Part III

Multi-way metrics

CHAPTER 11

Axiom systems for two-way, three-way and multi-way dissimilarities

Dissimilarities are functions that are used with various multivariate data analysis techniques. Well-known examples are multidimensional scaling and cluster analysis. A function is called a dissimilarity if it satisfies certain axioms, that is, it is nonnegative and symmetric, and it satisfies the axiom of minimality. In addition, a dissimilarity may satisfy axioms like the triangle inequality or the ultrametric inequality. Dependencies between certain axioms have been noted by various authors (see, for example, Gower and Legendre (1986), Van Cutsem (1994) or Batagelj and Bren (1995) for the two-way case, and Joly and Le Calvé (1995), Bennani-Dosse (1993) and Heiser and Bennani (1997) for the three-way case).

Although many authors (including the above-mentioned) point out that the used set of axioms do not form a system with a minimum number of axioms (due to dependencies between axioms), it remains (sometimes) unclear what this minimum set looks like. An axiom system can be a minimum set of axioms if it forms an independent system of axioms. Within an axiom system an axiom is called independent if it cannot be derived from the other axioms in the system. Another (perhaps more) important property of an axiom system is consistency. An axiom system is consistent if it lacks contradiction, that is, the ability to derive both a statement and its negation from a set of axioms.

In this chapter the axiom systems for two-way and three-way dissimilarities are studied. Some axioms for two-way dissimilarities were briefly considered in Section 1.2 and Section 10.1. To obtain axiom systems with a minimum number of axioms, the (known) dependencies between various axioms are reviewed. Next, consistency and independence of several axiom systems are established by means of simple models. The remainder of the chapter is used to explore how basic axioms for multi-way dissimilarities, like nonnegativity, minimality and symmetry, may be defined. Generalizations of the two-way metric and the three-way metrics are further studied in Chapter 12. Multi-way extensions of the three-way ultrametric inequalities are investigated in Chapter 13. Using the tools for the axioms for three-way dissimilarities, independence and consistency may be established for the multi-way case.

11.1 Two-way dissimilarities

Let the function $d(x_1, x_2) : E \times E \to \mathbb{R}$ assign a real number to each pair (x_1, x_2), elements of the nonempty set E. The function $d(x_1, x_2)$ is called a two-way dissimilarity between objects x_1 and x_2 if it satisfies the axioms

$$
\begin{array}{lll}
(A1) & d(x_1, x_2) \geq 0 & \text{(nonnegativity)} \\
(A2) & d(x_1, x_1) = 0 & \text{(minimality)} \\
(A3) & d(x_1, x_2) = d(x_2, x_1) & \text{(symmetry).}
\end{array}
$$

In the French literature, a dissimilarity $d(x_1, x_2)$ is called respectively semi-proper and proper if it satisfies

$$
\begin{array}{lll}
(A4) & d(x_1, x_2) = 0 \ \Rightarrow \ d(x_1, x_3) = d(x_2, x_3) & \text{(evenness)} \\
(A5) & d(x_1, x_2) = 0 \ \Rightarrow \ x_1 = x_2 & \text{(definiteness).}
\end{array}
$$

Let

$$p_{123}^{111} = P\left(\overset{1}{x}_1, \overset{1}{x}_2, \overset{1}{x}_3\right)$$

denote the proportion of 1s shared by variables x_1, x_2 and x_3 in the same positions, let

$$p_{123}^{110} = P\left(\overset{1}{x}_1, \overset{1}{x}_2, \overset{0}{x}_3\right)$$

denote the proportion of 1s shared by variables x_1 and x_2, and 0s by variable x_3 in the same positions, and let

$$p_1^1 = P\left(\overset{1}{x}_1\right)$$

denote the proportion of 1s in variable x_1. For example, it holds that

$$p_1^1 = p_{12}^{10} + p_{12}^{11} \quad \text{and} \quad p_{12}^{10} = p_{123}^{100} + p_{123}^{101}.$$

Proposition 11.1. $(A1)$, $(A2)$, $(A3)$ and $(A4)$ *form a consistent and independent system of axioms.* $(A1)$, $(A2)$, $(A3)$ and $(A5)$ *form a consistent and independent system of axioms.*
Proof: First, note that $(A5) \Rightarrow (A4)$. Consistency of the two axiom systems is established by the first example of $d(x_1, x_2)$ in the table below. The independence of $(A1)$, $(A2)$ and $(A3)$ with respect to the remaining four axioms is established with the bottom three examples of $d(x_1, x_2)$ in the table below.

	Is the axiom valid?				
$d(x_1, x_2)$	$(A1)$	$(A2)$	$(A3)$	$(A4)$	$(A5)$
$p_1^1 + p_2^1 - 2p_{12}^{11}$	Yes	Yes	Yes	Yes	Yes
$2p_{12}^{11} - p_1^1 - p_2^1$	No	Yes	Yes	Yes	Yes
$p_1^1 + p_2^1 - p_{12}^{11}$	Yes	No	Yes	Yes	Yes
$2p_1^1 + p_2^1 - 3p_{12}^{11}$	Yes	Yes	No	Yes	Yes

Next, consider the function $d(x_1, x_2) = \min(p_1^1, p_2^1) - p_{12}^{11}$. It is readily verified that $d(x_1, x_2)$ satisfies $(A1)$, $(A2)$ and $(A3)$. However, $(A4)$ and $(A5)$ are not valid if there is a pair (x_1, x_2) for which $p_{12}^{11} = \min(p_1^1, p_2^1)$. $\square$

A two-way dissimilarity $d(x_1, x_2)$ is called a distance if it satisfies definiteness and

$$(A6) \quad d(x_1, x_2) \ \leq \ d(x_1, x_3) + d(x_2, x_3) \quad \text{(triangle inequality).}$$

A dissimilarity may also satisfy one of two axioms that define properties of trees, that is, an inequality by Buneman (1974)

$$(A7) \quad d(x_1, x_2) + d(x_3, x_4) \ \leq \ \max[d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)]$$

(additive tree) or

$$(A8) \quad d(x_1, x_2) \ \leq \ \max[d(x_1, x_3), d(x_2, x_3)] \quad \text{(ultrametric inequality).}$$

Proposition 11.2.

(i) $(A6)$ *together with* $(A2) \Rightarrow (A1)$, $(A3)$ *and* $(A4)$
(ii) $(A7)$ *together with* $(A2) \Rightarrow (A1)$, $(A3)$, $(A4)$ *and* $(A6)$
(iii) $(A8)$ *together with* $(A2) \Rightarrow (A1)$, $(A3)$, $(A4)$ *and* $(A6)$.

Proof: The proof of (i) can be found in Gower and Legendre (1986, p. 6). For (ii) setting x_3 equal to x_4 in $(A7)$ and applying $(A2)$, we obtain $(A6)$. For (iii), for triplet (x_1, x_1, x_2) we obtain $d(x_1, x_2) \geq 0$, that is (A1). Moreover, $(A8)$ together with $(A1) \Rightarrow (A6)$. $\square$

Proposition 11.3. *(A2), (A5) and (A6) (or (A7) or (A8)) form a consistent and independent system of axioms.*

Proof: Consider the assertion with respect to $(A6)$ first. An example for consistency is the function given by

$$d(x_1, x_2) = 1 - p_{12}^{11} - p_{12}^{00}.$$

Validity of $(A2)$ and $(A5)$ is readily verified. Using $d(x_1, x_2)$ in $(A6)$ we obtain

$$1 + p_{12}^{11} + p_{12}^{00} \geq p_{13}^{11} + p_{13}^{00} + p_{23}^{11} + p_{23}^{00} \qquad \text{if and only if} \qquad 2p_{123}^{110} + 2p_{123}^{001} \geq 0.$$

With respect to independence, consider the function $d(x_1, x_2) = 1 - p_{12}^{11}$. Using $d(x_1, x_2)$ in $(A6)$ we obtain

$$1 + p_{12}^{11} \geq p_{13}^{11} + p_{23}^{11} \qquad \text{if and only if} \qquad p_{123}^{000} + p_{123}^{100} + p_{123}^{010} + p_{123}^{001} + 2p_{123}^{110} \geq 0.$$

Hence, $d(x_1, x_2)$ satisfies $(A6)$. Moreover, axiom $(A5)$ is not violated. However, as long as $p_1^1 \neq 1$, $d(x_1, x_2)$ does not satisfy $(A2)$. Hence, $(A2)$ is independent from $(A5)$ and $(A6)$.

Second, consider the function $d(x_1, x_2) = \min(p_1^1, p_2^1) - p_{12}^{11}$. Axiom $(A2)$ is valid. Assuming $p_1^1 \geq p_2^1 \geq p_3^1$ and Using $d(x_1, x_2)$ in $(A6)$, we obtain

$$2p_3^1 + p_{12}^{11} \geq p_1^1 + p_{13}^{11} + p_{23}^{11} \qquad \text{if and only if} \qquad 2p_{123}^{001} + p_{123}^{101} \geq p_{123}^{010}.$$

Furthermore, $(A5)$ is not valid if $p_{12}^{11} = \min(p_1^1, p_2^1) = p_2^1$ if and only if p_{12}^{01} equals 0. Thus, $(A2)$ and $(A6)$ may be valid, while $(A5)$ is not.

Third, consider the function $d(x_1, x_2) = 2p_{12}^{11} - p_1^1 - p_2^1$. It is readily verified that for this function $(A2)$ and $(A5)$ are valid. However, $(A6)$ is only valid if $p_{123}^{110} + p_{123}^{001} \leq 0$ if and only if $p_{123}^{110} = p_{123}^{001} = 0$, since p_{123}^{110} and p_{123}^{001} are nonnegative quantities.

The proofs of the assertion with respect to $(A7)$ and $(A8)$ are very similar to that of $(A6)$. Furthermore, suppose $d(x_1, x_2)$ satisfies $(A8)$. Then for the three two-way dissimilarities defined on the same three objects, the largest two are equal. This property is unrelated to the value of $d(x_1, x_2)$. □

11.2 Three-way dissimilarities

Axioms for three-way dissimilarities and distances can be found in Bennani-Dosse (1993), Heiser and Bennani (1997) and Chepoi and Fichet (2007). In addition, three-way distances are considered in Joly and Le Calvé (1995). Let $d_3(x_1, x_2, x_3) : E \times E \times E \to \mathbb{R}$ be a function that assigns a real number to each triplet (x_1, x_2, x_3). Heiser and Bennani (1997, p. 191) call $d_3(x_1, x_2, x_3)$ a three-way dissimilarity if it satisfies the axioms

$(B1a)$	$d_3(x_1, x_2, x_3) \geq 0$	(nonnegativity)
$(B2a)$	$d_3(x_1, x_1, x_1) = 0$	(minimality)
$(B3)$	$d_3(x_1, x_2, x_3) = d_3(x_1, x_3, x_2) = d_3(x_2, x_1, x_3) =$ $d_3(x_2, x_3, x_1) = d_3(x_3, x_1, x_2) = d_3(x_3, x_2, x_1)$	(symmetry),

the three-way generalizations of $(A1)$, $(A2)$ and $(A3)$, and in addition

$$d_3(x_1, x_1, x_2) = d_3(x_1, x_2, x_2). \tag{11.1}$$

Equality (11.1) is referred to as the diagonal-plane equality by Heiser and Bennani (1997), and is also proposed in Joly and Le Calvé (1995).

