Available online at www.sciencedirect.com

ScienceDirect

Theoretical Population Biology 70 (2006) 237–243

Theoretical Population Biology

www.elsevier.com/locate/tpb

Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index

Carlo Ricotta[a,*], Laszlo Szeidl[b]

[a]University of Rome ''La Sapienza'', Department of Plant Biology, Piazzale Aldo Moro 5, 00185 Rome, Italy
[b]Budapest Tech, Bécsi út 96/B, H-1034 Budapest, Hungary and Eötvös Loránd University, Pázmány P. st. 1/C, H-1117 Budapest, Hungary

Received 7 July 2004
Available online 13 June 2006

Abstract

The diversity of a species assemblage has been studied extensively for many decades in relation to its possible connection with ecosystem functioning and organization. In this view most diversity measures, such as Shannon's entropy, rely upon information theory as a basis for the quantification of diversity. Also, traditional diversity measures are computed using species relative abundances and cannot account for the ecological differences between species. Rao first proposed a diversity index, termed quadratic diversity (Q) that incorporates both species relative abundances and pairwise distances between species. Quadratic diversity is traditionally defined as the expected distance between two randomly selected individuals. In this paper, we show that quadratic diversity can be interpreted as the expected conflict among the species of a given assemblage. From this unusual interpretation, it naturally follows that Rao's Q can be related to the Shannon entropy through a generalized version of the Tsallis parametric entropy.

Keywords: Concavity; Conflict; Information theory; Pairwise species distances; Parametric diversity; Rarity

1. Introduction

Traditional diversity measures, such as the Shannon index or the Simpson index, summarize the information about the relative abundances of species without incorporating ecological differences between species. However, a community composed of ecologically dissimilar taxa is intuitively more diverse than a community composed of more similar taxa. Therefore, useful indices of diversity should account for inter-species differences related to functional type, morphology, taxonomic relatedness, or genetic distances, as ecological differences between species are believed to be reflected in each of these (Izsák and Papp, 1995, 2000; Warwick and Clarke, 1995; Clarke and Warwick, 1998; Shimatani, 2001; Ricotta, 2004).

Imagine a species assemblage or community composed of S species, where p_j is the relative abundance (measured as number of individuals, biomass or any other biological parameter that is thought to influence ecosystem functioning) of the jth species ($j = 1, 2, ..., S$) such that $0 \leqslant p_j \leqslant 1$ and $\sum_{j=1}^{S} p_j = 1$.

In his seminal paper, Rao (1982) first proposed a diversity index that incorporates both species relative abundances and a measure of the pairwise distances between species. Quadratic diversity, or quadratic entropy (Q) is defined as the expected distance between two randomly selected individuals:

$$Q = \sum_{i=1}^{S} \sum_{j=1}^{S} d_{ij} p_i p_j = p' \Delta p, \tag{1}$$

where d_{ij} is the distance between species i and j, p' denotes the transpose of the vector $p = (p_1, p_2, \ldots, p_S)$, and Δ is the quadratic distance matrix with elements d_{ij}.

To apply Q in practice, the pairwise species distances d_{ij} need to be specified. There are already several methods offered for numerically expressing pairwise interspecies distances. For example, these distances can be based on morphological or functional differences (Izsák and Papp, 1995), on phylogenetic relationships among species (Webb,

---

*Corresponding author. Fax: +3906 4457 540.

E-mail addresses: carlo.ricotta@uniroma1.it (C. Ricotta), szeidl@bmf.hu (L. Szeidl).

0040-5809/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2006.06.003

2000), on the number of nodes of a Linnaean taxonomic tree separating two species (Izsák and Papp, 1995, 2000; Warwick and Clarke, 1995, 1998; Ricotta, 2004), or on more refined molecular biological methods (Solow et al., 1993; Shimatani, 2001).

Using pairwise species distances obtained from the topology of Linnaean classifications, Warwick and Clarke (1995) demonstrated a continuous decrease in the taxonomic diversity of a marine assemblage along a gradient of increasing environmental contamination in a situation where species diversity remains constant. Shimatani (2001) examining the effects of different forest management techniques in southwest Michigan (USA) further showed that, if the pairwise species distances are obtained from taxonomic trees as proposed by Warwick and Clarke (1995), the thinning operations for promoting survival of specific species increased the taxonomic diversity of the vegetation.

