IMA Journal of Applied Mathematics (2017) **82**, 548–560
doi:10.1093/imamat/hxx002
Advance Access Publication on 19 April 2017

Animal population social structure models

THOMAS P. WIHLER*
Mathematics Institute, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland
*Corresponding author: wihler@math.unibe.ch

[Received on 20 November 2015; revised on 2 August 2016; accepted on 27 January 2017]

We develop a new class of mathematical ranking models for the purpose of quantifying the social structure in animal populations. Our approach is based on taking into account both the interaction between single individuals as well as their role within their community as a whole. From a mathematical point of view, these models are (possibly nonlinear) eigenvalue problems for column stochastic matrices. In order to provide a procedure for their computational treatment, we derive a suitable Newton iteration method on simplexes.

Keywords: animal population biology; animal social networks; social structure models; ranking; Newton method on a simplex; numerical solution of fixed point equations; nonlinear matrix eigenvalue problems.

1. Introduction

Ranking of (finitely many) interacting individuals or objects by means of mathematical models is traditionally based on identifying the underlying network structure with a directed graph. Information is passed from one individual or object to another by means of some random walk process along the graph's edges. In the simplest case, for each node, the transition to neighbouring nodes is chosen with the same probability, and the system (or transition) matrix of the associated Markov process is called the *network (or hyperlink) matrix* $\boldsymbol{H}$. The ranking vector $\boldsymbol{r}$ for the individual nodes is (if $\boldsymbol{H}$ is irreducible) given as the Perron vector of $\boldsymbol{H}$, i.e. the unique positive eigenvector of $\boldsymbol{H}$ (scaled such that the components add to 1) corresponding to the eigenvalue 1: $\boldsymbol{Hr} = \boldsymbol{r}$; see, e.g. (Meyer, 2000, Section 8). In the famous PageRank model by Brin & Page (1998) for electronic networks this basic approach has been extended by taking into account a (uniformly distributed) stochastic component between *all* the nodes of the graph. Over the last years, this idea has been applied in many different areas, and numerous modifications have been proposed; see the overview paper by Gleich (2015) (cf. also Krause *et al.* (2014) for animal social networks). We note that, whilst the PageRank model and its variants are based on tracing the flow along paths of a network, other approaches (including eigenvector centrality) directly employ the underlying adjacency matrix of the graph.

The focus of this work is on providing a methodology by which hierarchies within animal populations can be quantified. This can be of particular interest, for instance, to behavioural biologists who study different ways of interaction in (smaller) animal groups. Whilst the models to be presented in this article are based on the network idea as well, they specifically take into account the social role of each individual. The key idea in deriving our ranking models consists in the introduction of an additional individual whose role is to represent the community of all individuals as a whole. This virtual individual can be interpreted, for instance, as a joint effort put forth by everybody for the well-being, protection, or conservation of the entire population. On a closely related note, Allesina & Pascual (2009) propose the application of a 'root node' in the context of food webs. From the view point of Markov processes, such approaches can be seen as *censored node constructions* (see, e.g. Eiron *et al.*, 2004; Lee *et al.*, 2007, or Gleich, 2015, Section 5.5).

In this article, we will study different ways of interaction between the community and single individuals, and thereby, obtain linear as well as nonlinear ranking models. They are expressed in terms of eigenvalue problems for the associated social structure matrices, which, depending on the type of interaction, may be linear or even nonlinear (Section 2). In the latter case, we propose a Newton iteration for the numerical approximation of the ranking vector (Section 3). In order to provide some specific computations, we consider an example of a real-life (small) animal group, and use linear and nonlinear ranking to test the proposed ideas (Section 4).

2. Social structure models

Let us consider a network of m individuals, $I_1, \ldots, I_m$, which may or may not share their resources (such as, for instance, natural goods, time, etc.) with each other. To design a mathematical model for the purpose of describing a ranking structure amongst individuals, we proceed in a standard way and identify the network with a directed graph, whereby the individuals take the role of the nodes of the graph; furthermore, the graph has a directed edge from node i to node j whenever individual I_i shares its resources with individual I_j. We let

$$\sigma_j = \begin{cases} 1 & \text{if individual } I_j \text{ shares some of its resources with other individuals,} \\ 0 & \text{if individual } I_j \text{ does not share any of its resources,} \end{cases}$$

and, more specifically, if $\sigma_j = 1$, denote by l_j the number of outgoing edges from node j, i.e. the number of individuals being supplied with the resources of individual I_j. The matrix $\boldsymbol{H}$ of such a network is an $m \times m$-matrix whose entries are defined by

$$H_{ij} = \begin{cases} 1/l_j & \text{if individual } I_j \text{ shares some of its resources with individual } I_i, \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

for $1 \leq i,j \leq m$. Here, for later purposes, we do not consider self-links in $\boldsymbol{H}$ (i.e. the diagonal entries of $\boldsymbol{H}$ are assumed to be zero). We quantify the resource of each individual I_i by a number $0 \leq r_i \leq 1$, and suppose that this resource is shared equally amongst the l_i individuals it points to. The resource r_i, in turn, is obtained as the sum of all the resources received by individual I_i from other individuals. To strive towards uniqueness, the resource vector $\boldsymbol{r} = (r_1, \ldots, r_m)^\top$ is normalized so that $\boldsymbol{r} \in S_1^m$, where, for the vector $\boldsymbol{q} := (1, \ldots, 1)^\top \in \mathbb{R}^m$, we define the simplex