Equality (11.1) is an answer to a complication that arises with three-way dissimilarities, not encountered with two-way dissimilarities, when one of three variables or entities is identical to one of the others. For this reason, Chepoi and Fichet (2007) studied explicitly the case of three-way dissimilarities for which all entities are different. The lack of resemblance between the two nonidentical entities should, according to Heiser and Bennani (1997), remain invariant regardless of which two entities are the same:

$$d_3(x_1, x_1, x_2) = d_3(x_1, x_2, x_2) = d_3(x_1, x_2, x_1) =$$
$$d_3(x_2, x_1, x_1) = d_3(x_2, x_1, x_2) = d_3(x_2, x_2, x_1).$$

Equality (11.1) is referred to as the diagonal-plane equality in Heiser and Bennani (1997), because it requires equality of the three matrices

$$\{d_3(x_1, x_1, x_2)\}\,,\ \{d_3(x_1, x_2, x_2)\}\ \text{and}\ \{d_3(x_1, x_2, x_1)\}$$

which are formed by cutting the three-way cube or block diagonally, starting at one of the three edges joining at the node or corner $d(1, 1, 1)$. This seems to be a misnomer, since equality (11.1) only requires equality of the first two matrices. Equality (11.1) together with three-way symmetry $(B3)$ implies the stronger equality

$$(B4) \quad d_3(x_1, x_1, x_2) = d_3(x_1, x_2, x_2) = d_3(x_1, x_2, x_1).$$

Proposition 11.4. $(B1a)$, $(B2a)$, $(B3)$ *and* $(B4)$ *form a consistent and independent system of axioms.*
Proof: Consistency of the axiom system is shown with the first example of $d_3(x_1, x_2, x_3)$ in the table below.

	Is the axiom valid?			
$d_3(x_1, x_2, x_3)$	$(B1a)$	$(B2a)$	$(B3)$	$(B4)$
$1 - p_{123}^{111} - p_{123}^{000}$	Yes	Yes	Yes	Yes
$p_{123}^{111} + p_{123}^{000} - 1$	No	Yes	Yes	Yes
$1 - p_{123}^{111}$	Yes	No	Yes	Yes
$p_1^1 - p_{123}^{111}$	Yes	Yes	No	Yes
$p_1^1 + p_2^1 + p_3^1 - 3p_{123}^{111}$	Yes	Yes	Yes	No

Independence is established with the bottom four examples of $d_3(x_1, x_2, x_3)$ in the table. Each function satisfies three out of four axioms. □

At this point it should be noted that there exists mathematical literature on multi-way concepts, including distances and metrics, that is older that the above mentioned literature. Some of the references from this literature may be found in Deza and Rosenberg (2000, 2005). Characteristic of this literature are the extensions of axioms $(A1)$ and $(A2)$ given by

$$(B1b) \qquad x_1 \neq x_2 \Rightarrow d_3(x_1,x_2,x_3) > 0 \text{ for some } x_3 \in E$$
$$(B2b) \qquad d_3(x_1,x_1,x_2) = 0$$

and axiom $(B6c)$ presented below. Axiom $(B2b)$ makes perfect sense in geometry where $d_3(x_1,x_1,x_2)$ is, for example, the area of the triangle with vertices x_1, x_2, and x_3. Deza and Rosenberg (2000, 2005) find axioms $(B1b)$ and $(B2b)$ too restrictive and drop them. The two axioms are also ignored in this chapter.

A three-way dissimilarity $d_3(x_1,x_2,x_3)$ is called a three-way distance in Heiser and Bennani (1997, p. 191) if it satisfies

$$(B5) \qquad d_3(x_1,x_2,x_3) = 0 \quad \Rightarrow \quad x_1 = x_2 = x_3 \qquad \text{(definiteness)}$$

and the so-called tetrahedral inequality

$$(B6a) \qquad 2d_3(x_1,x_2,x_3) \leq d_3(x_2,x_3,x_4) + d_3(x_1,x_3,x_4) + d_3(x_1,x_2,x_4).$$

Alternatively, Joly and Le Calvé (1995) call $d(x_1,x_2,x_3)$ a three-way distance if it satisfies

$$(B6b) \qquad d_3(x_1,x_2,x_3) \leq d_3(x_2,x_3,x_4) + d_3(x_1,x_3,x_4)$$
$$(B7) \qquad d_3(x_1,x_2,x_3) \geq d_3(x_1,x_1,x_3)$$

and a proper three-way distance if it, in addition, satisfies $(B5)$. Axioms $(B6a)$ and $(B6b)$ are called respectively strong and weak metrics in Chepoi and Fichet (2007). Deza and Rosenberg (2000, 2005) present yet another extension of the triangle inequality. The so-called tetrahedron inequality is given by

$$(B6c) \quad d_3(x_1,x_2,x_3) \leq d_3(x_2,x_3,x_4) + d_3(x_1,x_3,x_4) + d_3(x_1,x_2,x_4).$$

Axiom $(B6c)$ is not studied further in this chapter (but see Chapter 12).

Three-way generalizations of two-way ultrametric inequality $(A8)$ are considered in Joly and Le Calvé (1995, p. 195) and Bennani-Dosse (1993, p. 99-110):

$$(B8a) \qquad d_3(x_1,x_2,x_3) \leq \max\left[d_3(x_2,x_3,x_4), d_3(x_1,x_3,x_4)\right]$$
$$(B8b) \qquad d_3(x_1,x_2,x_3) \leq \max\left[d_3(x_2,x_3,x_4), d_3(x_1,x_3,x_4), d_3(x_1,x_2,x_4)\right].$$

Axioms $(B8a)$ and $(B8a)$ are called respectively strong and weak ultrametrics in Chepoi and Fichet (2007).

As noted in Bennani-Dosse (1993, p. 20), the dependencies between $(B1)$ to $(B8)$ are not as straightforward as the dependencies between $(A1)$ to $(A8)$ given in Proposition 11.2.

Proposition 11.5.

$$\begin{array}{ll} & (B6b) \text{ together with } (B7) \text{ and } (B2a) \Rightarrow (B1a) \\ (i) & (B6b) \text{ together with } (B3) \Rightarrow (B1a) \\ & (B6a) \text{ together with } (B3) \Rightarrow (B1a) \text{ and } (B6b) \\ & (B7) \text{ together with } (B3) \Rightarrow (B4) \\ (ii) & (B8a) \Rightarrow (B6a),\ (B7) \text{ and } (B8b). \end{array}$$

The proofs for (i) and (ii) are presented below. The proofs of the other assertions can be found in Joly and Le Calvé (1995, p. 193) and Heiser and Bennani (1997, p. 192).

Proof: For (i), adding the two variants of $(B6b)$

$$\begin{array}{rl} & d_3(x_1,x_2,x_3) \leq d_3(x_2,x_3,x_4) + d_3(x_1,x_3,x_4) \\ \text{and} & d_3(x_2,x_3,x_4) \leq d_3(x_1,x_2,x_3) + d_3(x_1,x_3,x_4) \end{array}$$

we obtain $2d_3(x_1,x_3,x_4) \geq 0$. With respect to (ii), note that, if $d(x_1,x_2,x_3)$ satisfies $(B8a)$, then for any four three-way dissimilarities the largest three are equal. □

The dependencies in Proposition 11.5 suggest the independence of various axiom systems. First, we consider a system of structural, that is, non-metric axioms.

Proposition 11.6. *$(B1a)$, $(B2a)$, $(B3)$, $(B5)$ and $(B7)$ form a consistent and independent system of axioms.*

Proof: An example of consistency of the axiom system is the function $d_3(x_1,x_2,x_3) = 1 - p_{123}^{111} - p_{123}^{000}$. It is readily verified that $(B1a)$, $(B2a)$, $(B3)$ and $(B5)$ are valid. Using $d_3(x_1,x_2,x_3)$ in $(B7)$ we obtain

$$p_{13}^{11} + p_{13}^{00} \geq p_{123}^{111} + p_{123}^{000} \quad \text{if and only if} \quad p_{123}^{101} + p_{123}^{010} \geq 0.$$

With respect to independence, consider the function $d_3(x_1,x_2,x_3) = 3p_{123}^{111} - p_1^1 - p_2^1 - p_3^1$. Axioms $(B2a)$, $(B3)$ and $(B5)$ are valid, but $(B1a)$ is not. Using the function in $(B7)$ we obtain

$$\begin{aligned} 3p_{123}^{111} + p_1^1 &\geq 3p_{13}^{11} + p_3^1 \\ p_{123}^{100} + p_{123}^{110} &\geq 3p_{123}^{101} + p_{123}^{001} + p_{123}^{011} \\ p_{13}^{10} &\geq 3p_{123}^{101} + p_{13}^{01}. \end{aligned}$$

Thus, $(B1a)$ is independent from $(B2a)$, $(B3)$, $(B5)$ and $(B7)$.

Second, consider the function $d_3(x_1, x_2, x_3) = p_1^1 + p_2^1 + p_3^1 - 2p_{123}^{111}$. Axioms $(B1a)$, $(B3)$ and $(B5)$ are valid, but $(B2a)$ is not. The function satisfies $(B7)$ if and only if $p_{12}^{01} + 2p_{123}^{101} \geq p_{12}^{10}$. Thus, axiom $(B2a)$ is independent from $(B1a)$, $(B3)$, $(B5)$ and $(B7)$.

Third, consider the function $d_3(x_1, x_2, x_3) = 2p_1^1 + p_2^1 + p_3^1 - 4p_{123}^{111}$. Axioms $(B1a)$, $(B2a)$ and $(B5)$ are valid, but $(B3)$ is not. The function satisfies $(B7)$ if and only if $p_{12}^{01} + 4p_{123}^{101} \geq p_{12}^{10}$, which shows that $(B3)$ is independent from the remaining four axioms.

Next, consider the function

$$d_3(x_1, x_2, x_3) = \min(p_{12}^{11}, p_{13}^{11}, p_{23}^{11}) - p_{123}^{111}.$$

It is readily verified that $(B1a)$, $(B2a)$, $(B3)$ and $(B7)$ are valid. However, if there is a triple (x_1, x_2, x_3) for which $p_{123}^{111} = \min(p_{12}^{11}, p_{13}^{11}, p_{23}^{11})$, then $(B5)$ does not hold.

Finally, consider the function $d_3(x_1, x_2, x_3) = p_1^1 + p_2^1 + p_3^1 - 3p_{123}^{111}$. It is readily verified that $(B1a)$, $(B2a)$, $(B3)$ and $(B5)$ are valid. Furthermore, we have $d_3(x_1, x_2, x_3) \leq d_3(x_1, x_1, x_2)$ if and only if $p_{12}^{01} + 3p_{123}^{101} \leq p_{12}^{10}$, which show the independence of $(B7)$ with respect to the remaining four axioms. □

Finally, we consider an axiom system with a minimum number of axioms.

Proposition 11.7. $(B2a)$, $(B3)$, $(B5)$, $(B6a)$ *and* $(B7)$ *form a consistent and independent system of axioms.*
Proof: An example for the consistency of the axiom system is the function $d_3(x_1, x_2, x_3) = 1 - p_{123}^{111} - p_{123}^{000}$. It is readily verified that $(B2a)$, $(B3)$, $(B5)$ and $(B7)$ are valid. Using $d_3(x_1, x_2, x_3)$ in $(B6a)$ we obtain

$$1 - (p_{234}^{111} + p_{134}^{111} + p_{124}^{111} + p_{234}^{000} + p_{134}^{000} + p_{124}^{000}) + 2p_{123}^{111} + 2p_{123}^{000} \geq 0. \tag{11.2}$$

Since the quantity in between brackets in (11.2) is smaller than unity, $(B6a)$ is valid.

