Economiccomplexity: Computational Methods for Economic Complexity

23 August 2018

Summary

Economic complexity is a state-of-the-art line of research that started back nearly ten years ago and introduces network theory concepts to study different social sciences issues related to international trade and income inequality. Its approach starts from representing international trade data as a bipartite network that connects countries to the products that they export.

Seminal papers in economic complexity are Hidalgo et al. (2007), that introduces graphs to explore trade diversification and export opportunities, and Hidalgo and Hausmann (2009), that develops dedicated metrics of complexity. These two articles are expanded and interpreted in Hausmann et al. (2014), whose typesetted equations were translated to code in the R package **economiccomplexity**.

This R package provides different methods to compute complexity metrics that ease access to this line of research for social scientists. `economiccomplexity` depends on the `Matrix` (Bates and Maechler 2019) and `igraph` (Csardi and Nepusz 2006) packages, and follows the same design philosophy, grammar, and data structures from it.

Recent articles such as Hartmann et al. (2017) introduce the question whereas a country's mix of products could predict its pattern of diversification and income inequality, but do not include linked and executable code and data. This separation between the research, the complete process that produced the results, and its presentation, makes it difficult for others to verify the findings in the study.

`economiccomplexity` might help to evaluate research findings in this particular area, helping to produce less studies that are not reproducible, or partially reproducible with some discrepancies, conditional on the availability of data, metadata, and computing power that may be unavailable to all researchers.

The central contribution of this package is to be the first that provides functions to use the recursive linear algebra methods from Hidalgo and Hausmann (2009) in R and adds tests and full documentation to offer the best possible package on CRAN.

Tacchella et al. (2012) presents non-linear iterative methods that extend the linear approach from Hidalgo and Hausmann (2009) in order to capture the link between the export basket of different countries and their industrial competitiveness. This approach is also implemented in `economiccomplexity` but following the formulation in Mariani et al. (2015) that introduces extremality parameters that generalize the original formulation.

Peng (2011) states that reproducibility has the potential to serve as a minimum standard when full independent replication of a study is not possible, and that becomes even more important to evaluate scientific claims in studies with public policy implications. A tenet of the scientific method holds that every research finding should be reproducible before it becomes accepted as a genuine contribution to human knowledge (Stodden (2009)).

Complexity methods are also used outside its original research area. In particular, the fitness method from Tacchella et al. (2012) has been used in ecology to study species interaction (see Domínguez-García and

Munoz (2015)) and, more directly related, to study the scientific competitiveness of nations (See Cimini, Gabrielli, and Sylos Labini (2014)).

The application widespread of economic complexity shall be looked with a bit of caution. Any theoretical or computational implementation shall be far from becoming a keystone if reproducibility is ignored. The extent to which code in computational research would build with reasonable effort is lower than 20% (Collberg et al. 2014). A desirable growth pattern should focus on reproducibility, and I hope this package means a contribution to transparent research practices.

References

Bates, Douglas, and Martin Maechler. 2019. *Matrix: Sparse and Dense Matrix Classes and Methods*. https://CRAN.R-project.org/package=Matrix.

Cimini, Giulio, Andrea Gabrielli, and Francesco Sylos Labini. 2014. "The Scientific Competitiveness of Nations." *PLOS ONE* 9 (12). Public Library of Science: 1–11. https://doi.org/10.1371/journal.pone.0113470.

Collberg, Christian, Todd Proebsting, Gina Moraila, Akash Shankaran, Zuoming Shi, and Alex M Warren. 2014. "Measuring Reproducibility in Computer Systems Research." *Department of Computer Science, University of Arizona, Tech. Rep* 37.

Csardi, Gabor, and Tamas Nepusz. 2006. "The Igraph Software Package for Complex Network Research." *InterJournal* Complex Systems: 1695. http://igraph.org.

Domínguez-García, Virginia, and Miguel A Munoz. 2015. "Ranking Species in Mutualistic Networks." *Scientific Reports* 5. Nature Publishing Group: 8182.

Hartmann, Dominik, Miguel R. Guevara, Cristian Jara-Figueroa, Manuel Aristarán, and César A. Hidalgo. 2017. "Linking Economic Complexity, Institutions, and Income Inequality." *World Development* 93: 75–93. https://doi.org/10.1016/j.worlddev.2016.12.020.

Hausmann, Ricardo, César Hidalgo, Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed Yildirim. 2014. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. MIT Press. https://doi.org/10.7551/mitpress/9647.001.0001.

Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann. 2007. "The Product Space Conditions the Development of Nations." *Science* 317 (5837). American Association for the Advancement of Science: 482–87. https://doi.org/10.1126/science.1144581.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences* 106 (26). National Academy of Sciences: 10570–5. https://doi.org/10.1073/pnas.0900943106.

Mariani, Manuel, Alexandre Vidmer, Matsúš Medo, and Yi-Cheng Zhang. 2015. "Measuring Economic Complexity of Countries and Products: Which Metric to Use?" *The European Physical Journal B* 88 (11). Springer: 293. https://doi.org/10.1140/epjb/e2015-60298-7.

Peng, Roger D. 2011. "Reproducible Research in Computational Science." *Science* 334 (6060). American Association for the Advancement of Science: 1226–7. https://doi.org/10.1126/science.1213847.

Stodden, Victoria. 2009. "The Legal Framework for Reproducible Scientific Research: Licensing and Copyright." *Computing in Science Engineering* 11 (1): 35–40. https://doi.org/10.1109/MCSE.2009.19.

Tacchella, Andrea, Matthieu Cristelli, Guido Caldarelli, Andrea Gabrielli, and Luciano Pietronero. 2012. "A New Metrics for Countries' Fitness and Products' Complexity." *Scientific Reports* 2. Nature Publishing Group: 723.