$$S_1^m := \left\{ \boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m_{\geq 0} : \boldsymbol{q}^\top \boldsymbol{x} = \sum_{i=1}^m x_i = 1 \right\};$$

here, $\mathbb{R}_{\geq 0}$ signifies the set of all nonnegative real numbers. This model corresponds to a linear eigenvalue problem,

$$\boldsymbol{r} \in S_1^m : \qquad \boldsymbol{H}\boldsymbol{r} = \boldsymbol{r}, \tag{2.2}$$

where the solution $\boldsymbol{r}$ is the Perron vector of $\boldsymbol{H}$ (if it exists); cf. (Meyer, 2000, Section 8).

In our models, the status of each individual is determined by the resource it owns, i.e. the resource vector $\boldsymbol{r}$ can be seen as ranking vector. It is well known that, in general, the solution vector $\boldsymbol{r}$ of the above equation (2.2) may contain meaningless zeros due to the possible appearance of dangling nodes in the graph of $\boldsymbol{H}$; specifically, this may happen if the network graph is not strongly connected, or, equivalently, if the network matrix is reducible (see, e.g. Austin, 2008; Meyer, 2000). In the PageRank model by Brin & Page (1998), this problem is circumvented by perturbing the network matrix by a strictly positive stochastic part, which mimics a random switch between network nodes.

2.1. *A linear social structure model*

Following the basic principle of *natura non facit saltus*, the concept of using random jumps between individuals, as proposed in the PageRank model, seems rather unrealistic in real-life animal populations (e.g. if the exchange of resources always happens according to the same rules, or, if individuals demonstrate predominantly deterministic behavioural patterns). Nevertheless, a mathematically equivalent idea, by which a small linking probability between individuals with infrequent (i.e. practically unobservable) interactions is taken into account, could still be applied. In this article, we focus on yet another aspect of interaction: community. Here, our aim is to quantify (in an averaged sense) all efforts (going beyond individual interests and connections between group members) undertaken to maintain the common weal of a population. To this end, the key idea is to introduce a virtual individual $\mathscr{C} = I_{m+1}$, which represents the community and its activities as a whole, and that everybody contributes to and benefits from. The model is built upon the following rules:

(a) Every individual (including the community node $\mathscr{C}$) owns a resource $r_i \geq 0$; the resource (or ranking) vector

$$\widehat{\boldsymbol{r}} = \begin{pmatrix} \boldsymbol{r} \\ r_{\mathscr{C}} \end{pmatrix} = (r_1, \ldots, r_m, r_{\mathscr{C}})^\top$$

is normalized so that $\widehat{\boldsymbol{r}} \in S_1^{m+1}$;

(b) Every individual I_i, $i = 1, \ldots, m$, keeps a certain fixed proportion $\alpha_s r_i$ of its resource r_i, where $0 \leq \alpha_s < 1$ is a fixed parameter, to itself, and transfers another fixed portion $\alpha_{\mathscr{C}} r_i$, with $0 < \alpha_{\mathscr{C}} \leq 1 - \alpha_s$ (taking the role of a flat rate), to the community $\mathscr{C}$; the latter contribution can be thought of as a common purpose rate required to sustain the entire community (e.g. joint hunting or defence activities, etc.). The remaining proportion, $(1 - \alpha_s - \alpha_{\mathscr{C}})r_i$, of the resource of I_i is shared equally amongst the l_i individuals it points to as described earlier.

(c) The community splits its resource $r_{\mathscr{C}}$ equally amongst all individuals in the network (excluding itself).

This leads to a linear eigenvalue problem

$$\widehat{\boldsymbol{r}} \in S_1^{m+1}: \qquad \boldsymbol{S}\widehat{\boldsymbol{r}} = \widehat{\boldsymbol{r}}, \tag{2.3}$$

where

$$\boldsymbol{S} = \left(\begin{array}{c|c} & 1/m \\ \boldsymbol{P} & \vdots \\ & 1/m \\ \hline \alpha_{\mathscr{C}} \cdots \alpha_{\mathscr{C}} & 0 \end{array} \right) \tag{2.4}$$

is the *social structure matrix*. Here, we let

$$\boldsymbol{P} = \alpha_s \boldsymbol{I}_{m\times m} + (1 - \alpha_s - \alpha_{\mathscr{C}})(\boldsymbol{H} + \boldsymbol{\Sigma}),$$

where $\boldsymbol{I}_{m\times m}$ denotes the $m \times m$-identity matrix, $\boldsymbol{H}$ is the network matrix from (2.1), and $\boldsymbol{\Sigma}$ is an $m \times m$-matrix defined by

$$\Sigma_{ij} = (1 - \sigma_j)\delta_{ij}, \tag{2.5}$$

where δ_{ij} is Kronecker's delta; the purpose of the matrix $\boldsymbol{\Sigma}$ is to guarantee that the columns of $\boldsymbol{S}$ all add to 1 even if there are individuals without any outgoing links. Evidently, since the graph of $\boldsymbol{S}$ is strongly connected through the virtual node $\mathscr{C}$, it follows that $\boldsymbol{S}$ is irreducible. By the Perron–Frobenius theorem, this implies that the resource vector $\widehat{\boldsymbol{r}}$ from (2.3) is uniquely determined in S_1^{m+1}, and contains only positive entries (i.e. $\widehat{\boldsymbol{r}}$ is the Perron vector of $\boldsymbol{S}$); see Meyer (2000, Section 8) for details. Upon defining the normalized resource (or ranking) vector