With respect to independence, consider the function $d_3(x_1, x_2, x_3) = p_1^1 + p_2^1 + p_3^1 - 2p_{123}^{111}$. Axioms $(B3)$ and $(B5)$ are valid, and $(B2a)$ is not. Using the function in $(B6a)$ we obtain

$$3p_4^1 + 4p_{123}^{111} \geq p_{234}^{111} + p_{134}^{111} + p_{124}^{111}$$

which holds if and only if

$$3p_{1234}^{0001} + 3p_{1234}^{1001} + 3p_{1234}^{0101} + 3p_{1234}^{0011} + p_{1234}^{1101} + p_{1234}^{1011} + p_{1234}^{0111} + p_{1234}^{1111} + 4p_{1234}^{1110} \geq 0.$$

Furthermore, axiom $(B7)$ is valid if and only if

$$p_2^1 + 2p_{12}^{11} \geq p_1^1 + 2p_{123}^{111} \quad \text{if and only if} \quad p_{12}^{01} + 2p_{123}^{110} \geq p_{12}^{10}.$$

Thus, $(B2a)$ is independent from the remaining four axioms.

Second, consider the function $d_3(x_1,x_2,x_3) = 2p_1^1 + p_2^1 + p_3^1 - 4p_{123}^{111}$. Axioms ($B2a$), ($B5$) and ($B7$) are valid, but ($B3$) is not. Using the function in ($B6a$), we obtain the inequality

$$p_2^1 + 3p_4^1 + 8p_{123}^{111} \geq 4p_{234}^{111} + 4p_{134}^{111} + 4p_{124}^{111}$$

which holds if and only

$$p_{1234}^{0100} + p_{1234}^{1100} + p_{1234}^{0110} + 4p_{1234}^{0101} + 8p_{1234}^{1110} + 3p_{1234}^{0001} + 3p_{1234}^{1001} + 3p_{1234}^{0011} \geq p_{1234}^{1011}$$

which shows that ($B3$) is independent from the remaining four axioms.

Third, consider the function

$$d(x_1,x_2,x_3) = \min(p_{12}^{11}, p_{13}^{11}, p_{23}^{11}) - p_{123}^{111}$$

Axioms ($B2a$), ($B3$) and ($B7$) are valid. Assuming $p_{12}^{11} \geq p_{13}^{11} \geq p_{14}^{11} \geq p_{23}^{11} \geq p_{24}^{11} \geq p_{34}^{11}$ and Using $d(x_1,x_2,x_3)$ in ($B6a$), we obtain

$$2p_{34}^{11} + p_{24}^{11} + 2p_{123}^{111} \geq 2p_{23}^{11} + p_{234}^{111} + p_{134}^{111} + p_{124}^{111}$$

if and only if

$$2p_{1234}^{0011} + p_{1234}^{1011} + p_{1234}^{0101} \geq 2p_{1234}^{0110}.$$

Note that axiom ($B5$) is not valid if $p_{123}^{111} = \min(p_{12}^{11}, p_{13}^{11}, p_{23}^{11}) = p_{23}^{11}$ if and only if $p_{123}^{011} = 0$. The latter implies that $p_{1234}^{0110} = 0$, from which it follows that ($B6a$) holds. Thus, ($B5$) is independent from the remaining four axioms.

Next, consider the function $d_3(x_1,x_2,x_3) = 3p_{123}^{111} - p_1^1 - p_2^1 - p_3^1$. Axioms ($B2a$), ($B3$) and ($B5$) are valid for both $d_3(x_1,x_2,x_3)$ and $-d_3(x_1,x_2,x_3)$. Axiom ($B6a$) is valid for $-d_3(x_1,x_2,x_3)$, since filling in $-d_3(x_1,x_2,x_3)$ in ($B6a$) gives

$$p_4^1 + 2p_{123}^{111} \geq p_{234}^{111} + p_{134}^{111} + p_{124}^{111}$$

if and only if

$$2p_{1234}^{1110} + p_{1234}^{0001} + p_{1234}^{1001} + p_{1234}^{0101} + p_{1234}^{0011} \geq 0.$$

Using similar arguments it is clear that ($B6a$) is not valid for $d_3(x_1,x_2,x_3)$. Finally, ($A7$) is valid for $d_3(x_1,x_2,x_3)$ not valid for $-d_3(x_1,x_2,x_3)$ if and only if $p_{12}^{01} + 2p_{123}^{101} \leq p_{123}^{100}$. Hence, ($B6a$) and ($B7$) are independent from the remaining four axioms. □

11.3 Multi-way dissimilarities

In this final section it is explored how basic axioms for multi-way dissimilarities, like nonnegativity, minimality and symmetry, may be defined. However, axioms for the four-way and five-way case are considered first. Generalizations of the two-way metric and the three-way metrics to k-way metrics are further studied in the next chapter (Chapter 12). Multi-way formulations of the three-way ultrametrics are explored in Chapter 13. Independence and consistency of axioms for multi-way dissimilarities may be established using the tools from the previous section.

As it turns out, definitions of some axioms are considerably more complicated in the four-way case compared to the three-way case. Let

$$d_4(x_1, x_2, x_3, x_4) : E^4 \to \mathbb{R} \quad \text{or} \quad d_{1234} : E^4 \to \mathbb{R}$$

be a function that assigns a real number to each quadruplet (x_1, x_2, x_3, x_4). Formulations of nonnegativity and minimality are straightforward:

$$(C1) \quad d_4(x_1, x_2, x_3, x_4) \geq 0 \qquad \text{(nonnegativity)}$$
$$(C2) \quad d_4(x_1, x_1, x_1, x_1) = 0 \qquad \text{(minimality).}$$

The definition of four-way symmetry is somewhat more involved. Four-way symmetry is given by

$$\begin{aligned} d_{1234} &= d_{1243} = d_{1324} = d_{1342} = d_{1423} = d_{1432} = \\ d_{2134} &= d_{2143} = d_{2314} = d_{2341} = d_{2413} = d_{2431} = \\ d_{3124} &= d_{3142} = d_{3214} = d_{3241} = d_{3412} = d_{3421} = \\ d_{4123} &= d_{4132} = d_{4213} = d_{4231} = d_{4312} = d_{4321}. \end{aligned}$$

If $d_4(x_1, x_2, x_3, x_4)$ is four-way symmetric, then for all $x_1, x_2, x_3, x_4 \in E$ and every permutation π of $\{1, 2, 3, 4\}$

$$(C3) \qquad d_4(x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)}, x_{\pi(4)}) = d_4(x_1, x_2, x_3, x_4).$$

Similar to the three-way case, the four-way function can be defined on a quadruplet or four-tuple of which some entities are identical. Following the reasoning in Heiser and Bennani (1997), it seems reasonable to require that when one of four variables or entities is identical to one of the others, then the lack of resemblance between the three nonidentical entities should remain invariant regardless of which two entities are the same. A generalization of equality (11.1) is given by

$$d_4(x_1, x_1, x_2, x_3) = d_4(x_1, x_2, x_2, x_3) = d_4(x_1, x_2, x_3, x_3) \tag{11.3}$$

or $d_{1123} = d_{1223} = d_{1233}$. Equality (11.3) together with four-way symmetry, implies

$$\begin{aligned} d_{1123} &= d_{1132} = d_{1213} = d_{1312} = d_{1231} = d_{1321} = \\ d_{2113} &= d_{3112} = d_{2131} = d_{3121} = d_{2311} = d_{3211} = \\ d_{2213} &= d_{2231} = d_{2123} = d_{2321} = d_{2132} = d_{2312} = \\ d_{1223} &= d_{3221} = d_{1232} = d_{3212} = d_{1322} = d_{3122} = \\ d_{3312} &= d_{3321} = d_{3132} = d_{3231} = d_{3123} = d_{3213} = \\ d_{1332} &= d_{2331} = d_{1323} = d_{2313} = d_{1233} = d_{2133}. \end{aligned}$$

The latter equality is the mathematical formulation of the requirement that, when one of four vectors or entities is identical to one of the others, then the lack of similarity between the three nonidentical entities should remain invariant regardless of which two entities are the same.

Apart from the possibility that two entities are identical, up to two additional possibilities may be encountered in the four-way case. First of all, the four-way function may be defined on a quadruplet of which three entities are identical. Secondly, the four-way function may be defined on two pairs of identical entities. Following the above reasoning, we require that if the resemblance between two groups of identical entities is measured, then the lack of resemblance between the two nonidentical groups should remain invariant regardless of the group sizes. The requirement may be formalized with the definition of equality

$$d_4(x_1, x_1, x_1, x_2) = d_4(x_1, x_1, x_2, x_2) = d_4(x_1, x_2, x_2, x_2) \tag{11.4}$$

or $d_{1112} = d_{1122} = d_{1222}$. Equality (11.4), together with four-way symmetry, implies

$$\begin{aligned} &d_{1112} = d_{1121} = d_{1211} = d_{2111} \\ =&d_{1122} = d_{1212} = d_{1221} = d_{2112} = d_{2121} = d_{2211} \\ =&d_{1222} = d_{2122} = d_{2212} = d_{2221}. \end{aligned}$$

The definitions of axioms for five-way dissimilarities are now straightforward. Let

$$d_5(x_1, x_2, x_3, x_4, x_5) : E^5 \to \mathbb{R} \quad \text{or} \quad d_{12345} : E^5 \to \mathbb{R}$$

be a function that assigns a real number to each tuple $(x_1, x_2, x_3, x_4, x_5)$. The basic axioms for the five-way case are

$$\begin{array}{lll} (D1) & d_5(x_1, x_2, x_3, x_4, x_5) \geq 0 & \text{(nonnegativity)} \\ (D2) & d_5(x_1, x_1, x_1, x_1, x_1) = 0 & \text{(minimality)} \\ (D3) & d_5(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(5)}) = d_5(x_1, x_2, ..., x_5) & \text{(symmetry).} \end{array}$$

In the case that two out of five entities are identical, the first additional requirement is given by

$$d_{11234} = d_{12234} = d_{12334} = d_{12344}.$$

If there are three sets of identical entities (size of the set unspecified), the second additional requirement is given by

$$d_{11123} = d_{12223} = d_{12333} = d_{11223} = d_{11233} = d_{11233}.$$

When there are two sets of identical entities (size of the set unspecified), the third additional requirement is given by

$$d_{11112} = d_{11122} = d_{11222} = d_{12222}.$$

Thus, for the k-way case up to $(k-2)$ additional requirements must be specified to cover all the cases of identical entities or objects.

For the definition of the axioms for general multi-way dissimilarities the following notation is used. Let $x_{1,k} = \{x_1, x_2, ..., x_k\}$ be a k-tuple and let

$$d_k(x_{1,k}) : E^k \rightarrow \mathbb{R}$$

denote the multi-way dissimilarity for k objects or variables. The basic axioms for the measure $d_k(x_{1,k})$ are given by

$(K1)$	$d_k(x_{1,k}) \geq 0$	(nonnegativity)
$(K2)$	$d_k(\mathbf{x}_1) = 0$	(minimality)
$(K3)$	$d_k(x_{1,k}) = d_k(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(k)})$	(symmetry)

where $\mathbf{x}_1$ is a k-tuple with elements x_1.

