Journal of Informetrics 11 (2017) 257–274

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Regular article

Partial orders for zero-sum arrays with applications to network theory

Yuxian Liu[a], Ronald Rousseau[b,c], Leo Egghe[d,*]

[a] Tongji University Library, Tongji University, Siping Street 1239, Yangpu District, Shanghai 200092, China
[b] University of Antwerp, Antwerpen, Belgium
[c] KU Leuven, Department of Mathematics, Leuven, Belgium
[d] University of Hasselt (UHasselt), Campus Diepenbeek, Agoralaan, B-3590 Diepenbeek, Belgium

ARTICLE INFO

Article history:
Received 6 September 2016
Received in revised form
17 November 2016
Accepted 22 November 2016
Available online 21 January 2017

Keywords:
Zero-sum arrays
Dominance
Hierarchy
Power structures
Pseudo-Lorenz curves
Acyclic digraphs

ABSTRACT

In this contribution we study partial orders in the set of zero-sum arrays. Concretely, these partial orders relate to local and global hierarchy and dominance theories. The exact relation between hierarchy and dominance curves is explained. Based on this investigation we design a new approach for measuring dominance or stated otherwise, power structures, in networks. A new type of Lorenz curve to measure dominance or power is proposed, and used to illustrate intrinsic characteristics of networks. The new curves, referred to as D-curves are partly concave and partly convex. As such they do not satisfy Dalton's transfer principle. Most importantly, this article introduces a framework to compare different power structures as a whole.

It is shown that D-curves have several properties making them suitable to measure dominance. If dominance and being a subordinate are reversed, the dominance structure in a network is also reversed.

1. Introduction

Lorenz curves were introduced at the beginning of the 20th century as a graphical device to show the intrinsic inequality among a set of cells (Lorenz, 1905). Since their introduction, it has become clear that they constitute a powerful device that can be used in many fields and for many applications (Marshall, Olkin & Arnold, 2011). Moreover, many variations on the basic construction have been introduced. For an overview we refer to (Rousseau, 2011). In such applications one studies individuals or households (income inequality studies), species (in ecology), authors (informetric concentration) and so on. In general terms these are called cells. The number of cells in ecological diversity studies are often referred to as species richness, i.e. the number of observed species; in informetric studies the number of cells is just the number of authors considered in an investigation. Classically the values associated with cells used in the construction of a Lorenz curve are positive. Yet in applications cells may have positive as well as negative values. For instance, Lorenz curves representing wealth may include people who are in debt, hence have negative wealth. For this reason Lorenz curves are sometimes allowed to include cells with negative values. This case is studied, e.g. in (Cowell & Van Kerm, 2015; Egghe, 2002; section 2.1). Yet, when the sum of all data is zero this approach cannot be used. This is the main motivation for this study of zero-sum arrays.

* Corresponding author.
E-mail addresses: yxliu@tongji.edu.cn (Y. Liu), ronald.rousseau@uantwerpen.be, ronald.rousseau@kuleuven.be (R. Rousseau), leo.egghe@uhasselt.be (L. Egghe).

http://dx.doi.org/10.1016/j.joi.2016.11.006

This contribution consists of two main parts: one studying zero-sum arrays in general (part A) and one about applications in directed networks (part B). We make a distinction between hierarchical structures and dominance structures. Besides in information science where one studies networks such as article citation networks, we also have in mind future applications in food webs (Cohen, 1978; Garlaschelli et al., 2003), disease networks (Goh et al., 2007), social networks (Shizuka & McDonald, 2012), innovation networks (Guan et al., 2015), river flows (Poff et al., 2003), epidemiology (Moslonka-Lefebvre et al., 2011), management (Montgomery & Oliver, 2007) and complex networks in general (Strogatz, 2001; Newman, 2003). We like to point out that in this contribution we study structures, not single elements. Concretely, we introduce a framework to compare power or dominance structures. For instance, when applied to networks we attach a value to the dominance structure as a whole and do not intend to determine a value representing the power of a single node, as done e.g. in (Bonacich, 1987; van den Brink & Gilles, 2000).