$$\boldsymbol{\rho} = (\rho_1, \ldots, \rho_m), \qquad \boldsymbol{\rho} = \left(\boldsymbol{q}^\top \boldsymbol{r}\right)^{-1} \boldsymbol{r} \in S_1^m, \tag{2.6}$$

we proceed as in Gleich (2015, Section 5.5), and note that

$$\boldsymbol{P}\boldsymbol{r} + \frac{r_{\mathscr{C}}}{m}\boldsymbol{q} = \boldsymbol{r}, \qquad \alpha_{\mathscr{C}}\boldsymbol{q}^\top \boldsymbol{r} = r_{\mathscr{C}}.$$

Hence,

$$\boldsymbol{r} = \boldsymbol{P}\boldsymbol{r} + \frac{\alpha_{\mathscr{C}}}{m}\left(\boldsymbol{q}^\top \boldsymbol{r}\right)\boldsymbol{q} = \left(\boldsymbol{P} + \frac{\alpha_{\mathscr{C}}}{m}\boldsymbol{q}\boldsymbol{q}^\top\right)\boldsymbol{r},$$

and therefore,

$$\boldsymbol{\rho} = \left(\boldsymbol{P} + \frac{\alpha_{\mathscr{C}}}{m}\boldsymbol{q}\boldsymbol{q}^\top\right)\boldsymbol{\rho}. \tag{2.7}$$

We observe that the matrix

$$\mathfrak{S} = \boldsymbol{P} + \frac{\alpha_{\mathscr{C}}}{m}\boldsymbol{q}\boldsymbol{q}^\top$$

is column stochastic, and, in addition, has only positive entries. Therefore, $\mathfrak{S}$ is primitive, and a power iteration method for the numerical approximation of the resource vector $\boldsymbol{\rho}$ may be applied (see Meyer, 2000, Eq. 8.3.10). Moreover, from a mathematical point of view, we observe that the term $\alpha_{\mathscr{C}}/m\,\boldsymbol{q}\boldsymbol{q}^\top$ can be interpreted as a stochastic jump part as proposed in the PageRank model.

2.2. *A nonlinear approach*

In the model assumption (b) above each individual transfers *the same proportion* $\alpha_{\mathscr{C}}$ of its resource to the community. It is interesting to generalize this approach, for instance, by introducing a 'tax rate' for each individual I_j, $j = 1, \dots, m$, that depends on the resource ρ_j it owns. In analogy with (2.4), this results in a social structure matrix $\boldsymbol{S} = \boldsymbol{S}(\boldsymbol{\rho}) \in \mathbb{R}^{(m+1)\times(m+1)}$, with the normalized resource vector $\boldsymbol{\rho}$ given as in (2.6), which is defined by

$$\boldsymbol{S}(\boldsymbol{\rho}) = \left(\begin{array}{c|c} & 1/m \\ \boldsymbol{P}(\boldsymbol{\rho}) & \vdots \\ & 1/m \\ \hline \alpha_{\mathscr{C}}(\rho_1)\cdots\alpha_{\mathscr{C}}(\rho_m) & 0 \end{array}\right).$$

The entries of the matrix $\boldsymbol{P}(\boldsymbol{\rho})$ are given by

$$P_{ij}(\boldsymbol{\rho}) = \alpha_s\delta_{ij} + (1 - \alpha_s - \alpha_{\mathscr{C}}(\rho_j))(H_{ij} + \Sigma_{ij}), \qquad 1 \le i, j \le m,$$

where H_{ij} are the entries of the network matrix from (2.1), and Σ_{ij} is defined in (2.5). Furthermore,

$$\alpha_{\mathscr{C}} : [0, 1] \to [0, 1], \qquad \rho \mapsto \alpha_{\mathscr{C}}(\rho) \tag{2.8}$$

is a 'tax rate' function. The corresponding eigenvalue problem,

$$\widehat{\boldsymbol{r}} = \begin{pmatrix} \boldsymbol{r} \\ r_{\mathscr{C}} \end{pmatrix} \in S_1^{m+1},\ \boldsymbol{\rho} = \left(\boldsymbol{q}^\top \boldsymbol{r}\right)^{-1} \boldsymbol{r} \in S_1^m : \qquad \boldsymbol{S}(\boldsymbol{\rho})\widehat{\boldsymbol{r}} = \widehat{\boldsymbol{r}}$$

is nonlinear. Introducing the vector function

$$\boldsymbol{\alpha}_{\mathscr{C}}(\boldsymbol{\rho}) = (\alpha_{\mathscr{C}}(\rho_1), \dots, \alpha_{\mathscr{C}}(\rho_m))^\top$$

and proceeding as before in Section 2.1 (see, in particular, Gleich, 2015, Section 5.5), yields

$$\boldsymbol{P}(\boldsymbol{\rho})\boldsymbol{r} + \frac{r_{\mathscr{C}}}{m}\boldsymbol{q} = \boldsymbol{r}, \qquad \boldsymbol{\alpha}_{\mathscr{C}}(\boldsymbol{\rho})^\top \boldsymbol{r} = r_{\mathscr{C}}.$$