11.4 Epilogue

The topic of this chapter was axioms, like nonnegativity, minimality and symmetry, for two-way, three-way and general multi-way dissimilarities. Generalizations of the triangle inequality are studied in the next chapter, Chapter 12. For the axioms of two-way and three-way dissimilarities several axiom systems were studied. Using simple models, the consistency and independence of these axiom systems were established.

In the final section of the chapter axioms of multi-way dissimilarities were considered. Multi-way axioms are already quite complicated for the four-way and five-way case. Multi-way definitions of nonnegativity, minimality and symmetry are straightforward. If $x_{1,k}$ is a k-tuple, then $d(x_{1,k}) = 0$ if all elements in $x_{1,k}$ are identical. However, for $k \geq 3$ it may occur that not all but some elements in $x_{1,k}$ are identical. Additional axioms are required to deal with these new possibilities. For the three-way case Heiser and Bennani (1997) required that when one of three variables is identical to one of the others, then the lack of resemblance between the two non-identical entities should remain invariant regardless of which two entities are the same. Following this line of reasoning, additional axioms may be formulated for the four-way case, the five-way case, and the general multi-way case.

CHAPTER 12

Multi-way metrics

Measures of resemblance play an important role in many domains of data analysis. However, similarity coefficients often only allow pairwise or two-way comparison of objects or entities. An alternative to two-way resemblance measures is to formulate multi-way coefficients (see, for example, Diatta, 2006, 2007). Several authors have studied three-way dissimilarities and generalized various concepts defined for the two-way case to the three-way case (see, for example, Bennani-Dosse, 1993; Joly and Le Calvé, 1995; Heiser and Bennani, 1997). Axioms for two-way and three-way dissimilarities were reviewed in the previous chapter. Chapter 11 was also used to investigate and formulate basic axioms, like nonnegativity, minimality and symmetry for multi-way dissimilarities. In the present chapter extensions of the two-way metric and the three-way metric axioms are explored. Chapter 13 is concerned with extensions of the two three-way ultrametric axioms.

In mathematics, a metric space is a set where a notion of distance between elements of the set is defined. A two-way dissimilarity is called a metric if it is nonnegative, symmetric, satisfies minimality, and (most importantly) if it satisfies the triangle inequality. Both Joly and Le Calvé (1995) and Heiser and Bennani (1997) have considered three-way generalizations of the triangle inequality, defined for the two-way case. The two different metrics are called weak and strong in Chepoi and Fichet (2007). In this chapter the ideas on three-way metrics presented in Joly and Le Calvé (1995) and Heiser and Bennani (1997) are adopted and extended to multi-way metrics.

The inspiration for this chapter on multi-way metricity comes from the paper by Heiser and Bennani (1997). Various ideas on, and properties of, the three-way tetrahedral inequality presented in their paper, are extended in this chapter for a broad class of inequalities that generalize the triangle inequality. An important topic is how the k-way inequalities are related to the $(k-1)$-way inequalities.

12.1 Definitions

In this chapter we study a family of k-way metrics that generalize the two-way metric. Let $x_{1,k}$ denote the k-tuple $(x_1, x_2, ..., x_k)$ and let $x_{1,k}^{-i}$ denote the $(k-1)$-tuple $(x_1, ..., x_{i-1}, x_{i+1}, ..., x_k)$ where the minus in the superscript of $x_{1,k}^{-i}$ is used to indicate that element x_i drops out. In the following the elements of tuple $x_{1,k}$ will be referred to as objects.

A dissimilarity $d_k : E^k \to \mathbb{R}_+$ is totally symmetric if for all $x_1, x_2, ..., x_k \in E$ and every permutation π of $\{1, 2, ..., k\}$

$$d_k(x_{\pi(1)}, ..., x_{\pi(k)}) = d_k(x_1, ..., x_k).$$

As a generalization of minimality we define $d_k(x_1, ..., x_1) = 0$. It is assumed throughout the chapter that the equations hold for all objects in E that are involved in a definition.

Both Joly and Le Calvé (1995) and Heiser and Bennani (1997) introduced three-way generalizations of the triangle inequality. The two inequalities are given by respectively

$$d_3(x_{1,3}) \leq d_3(x_{2,4}) + d_3(x_{1,4}^{-2}) \tag{12.1}$$
$$2d_3(x_{1,3}) \leq d_3(x_{2,4}) + d_3(x_{1,4}^{-2}) + d_3(x_{1,4}^{-3}). \tag{12.2}$$

Inequalities (12.1) and (12.2) are called respectively weak and strong metrics in Chepoi and Fichet (2007). Deza and Rosenberg (2000, 2005) generalize (12.1) to

$$d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}). \tag{12.3}$$

De Rooij (2001, p. 128) noted that inequality (12.2) can be generalized to

$$(k-1) \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}) \quad \text{(the polyhedral inequality)}. \tag{12.4}$$

We may generalize (12.3) and (12.4) to

$$u \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}) \tag{12.5}$$

where u is a positive real number. We can further generalize (12.5) to

$$u \times d_k(x_{1,k}) \leq \sum_{i=1}^{v} d_k(x_{1,n+1}^{-i}) \tag{12.6}$$

where v is a positive integer bounded by $2 \leq v \leq k$. Note that the number of linear terms on the right-hand side of (12.5) is determined by k, whereas the number of linear terms on the right-hand side of (12.6) is determined by v.

If u^* is a positive integer and $u \geq u^*$, then (12.6) implies

$$u^* \times d_k(x_{1,k}) \leq \sum_{i=1}^{v} d_k(x_{1,k+1}^{-i}).$$

Furthermore, if $v \leq v^*$, then (12.6) implies

$$u \times d_k(x_{1,k}) \leq \sum_{i=1}^{v^*} d_k(x_{1,k+1}^{-i}).$$

Moreover, for $u = 1$ and $k = 1$, adding the two inequalities

$$\begin{aligned} & d_k(x_{1,k}) \leq d_k(x_{2,k+1}) + d_k(x_{1,k+1}^{-2}) \\ \text{and} \quad & d_k(x_{2,k+1}) \leq d_k(x_{1,k}) + d_k(x_{1,k+1}^{-2}) \end{aligned}$$

shows that dissimilarity $d_k(x_{1,k}) \geq 0$. In addition, we have the following property.

Proposition 12.1. *For* $u > 1$, (12.6) *implies*

$$(u-1) \times d_k(x_{1,k}) \leq \sum_{i=2}^{v} d_k(x_{1,k+1}^{-i}). \tag{12.7}$$

Proof: Interchanging the roles of x_1 and x_{k+1} in (12.6) and dividing the result by u, we obtain

$$d_k(x_{2,k+1}) \leq \frac{1}{u} d_k(x_{1,k}) + \frac{1}{u} \sum_{i=2}^{v} d_k(x_{1,k+1}^{-i}). \tag{12.8}$$

Adding (12.8) to (12.6) we obtain

$$\frac{u^2-1}{u} \times d_k(x_{1,k}) \leq \frac{u+1}{u} \sum_{i=2}^{v} d_k(x_{1,k+1}^{-i}). \tag{12.9}$$

Using $u^2 - 1 = (u+1)(u-1)$, multiplication of (12.9) by $u/(u+1)$ yields (12.7). □

12.2 Two identical objects

In the remainder of the chapter we are interested in how dissimilarity d_k is related to d_{k-1}. In Section 12.3 we consider lower and upper bounds of d_k in terms of d_{k-1}. Furthermore, in Section 12.4 we study what $(k-1)$-way metrics are implied by (12.6). Apart from minimality, symmetry and (12.6), we discuss below several additional requirements that specify how d_k and d_{k-1} are related when two objects of d_k are identical.

A first requirement is the following condition. Following Heiser and Bennani (1997) for the three-way case and Deza and Rosenberg (2000, 2005) for the k-way case, we require that, if two objects are identical then d_k should remain invariant regardless which two objects are the same, that is,

$$d_k(x_1, x_{1,k-1}) = d_k(x_{1,2}, x_{2,k-1}) = ... = d_k(x_{1,k-1}, x_{k-1}). \tag{12.10}$$

In view of the total symmetry, (12.10) implies that $d_k(x_1, ..., x_k)$ only depends on the h-element set $\{x_{i_1}, ..., x_{i_h}\}$ such that $\{x_1, ..., x_k\} = \{x_{i_1}, ..., x_{i_h}\}$ where $1 \leq i_1 \leq i_h \leq k$. We consider the following example that satisfies (12.10).

Deza and Rosenberg (2000, p. 803) introduced the k-way extension of the three-way star distance discussed in Joly and Le Calvé (1995). Let $|\{x_1, ..., x_n\}|$ denote the cardinality of set $\{x_1, ..., x_k\}$. Let $\alpha : E \to \mathbb{R}_+$ and $k \geq 3$. The star k-distance $d_k^\alpha : E^k \to \mathbb{R}_+$ is defined as follows. Let $x_1, ..., x_k \in E$ and let $0 \leq i_1 \leq ... \leq i_h \leq k$ be such that $|\{x_1, ..., x_k\}| = |\{x_{i_1}, ..., x_{i_h}\}| = h$. Set

$$d_k^\alpha(x_{1,k}) = \begin{cases} \sum_{j=1}^h \alpha(x_{i_j}) & \text{if } h > 1, \\ 0 & \text{if } h = 1. \end{cases}$$

Deza and Rosenberg (2000, p. 803) showed that the star k-distance d_k^α satisfies (12.10).

Condition (12.10) is perhaps not an intuitive requirement, since it may not hold for certain functions. For example, the perimeter distance gives a geometrical interpretation of the concept "average distance" between objects. Heiser and Bennani (1997) and De Rooij and Gower (2003) study the three-way perimeter distance function

$$d_3^p(x_{1,3}) = d(x_1, x_2) + d(x_1, x_3) + d(x_2, x_3). \tag{12.11}$$

A possible k-way extension of (12.11) is

$$d_k^p(x_{1,k}) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} d(x_i, x_j).$$

Perimeter distance d_k^p is the sum of all pairwise distances between the objects involved. It may be verified that d_k^p does not satisfy (12.10) for $k \geq 4$.

In the remainder of this chapter it is assumed that $d_k(x_{1,k})$ satisfies (12.10). To relate a k-way dissimilarity d_k to a $(k-1)$-way dissimilarity d_{k-1}, we study two additional restrictions. Let p be a real positive value. Suppose that, if two objects of the k-way dissimilarity are identical, d_k and d_{k-1} are equal up to multiplication by a factor p, that is,

$$d_{k-1}(x_{1,k-1}) = \frac{1}{p}\, d_k(x_1, x_{1,k-1}). \tag{12.12}$$

The value of p in (12.12) may depend on the particular distance model or function that is used. For example, Joly and Le Calvé (1995) introduce the three-way semi-perimeter distance

$$d_3^{sp}(x_{1,3}) = \frac{d(x_1,x_2) + d(x_1,x_3) + d(x_2,x_3)}{2}. \tag{12.13}$$

Applying (12.11) with tuple (x_1, x_1, x_2) we obtain $d_3^p(x_1, x_1, x_2) = 2d(x_1, x_2)$. However, applying (12.13) with tuple (x_1, x_1, x_2) we obtain $d_3^{sp}(x_1, x_1, x_2) = d(x_1, x_2)$.