Regardless of how pairwise species distances are constructed, some restrictions should be imposed on the d_{ij} elements of Δ to gain advantageous properties of Q. For instance, d_{ij} should be a metric on the set of S species, thus conforming to the following requirements: (a) symmetry ($d_{ij} = d_{ji}$), (b) non-negativity ($d_{ij} \geqslant 0$), (c) zero self-difference ($d_{ii} = 0$), and (d) triangle inequality ($d_{ij} \leqslant d_{ik} + d_{kj}$). Properties (a), (c), and (d) imply (b). Also, the distances d_{ij} should be standardized such that $0 \leqslant d_{ij} \leqslant 1$. The reason for imposing these restrictions will be clear in the following sections when dealing with measures of conflict. Notice that in relevant cases, for example in the case of resource-type differences (see Izsák and Papp, 1995), the pairwise species differences d_{ij} do not necessarily fulfill the distance axioms.

It is easily shown that, for any multivariate distance d that is metric, if $d_{ij} = 1$, Q reduces to the Simpson (1949) index $D = 1 - \sum_{j=1}^{S} p_j^2$. For a parallelism between D and Q within the framework of interspecific encounter theory, see Izsák and Szeidl (2002).

In this paper, we show that Q can be interpreted as the average conflict among the species of a given assemblage. Based on this unusual way of looking at Q, we further define a parametric version of quadratic diversity Q_α as an extension of the Tsallis (1998) generalized entropy.

2. A parametric extension of the Shannon entropy

According to Shannon (1948), the amount of statistical information associated to a given relative abundance vector p (i.e., its Shannon entropy) is measured by the quantity:

$$H = -\sum_{j=1}^{S} p_j \ln p_j. \tag{2}$$

The above equation is Shannon's fundamental expression for the description of the average information content per symbol in signals that carry a message from source to destination. Entropy is never negative; in Shannon's definition, base 2 logarithm is used to measure information content in bits, while in ecological applications the natural logarithm is traditionally used. The Shannon entropy H of a given community is basically a measure of uncertainty in predicting the relative abundance of the species that compose the relative abundance vector $p = (p_1, p_2, \ldots, p_S)$. The amount of information obtained from observing the result of an experiment depending on chance can be taken to be numerically equal to the amount of uncertainty in the outcome of the experiment before carrying it out. Therefore, Shannon's entropy can also be viewed as a measure of information (Aczél and Daróczy, 1975).

It is easily shown that the maximum value of Shannon's entropy $H^{\max} = \ln S$ is obtained in case of equiprobability (i.e., if $p_j = 1/S$ for all j). By contrast, minimum entropy is obtained if the relative abundance of a given species approaches 1 (the abundances of all other species approaching null), which implies $H^{\min} = 0$.

In statistical parlance, the quantity $(-\ln p_j)$ in (2) is sometimes referred to as the surprise $\sigma(p_j)$ associated to species j. If p_j is small, we would be quite surprised in finding species j. Accordingly, $(-\ln p_j)$ is large for small relative abundances p_j. To the contrary, if p_j is large, the corresponding surprise is small. Thus, (2) may be viewed as the mathematical expectation value of the surprise $H = \sum_{j=1}^{S} p_j \sigma(p_j)$. This way of looking at Shannon's entropy reinforces the observation that H is a measure of the uncertainty associated with the species relative abundance vector p. Since uncertainty is maximal when entropy is highest, information-theoretical measures seem an adequate choice for providing a statistical base for diversity analysis (He and Orlóci, 1993). It is therefore not surprising that Shannon's entropy appeared in early works on community structure (McArthur, 1955; Pielou, 1966) and has rapidly become the most widely used index for summarizing the uncertainty associated to a species relative abundance vector.

Almost two decades later, Patil and Taillie (1982), equating surprise $\sigma(p_j)$ with rarity $\rho(p_j)$, generalized the above observations and defined a diversity index δ as the average rarity of a given species assemblage: $\delta = \sum_{j=1}^{S} p_j \rho(p_j)$. In this view, the Simpson diversity can be rewritten as $D = \sum_{j=1}^{S} p_j (1 - p_j)$, where the quantity $(1-p_j)$ is the rarity function of D.

Patil and Taillie (1982) further defined a parametric diversity measure δ_β, known as the diversity index of degree β as $\delta_\beta = (1 - \sum_{j=1}^{S} p_j^{\beta+1})/\beta$.