Therefore,

$$\boldsymbol{r} = \left(\boldsymbol{P}(\boldsymbol{\rho}) + \frac{1}{m}\boldsymbol{q}\boldsymbol{\alpha}_{\mathscr{C}}(\boldsymbol{\rho})^\top\right)\boldsymbol{r},$$

and, thus,

$$\boldsymbol{\rho} = \mathfrak{S}(\boldsymbol{\rho})\boldsymbol{\rho} \tag{2.9}$$

with

$$\mathfrak{S}(\boldsymbol{\rho}) = \boldsymbol{P}(\boldsymbol{\rho}) + \frac{1}{m}\boldsymbol{q}\boldsymbol{\alpha}_{\mathscr{C}}(\boldsymbol{\rho})^\top. \tag{2.10}$$

In Section 3, we will discuss how this problem can be dealt with by means of a suitable Newton iteration scheme, and some examples will be provided in Section 4.

2.3. *More general models*

Evidently, depending on a specific application, the nonlinear model in Section 2.2 can be modified and extended in various ways. For instance, a nonlinear distribution of the resources leaving from the community node (paying more attention, e.g. to individuals with special needs), or a weighted transfer of an individual's resources (being more or less generous to individuals with higher ranks) could be taken into account. Moreover, more than one community node could be introduced if, for instance, several behavioural aspects (such as playing, sharing, grooming, etc.) of the community are modelled separately.

3. A Newton approach on simplexes

The nonlinear eigenvalue problem (2.9) can be cast into a general framework. Indeed, introducing the function $\boldsymbol{f}(\boldsymbol{\rho}) := \mathfrak{S}(\boldsymbol{\rho})\boldsymbol{\rho}$, and noticing that $\mathfrak{S}(\boldsymbol{\rho})$ is column stochastic, it is straightforward to verify that $\boldsymbol{f} : S_1^m \to S_1^m$ is a self-mapping, and that (2.9) is equivalent to solving $\boldsymbol{f}(\boldsymbol{\rho}) = \boldsymbol{\rho}$ on S_1^m.

3.1. *Fixed point formulation*

More generally, let us consider a differentiable self-mapping

$$\boldsymbol{f} : S_1^m \to S_1^m \tag{3.1}$$

on the simplex S_1^m. Then, the fixed point equation,

$$\boldsymbol{x} \in S_1^m : \qquad \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x} \tag{3.2}$$

has at least one solution by Brouwer's fixed point theorem (see, e.g. Kellogg *et al.*, 1976). In the sequel, we will define a Newton-type approximation procedure on S_1^m for the numerical solution of (3.2). We emphasize that the development of numerical procedures for fixed point problems of the form (3.2), in particular, in the context of Brouwer's fixed point theorem, dates back quite a few decades. For example, several works have pursued a geometric approach, whereby appropriate space partition strategies are employed; we mention the seminal paper by Scarf (1967) (see also Kuhn, 1969), where a numerical scheme for the solution of fixed point equations for continuous functions on S_1^m in low dimensions, based on a variant of Sperner's lemma, has been presented. A completely different idea is the development of dynamical system formulations that allow to transport suitable initial guesses iteratively along discrete trajectories to a fixed point; see, e.g. the article Kellogg *et al.* (1976), which includes a constructive proof of Brouwer's fixed point theorem (based on the retraction principle) and enables the derivation of an iterative Newton-type path-following formulation. In addition, we mention the homotopy-based methods by Chow *et al.* (1978) and Eaves (1972). Moreover, let us refer to Allgower & Georg (2003, Section 11) for further references and details on numerical considerations on Brouwer's fixed point theorem.

The approach taken in this article is based on formulating a Newton method as applied to (3.2) which operates *within* the simplex S_1^m. This is accomplished by formulating the Newton iteration with respect to a suitable coordinate system on the linear subspace

$$V_0 := \left\{ \boldsymbol{x} \in \mathbb{R}^m : \boldsymbol{q}^\top \boldsymbol{x} = 0 \right\}, \qquad \boldsymbol{q} = (1, \dots, 1)^\top \in \mathbb{R}^m,$$

which contains the translated simplex

$$W_0 := S_1^m - \frac{1}{m}\boldsymbol{q} \subset V_0,$$

and, for this purpose, by deriving a corresponding representation of the Fréchet derivative of $\boldsymbol{f}$. We remark that this approach is related to the work (Deuflhard, 2004, Section 6.4) on pseudo-transient continuation in the context of dynamical invariants. Pursuing the standard idea of writing a fixed point equation as a root problem, we notice that $\overline{\boldsymbol{x}} \in S_1^m$ is a fixed point of $\boldsymbol{f}$ in (3.1) if and only if the function $\boldsymbol{F}(\boldsymbol{z}) := \boldsymbol{f}(\boldsymbol{z} + {}^1\!/_m\boldsymbol{q}) - \boldsymbol{z} - {}^1\!/_m\boldsymbol{q}$ has a zero at $\overline{\boldsymbol{z}} = \overline{\boldsymbol{x}} - {}^1\!/_m\boldsymbol{q} \in W_0$, that is,

$$\overline{\boldsymbol{z}} \in W_0 : \qquad \boldsymbol{F}(\overline{\boldsymbol{z}}) = \boldsymbol{0}. \tag{3.3}$$

In the following, we will aim to formulate a Newton method for the iterative solution of (3.3) in the subspace V_0. To this end, we will introduce a suitable basis in V_0, and derive a matrix expression of the Jacobian $\boldsymbol{F}'(\boldsymbol{z})$, with $\boldsymbol{z} \in W_0$, as acting on V_0.