For generality we let p in (12.12) be a positive real number. Of course, it may be argued that $p \geq 1$. The bounds studied in the Section 12.3 depend on the value of p. The bounds of d_k in terms of the d_{k-1} therefore depend on the distance function that is used to relate the k-way dissimilarity and $(k-1)$-way dissimilarity. The results in Section 12.4 however, do not depend on the value of p.

The final requirement we discuss in this section is given by

$$d_k(x_1, x_{1,k-1}) \leq d_k(x_{1,k}). \tag{12.14}$$

In (12.14), the k-way dissimilarity without identical objects is equal to or greater than the k-way dissimilarity with two identical objects. Condition (12.14) seems to be a natural requirement for a multi-way dissimilarity. Combining (12.12) and (12.14) we obtain

$$p\, d_{k-1}(x_{1,k-1}) \leq d_k(x_{1,k}). \tag{12.15}$$

12.3 Bounds

In this section we study the lower and upper bounds of dissimilarity d_k in terms of the d_{k-1}. We first turn our attention to the lower bound of k-way dissimilarity $d_k(x_{1,k})$ that satisfies minimality, total symmetry, and (12.10).

Proposition 12.2. *If* (12.12) *and* (12.14) *hold, then for k-way dissimilarity $d_k(x_{1,k})$ we have*

$$\frac{p}{k}\sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \leq d_k(x_{1,k}). \tag{12.16}$$

Proof: For given k, there are k variants of $d_{k-1}(x_{1,k-1})$, which are given by $d_{k-1}(x_{1,k}^{-i})$ for $i = 1, 2, ..., k$. We obtain k variants of (12.15) by substituting $d_{k-1}(x_{1,k-1})$ on the left-hand side of (12.15) by one of its variants. Adding up all k variants of (12.15), that is, adding inequalities

$$\begin{aligned} p\, d_{k-1}(x_{1,k}^{-k}) &\leq d_k(x_{1,k}) \\ p\, d_{k-1}(x_{1,k}^{-(k-1)}) &\leq d_k(x_{1,k}) \\ &\vdots \\ p\, d_{k-1}(x_{1,k}^{-3}) &\leq d_k(x_{1,k}) \\ p\, d_{k-1}(x_{1,k}^{-2}) &\leq d_k(x_{1,k}) \\ p\, d_{k-1}(x_{2,k}) &\leq d_k(x_{1,k}) \end{aligned}$$

followed by division by k, we obtain (12.16). $\square$

For $p = 1$, lower bound (12.16) is equivalent to the arithmetic mean of the $(k-1)$-way dissimilarities $d_{k-1}(x_{1,k}^{-i})$.

For the case $(u - v + 2) > 0$, we have the following lower bound for a k-way distance (that is, $d_k(x_{1,n})$ satisfies minimality, total symmetry, (12.6) and (12.10)). In contrast to Proposition 12.2, we only require validity of (12.12), not (12.14), for this lower bound.

Proposition 12.3. *Suppose* (12.12) *holds and* $(u - v + 2) > 0$. *Then for* k*-way distance* $d_k(x_{1,k})$ *we have*

$$\frac{p(u-v+2)}{2k} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \leq d_k(x_{1,k}). \tag{12.17}$$

Proof: Applying (12.6) with $(k+1)$-tuple $(x_1, x_1, x_3, ..., x_{k+1})$, and replacing x_{k+1} by x_2 in the result, we obtain

$$p\,u \times d_{k-1}(x_{1,k}^{-2}) \leq 2d_k(x_{1,k}) + p \sum_{i=3}^{v} d_{k-1}(x_1, x_2, x_{3,k}^{-i}) \quad \text{for} \quad v \geq 3 \tag{12.18}$$

$$p\,u \times d_{k-1}(x_{1,k}^{-2}) \leq 2d_k(x_{1,k}) \quad \text{for} \quad v = 2. \tag{12.19}$$

We have k variants of d_{k-1} for given k, for example $d_{k-1}(x_{1,k}^{-2})$ in left-hand side of (12.19). We may obtain k variants of (12.19) by replacing $d_{k-1}(x_{1,k}^{-2})$ by one of the other $(k-1)$ variants. Adding up all k variants of (12.19), followed by division by $2k$, we obtain

$$\frac{p\,u}{2k} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \leq d_k(x_{1,k})$$

which is the inequality that is obtained by using $v = 2$ in (12.17).

We may obtain k variants of (12.18) by replacing $d_{k-1}(x_{1,k}^{-2})$ in the left-hand side of (12.18) by one of the other $(k-1)$ variants. Considering all k variants of (12.18), the k variants of d_{k-1} on the right-hand side each occur a total of $(v-2)$ times. Adding up all k variants of (12.18), followed by division by $2k$, we obtain (12.17). □

If (12.12) and (12.4) hold, then $d_k(x_{1,k})$ has a lower bound

$$\frac{p}{2k} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \leq d_k(x_{1,k}). \tag{12.20}$$

We obtain (12.20) by using $u = k - 1$ and $v = k$ in (12.17). For $p = 2$ the lower bound of $d_k(x_{1,k})$ is equivalent to the arithmetic mean of the $(k-1)$-way dissimilarities $d_{k-1}(x_{1,k}^{-i})$. If not only (12.12) but also (12.14) is valid, then (12.16) is the lower bound of $d_k(x_{1,k})$. Note that (12.16) is sharper than (12.20).

Next, we focus on the upper bound of k-way distance $d_k(x_{1,k})$.

Proposition 12.4. *If* (12.12) *holds, then for* k*-way distance* $d_k(x_{1,k})$ *we have*

$$d_k(x_{1,k}) \leq \frac{vp}{ku} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \quad \textit{for} \quad 2 \leq v \leq k-1 \tag{12.21}$$

$$d_k(x_{1,k}) \leq \frac{(k-1)p}{k(u-1)} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \quad \textit{for} \quad v = k. \tag{12.22}$$

Proof: Applying (12.6) with $(k+1)$-tuple $(x_1, ..., x_k, x_k)$ we obtain

$$u \times d_k(x_{1,k}) \leq p \sum_{i=1}^{v} d_{k-1}(x_{1,k}^{-i}) \quad \text{for} \quad 2 \leq v \leq k-1 \tag{12.23}$$

$$(u-1) \times d_k(x_{1,k}) \leq p \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}) \quad \text{for} \quad v = k. \tag{12.24}$$

We have k variants of $d_{k-1}(x_{1,k}^{-i})$ in (12.23) and (12.24). Considering all k variants of (12.23) and (12.24), each $d_{k-1}(x_{1,k}^{-i})$ occurs a total of v times. Adding up all k variants of (12.23) and (12.24), followed by division by ku, respectively $k(u-1)$, we obtain (12.21) and (12.22). □

Using $u = k$ and $v = k$ in (12.6) yields

$$k \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}). \tag{12.25}$$

If (12.12) and (12.25) hold, then the k-way distance $d_k(x_{1,k})$ is bounded from above by

$$d_k(x_{1,k}) \leq \frac{p}{k} \sum_{i=1}^{k} d_k(x_{1,k}^{-i}). \tag{12.26}$$

We obtain (12.26) by using $u = k$ in (12.22). For $p = 1$ the upper bound of $d_k(x_{1,k})$ is equivalent to the arithmetic mean of the $(k-1)$-way distances $d_{k-1}(x_{1,k}^{-i})$.

12.4 $(k-1)$-Way metrics implied by k-way metrics

In this section we study what $(k-1)$-way metrics are implied by the family of k-way metrics defined in (12.6). Again k-way dissimilarity $d_k(x_{1,k})$ satisfies minimality, total symmetry, and (12.10). It is interesting to note that, although we use condition (12.12) throughout this section, the results do not depend on the value of p in (12.12). Unless stated otherwise we assume $k \geq 3$ throughout this section.

Proposition 12.5. *If* (12.12) *and* (12.14) *hold, then* (12.6) *implies*

$$u \times d_{k-1}(x_{1,k-1}) \leq \sum_{i=1}^{v} d_{k-1}(x_{1,k}^{-i}) \quad \textit{for} \quad 2 \leq v \leq k-1 \tag{12.27}$$

$$(u-1) \times d_{k-1}(x_{1,k-1}) \leq \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}) \quad \textit{for} \quad v = k,\ k > 1. \tag{12.28}$$

Proof: Inequalities (12.27) and (12.28) are obtained from combining (12.15) with (12.23), respectively (12.24). □

As it turns out, condition (12.14) is not required to obtain (12.27). We first show that if (12.12) holds, then (12.6) implies (12.27) for $k \geq 4$ and $2 \leq v \leq k-2$.

Proposition 12.6. *If* (12.12) *holds, then* (12.6) *implies* (12.27) *for* $k \geq 4$ *and* $2 \leq v \leq k-2$.
Proof: Applying (12.6) with $(k+1)$-tuple $(x_1, ..., x_{k-1}, x_{k-1}, x_{k+1})$ and replacing x_{k+1} by x_k in the result, we obtain (12.27). □

Using $v = k-1$ in (12.6) we obtain

$$u \times d_k(x_{1,k}) \leq \sum_{i=1}^{k-1} d_k(x_{1,k+1}^{-i}). \tag{12.29}$$

Using $v = k-1$ in (12.27) we obtain

$$u \times d_{k-1}(x_{1,k-1}) \leq \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{12.30}$$

Next, we show that if (12.12) holds, then (12.29) implies (12.30) for $u \geq 1$.

Proposition 12.7. *If* (12.12) *holds, then for* $u \geq 1$, (12.29) *implies* (12.30).
Proof: Applying (12.29) with $(k+1)$-tuple $(x_1, ..., x_{k-1}, x_{k-1}, x_{k+1})$ and replacing x_{k+1} by x_k in the result, we obtain

$$p\, u \times d_{k-1}(x_{1,k-1}) \leq p \sum_{i=1}^{k-2} d_{k-1}(x_{1,k}^{-i}) + d_k(x_{1,k}). \tag{12.31}$$

Using $v = k - 1$ in (12.23) we obtain

$$u \times d_k(x_{1,k}) \le p \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{12.32}$$

Adding (12.33) to $u \times$ (12.31) yields

$$u^2 \times d_{k-1}(x_{1,k-1}) \le u \sum_{i=1}^{k-2} d_{k-1}(x_{1,k}^{-i}) + d_{k-1}(x_{1,k}^{-(k-1)}). \tag{12.33}$$

Apart from variant $d_{k-1}(x_{1,k-1})$ on the left-hand side of (12.33), there are $(k-1)$ variants of d_{k-1}, for example, variant $d_{k-1}(x_{1,k}^{-(k-1)})$, on the right-hand side of (12.33). We have $(k-1)$ variants of (12.33) by varying all $(k-1)$ variants of d_{k-1} on the right-hand side of (12.33). Adding up all $(k-1)$ variants of (12.33), followed by division by $(k-1)u$, yields

$$u \times d_{k-1}(x_{1,k-1}) \le \left[\frac{(k-2)u+1}{(k-1)u}\right] \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{12.34}$$

To complete the proof, it must be shown that parametrized inequality (12.34) is stronger than (12.30). We have

$$\frac{(k-2)u+1}{(k-1)u} \le 1$$

if and only if $u \ge 1$. The latter requirement is true under the conditions of the theorem. This completes the proof. □