Although δ_β was proposed in the ecological literature without any information-theoretical interpretation (see Patil and Taillie, 1982), nevertheless putting $\alpha = \beta + 1$, the diversity index of degree β is identical to Tsallis (1998, 2002) non-extensive entropy:

$$H_\alpha = \left(1 - \sum_{j=1}^{S} p_j^\alpha\right) \Big/ \alpha - 1,$$

where

$$\rho(p_j) = \begin{cases} (1 - p_j^{\alpha-1}) \big/ \alpha - 1 & \text{if} \quad \alpha \neq 1, \\ -\ln p_j & \text{if} \quad \alpha = 1. \end{cases} \tag{3}$$

H_α is a rediscovery of the generalized entropy of type α (Aczél and Daróczy, 1975; Tóthmérész, 1995), though with a different prefactor (fitted for base e variables instead of binary variables).

For a thorough examination of the mathematical properties of H_α, see Aczél and Daróczy (1975), Tsallis (1998, 2002). From a biological viewpoint, one particularly convenient property of H_α is that both the Shannon and the Simpson diversities are special cases of this function. For $\alpha = 2$, $H_2 = D$, while for $\alpha = 1$, H_α is defined in the limiting sense using l'Hospital's rule of calculus, and H_1 equals Shannon's entropy. Finally, for $\alpha = 0$, $H_0 = S-1$ (i.e., a linear function of species richness that sets the diversity of monospecific communities equal to zero).

A well-known property of these measures is that they have different sensitivities to the presence of rare and abundant species in the community (Hill, 1973; Ricotta, 2000). For instance, the Simpson index is less affected by the relative abundances of rare species than the species richness or the Shannon's entropy in the computation of diversity (i.e., if $p_i \gg p_j$, $p_i^2 \gg p_j^2$).

Therefore, while traditional diversity indices supply point descriptions of community structure, according to Patil and Taille (1982) definition of parametric diversity, there is a continuum of possible diversity measures that differ in their sensitivity to the rare and abundant species, becoming increasingly dominated by the commonest species for increasing values of parameter α.

For a given community, H_α is a decreasing function of α, and the various diversity measures obtained by changing α are in fact different moments of the same basic function. In this view, changing α can be considered as a scaling operation that takes place not in the real, but in the data space (Podani, 1992). This scaling operation allows for vector description of community structure that can be represented by its diversity profile of H_α vs. α (Patil and Taillie, 1982). All this is very classic.

3. Extending the meaning of quadratic diversity

To make clear the type of uncertainty measured by the Shannon entropy, it is convenient to rewrite H in the form:

$$H = -\sum_{j=1}^{S} p_j \ln\left(1 - \sum_{i \neq j}^{S} p_i\right), \tag{4}$$

where the term $C(p_j) = \sum_{i \neq j}^{S} p_i$ measures the relative abundance of all species that differ from species j. That is, the probabilities of all events that conflict with p_j.

While $C(p_j) \in [0, 1]$, the rarity function $-\ln[1 - C(p_j)]$ is monotonic increasing with $C(p_j)$ and extends its range from $[0, 1]$ to $[0, \infty]$. It follows that the Shannon entropy basically measures the expected value of the conflict among the different elements in p.

If we stay with a rigorous probabilistic approach to the measurement of uncertainty, Eq. (4) implicitly states that the elements in p fully conflict with each other. On the other hand, imagine a species assemblage in which all pairwise species distances are measured using a standardized coefficient $0 \leqslant d_{ij} \leqslant 1$. In this case, it seems natural to summarize the conflict between species j and the remaining species as $C_d(p_j) = \sum_{i \neq j}^{S} d_{ij} p_i$, such that the Shannon-like expression of average conflict becomes

$$H_d = -\sum_{j=1}^{S} p_j \ln\left(1 - \sum_{i \neq j}^{S} d_{ij} p_i\right). \tag{5}$$

Therefore, the above equation enables to quantify a more general form of the conflict among the species of a given community taking into account the pairwise species distances d_{ij}.

Notice that a more restrictive formulation of Eq. (5) with an asymmetric measure of dissimilarity d_{ij} is extensively used within the context of the Dempster–Shafer theory of evidence for deriving information-theoretical measures of uncertainty if the conflict between evidential claims is not maximal (see Klir and Wierman, 1999).

This approach can be easily generalized to virtually any diversity measure δ that is expressed as $\delta = \sum_{j=1}^{S} p_j \rho(p_j)$, as suggested by Patil and Taillie (1982).

Take for example the case of the Simpson index $D = \sum_{j=1}^{S} p_j [1 - (1 - \sum_{i \neq j}^{S} p_i)]$. Incorporating a measure of pairwise species distances d_{ij} in D, we obtain

$$D_d = \sum_{j=1}^{S} p_j \sum_{i \neq j}^{S} \left(d_{ij} p_i\right), \tag{6}$$

where D_d is identical to Rao's quadratic diversity Q. Hence, Rao's Q can be interpreted as a measure of average conflict among species. This is a very interesting outcome of Patil and Taille (1982) proposal of expressing diversity as the average rarity of a given species assemblage.

