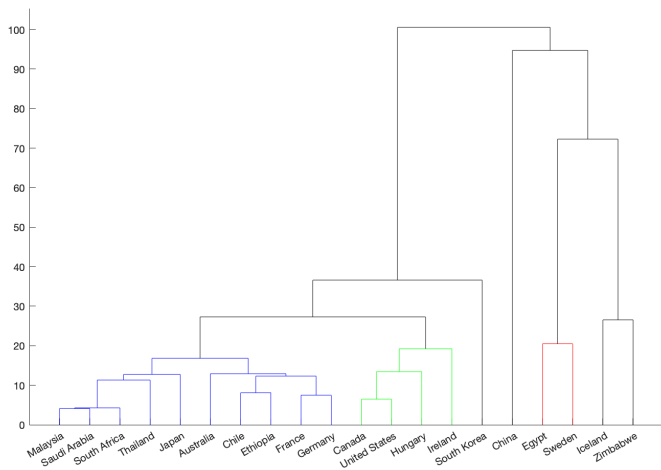
短期性分析 (地区数据的 k -mean 聚类)

主成分分析. 肘部原理 $\Rightarrow k = 4$.



短期性分析（时间序列特征的谱系聚类）

某时期各国家的：(峰峰值、平均值、标准差)/人口，偏度和峰度。

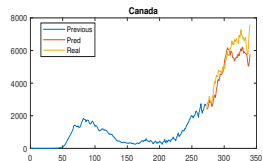
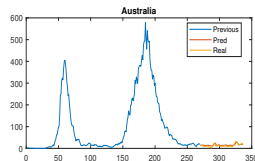
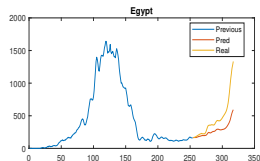
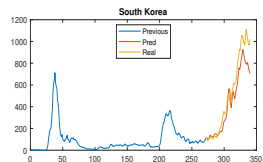
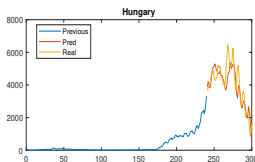
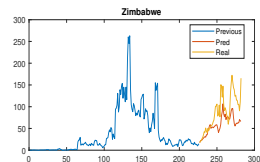
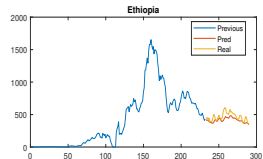
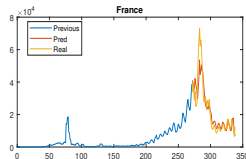
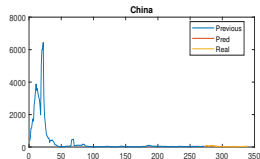


短期性分析（基于 LSTM 模型的预测分析）

MATLAB 的神经网络工具箱. 简单情形 \Rightarrow LSTM.

```
sequenceInputLayer(8)
→ fullyConnectedLayer(30)
→ reluLayer
→ lstmLayer(400)
→ reluLayer
→ fullyConnectedLayer(30)
→ reluLayer
→ fullyConnectedLayer(numResponses)
→ regressionLayer
```


短期性分析（基于 LSTM 模型的预测分析）



以上就是我们对 COVID-19 数据集的分析，
水平有限，多有疏漏，恳请各位批评指正。

最后，感谢陈阳老师和两位助教一学期以来的教导与帮助！