Using $v = k$ in (12.6) we obtain (12.5). From Proposition 12.5 we know that if both (12.12) and (12.14) hold, then (12.5) implies (12.28). If only (12.12) is valid, (12.5) implies the parametrized inequality

$$(u-1) \times d_{k-1}(x_{1,k-1}) \le \left[1 + \frac{k-u}{(k-1)u}\right] \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{12.35}$$

Proposition 12.8. *If* (12.12) *holds, then for* $u > 1$, (12.5) *implies* (12.35).
Proof: Applying (12.5) with $(k+1)$-tuple $(x_1, ..., x_{k-1}, x_{k-1}, x_{k+1})$ and replacing x_{k+1} by x_k in the result, we obtain

$$p\,u \times d_{k-1}(x_{1,k-1}) \le p \sum_{i=1}^{k-2} d_{k-1}(x_{1,k}^{-i}) + 2d_k(x_{1,k}). \tag{12.36}$$

Adding $2 \times$ (12.24) to $(u-1) \times$ (12.36) we obtain

$$u(u-1) \times d_{k-1}(x_{1,k-1}) \le (u+1) \sum_{i=1}^{k-2} d_{k-1}(x_{1,k}^{-i}) + 2d_{k-1}(x_{1,k}^{-(k-1)}). \tag{12.37}$$

Apart from variant $d_{k-1}(x_{1,k-1})$ on the left-hand side of (12.37), there are $(k-1)$ variants of d_{k-1} on the right-hand side of (12.37). We have $(k-1)$ variants of (12.37) by varying all $(k-1)$ variants of d_{k-1} on the right-hand side of (12.37). Adding up these $(k-1)$ variants of (12.37), followed by division by $(k-1)u$, yields (12.35). □

The parametrized inequality (12.35) is weaker than (12.28) for $k > u$, and stronger than (12.28) for $3 \leq k < u$. With respect to quantity

$$1 + \frac{k-u}{(k-1)u} \tag{12.38}$$

in (12.35) we have limits

$$\lim_{k\to\infty}\left[1 + \frac{k-u}{(k-1)u}\right] = 1 + \frac{1}{u}, \qquad \lim_{u\to\infty}\left[1 + \frac{k-u}{(k-1)u}\right] = 1 - \frac{1}{k}$$

and

$$\lim_{k,u\to\infty}\left[1 + \frac{k-u}{(k-1)u}\right] = 1.$$

Because of these limits it may be argued that (12.38) and (12.35) are only interesting for small k and u. Furthermore, if $k = u$, then (12.39) = 1, and (12.35) is equivalent to (12.28).

Using $u = k-1$ in (12.5) we obtain the polyhedral inequality (12.4). If (12.12) holds, then for $k \geq 3$ the polyhedral inequality (12.4) implies

$$(u-2) \times d_{k-1}(x_{1,k-1}) \leq \left[1 + \frac{1}{(k-1)^2}\right] \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{12.39}$$

We obtain (12.39) by using $u = k-1$ in (12.35) and noting that $k^2 - 2k + 2 = (k-1)^2 + 1$. The quantity

$$1 + \frac{1}{(k-1)^2} \quad \text{in (12.39) with limit} \quad \lim_{k\to\infty}\left[1 + \frac{1}{(k-1)^2}\right] = 1$$

approximates 1 rapidly as k increases. As shown in Heiser and Bennani (1997, p. 192), if (12.12) holds then the tetrahedral inequality (12.2) does not imply the triangle inequality, but the weaker parametrized triangle inequality

$$d(x_1, x_2) \leq \frac{5}{4}\left[d(x_2, x_3) + d(x_1, x_3)\right].$$

Furthermore, if (12.12) holds, then

$$3d_4(x_{1,4}) \leq d_4(x_{2,5}) + d_4(x_{1,5}^{-2}) + d_4(x_{1,5}^{-3}) + d_4(x_{1,5}^{-4})$$

does not imply the tetrahedral inequality (12.2), but the weaker parametrized inequality

$$2d_3(x_{1,3}) \leq \frac{10}{9}\left[d_3(x_{2,4}) + d_3(x_{1,4}^{-2}) + d_3(x_{1,4}^{-3})\right].$$

12.5 Epilogue

In this chapter a family of k-way metrics that extend the usual two-way metric was studied. The three-way metrics introduced by Joly and Le Calvé (1995) and Heiser and Bennani (1997) and the k-way metrics studied in Deza and Rosenberg (2000) are in the family. The family gives an indication of the many possible extensions for introducing k-way metricity. It was shown how k-way metrics and k-way dissimilarities are related to their $(k-1)$-way counterparts under different set of axioms.

Validity of a metric axiom for $k \geq 3$ appears not to be important for methods used in applied multi-way data analysis, such as multi-way principal component and factor analysis (Kroonenberg, 2008), or multi-way dimensional scaling (Gower and De Rooij, 2003; Heiser and Bennani, 1997). For example, the three-way multidimensional scaling done in Gower and De Rooij (2003) merely required that the underlying two-way coefficients satisfied the triangle inequality, since the three-way dissimilarities are linear transformations of the two-way information. The multi-way procedure based on the gradient method used in Cox, Cox and Branco (1991) and the three-way least squares procedure used in Heiser and Bennani (1997) do not require that the dissimilarities satisfy stronger conditions. At this point the formulations and properties presented in this chapter appear to be of theoretical interest only. From a theoretical point of view it is unfortunate that no well-established basic multi-way metric structure emerged from the study.

CHAPTER 13

Multi-way ultrametrics

Multi-way dissimilarities are natural generalizations of pairwise dissimilarities, that allow global comparison of more than two objects or variables. Various authors have studied three-way dissimilarities and generalized various concepts defined for the two-way case to the three-way case (see, for example, Bennani-Dosse, 1993; Joly and Le Calvé, 1995; Heiser and Bennani, 1997). One of these topics is ultrametric dissimilarities (Diatta and Fichet, 1998; Murtagh, 2004; Diatta, 2007). A two-way dissimilarity $d(x_1, x_2)$ is called a two-way ultrametric if it satisfies the ultrametric inequality, which is given by

$$d(x_1, x_2) \leq \max[d(x_1, x_3), d(x_2, x_3)].$$

The two-way ultrametric inequality implies that the triangle formed by the three points x_1, x_2 and x_3 is isosceles, that is, at least the largest two sides are of equal length. A recent review on where ultrametricity may be encountered is given by Murtagh (2004). Diatta and Fichet (1998) and Diatta (2006, 2007) consider a class of multi-way quasi-ultrametrics that extend the fundamental bijection in classification between ultrametric dissimilarities and indexed hierarchies.

Joly and Le Calvé (1995) and Bennani-Dosse (1993) describe three-way generalizations of the ultrametric inequality, defined for the two-way case. The two different ultrametrics are called weak and strong in Chepoi and Fichet (2007). In this chapter the ideas on three-way ultrametrics presented in Joly and Le Calvé (1995) and Bennani-Dosse (1993) are adopted and extended to multi-way ultrametrics. For the two-way case we have the ultrametric inequality; for the three-way case two equalities have been proposed; for the four-way case three inequalities are presented; and for the multi-way case $(k-1)$ inequalities may be defined. The inspiration for this chapter comes from the thesis by Bennani-Dosse (1993). Some ideas on the three-way ultrametrics presented in that thesis, are explored in this chapter for multi-way dissimilarities.

13.1 Definitions

Let $x_{1,k} = \{x_1, x_2, ..., x_k\}$ be a k-tuple and let $x_{1,k}^{-i}$ be a $(k-1)$-tuple with elements x_1 to x_k where the minus in $x_{1,k}^{-i}$ is used to indicate that element x_i drops out. Both Bennani-Dosse (1993) and Chepoi and Fichet (2007, p. 5) consider two three-way generalizations of the ultrametric inequality, namely

$$\begin{aligned} d(x_{1,3}) &\leq \max\left[d(x_{2,4}), d(x_{1,4}^{-2}), d(x_{1,4}^{-3})\right] \\ d(x_{1,3}) &\leq \max\left[d(x_{2,4}), d(x_{1,4}^{-2})\right] . \end{aligned}$$

These inequalities are called respectively weak and strong ultrametrics in Chepoi and Fichet (2007). For groups of size $k = 4$ it is possible to formulate three ultrametric inequalities. From weak to strong, the three ultrametrics are given by

$$\begin{aligned} d(x_{1,4}) &\leq \max\left[d(x_{2,5}), d(x_{1,5}^{-2}), d(x_{1,5}^{-3}), d(x_{1,5}^{-4})\right] \\ d(x_{1,4}) &\leq \max\left[d(x_{2,5}), d(x_{1,5}^{-2}), d(x_{1,5}^{-3})\right] \\ d(x_{1,4}) &\leq \max\left[d(x_{2,5}), d(x_{1,5}^{-2})\right] . \end{aligned}$$

We may thus formulate $(k-1)$ ultrametrics for a group of k objects.

For the properties in this chapter it is more convenient to define an ultrametric on the number of dissimilarities involved. For example, the inequality $d_3 \leq \max(d_1, d_2)$ represents all metrics of which the definition involves three multi-way dissimilarities, that is,

$$\begin{aligned} d(x_{1,2}) &\leq \max\left[d(x_{2,3}), d(x_{1,3}^{-2})\right] \\ d(x_{1,3}) &\leq \max\left[d(x_{2,4}), d(x_{1,4}^{-2})\right] \\ d(x_{1,4}) &\leq \max\left[d(x_{2,5}), d(x_{1,5}^{-2})\right] \\ d(x_{1,5}) &\leq \max\left[d(x_{2,6}), d(x_{1,6}^{-2})\right] \\ d(x_{1,6}) &\leq \max\left[d(x_{2,7}), d(x_{1,7}^{-2})\right] \\ &\text{etc. ...} \end{aligned}$$

The inequality

$$d_3 \leq \max(d_1, d_2) \tag{13.1}$$

defines the strongest class of ultrametrics, whereas

$$d_4 \leq \max(d_1, d_2, d_3) \tag{13.2}$$

defines the second strongest class. To see that inequality (13.1) defines a stronger ultrametric compared to inequality (13.2), suppose the multi-way dissimilarities are given by

$$d_1 = d_2 = 5 \qquad d_3 = 3 \qquad \text{and} \qquad d_4 = 2.$$

These multi-way dissimilarities satisfy (13.2), since $5 \leq \max(2, 3, 5)$, but not (13.1), because $5 \neq \max(2, 3)$. As a second example, the multi-way dissimilarities given by

$$d_1 = d_2 = 5 \qquad d_3 = 3 \qquad d_4 = 4 \qquad \text{and} \qquad d_5 = 2$$

do not satisfy either (13.1) or (13.2). However, these multi-way dissimilarities do satisfy the weaker ultrametric inequality

$$d_5 \leq \max(d_1, d_2, d_3, d_4) \qquad \text{(for example, } 5 \leq \max(2, 3, 4, 5)).$$

Following this line of reasoning we may conclude that a multi-way ultrametric implies all (possible) weaker ultrametrics.

Proposition 13.1. *Let $d_1, d_2, ..., d_n$ be n multi-way dissimilarities. Then*

$$d_{n-1} \leq \max(d_1, d_2, ..., d_{n-2}) \qquad \Rightarrow \qquad d_n \leq \max(d_1, d_2, ..., d_{n-1}).$$

Let $d_{1,k} = \{d_1, d_2, ..., d_k\}$ be a k-tuple. Then

$$d_{k+1} \leq \max(d_{1,k})$$

defines the weakest class of ultrametrics.