Likewise, adding pairwise species distances to Patil and Taillie parametric diversity measure H_α, we obtain

$$Q_\alpha = \frac{1 - \sum_{j=1}^{S} p_j \left(1 - \sum_{i \neq j}^{S} d_{ij} p_i\right)^{\alpha-1}}{\alpha - 1}. \tag{7}$$

The above equation represents a parametric counterpart of Rao's quadratic diversity that bridges the gap with traditional diversity measures, such as the Shannon entropy or the Simpson diversity. In this sense, expressing diversity as the mean conflict among the elements in p, it is possible to frame Rao's quadratic diversity within a larger family of measures that summarize diversity including both species relative abundances and pairwise species dissimilarities in the computation.

4. Final remarks

Regarding index properties, since Q_α is not a traditional diversity measure computed uniquely from species relative abundances, it violates the usual diversity axiom that the maximum value $Q_\alpha^{\max}$ is obtained at the uniform distribution (Shimatani, 2001). Rather, very often only few points are given non-zero abundances at $Q_\alpha^{\max}$. Another diversity axiom that is violated by Q_α is the permutation invariance (Pielou, 1975). This postulates that the diversity values corresponding to the relative abundances $p_1, p_2, \ldots, p_S$ and to a $p_1', p_2', \ldots, p_S'$ permutation of those are identical. These properties imply that the diversity measures do not depend on the relations between species and so all species play the same role in the formation of the diversity measures. In a previous paper (Ricotta, 2002), the diversity indices such as Q_α that have their greatest value for non-completely even communities were defined 'weak diversity indices'.

Finally, as suggested by Rao (1986); see also Lau (1985), a desirable property for a diversity measure is concavity, which means that the total diversity in a pooled set of communities should not be lower to the average diversity within communities. Given two relative abundance vectors p and q and two weights $0 \leqslant \mu_1 \leqslant 1$, $0 \leqslant \mu_2 \leqslant 1$, $\mu_1 + \mu_2 = 1$,

$$\delta(\mu_1 p + \mu_2 q) \geqslant \mu_1 \delta(p) + \mu_2 \delta(q), \tag{8}$$

where weights μ_1 and μ_2 associated to p and q may be equal (i.e., $\mu_1 = \mu_2 = 1/2$), or may reflect properties as diverse as community size, their conservation value, etc. In concrete terms, the concavity of δ means that diversity increases by mixing (Pavoine et al., 2004, 2005).

Patil and Taillie (1982) demonstrated the concavity of parametric diversity H_α in the interval $0 \leqslant \alpha \leqslant \infty$, while Champely and Chessel (2002) demonstrated that, if the species distance matrix Δ is Euclidean, then Rao's quadratic diversity Q is concave.

More specifically, a matrix $\Delta(S \times S)$ of distances is said to be Euclidean if and only if S points can be embedded in a Euclidean space such that the Euclidean distance between S_i and S_j is d_{ij} (Gower and Legendre, 1986).

In the following appendix, we demonstrate that Rao's quadratic diversity Q ($Q = D_d = Q_2$) is concave if and only if the matrix $H = (h_{ij})_{i,j=1}^{S-1}$, $h_{ij} = d_{iS} + d_{jS} - d_{ij}$, $1 \leqslant i,j \leqslant S-1$ is positive semidefinite. This property holds for every distance matrix for $S = 2, 3, 4$. For $S \geqslant 5$ it remains an open question. In the special case of an Euclidean distance matrix Δ, then the matrix H is positive semidefinite; this follows from the results of Champely and Chessel (2002).

More generally, for $0 < \alpha < 2$ the concavity supposedly is not fulfilled and this can be demonstrated by given calculations. Finally, we demonstrate that the diversity index Q_α, $\alpha \geqslant 2$ is concave for a special class of distance matrices Δ: if the matrix $C = (c_{ij})$, $c_{ij} = 1 - d_{ij}$ is positive semidefinite, then Q_α is concave.

Although these results may be disappointing, we do not think they should be a cause for pessimism. For instance, in a recent review on diversity indices, Ricotta (2005) showed that there exists a clear disagreement among ecologists as far as biodiversity conceptualization and evaluation is concerned. This fundamental ambiguity has ensured that the measurement of diversity remains "capricious" (Sarkar and Margules, 2002) and that even a property as general as concavity is not unanimously accepted among ecologists. For example, one very common diversity index that is not concave is the Simpson diversity in the form: $D = 1/\sum_{j=1}^{S} p_j^2$.

