Economiccomplexity: Computational Methods for Economic Complexity

Mauricio Vargas[1]

1 Pontifical Catholic University of Chile

DOI: 10.21105/joss.01866

Software

- Review
- Repository
- Archive

Editor: Olivia Guest

Reviewers:

- @JohnCoene
- @mpadge

Submitted: 13 September 2019
Published: 19 February 2020

Summary

Economic complexity introduces network theory concepts to different social science considerations related to international trade and income inequality. With the bulk of literature established in the last decade, the field of economic complexity is relatively new. 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, Klinger, Barabási, & Hausmann (2007), which introduces graphs to explore trade diversification and export opportunities; and Hidalgo & Hausmann (2009), which develops dedicated metrics of complexity. These two articles are expanded and interpreted by 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 `Matrix::` and `igraph::`, and follows the same design philosophy, grammar, and data structures from it (Bates & Maechler, 2019; Csardi & Nepusz, 2006).

Recent articles such as Hartmann, Guevara, Jara-Figueroa, Aristarán, & Hidalgo (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, reducing the number of studies that are not reproducible, or only 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 `economiccomplexity::` is to provide functions that use recursive linear algebra methods from Hidalgo & Hausmann (2009) in R, with tests and full documentation for release on CRAN, the dominant repository of R software.

Tacchella, Cristelli, Caldarelli, Gabrielli, & Pietronero (2012) presents non-linear iterative methods that extend the linear approach from Hidalgo & 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 from Mariani, Vidmer, Medo, & Zhang (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 (Domínguez-García & Munoz, 2015) and, more directly related, to study the scientific competitiveness of nations (Cimini, Gabrielli, & Sylos Labini, 2014).

The application widespread of economic complexity should be considered with 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.

Hausmann et al. (2014) proposes the equivalency between the reflections and the eigenvalues methods that they present. My code on GitHub checks for the sign of the correlation for the vector output of the two methods and corrects the eigenvalues output when necessary. What motivated those additional steps is that, what I initially tought it was an error, turned out to be a particular case that emerges with some datasets. Shortly after implementing that, both Python and Stata users started sending me emails when they, independently, found the same problem and then search engines led them to my package.

References

Bates, D., & Maechler, M. (2019). *Matrix: Sparse and dense matrix classes and methods*. Retrieved from https://CRAN.R-project.org/package=Matrix

Cimini, G., Gabrielli, A., & Sylos Labini, F. (2014). The scientific competitiveness of nations. *PLOS ONE*, *9*(12), 1–11. doi:10.1371/journal.pone.0113470

Collberg, C., Proebsting, T., Moraila, G., Shankaran, A., Shi, Z., & Warren, A. M. (2014). Measuring reproducibility in computer systems research. *Department of Computer Science, University of Arizona, Tech. Rep*, *37*.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695. Retrieved from http://igraph.org

Domínguez-García, V., & Munoz, M. A. (2015). Ranking species in mutualistic networks. *Scientific reports*, *5*, 8182. doi:10.1038/srep08182

Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, *93*, 75–93. doi:10.1016/j.worlddev.2016.12.020

Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press. doi:10.7551/mitpress/9647.001.0001

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575. doi:10.1073/pnas.0900943106

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482–487. doi:10.1126/science.1144581

Mariani, M., Vidmer, A., Medo, M., & Zhang, Y.-C. (2015). Measuring economic complexity of countries and products: Which metric to use? *The European Physical Journal B*, *88*(11), 293. doi:10.1140/epjb/e2015-60298-7

Peng, R. D. (2011). Reproducible research in computational science. *Science*, *334*(6060), 1226–1227. doi:10.1126/science.1213847

Stodden, V. (2009). The legal framework for reproducible scientific research: Licensing and copyright. *Computing in Science Engineering*, *11*(1), 35–40. doi:10.1109/MCSE.2009.19

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific reports*, *2*, 723. doi:10.1038/srep00723