13.2 Strong ultrametrics

The strongest class of ultrametrics is characterized by inequality (13.1). It turns out that, if n multi-way dissimilarities satisfy inequality (13.1), then the $(n-1)$ largest dissimilarities are equal. The sufficiency of this statement is clear from the definition of the class of ultrametrics in inequality (13.1). The proof of necessity goes as follows. We first consider the proof for $n = 3, 4, 5$. The proof for $n = 4$ was already presented in Bennani-Dosse (1993). Furthermore, for $n = 4, 5$ alternative proofs are presented, where the fact is used that the assertion is true for $n - 1$. Finally, the proof is completed by means of induction.

Proposition 13.2. *Let $d_1, d_2, ..., d_n$ be n multi-way dissimilarities. If the n dissimilarities satisfy inequality* (13.1), *then the largest $n-1$ dissimilarities are equal.*
Proof for $n=3$: Assume $d_2 \leq d_1$. From $d_3 \leq \max(d_1, d_2)$ we obtain $d_3 \leq d_1$. Then

$$\begin{aligned} d_3 \leq d_2 \text{ and } d_1 \leq \max(d_2, d_3) \;\Rightarrow\; d_1 \leq d_2 \;\Rightarrow\; d_3 \leq d_1 = d_2 \\ d_2 \leq d_3 \text{ and } d_1 \leq \max(d_2, d_3) \;\Rightarrow\; d_1 \leq d_3 \;\Rightarrow\; d_2 \leq d_1 = d_3. \end{aligned}$$

Proof for $n=4$: Assume $d_2 \leq d_1$. From $d_3 \leq \max(d_1, d_2)$ we obtain $d_3 \leq d_1$.
First, if $d_3 \leq d_2$

$$\begin{aligned} \text{then } d_1 \leq \max(d_2, d_3) \;\Rightarrow\; d_1 \leq d_2 \;\Rightarrow\; d_3 \leq d_1 = d_2 \\ \text{and } d_4 \leq \max(d_2, d_3) \;\Rightarrow\; d_4 \leq d_2. \end{aligned}$$

Then

$$\begin{aligned} d_4 \leq d_3 \text{ and } d_2 \leq \max(d_3, d_4) \;\Rightarrow\; d_2 \leq d_3 \;\Rightarrow\; d_4 \leq d_1 = d_2 = d_3 \\ d_3 \leq d_4 \text{ and } d_2 \leq \max(d_3, d_4) \;\Rightarrow\; d_2 \leq d_4 \;\Rightarrow\; d_3 \leq d_1 = d_2 = d_4. \end{aligned}$$

Alternatively, if $d_2 \leq d_3$

$$\begin{aligned} \text{then } d_1 \leq \max(d_2, d_3) \;\Rightarrow\; d_1 \leq d_3 \;\Rightarrow\; d_2 \leq d_1 = d_3 \\ \text{and } d_4 \leq \max(d_2, d_3) \;\Rightarrow\; d_4 \leq d_3. \end{aligned}$$

Then

$$\begin{aligned} d_4 \leq d_2 \text{ and } d_3 \leq \max(d_2, d_4) \;\Rightarrow\; d_3 \leq d_2 \;\Rightarrow\; d_4 \leq d_1 = d_2 = d_3 \\ d_2 \leq d_4 \text{ and } d_3 \leq \max(d_2, d_4) \;\Rightarrow\; d_3 \leq d_4 \;\Rightarrow\; d_2 \leq d_1 = d_3 = d_4. \end{aligned}$$

This completes the proof for $n=4$.
Alternative proof for $n=4$: Assume that the assertion is true for $n=3$. If $d_3 \leq d_2 \leq d_1$, then $d_3 \leq d_1 = d_2$ and $d_4 \leq d_2$. Then

$$\begin{aligned} d_4 \leq d_3 \text{ and } d_2 \leq \max(d_3, d_4) \;\Rightarrow\; d_2 \leq d_3 \;\Rightarrow\; d_4 \leq d_1 = d_2 = d_3 \\ d_3 \leq d_4 \text{ and } d_2 \leq \max(d_3, d_4) \;\Rightarrow\; d_2 \leq d_4 \;\Rightarrow\; d_3 \leq d_1 = d_3 = d_4. \end{aligned}$$

This completes the alternative proof for $n=4$.

Proof for $n=5$: Assume $d_2 \leq d_1$. From $d_3 \leq \max(d_1, d_2)$ we obtain $d_3 \leq d_1$.
First, if $d_3 \leq d_2$

$$\begin{aligned} \text{then } d_1 \leq \max(d_2, d_3) \;\Rightarrow\; d_1 \leq d_2 \;\Rightarrow\; d_3 \leq d_1 = d_2 \\ \text{and } d_4 \leq \max(d_2, d_3) \;\Rightarrow\; d_4 \leq d_2. \end{aligned}$$

Furthermore, if $d_4 \leq d_3$

$$\begin{aligned} \text{then } d_2 \leq \max(d_3, d_4) \;\Rightarrow\; d_2 \leq d_3 \;\Rightarrow\; d_4 \leq d_1 = d_2 = d_3 \\ \text{and } d_5 \leq \max(d_3, d_4) \;\Rightarrow\; d_5 \leq d_3. \end{aligned}$$

Then

$$d_5 \le d_4 \text{ and } d_3 \le \max(d_4, d_5) \ \Rightarrow\ d_3 \le d_4 \ \Rightarrow\ d_5 \le d_1 = d_2 = d_3 = d_4$$
$$d_4 \le d_5 \text{ and } d_3 \le \max(d_4, d_5) \ \Rightarrow\ d_3 \le d_5 \ \Rightarrow\ d_4 \le d_1 = d_2 = d_3 = d_5.$$

Alternatively, if $d_3 \le d_4$

$$\text{then } d_2 \le \max(d_3, d_4) \ \Rightarrow\ d_2 \le d_4 \ \Rightarrow\ d_3 \le d_1 = d_2 = d_4$$
$$\text{and } d_5 \le \max(d_3, d_4) \ \Rightarrow\ d_5 \le d_4.$$

Then

$$d_5 \le d_3 \text{ and } d_4 \le \max(d_3, d_5) \ \Rightarrow\ d_4 \le d_3 \ \Rightarrow\ d_5 \le d_1 = d_2 = d_3 = d_4$$
$$d_3 \le d_5 \text{ and } d_4 \le \max(d_3, d_5) \ \Rightarrow\ d_4 \le d_5 \ \Rightarrow\ d_3 \le d_1 = d_2 = d_4 = d_5.$$

Second, if $d_2 \le d_3$

$$\text{then } d_1 \le \max(d_2, d_3) \ \Rightarrow\ d_1 \le d_3 \ \Rightarrow\ d_2 \le d_1 = d_3$$
$$\text{and } d_4 \le \max(d_2, d_3) \ \Rightarrow\ d_4 \le d_3.$$

Furthermore, if $d_4 \le d_2$

$$\text{then } d_3 \le \max(d_2, d_4) \ \Rightarrow\ d_3 \le d_2 \ \Rightarrow\ d_4 \le d_1 = d_2 = d_3$$
$$\text{and } d_5 \le \max(d_2, d_4) \ \Rightarrow\ d_5 \le d_2.$$

Then

$$d_5 \le d_4 \text{ and } d_2 \le \max(d_4, d_5) \ \Rightarrow\ d_2 \le d_4 \ \Rightarrow\ d_5 \le d_1 = d_2 = d_3 = d_4$$
$$d_4 \le d_5 \text{ and } d_2 \le \max(d_4, d_5) \ \Rightarrow\ d_2 \le d_5 \ \Rightarrow\ d_4 \le d_1 = d_2 = d_3 = d_5.$$

Alternatively, if $d_2 \le d_4$

$$\text{then } d_3 \le \max(d_2, d_4) \ \Rightarrow\ d_3 \le d_4 \ \Rightarrow\ d_2 \le d_1 = d_3 = d_4$$
$$\text{and } d_5 \le \max(d_2, d_4) \ \Rightarrow\ d_5 \le d_4.$$

Then

$$d_5 \le d_2 \text{ and } d_4 \le \max(d_2, d_5) \ \Rightarrow\ d_4 \le d_2 \ \Rightarrow\ d_5 \le d_1 = d_2 = d_3 = d_4$$
$$d_2 \le d_5 \text{ and } d_4 \le \max(d_2, d_5) \ \Rightarrow\ d_4 \le d_5 \ \Rightarrow\ d_2 \le d_1 = d_3 = d_4 = d_5.$$

This completes the proof for $n = 5$.
Alternative proof for $n = 5$: Assume that the assertion is true for $n = 4$. If $d_4 \le d_3 \le d_2 \le d_1$, then $d_4 \le d_1 = d_2 = d_3$ and $d_5 \le d_3$. Then

$$d_5 \le d_4 \text{ and } d_3 \le \max(d_4, d_5) \ \Rightarrow\ d_3 \le d_4 \ \Rightarrow\ d_5 \le d_1 = d_2 = d_3 = d_4$$
$$d_4 \le d_5 \text{ and } d_3 \le \max(d_4, d_5) \ \Rightarrow\ d_3 \le d_5 \ \Rightarrow\ d_4 \le d_1 = d_2 = d_3 = d_5.$$

This completes the alternative proof for $n = 5$.
General proof: Assume that the assertion is true for $n = m$. If $d_m \leq d_{m-1} \leq ... \leq d_2 \leq d_1$, then $d_m \leq d_1 = d_2 = ... = d_{m-2} = d_{m-1}$ and $d_{m+1} \leq d_{m-1}$. Then $d_{m+1} \leq d_m$ and $d_{m-1} \leq \max(d_m, d_{m+1})$ lead to

$$d_{m-1} \leq d_m \;\Rightarrow\; d_{m+1} \leq d_1 = d_2 = ... = d_{m-1} = d_m$$

and $d_m \leq d_{m+1}$ and $d_{m-1} \leq \max(d_m, d_{m+1})$ lead to

$$d_{m-1} \leq d_{m+1} \;\Rightarrow\; d_m \leq d_1 = d_2 = ... = d_{m-1} = d_{m+1}.$$

Hence, the assertion is true for $n = m + 1$. □

13.3 More strong ultrametrics

The second strongest class of ultrametrics is characterized by inequality (13.2). As it turns out, if n multi-way dissimilarities satisfy inequality (13.2), then the $(n-2)$ largest dissimilarities are equal. Similar to Proposition 13.2, sufficiency follows from the definition of ultrametric inequality (13.2). The proof of necessity is slightly more involved compared to the proof of Proposition 13.2. We only consider the proof for $n = 4$ of the assertion, and therefore refer to it as a conjecture.