More generally, we think that the major result of this paper does not consist in the proposal of an additional measure of parametric diversity, but rather in the interpretation of quadratic diversity as a measure of average (distance-dependent) conflict among species. Based on this unusual definition, parametric diversity Q_α emerges as a natural way for bridging the gap between quadratic diversity and traditional information-theoretical measures such as the Shannon entropy. We hope, this connection will prove fruitful in future ecological research.

Acknowledgments

The authors wish to thank the anonymous referees for the stimulating comments that substantially helped improving this paper.

Appendix A

Let $\Delta = (d_{ij})$ be a standardized distance matrix, i.e.

$$0 \leqslant d_{ij} \leqslant 1,\ d_{ii} = 0,\ d_{ij} = d_{ji},\ d_{ij} \leqslant d_{ik} + d_{kj},\ 1 \leqslant i,j,k \leqslant S.$$

Denote by $\Pi_S = \{p = (p_1, \ldots, p_S):\ p_i \geqslant 0,\ p_1 + \cdots + p_S = 1\}$ the S-dimensional simplex and let us define the symmetric matrix $C = (c_{ij})_{i,j=1}^{S}$ with elements

$$c_{ij} = 1 - d_{ij},\ 1 \leqslant i, j \leqslant S.$$

Since $d_{ii} = 0$, $1 \leqslant i \leqslant S$, then for $0 < \alpha < \infty$, $\alpha \neq 1$ we have

$$\begin{aligned} Q_\alpha(p) &= \frac{1}{\alpha - 1}\left[1 - \sum_{j=1}^{S} p_j \left(1 - \sum_{1 \leqslant i \leqslant S,\, i \neq j} d_{ij} p_i\right)^{\alpha-1}\right] \\ &= \frac{1}{\alpha - 1}\left[1 - \sum_{j=1}^{S} p_j \left(\sum_{i=1}^{S} p_i - \sum_{i=1}^{S} d_{ij} p_i\right)^{\alpha-1}\right] \\ &= \frac{1}{\alpha - 1}\left[1 - \sum_{j=1}^{S} p_j \left(\sum_{i=1}^{S} c_{ij} p_i\right)^{\alpha-1}\right] \end{aligned}$$

and for $\alpha = 1$ we define $Q_1(p)$ as the limit

$$Q_1(p) = \lim_{\alpha \to 1} Q_\alpha(p) = -\sum_{j=1}^{S} p_j \ln\left(\sum_{i=1}^{S} c_{ij} p_i\right).$$

The function $Q_\alpha(p)$ is concave if the following inequality holds for every $p, q \in \Pi_S$

$$Q_\alpha(\lambda p + (1-\lambda)q) \leqslant \lambda Q_\alpha(p) + (1-\lambda)Q_\alpha(q), \quad 0 \leqslant \lambda \leqslant 1. \tag{A.1}$$

The function $f_\alpha(\lambda) = Q(\lambda p + (1-\lambda))$, $0 < \lambda < 1$ has continuous second derivative f''_α. Therefore the function $Q_\alpha(p)$ is concave if and only if

$$f''_\alpha(\lambda) = \frac{\partial^2}{\partial \lambda^2} Q_\alpha(\lambda p + (1-\lambda)q) \leqslant 0, \ 0 < \lambda < 1.$$

Let us introduce the notations

$$a_j = \sum_{i=1}^{S} c_{ij} p_i, \ b_j = \sum_{i=1}^{S} c_{ij} q_i.$$

We can get for $0 < \alpha < \infty$, $\alpha \neq 1$ with simple calculation

$$\begin{aligned} f''_\alpha(\lambda) = & -(\alpha-1)\sum_{j=1}^{S} 2(p_j - q_j)(a_j - b_j)(\lambda(a_j - b_j) + b_j)^{\alpha-2} \\ & -(\alpha-1)(\alpha-2)\sum_{j=1}^{S}(\lambda(p_j - q_j) + q_j)(a_j - b_j)^2 \\ & \times(\lambda(a_j - b_j) + b_j)^{\alpha-3} = -I_{1,\alpha}(\lambda) - I_{2,\alpha}(\lambda) \end{aligned} \tag{A.2}$$

and for $\alpha = 1$:

$$\begin{aligned} f''_1(\lambda) = & -\sum_{j=1}^{S} 2(p_j - q_j)(a_j - b_j)\left(\sum_{i=1}^{S} c_{ij}(\lambda(p_i - q_i)\right)^{-1} \\ & -(-1)\sum_{j=1}^{S}(\lambda(a_j - b_j) + b_j)(a_j - b_j)^2\left(\sum_{i=1}^{S} c_{ij}(\lambda(p_i - q_i)\right)^{-2} \\ = & -I_{1,1}(\lambda) - I_{2,1}(\lambda). \end{aligned} \tag{A.3}$$