2. Zero-sum arrays

2.1. Basic definitions

If X is a (finite) array, i.e. an N-tuple, then the j-th element of X is denoted as $(X)_j = x_j$, where x_j is a real number. In this investigation the components of any array are assumed to be ranked in decreasing order. If X is an array then $-$ X denotes the array where every component x_j is replaced by $-x_j$. Note that also the components of $-X$ are ranked in decreasing order. Hence: $(-X)_j = -(X)_{N-j+1}$. This array is called the opposite array of X. Clearly, the opposite of the opposite is again the original. Hence, taking the opposite array is an involutive operation, e.g. when applied twice it returns the original object. An array is said to be symmetric if, after re-ranking from largest to smallest, $X = -X$.

Examples. If X = (3,2,−6) then −X = (6,−2,−3). The array S = (3,1,0,−1,−3) is symmetric.

Definition. Zero-sum array

If $X = (x_1, \ldots, x_N)$ is a real-valued array (N-tuple), such that $\sum_{i=1}^{N} x_i = 0$, then X is called a zero-sum array. The set of all zero-sum arrays is denoted as **Z;** its subset of arrays of length N, namely $Z \cap \mathbb{R}^N$ is denoted as $\mathbf{Z}_N$. The symmetric array S from the previous example is also an example of a zero-sum array. Actually all symmetric arrays are zero-sum arrays. $Y_1 = (1,1,-2)$ and Y2 = (0.4, 0.3, 0, 0, −0.7) are examples of a non-symmetric zero-sum arrays.

Operations on zero-sum arrays: elementary properties.

1 If X is a zero-sum array of length N, then cX, with c a real number, is again a zero-sum array of length N.
2 $\mathbf{Z}_N$ is a convex cone over the real numbers. This means that if X and Y are zero-sum arrays of length N and α and β are positive real numbers, then αX + βY is again a zero-sum array of length N.

2.2. Two types of pseudo-Lorenz curves for zero-sum arrays

Given a zero-sum array X, we set

$$I_+(X) = \left\{ i \in \left\{1, \ldots, N\right\} \text{ such that } x_i > 0 \right\}$$
$$I_0(X) = \left\{ i \in \left\{1, \ldots, N\right\} \text{ such that } x_i = 0 \right\} \ \textit{and}$$
$$I_-(X) = \left\{ i \in \left\{1, \ldots, N\right\} \text{ such that } x_i < 0 \right\}.$$

We assume from now on that X is not the trivial zero-array, hence $I_0 \neq \{1, \ldots, N\}$. This implies that $I_+(X)$ and $I_-(X)$ are always non-empty, but they may have different numbers of elements. When it is clear about which array we are talking or when it does not matter we simply write I_+, I_0 or I_-.

As $\sum_{i=1}^{N} x_i = 0$ we see that $\sum_{i \in I_+} x_i + \sum_{i \in I_0} x_i + \sum_{i \in I_-} x_i = \sum_{i \in I_+} x_i + \sum_{i \in I_-} x_i = 0$. Hence $\sum_{i \in I_+} x_i = -\sum_{i \in I_-} x_i$.

Next we put $\Sigma_+ = \sum_{i \in I_+} x_i$ and $\forall i = 1, \ldots, N : a_i = \frac{x_i}{\Sigma_+}$. With each zero-sum array X, we associate a corresponding A-array, denoted A_X, and equal to $A_X = (a_1, \ldots, a_N)$. Of course, also A_X is a zero-sum array and $A_{A_X} = A_X$ (this is a projection property). Next we form the array P_X of partial sums: $(P_X)_j = p_j = \sum_{k=1}^{j} a_k$. Clearly $p_N = 0$ and $p_{N-|I_-|} = 1$. Similarly we form the array Q_X, with $(Q_X)_j = q_j = \sum_{k=1}^{j} |a_k|$. For $i = 1, \ldots, N-|I_-|$: $p_i = q_i$.