

利用动力系统的方法预测GDP

将GDP看作人均GDP和市场fitness作为动力的动力系统，可以得到比较好的预测值。

新方法可以补充国际货币基金组织（IMF）的预测，准确度甚至优于IMF。这项全新的预测方法把经济增长视为一套物理系统，通过复杂系统物理建模对产品出口数据进行分析，并预测该系统的动态。

对一个国家的复杂经济系统建模是极大的挑战。尽管经济学家可以获取的数据越来越多，但从海量的数据中得出可靠的、重现性好的结果远不是件容易的事。这或许需要复杂系统物理学的帮忙：这一领域的技术优越性在于可以对难以从单个要素中分析的系统进行建模，例如流行病蔓延、交通流量变化等。

意大利国家研究委员会（CNR）的Andrea Tacchella和同事开发的这套GDP预测方法的基本理念是：在已知的历史数据中找到未知的复杂系统的近似模拟，并观察模拟的时间演化，从而对复杂系统的行为进行预测。

不过，这种方法只有在使用低维模型（维度即输入变量的个数）时才可靠，而任意添加数据反而会降低预测可靠性。因此，作者只用两个变量进行了有效的GDP预测——国家人均GDP和“健康度”（fitness）。其中，“健康度”是作者对出口数据运行大量计算后构建的一个衡量国家竞争力的单量。

研究小组将自己的预测与IMF基于既往数据的预测进行对比，结果显示，他们的预测平均准确度高了25%。同时，两套预测模型误差无相关性，意味着两种方法对GDP增长的预期具有互补性，预示了未来两者连用以提高预测准确性的可能。

数据集：COMTRADE

有各个国家进出口超过4000种商品的数据。

数据处理

有的商品在进口处与出口处的记录不一样。作者在未来的一篇文章中，利用贝叶斯的方法来进行同意了。

统计指标

某国家某产品的竞争力：RCA

定义为 $RCA_{c,p} = \frac{E_{cp}}{\sum_p E_{cp}} / \frac{\sum_c E_{cp}}{\sum_c \sum_p E_{cp}}$ 。其中 E_{cp} 代表国家C出口产品P的量。RCA指数去除了国家规模的影响。

$$M_{cp}$$

1. binary

在很多之前的研究中，这个值直接被定义为 $M_{cp} = \delta_{\{RCA_{cp} \geq 1\}}$ 。但是这种方法带来的问题是这个矩阵会受噪音影响非常大。

2. 隐马尔科夫模型

隐马氏模型可以适应随机情况。

具体做法：

- 1. 假设总共有4种隐藏状态。
- 2. 将RCA矩阵的值改为1, 2, 3, 4. 离散化有助于数据处理。
- 3. 对于每个国家的时间序列训练一个模型。

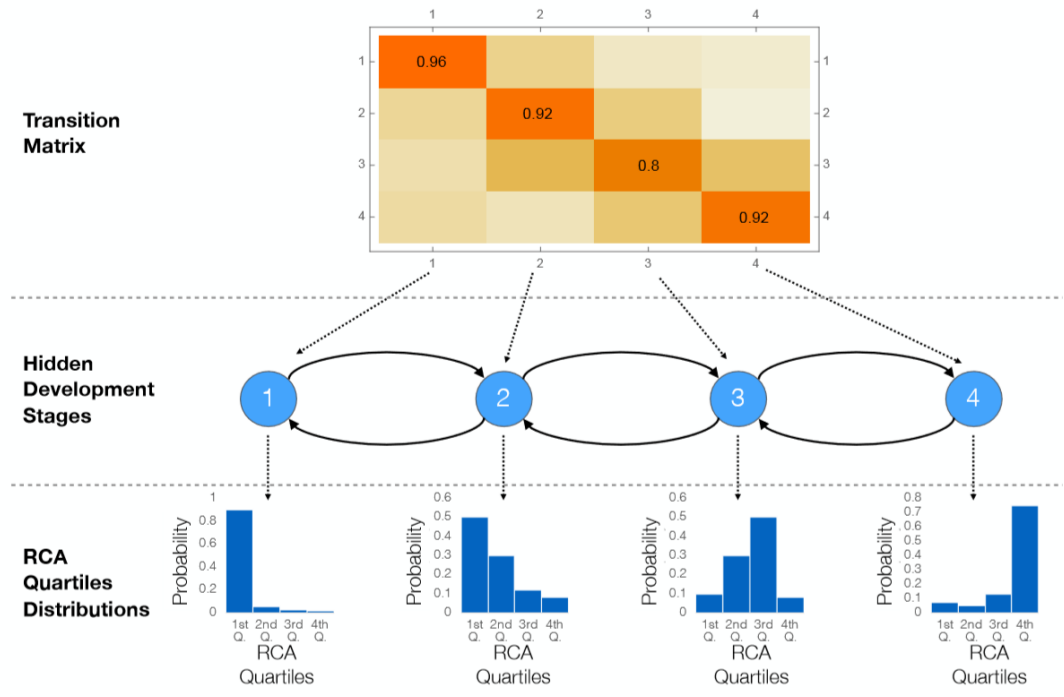


Figure S1: Schematic representation of the Hidden Markov Model used for export data regularization of one particular country. The top panel shows the optimized transition probabilities. The middle panel shows a simplified representation of the topology of the transitions among the hidden states. The bottom panel shows the RCA quartiles distribution associated to each hidden state.

可以看出，这种处理对鲁棒性有着很好的提高。

而后构建 M 矩阵的方法就很简单了，如果隐藏状态为1, 2就设为0；3, 4就设为1.