I Case $0 < \alpha < 2$

With the procedure proving Theorem 2 (see below) we can verify that for a positive semidefinite matrix C and $0 < \alpha < 2$, the signs of $I_{1,\alpha}$ and $I_{2,\alpha}$ differ from each other in formulae (A.2) and (A.3). Therefore, the sign of $I_{1,\alpha} + I_{2,\alpha}$ depends on the parameters p and q. This is the reason why in the case $0 < \alpha < 2$ we cannot prove the concavity of Q_α. In this case, the concavity supposedly is not fulfilled and this can be demonstrated by given calculations.

II Case $\alpha = 2$ *(the case of Rao's quadratic diversity* $Q_2 = Q$*)*

It is easy to see that the second derivative f''_2 has the form

$$\begin{aligned} f''_2(\lambda) = & -I_{1,\alpha}(\lambda) = -2\sum_{j=1}^{S}(p_j - q_j)(a_j - b_j) \\ = & -2\sum_{j=1}^{S}\sum_{i=1}^{S}(p_j - q_j)c_{ij}(p_i - q_i) = -(p-q)'C(p-q). \end{aligned} \tag{A.4}$$

Let us introduce the symmetric matrix

$$H = (h_{ij})_{i,\,j=1}^{S-1}, \ h_{ij} = d_{iS} + d_{jS} - d_{ij}, \ 1 \leqslant i, \ j \leqslant S-1.$$

Theorem 1. *The quadratic diversity* $Q_2 = Q$ *is concave if and only if the matrix* H *is positive semidefinite.*

Proof. From formula (A.3) it is clear that relation (A.1) is equivalent to the following:

$$(p-q)'C(p-q) \geqslant 0, \ p, q \in \Pi_S \tag{A.5}$$

which is a necessary and sufficient condition for the concavity of the function $Q_2(p)$. □

First we prove the following identity for all $p, q \in \Pi_S$:

$$(p-q)'C(p-q) = \sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(p_i - q_i)h_{ij}(p_j - q_j). \tag{A.6}$$

For $p, q \in \Pi_S$ the following equations hold:

$$p_S = 1 - p_1 - \cdots - p_{S-1}, \ q_S = 1 - q_1 - \cdots - q_{S-1}$$

and

$$p_S - q_S = -\sum_{i=1}^{S-1}(p_i - q_i).$$

It is not difficult to see that ($e = (1, \ldots, 1)'$)

$$\begin{aligned} (p-q)'C(p-q) &= (p-q)'(ee' - \Delta)(p-q) \\ &= -\sum_{i=1}^{S}\sum_{j=1}^{S}(p_i - q_i)d_{ij}(p_j - q_j) \\ &= -\sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(p_i - q_i)d_{ij}(p_j - q_j) - \sum_{j=1}^{S-1}(p_S - q_S)d_{Sj}(p_j - q_j) \\ &\quad -\sum_{i=1}^{S-1}(p_i - q_i)d_{iS}(p_S - q_S) \\ &= \sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(p_i - q_i)(d_{Si} + d_{Sj} - d_{ij})(p_j - q_j) \\ &= \sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(p_i - q_i)h_{ij}(p_j - q_j) \end{aligned}$$

which leads to identity (A.6).

We prove that relation (A.5) holds if and only if the matrix H is positive semidefinite, i.e.

$$x'Hx = \sum_{i=1}^{S-1}\sum_{j=1}^{S-1} x_i h_{ij} x_j \geqslant 0, \ x = (x_1, \ldots, x_{S-1})' \in R^{S-1}. \tag{A.7}$$

At first, we suppose that condition (A.5) holds, then we prove (A.7).

