On convexity in complex networks

Lovro Šubelj
University of Ljubljana, Faculty of Computer and Information Science
Ljubljana, Slovenia
lovro.subelj@fri.uni-lj.si

Tilen Marc
Institute of Mathematics, Physics and Mechanics
Ljubljana, Slovenia
tilen.marc@imfm.si

Metric graph theory is a study of geometric properties of graphs based on a notion of the shortest path between the nodes defined as the path through the smallest number of edges [2]. Metric graph properties lie in the heart of the analysis of complex networks. Classical examples include Milgram's experiment of degrees of separation [8], node index called betweenness centrality [5] and the small-world network model [11].

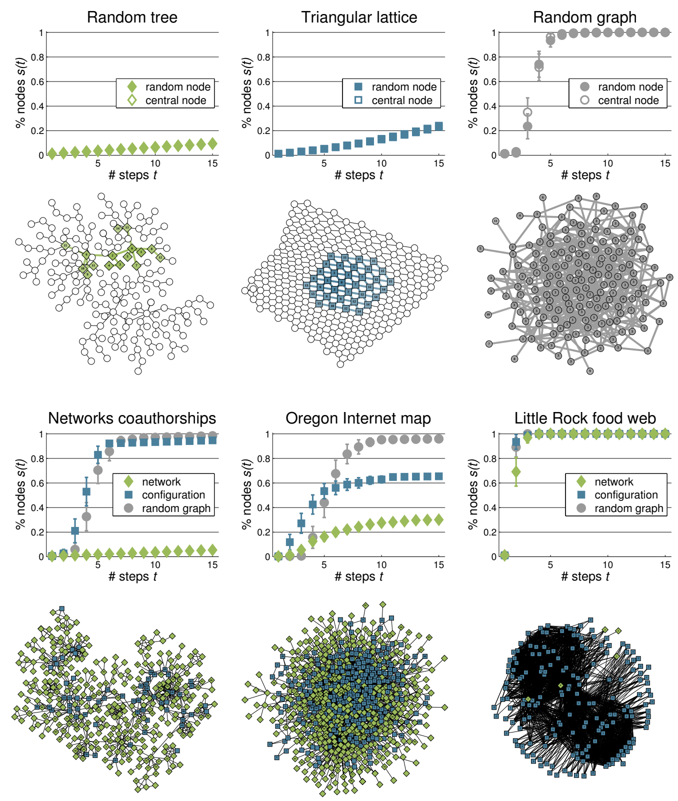
Independently of these efforts, graph theorists have been interested in understanding convexity in a given graph [6]. Consider a simple connected graph and a subgraph on some subset of nodes S. The subgraph is *induced* if all edges between the nodes in S in the graph are also included in the subgraph. Next, the subgraph is said to be *isometric* if at least one shortest path joining each two nodes in S is entirely included within S. Finally, the subgraph is a *convex subgraph* if all shortest paths between the nodes in S are entirely included within S. Notice that any convex subgraph is also isometric, while any isometric subgraph must necessarily be induced.

We study convexity in complex networks through the definition of a convex subgraph [7]. We explore convexity from a local and global perspective by analyzing the frequency of small convex subgraphs and the expansion of randomly grown convex subgraphs. In the case of the latter, we grow random connected subgraphs one node at a time and expand them to convex subgraphs if needed. For instance, every connected subgraph of a tree or a complete graph is convex and thus no expansion occurs. Hence, the expansion of convex subgraphs quantifies the presence of a tree-like or clique-like structure in a network.

We demonstrate three distinct forms of convexity in graphs and networks. *Global convexity* refers to a tree-like or clique-like structure of a network as a whole in which convex subgraphs grow very slowly and thus any connected subgraph is likely to be convex. Globally convex networks are spatial infrastructure networks and network science coauthorship graph. In random graphs [4], however, there is a sudden expansion of convex subgraphs when their size exceeds $\ln n/\ln\langle k\rangle$ nodes, where n is the number of nodes in a graph and $\langle k\rangle$ the average node degree. In fact, the only network studied that is globally less convex than a random graph is the Little Rock food web.

On the other hand, random graphs are *locally convex* meaning that any connected subgraph with up to $\ln n/\ln\langle k\rangle$ nodes is almost certainly convex. Globally convex networks are also fairly locally convex, or even more convex than random graphs under a loose definition of local convexity, whereas almost any other network studied is locally less convex than a random graph. Still, most of these networks are regionally convex.

Regional convexity refers to any type of heterogeneous network structure that is only partly convex. For instance, networks with core-periphery structure can be divided into a non-convex c-core surrounded by a convex periphery. Such are the Oregon Internet map and *C. elegans* protein network. Note that this type of regional convexity does not necessarily imply local convexity. This is because the nodes in convex periphery are generally disconnected and are connected only through the non-convex c-core.



Expansion of convex subgraphs in graphs and networks. (*top*) Expansion of convex subgraphs in a randomly grown tree (diamonds), triangular lattice of the same size (squares) and the corresponding random graph [4] (ellipses). Plots show the fractions of nodes $s(t)$ in the growing convex subgraphs at different steps t, $s(t) \geq (t+1)/n$. Graphs show particular realizations of convex subgraphs grown from the most central node for 15 steps. (*bottom*) Expansion of convex subgraphs in a globally convex coauthorship graph, regionally convex Internet map and non-convex food web. Plots show $s(t)$ for empirical networks (diamonds), randomly rewired networks or the configuration model graphs [10] (squares) and the corresponding random graphs [4] (ellipses). Networks show realizations of convex subgraphs, where diamonds represent the nodes included in the growing subgraphs by construction, while squares are the nodes included by expansion to convex subgraphs.

We propose different measures of local, regional and global convexity in networks. Among them, c-convexity can be used to assess global convexity and measures whether the structure of a network is either tree-like or clique-like, which is in contrast with the structure of a random graph. There are many measures that separate networks from random graphs like the average node clustering coefficient [11] and network modularity [9]. However, these clearly distinguish between the tree-like structure of infrastructure networks and the clique-like structure of coauthorship graphs. Yet, the two regimes are equivalent according to c-convexity. This is because they represent the border cases of networks with deterministic structure.

Convexity is thus an inherent structural property of some networks. Random graph models [4, 10] and also standard network models [11, 3] fail to reproduce convexity in networks. This is not surprising as most models are based on the existence of individual edges between the nodes and not on the inclusion of the entire shortest paths. Development of models of convex networks represents an important direction for future research.

Network convexity is an indication of the uniqueness of shortest paths in a network. The shortest paths are mostly unique in convex infrastructure networks due to high cost of connections, while largely redundant in a non-convex food web in order for the ecosystem to survive. Convex networks thus represent locally self-sufficient systems. As such convexity can be seen as a measure of network redundancy, a concept closely related to robustness and resilience [1].

[1] R. Albert, H. Jeong, and A. L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[2] H.-J. Bandelt and V. Chepoi. Metric graph theory and geometry: A survey. *Contemporary Mathematics*, 453:49–86, 2008.

[3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[4] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[5] L. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[6] F. Harary and J. Nieminen. Convexity in graphs. *Journal of Differential Geometry*, 16(2):185–190, 1981.

[7] T. Marc and L. Šubelj. Convexity in complex networks. *in submission to Network Science*, pages 1–25, 2016.

[8] S. Milgram. The small world problem. *Psychology Today*, 1(1):60–67, 1967.

[9] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004.

[10] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2):026118, 2001.

[11] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

This work has been supported in part by the Slovenian Research Agency under programs P1-0297 and P2-0359.