3.2. *Matrix representation of* $\boldsymbol{F}(\boldsymbol{z})|'_{V_0}$

We consider a set of linearly independent vectors $\{\boldsymbol{d}_1, \dots, \boldsymbol{d}_{m-1}\} \subset V_0$ given as follows:

$$\boldsymbol{d}_k \in \mathbb{R}^m : \quad (\boldsymbol{d}_k)_l = \begin{cases} {}^1\!/\sqrt{k(k+1)} & 1 \le l \le k \\ -{}^k\!/\sqrt{k(k+1)} & l = k+1 \\ 0 & k+1 < l \le m \end{cases}, \qquad k = 1, \dots, m-1.$$

It is elementary to verify that these vectors form an orthonormal basis of V_0 (with respect to the Euclidean product in $\mathbb{R}^m$). Therefore, defining

$$\boldsymbol{M} = (\boldsymbol{d}_1 | \boldsymbol{d}_2 | \cdots | \boldsymbol{d}_{m-1}) \in \mathbb{R}^{m \times (m-1)}$$

as the matrix whose columns are the above basis vectors, there holds $\boldsymbol{M}^\top \boldsymbol{M} = \boldsymbol{I}_{(m-1)\times(m-1)}$, where we write $\boldsymbol{I}_{(m-1)\times(m-1)}$ to signify the $(m-1) \times (m-1)$-identity matrix. In the sequel, we write $[\boldsymbol{v}] \in \mathbb{R}^{m-1}$ to denote the coordinate vector of a vector $\boldsymbol{v} \in V_0$ with respect to the basis $\boldsymbol{d}_1, \dots, \boldsymbol{d}_{m-1}$. Then, for $\boldsymbol{v} \in V_0$, we notice that $\boldsymbol{v} = \boldsymbol{M}[\boldsymbol{v}]$, and $[\boldsymbol{v}] = \boldsymbol{M}^\top \boldsymbol{v}$.

Given $\boldsymbol{z} \in W_0$, and using the coordinate representation in V_0, the Fréchet derivative of $\boldsymbol{F}(\boldsymbol{z})|_{V_0}$ is given by

$$\boldsymbol{F}(\boldsymbol{z})|'_{V_0} = \boldsymbol{M}^\top \boldsymbol{F}'(\boldsymbol{z}) \boldsymbol{M} \tag{3.4}$$

with respect to the coordinates in V_0.

3.3. *Newton iteration on* V_0

Starting from an initial guess $\boldsymbol{x}^{(0)} \in S_1^m$, and defining $\boldsymbol{z}^{(0)} := \boldsymbol{x}^{(0)} - {}^1\!/_m\boldsymbol{q} \in W_0$, a Newton iteration scheme for the numerical solution of (3.3) can be formulated by

$$[\boldsymbol{z}^{(n+1)}] = [\boldsymbol{z}^{(n)}] - \mu_n(\boldsymbol{F}(\boldsymbol{z}^{(n)})|'_{V_0})^{-1}[\boldsymbol{F}(\boldsymbol{z}^{(n)})] = [\boldsymbol{z}^{(n)}] - \mu_n(\boldsymbol{F}(\boldsymbol{z}^{(n)})|'_{V_0})^{-1}\boldsymbol{M}^\top\boldsymbol{F}(\boldsymbol{z}^{(n)})$$

for $n \geq 0$, where $\boldsymbol{F}(\boldsymbol{z}^{(n)})|'_{V_0}$ is defined in (3.4) and assumed invertible, and $\mu_n \geq 0$ will be specified below. Equivalently, for any $n \geq 0$, we can solve the linear system

$$\boldsymbol{F}(\boldsymbol{z}^{(n)})|'_{V_0}[\boldsymbol{\delta}^{(n)}] = -\boldsymbol{M}^\top\boldsymbol{F}(\boldsymbol{z}^{(n)}) \tag{3.5}$$

for $[\boldsymbol{\delta}^{(n)}]$, with the update formula $\boldsymbol{z}^{(n+1)} = \boldsymbol{z}^{(n)} + \mu_n\boldsymbol{M}[\boldsymbol{\delta}^{(n)}]$. In order to ensure that

$$\boldsymbol{x}^{(n+1)} = \boldsymbol{z}^{(n)} + \frac{1}{m}\boldsymbol{q} + \mu_n\boldsymbol{M}[\boldsymbol{\delta}^{(n)}] \in S_1^m,$$

or, equivalently $\boldsymbol{z}^{(n+1)} \in W_0$, we maximize $\mu_n \in [0, t_{\max}]$, for a prescribed parameter $t_{\max} > 0$, so that $\boldsymbol{z}^{(n)} + {}^1\!/_m\boldsymbol{q} + \mu_n\boldsymbol{M}[\boldsymbol{\delta}^{(n)}] \geq \boldsymbol{0}$, where '$\geq$' is understood componentwise, i.e.