Let $x = (x_1, \ldots, x_{S-1})' \in R^{S-1}$ be an arbitrarily fixed vector. Let us introduce the following notations:

$$s^+ = \max(s, 0), \ s^- = \max(-s, 0),$$

$$x^+ = (x_1^+, \dots, x_{S-1}^+), \quad x^- = (x_1^-, \dots, x_{S-1}^-),$$
$$x_0 = |x_1| + \cdots + |x_{S-1}|,$$

$$p_i = x_0^{-1} x_i^+,\ q_i = x_0^{-1} x_i^-,\ 1 \leqslant i \leqslant S-1,$$
$$p_S = 1 - p_1 - \cdots - p_{S-1},$$
$$q_S = 1 - q_1 - \cdots - q_{S-1}.$$

It is easy to verify that $p = (p_1, \dots, p_S)'$, $q = (q_1, \dots, q_S)' \in \Pi_S$ and

$$\begin{aligned}0 &\leqslant \sum_{i=1}^{S}\sum_{j=1}^{S}(p_i - q_i)c_{ij}(p_j - q_j)\\ &= \sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(p_i - q_i)h_{ij}(p_j - q_j)\\ &= \frac{1}{x_0^2}\sum_{i=1}^{S-1}\sum_{j=1}^{S-1}(x_i^+ - x_i^-)h_{ij}(x_i^+ - x_i^-)\\ &= \frac{1}{x_0^2}\sum_{i=1}^{S-1}\sum_{j=1}^{S-1}x_i h_{ij} x_j = \frac{1}{x_0^2}x'Hx.\end{aligned}$$

Conversely, if inequality (A.7) holds, we choose $x_i = p_i - q_i$, $1 \leqslant i \leqslant S-1$ and then by the use of identity (A.6), from relation (A.7) immediately follows (A.5).

Remark 1. It can be verified that for $S = 2,3,4$ the matrix H is positive semidefinite, therefore function $Q_\alpha(p)$ is concave on set Π_S.

For $S \geqslant 5$ the problem has not been solved.

Remark 2. Using the triangle inequality property of distances d_{ij}, we can observe that the elements of matrix H satisfy the following inequalities:

$$h_{ij} \geqslant 0,\ h_{ij} \leqslant h_{ii},\ h_{ij} \leqslant h_{jj},\ 1 \leqslant i,j \leqslant S-1, \tag{A.8}$$

i.e. H is a similarity matrix. Unfortunately, the positive semidefinite property is not guaranteed by inequalities (A.8), as the following example shows (Galántai, pers. commun.).

Let us define matrix H_0 as follows:

$$H_0 = \begin{pmatrix} 1.4501 & 0.4966 & 0.6111 & 0.4458 & 0.4746\\ 0.4966 & 0.9565 & 0.4052 & 0.8784 & 0.3988\\ 0.6111 & 0.4052 & 1.4218 & 0.8276 & 0.4947\\ 0.4458 & 0.8784 & 0.8276 & 0.9103 & 0.4518\\ 0.4746 & 0.3988 & 0.4947 & 0.4518 & 0.6389 \end{pmatrix}.$$

The eigenvalues of matrix H_0 are $\mathrm{eig}(H_0) = (–0.0469, 0.3297, 0.8019, 0.9521, 3.3408)$, therefore the matrix is not positive definite.

III Case $\alpha \geqslant 2$.

Theorem 2. *If matrix* C *is positive semidefinite* (*i.e.* $x'C_x \geqslant 0$ *for all* $x \in \mathbf{R}^S$,), *then the diversity index* $Q_\alpha(p)$ *is concave.*

Proof. At first we note that in case $\alpha = 2$ the statement of Theorem 2 is an immediate consequence of inequality (A.4) and the positive semidefinite property of matrix C. □

From formula (A.2) it can be seen that $I_{2,\alpha}(\lambda)$ is non-negative because all members in the expression $I_{2,\alpha}(\lambda)$ are non-negative. Thus, for inequality $f''_\alpha(p) \leqslant 0$ it is enough to prove inequality $I_{1,\alpha}(\lambda) \geqslant 0$.

It is clear that

$$\begin{aligned}I_{1,\alpha}(\lambda) &= 2(\alpha - 1)\sum_{j=1}^{S}(p_j - q_j)(a_j - b_j)\rho_j\\ &= 2(\alpha - 1)\sum_{j=1}^{S}\sum_{i=1}^{S}(p_j - q_j)\rho_j c_{ij}(p_i - q_i)\\ &= 2(\alpha - 1)(p - q)'RC(p - q),\end{aligned}$$

where

$$\rho_j = (\lambda(a_j - b_j) + b_j)^{a-2},\ 1 \leqslant j \leqslant S \quad \text{and}$$
$$R = \mathrm{diag}(\rho_1, \dots, \rho_S).$$

The matrices C and R are positive semidefinite, therefore the product matrix RC is also positive semidefinite. From this it follows that $I_{1,\alpha}(\lambda) \geqslant 0$.

References

Aczél, J., Daróczy, Z., 1975. On Measures of Information and their Characterizations. Academic Press, New York.