$$\mu_n = \max\left\{0 \leq t \leq t_{\max} : \boldsymbol{z}^{(n)} + \frac{1}{m}\boldsymbol{q} + t\boldsymbol{M}[\boldsymbol{\delta}^{(n)}] \geq \boldsymbol{0}\right\}. \tag{3.6}$$

Provided that the linear system (3.5) is uniquely solvable, and that $\mu_n > 0$, for all $n \geq 0$ up to a required number of iterations, the above procedure is repeated until the residual $\boldsymbol{F}(\boldsymbol{z}^{(n+1)})$ (or, alternatively, the relative difference of two consecutive iterates) is sufficiently small with respect to some appropriate norm. Incidentally, the solution vector $\boldsymbol{r}$ for a linear fixed-point problem, as in (2.3), is obtained exactly (up to rounding errors) after one step of the Newton method.

REMARK 3.1 We notice that the choice $t_{\max} = 1$ in (3.6) is most natural for the case of a root of $\boldsymbol{F}$ of algebraic multiplicity 1. Indeed, if $\mu_n = t_{\max} = 1$ is admissible, then quadratic convergence is generally expected close to the corresponding fixed point of $\boldsymbol{f}$. A more global convergence analysis of the proposed Newton method will require additional assumptions on $\boldsymbol{F}$ (including, for instance, some affine co-variant or contra-variant Lipschitz condition on the Jacobian $\boldsymbol{F}'$); see, e.g. the monograph (Deuflhard, 2004, Section 2) for details.

4. Example of a real-life animal population

Some animal groups are known to have quite a clear hierarchy structure. This becomes apparent in various types of interactions including, for instance, aggression or proximity. For the purpose of testing the proposed models, we study an example of a real group of $m = 10$ white-faced capuchins on Barro Colorado Island located in the Panama Canal. We shall focus on proximity, i.e. time spent (spatially) close to or interactively with another individual in the group. Our data is taken from (Crofoot *et al.*, 2011, Fig. 1, BLT group, proximity). The adjacency matrix $\boldsymbol{A}$ (which contains an entry 1 in row i and column j

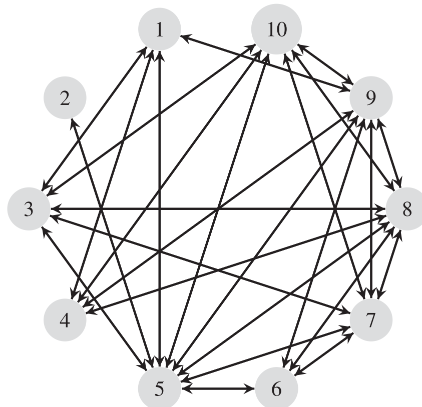

FIG. 1. Graph of a proximity interaction in a population of white-faced capuchins.

whenever individual j spends time with individual i, and an entry 0 otherwise) for this particular group is given by

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

Since proximity is considered a mutual interaction in Crofoot *et al.* (2011), the matrix $\boldsymbol{A}$ is symmetric in this case. Evidently, the network matrix $\boldsymbol{H}$ from (2.1) can be derived immediately from $\boldsymbol{A}$. Moreover, the associated network graph is displayed in Fig. 1. Throughout our experiments, we will use $\alpha_s = 0.2$.

For the nonlinear models, we apply the Newton method from Section 3 with $t_{\max} = 1$ in (3.6) for all computations (cf. Remark 3.1). We note that the Jacobian of the mapping $\boldsymbol{F}(\boldsymbol{\rho}) = \mathfrak{S}(\boldsymbol{\rho})\boldsymbol{\rho}$, with $\mathfrak{S}(\boldsymbol{\rho})$ from (2.10), is given componentwise by

$$F'_{ij}(\boldsymbol{\rho}) = \mathfrak{S}_{ij}(\boldsymbol{\rho}) + \left(\frac{1}{m} - (H_{ij} + \Sigma_{ij})\right)\rho_j\alpha'_{\mathscr{C}}(\rho_j), \qquad 1 \leq i,j \leq m,$$

where $\alpha'_{\mathscr{C}}$ signifies the derivative of the taxation function $\alpha_{\mathscr{C}}$. The starting vector for the Newton iteration (3.5) is chosen to be $(1/10, \ldots, 1/10) \in S_1^{10}$ in all examples, and the iteration is stopped once the residual $\|\mathfrak{S}(\boldsymbol{\rho})\boldsymbol{\rho} - \boldsymbol{\rho}\|_\infty$ is reasonably small (alternatively, the relative difference of two consecutive iterates could be monitored); in the context of ranking models, we emphasize that convergence to machine precision is typically not required.