Champely, S., Chessel, D., 2002. Measuring biological diversity using Euclidean metrics. Environ. Ecolog. Stat. 9, 167–177.

Clarke, K.R., Warwick, R.M., 1998. A taxonomic distinctness index and its statistical properties. J Appl. Ecol. 35, 523–531.

Gower, J.C., Legendre, P., 1986. Metric and Euclidean properties of dissimilarity coefficients. J. Classif. 3, 5–48.

He, X.S., Orlóci, L., 1993. Comparative diversity analysis of vegetation. Abstr. Bot. 17, 79–86.

Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54, 427–431.

Izsák, J., Papp, L., 1995. Application of the quadratic entropy index for diversity studies on drosophilid species assemblages. Environ. Ecol. Stat. 2, 213–224.

Izsák, J., Papp, L., 2000. A link between ecological diversity indices and measures of biodiversity. Ecol. Model. 130, 151–156.

Izsák, J., Szeidl, L., 2002. Quadratic diversity: its maximization can reduce the richness of species. Environ. Ecolog. Stat. 9, 423–430.

Klir, G.J., Wierman, M.J., 1999. Uncertainty-Based Information. Physica-Verlag, Heidelberg.

Lau, K.S., 1985. Characterization of Rao's quadratic entropies. Sankhya Ser. A 47, 295–309.

McArthur, R.H., 1955. Fluctuation of animal populations and a measure of community stability. Ecology 36, 533–536.

Patil, G.P., Taillie, C., 1982. Diversity as a concept and its measurement. J. Am. Stat. Assoc. 77, 548–567.

Pavoine, S., Dufour, A.B., Chessel, D., 2004. From dissimilarities among species to dissimilarities among communities: a double principal coordinate analysis. J. Theor. Biol. 228, 523–537.

Pavoine, S., Ollier, S., Pontier, D., 2005. Measuring diversity from dissimilarities with Rao's quadratic entropy: Are any dissimilarities suitable? Theor. Popul. Biol. 67, 231–239.

Pielou, E.C., 1966. Shannon's formula as a measure of species diversity: its use and misuse. Am. Nat. 100, 463–465.

Pielou, E.C., 1975. Ecological Diversity. Wiley Interscience, New York.

Podani, J., 1992. Space series analysis: processes reconsidered. Abstr. Bot. 16, 25–29.

Rao, C.R., 1982. Diversity and dissimilarity coefficients: a unified approach. Theor. Popul. Biol. 21, 24–43.

Rao, C.R., 1986. Rao's axiomatization of diversity measures. In: Kotz, S., Johnson, N.L. (Eds.), Encyclopedia of Statistical Sciences. Wiley, New York, pp. 614–617.

Ricotta, C., 2000. From theoretical ecology to statistical physics and back: self-similar landscape metrics as a synthesis of ecological diversity and geometrical complexity. Ecol. Model 125, 245–253.

Ricotta, C., 2002. Bridging the gap between ecological diversity indices and measures of biodiversity with Shannon's entropy: comment to Izsák and Papp. Ecol. Model. 152, 1–3.

Ricotta, C., 2004. A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. Divers. Distrib. 10, 143–146.

Ricotta, C., 2005. Through the jungle of biological diversity. Acta Biotheor. 53, 29–38.

Sarkar, S., Margules, C., 2002. Operationalizing biodiversity for conservation planning. J. Biosci. 27 (Suppl. 2), 299–308.

Shannon, C., 1948. A mathematical theory of communication. Bell Syst. Techn. J. 27, 379–423.

Shimatani, K., 2001. On the measurement of species diversity incorporating species differences. Oikos 93, 135–147.

Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.

Solow, A.R., Polasky, S., Brodaus, J., 1993. On the measurement of biological diversity. J. Environ. Econ. Manage. 24, 60–68.

Tóthmérész, B., 1995. Comparison of different methods for diversity ordering. J. Veg. Sci. 6, 283–290.

Tsallis, C., 1998. Possible generalization of Boltzmann–Gibbs statistics. J. Stat. Phys. 52, 479–487.

Tsallis, C., 2002. Entropic nonextensivity: a possible measure of complexity. Chaos Solitons Fract. 13, 371–391.

Warwick, R.M., Clarke, K.R., 1995. New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. Mar. Ecol. Prog. Ser. 129, 301–305.

Warwick, R.M., Clarke, K.R., 1998. Taxonomic distinctness and environmental assessment. J. Appl. Ecol. 35, 532–543.

Webb, C.O., 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. Am. Nat. 156, 145–155.