We begin by investigating the *linear model* from Section 2.1, and focus on low taxation ($\alpha_{\mathscr{C}} = 0.05$) as well as on high taxation ($\alpha_{\mathscr{C}} = 0.7$). The respective unique solutions of (2.7) together with the corresponding ranks are presented in Table 1. In the low-tax case, the rank of an individual is mainly

TABLE 1 *Performance data for different linear ranking models*

	Linear model for $\alpha_{\mathscr{C}} = 0.05$		Linear model for $\alpha_{\mathscr{C}} = 0.7$	
Individual	ρ	Rank	ρ	Rank
1	0.0783	7	0.0966	7
2	0.0244	10	0.0893	10
3	0.0959	6	0.0984	6
4	0.0778	8	0.0963	8
5	0.1545	1	0.1150	1
6	0.0775	9	0.0951	9
7	0.1137	4/5	0.1005	5
8	0.1320	3	0.1038	3
9	0.1323	2	0.1043	2
10	0.1137	4/5	0.1006	4

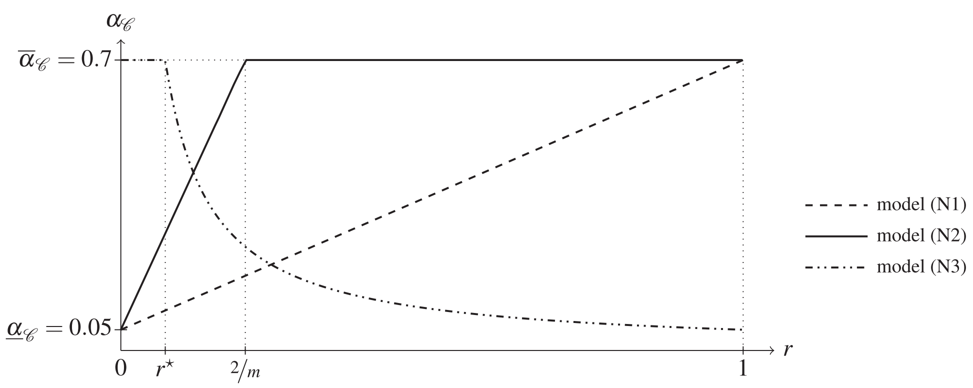

FIG. 2. Tax rate functions for $m = 10$.

determined by the network matrix, i.e. by the linking structure of the group. In contrast, when a high tax rate is imposed, then a high resource amount is less worthwhile, and although the overall ranking remains essentially the same in both examples, the individual resources in the network become notably more equidistributed in the latter case.

In addition, we study three different *nonlinear models* of the form (2.9), with a nonlinear tax rate function as in (2.8). The first two models are based on the assumption that more prominent individuals provide a larger contribution to the community than others. This is realized, for instance, by defining the linear tax rate function

$$\alpha_{\mathscr{C}}(r) = \underline{\alpha}_{\mathscr{C}}(1 - r) + \overline{\alpha}_{\mathscr{C}} r \tag{N1}$$

with $0 \leq r \leq 1$; see Fig. 2, model (N1). Here, $\overline{\alpha}_{\mathscr{C}}$ and $\underline{\alpha}_{\mathscr{C}}$ are upper and lower tax rate bounds, which, in our experiments, are chosen to be $\underline{\alpha}_{\mathscr{C}} = 0.05$ and $\overline{\alpha}_{\mathscr{C}} = 0.7$. The resulting resource vectors are displayed in Table 2. Since the individual resources are close to $1/m$, i.e. $\alpha_{\mathscr{C}}(1/m) \approx 0.1$, it is not surprising to see

TABLE 2 *Performance data for different nonlinear ranking models*

	Model (N1)		Model (N2)		Model (N3)	
Individual	ρ	Rank	ρ	Rank	ρ	Rank
1	0.0802	7	0.0863	8	0.0912	8
2	0.0311	10	0.0555	10	0.0697	10
3	0.0953	6	0.0937	6	0.0985	6
4	0.0792	8	0.0866	7	0.0911	9
5	0.1566	1	0.1577	1	0.1214	1
6	0.0775	9	0.0798	9	0.0913	7
7	0.1109	5	0.1025	5	0.1057	4
8	0.1285	3	0.1164	3	0.1128	2/3
9	0.1295	2	0.1186	2	0.1128	2/3
10	0.1112	4	0.1029	4	0.1056	5

that the computed numbers are fairly similar to the ones resulting from the linear model with $\alpha_{\mathscr{C}} = 0.05$ shown in Table 1.

We remark that the effective range of the tax rate function $\alpha_{\mathscr{C}}$ defined in (N1) depends strongly on the number m of individuals in the network. Indeed, suppose for a moment that all individuals have approximately the same rank $1/m$ (which is in fact the case for the given population); then, for large $m \gg 1$, the ranks will be situated in an interval $I_\epsilon = [0, \epsilon]$, for small $0 < \epsilon \ll 1$, and the tax rate function will only be evaluated on I_ϵ. In order to incorporate the scaling with respect to the size of the network, and thereby to exploit the entire domain [0, 1] of $\alpha_{\mathscr{C}}$, let us consider, for instance, the alternative function

$$\alpha_{\mathscr{C}}(r) = \underline{\alpha}_{\mathscr{C}}(1 - \min(rm/2, 1)) + \overline{\alpha}_{\mathscr{C}} \min(rm/2, 1); \tag{N2}$$

see Fig. 2, model (N2). Here, in comparison to the previous model, the results show that some resources are transferred from the medium ranking individuals to group members with lower ranks. This is due to the fact that, in the present model, taxation grows much faster as r increases. Moreover, we observe a switch of ranks of individuals 7 and 8.

Finally, we study a nonlinear model that is based on a constant tax rate $\widehat{\alpha}_{\mathscr{C}}$ for low-rank individuals, and, for all others, on transferring a fixed resource amount $\widehat{\beta}_{\mathscr{C}}$ to the community node. To this end, we define the taxation function

$$\alpha_{\mathscr{C}}(r) = \min\left(\widehat{\alpha}_{\mathscr{C}}, \widehat{\beta}_{\mathscr{C}}/r\right). \tag{N3}$$

We choose $\widehat{\alpha}_{\mathscr{C}} = 0.7$, and $\widehat{\beta}_{\mathscr{C}} = 0.5/m = 0.05$; see Fig. 2. We note that fixed tax deduction occurs for any $r \in [r^\star, 1]$, where $r^\star = \widehat{\beta}_{\mathscr{C}}/\widehat{\alpha}_{\mathscr{C}} \approx 0.0714$. In comparison with the linear model based on $\alpha_{\mathscr{C}} = 0.7$ (cf. Table 1), taxation becomes more favourable within this range. As a consequence, the resources of individuals with higher ranks are slightly increased at the cost of less resourceful group members, thereby leading to a change of ranks in the nonlinear model (N3); see Table 2.

REMARK 4.1 We note that the nonlinear tax rate functions (N2) and (N3) are not differentiable at some isolated points. Remarkably, the Newton method still worked without any problems for the examples presented above. Nevertheless, for differentiability purposes, it might be sensible to use suitable smooth approximations of the given taxation models.

REMARK 4.2 Whilst the resource vectors in all of the above experiments were found to be unique eigenvectors in S_1^m of the respective social hierarchy matrices, this is not true in general. Indeed, it is possible to find (academic) examples, where multiple ranking vectors for (2.9) exist.

5. Conclusions

In this article, we have introduced and discussed a class of linear and nonlinear matrix eigenvalue models for the purpose of quantifying hierarchies in animal populations. Our approach is based on adding a virtual individual (representing the community and its activities as a whole) to the underlying group of animals, and on identifying interactions between members of the population with a directed network graph. In case that the models are nonlinear, we have derived an iterative solution procedure which is based on a Newton scheme on simplexes. Whilst we have tested the proposed ideas in the specific context of a small white-faced capuchins population, more general scenarios could be considered as well:

- interactions between several animal populations (each of which taking the role of one 'individual') sharing a common habitat (representing the 'community');
- use of more than one virtual node in order to model different aspects of community;
- variable pay back rates from the community node to individual members of the group;
- nonlinear distribution of an individual's resources to other individuals.

Future work may include the development of especially tailored nonlinear models for particular types of animal groups and species, and the application and testing of our models in real-life situations (including the collection and evaluation of data).

REFERENCES

ALLESINA, S. & PASCUAL, M. (2009) Googling food webs: can an eigenvector measure species' importance for coextinctions?. *PLOS Comput Biol*, **5**, e1000494.

ALLGOWER, E. L. & GEORG, K. (2003) *Introduction to Numerical Continuation Methods*. Classics in Applied Mathematics, vol. 45. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

AUSTIN, D. (2008) How Google finds your needle in the Web's Haystack. AMS feature column (http://www.ams.org/samplings/feature-column/fcarc-pagerank).

BRIN, S. & PAGE, L. (1998) The anatomy of a large-scale hypertextual web search engine. *Comput. Networks ISDN*, **30**, 107–117.

CHOW, S. N., MALLET-PARET, J. & YORKE, J. A. (1978) Finding zeroes of maps: homotopy methods that are constructive with probability one. *Math. Comput.*, **32**, 887–899.

CROFOOT, M. C., RUBENSTEIN, D. I., MAIYA, A. S. & BERGER-WOLF, T. Y. (2011) Aggression, grooming and group-level cooperation in white-faced capuchins (Cebus capucinus): insights from social networks. *Am. J. Primatol.*, **73**, 821–833.

DEUFLHARD, P. (2004) *Newton Methods for Nonlinear Problems*. Springer Series in Computational Mathematics, vol. 35. Berlin: Springer. Affine invariance and adaptive algorithms.

EAVES, B. C. (1972) Homotopies for computation of fixed points. *Math. Program.*, **3**, 1–22.

EIRON, N., MCCURLEY, K. S. & TOMLIN, J. A. (2004) Ranking the web frontier. *Proceedings of the 13th International Conference on the World Wide Web (WWW2004)*. New York: ACM, pp. 309–318.

GLEICH, D. F. (2015) PageRank beyond the web. *SIAM Review*, **57**(3), 321–363.
KELLOGG, R. B., LI, T. Y. & YORKE, J. (1976) A constructive proof of the Brouwer fixed-point theorem and computational results. *SIAM J. Numer. Anal.*, **13**(4), 473–483.
KRAUSE, J., JAMES, R., FRANKS, D. & CROFT, D. (eds) (2014) *Animal Social Networks*. New York: Oxford University Press.
KUHN, H. W. (1969) Approximate search for fixed points. *Computing Methods in Optimization Problems, 2 (Proc. Conf., San Remo, 1968)*. New York: Academic Press, pp. 199–211.
LEE, C. P., GOLUB, G. H. & ZENIOS, S. A. (2007) A two-stage algorithm for computing PageRank and multistage generalizations. *Internet Mathematics*, **4**(4), 299–327.
MEYER, C. (2000) *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).
SCARF, H. (1967) The approximation of fixed points of a continuous mapping. *SIAM J. Appl. Math.*, **15**, 1328–1343.