Conjecture 13.1. *Let $d_1, d_2, ..., d_n$ be n multi-way dissimilarities. If* (13.2) *holds, then the largest $n-2$ dissimilarities are equal.*
Proof for $n = 4$: Assume $d_3 \leq d_4$.
First, if $d_2 \leq d_3$, then from $d_1 \leq \max(d_2, d_3, d_4)$ we obtain $d_1 \leq d_4$. Then

$$d_1 \leq d_3 \text{ and } d_4 \leq \max(d_1, d_2, d_3) \quad \Rightarrow\; d_4 \leq d_3 \;\Rightarrow\; \begin{cases} d_1 \leq d_3 = d_4 \\ d_2 \leq d_3 = d_4 \end{cases}$$

$$d_3 \leq d_1 \text{ and } d_4 \leq \max(d_1, d_2, d_3) \quad \Rightarrow\; d_4 \leq d_1 \;\Rightarrow\; d_2 \leq d_3 \leq d_1 = d_4.$$

Second, assume $d_3 \leq d_2$. If $d_2 \leq d_4$, then from $d_1 \leq \max(d_2, d_3, d_4)$ we obtain $d_1 \leq d_4$. Then

$$d_1 \leq d_3 \text{ and } d_4 \leq \max(d_1, d_2, d_3) \;\Rightarrow\; d_4 \leq d_2 \;\Rightarrow\; d_1 \leq d_3 \leq d_2 = d_4.$$

Alternatively, if $d_3 \leq d_1$, then

$$d_1 \leq d_2 \text{ and } d_4 \leq \max(d_1, d_2, d_3) \quad \Rightarrow\; d_4 \leq d_2 \;\Rightarrow\; d_3 \leq d_1 \leq d_2 = d_4$$

$$d_2 \leq d_1 \text{ and } d_4 \leq \max(d_1, d_2, d_3) \quad \Rightarrow\; d_4 \leq d_1 \;\Rightarrow\; d_3 \leq d_2 \leq d_1 = d_4.$$

Next, if $d_4 \leq d_2$, then

$$d_1 \leq d_3 \text{ and } d_2 \leq \max(d_1, d_3, d_4) \;\Rightarrow\; d_2 \leq d_4 \;\Rightarrow\; d_1 \leq d_3 \leq d_2 = d_4.$$

Alternatively, if $d_3 \leq d_1$, then from $d_1 \leq \max(d_2, d_3, d_4)$ we obtain $d_1 \leq d_2$. Then

$$d_1 \leq d_4 \text{ and } d_2 \leq \max(d_1, d_3, d_4) \quad \Rightarrow\; d_2 \leq d_4 \;\Rightarrow\; d_3 \leq d_1 \leq d_2 = d_4$$

$$d_4 \leq d_1 \text{ and } d_2 \leq \max(d_1, d_3, d_4) \quad \Rightarrow\; d_2 \leq d_1 \;\Rightarrow\; d_3 \leq d_4 \leq d_1 = d_2.$$

This completes the proof for $n = 4$.

13.4 Metrics implied by ultrametrics

In this section we apply the notation used in the first sections of this chapter to multi-way metrics, which were studied in Chapter 12. We are only concerned with the number of dissimilarities involved. For example, the inequality $d_3 \leq d_1 + d_2$ represents all metrics of which the definition involves three multi-way dissimilarities, that is,

$$\begin{aligned} d(x_{1,2}) &\leq d(x_{2,3}) + d(x_{1,3}^{-2}) \\ d(x_{1,3}) &\leq d(x_{2,4}) + d(x_{1,4}^{-2}) \\ d(x_{1,4}) &\leq d(x_{2,5}) + d(x_{1,5}^{-2}) \\ &\text{etc. ...} \end{aligned}$$

Three metric inequalities and two ultrametric inequalities for three-way dissimilarities were considered in Chapter 11. The strong metric $2d_1 \leq d_2 + d_3 + d_4$ introduced by Heiser and Bennani (1997) implies the metric $d_1 \leq d_2 + d_3$, introduced in Joly and Le Calvé (1995). The latter inequality in turn implies the weak metric $d_1 \leq d_2 + d_3 + d_4$. This metric is not considered by the above authors, nor is it considered a metric in Chepoi and Fichet (2007). Furthermore, the strong ultrametric $d_1 \leq \max(d_2, d_3)$ implies the weak ultrametric $d_1 \leq \max(d_2, d_3, d_4)$. The five inequalities are related as follows.

$$\begin{array}{ccc} d_1 \leq \max(d_2, d_3) & \Rightarrow & 2d_1 \leq d_2 + d_3 + d_4 \\ & & \Downarrow \\ \Downarrow & & d_1 \leq d_2 + d_3 \\ & & \Downarrow \\ d_1 \leq \max(d_2, d_3, d_4) & \Rightarrow & d_1 \leq d_2 + d_3 + d_4 \end{array}$$

For the four-way case we may formulate eight inequalities. The inequalities are related as follows.

$$\begin{array}{ccc} d_1 \leq \max(d_2, d_3) & \Rightarrow & 3d_1 \leq d_2 + d_3 + d_4 + d_5 \\ & & \Downarrow \\ & & 2d_1 \leq d_2 + d_3 + d_4 \\ \Downarrow & & \Downarrow \\ & & d_1 \leq d_2 + d_3 \\ & & \Downarrow \\ d_1 \leq \max(d_2, d_3, d_4) & \Rightarrow & d_1 \leq d_2 + d_3 + d_4 \\ \Downarrow & & \Downarrow \\ d_1 \leq \max(d_2, d_3, d_4, d_5) & \Rightarrow & d_1 \leq d_2 + d_3 + d_4 + d_5 \end{array}$$

A variety of properties can immediately be deduced from the above definitions of multi-way ultrametrics and metrics. First of all, the strongest ultrametric inequality for k-way dissimilarities implies the strongest metric inequality for k-way dissimilarities. Remember that, if the strongest k-way ultrametric inequality holds, then the k largest of the $(k+1)$ dissimilarities are equal. With respect to Proposition 13.3 and 13.4, let $d_{1,k} = \{d_1, d_2, ..., d_k\}$ be a k-tuple.

Proposition 13.3. *Let $d_1, d_2, ..., d_k, d_{k+1}$ be $(k+1)$ k-way dissimilarities. Then*

$$d_1 \leq \max(d_{2,k+1}) \qquad \Rightarrow \qquad (k-1)d_1 \leq \sum_{i=2}^{k+1} d_i.$$

Let $d_1, d_2, ..., d_n$ be n k-way dissimilarities $(n \leq k)$. All other multi-way ultrametric inequalities, other than the strongest, imply a metric inequality of the form

$$d_1 \leq \sum_{i=2}^{n} d_i.$$

Proposition 13.4. *Let $d_1, d_2, ..., d_n$ be n k-way dissimilarities $(n \leq k)$. Then*

$$d_1 \leq \max(d_{2,n}) \qquad \Rightarrow \qquad d_1 \leq \sum_{i=2}^{n} d_i.$$

13.5 Epilogue

Multi-way ultrametrics and some of their properties were the topic of investigation of this chapter. The tetrahedral inequality introduced in Heiser and Bennani (1997) is implied by the strong ultrametric inequality. Suppose we define "interesting" in the sense that a metric inequality is interesting if it is the strongest metric implied by an ultrametric inequality. Then we may say that the tetrahedral inequality (and its multi-way generalization) is more interesting compared to the three-way metric inequality introduced in Joly and Le Calvé (1995).

Some of the ultrametrics and corresponding properties discussed here may find their way into a procedure or algorithm. It is well known that a distance is an ultrametric if and only if it can be represented by a hierarchical tree. Joly and Le Calvé (1995) line out how a hierarchical algorithm may be adopted to the three-way case. First the triple corresponding to the smallest distance is aggregated and the new distances are computed involving this triple as defined in the specific algorithm. The resulting dendrogram has approximately the same properties as in the ordinary two-way case. The only difference is that there will be many levels with three clusters instead of two in the hierarchical tree representation. Applications of three-way ultrametrics and hierarchical trees can be found in Joly and Le Calvé (1995) and Bennani-Dosse (1993).

CHAPTER 14

Perimeter models

Dissimilarities are important tools in many domains of data analysis. Most dissimilarity analysis has however been limited to the two-way case. Multi-way dissimilarities may be used to evaluate complex relationships between three or more objects (see, for example, Diatta, 2006, 2007).

Perimeter models are linear functions that can be used to relate k-way dissimilarities of different degrees k. Their linear form makes perimeter functions simple models with a straightforward interpretation. For example, the three-way perimeter distance is equivalent to the sum of the three two-way distances formed between the three objects. This distance is equivalent to the sum of the three sides of the triangle formed by the three objects. The perimeter distance gives a geometrical interpretation of the concept "average distance" between objects.

The present chapter explores two extensions of the three-way perimeter model. Decompositions and metric properties of both generalizations are investigated. As an extra, the three-way maximum function, together with its multi-way extension and a metric property of the generalization, is studied in the last section.

14.1 Definitions

Let $x_{1,k}$ denote the k-tuple $(x_1, x_2, ..., x_k)$ and let $x_{1,k}^{-i}$ denote the $(k-1)$-tuple $(x_1, ..., x_{i-1}, x_{i+1}, ..., x_k)$ where the minus in the superscript of $x_{1,k}^{-i}$ is used to indicate that element x_i drops out. Let E be a nonempty set of n objects. A dissimilarity $d_k : E^k \to \mathbb{R}_+$ is totally symmetric if for all $x_1, x_2, ..., x_k \in E$ and every permutation π of $\{1, 2, ..., k\}$

$$d_k(x_{\pi(1)}, ..., x_{\pi(k)}) = d_k(x_1, ..., x_k).$$

Furthermore, as a generalization of minimality we define $d_k(x_1, ..., x_1) = 0$.

We define two types of k-way perimeter models. For $k \geq 3$ we define

$$d_k(x_{1,k}) = \frac{1}{p} \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \tag{14.1}$$

and

$$d_k(x_{1,k}) = \frac{1}{p} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} d(x_i, x_j) \tag{14.2}$$

where p is a positive real number. Dissimilarity $d_k(x_{1,k})$ in (14.1) is equivalent to the sum of the k dissimilarities $d_{k-1}(x_{1,k}^{-i})$ divided by a factor p. Distance measure $d_n(x_{1,n})$ in (14.2) may be interpreted as the sum of the sides of the polyhedron formed by the k objects in $\{x_1, x_2, ..., x_k\}$, rescaled by a factor p.

Using $k = 3$ in either (14.1) or (14.2) we obtain

$$d_3(x_{1,3}) = \frac{d(x_1, x_2) + d(x_1, x_3) + d(x_2, x_3)}{p}. \tag{14.3}$$

Using $p = 1$ in (14.3) we obtain the three-way perimeter model considered in Heiser and Bennani (1997), De Rooij and Gower (2003), and Chepoi and Fichet (2007). Using $p = 2$ in (14.3) we obtain the three-way semi-perimeter model which is studied in Bennani-Dosse (1993) and Joly and Le Calvé (1995).

Instead of the notation used in (14.3) we will use a shorter, more convenient notation in the next section on decompositions of perimeter models. We write (14.3) as

$$d_{ijl}^{(3)} = \frac{d_{ij} + d_{il} + d_{jl}}{p}. \tag{14.4}$$

Using $k = 4$ in (14.1) and (14.2) we obtain respectively

$$d_{ijlh}^{(4)} = \frac{d_{ijl}^{(3)} + d_{ijh}^{(3)} + d_{ilh}^{(3)} + d_{jlh}^{(3)}}{p} \tag{14.5}$$

and

$$d_{ijlh}^{(4)} = \frac{d_{ij} + d_{il} + d_{ih} + d_{jl} + d_{jh} + d_{lh}}{p}. \tag{14.6}$$

Note that we have expressed (14.4) and (14.5) in the same notation as (14.3).

14.2 Decompositions

The following theorem generalizes a result in Joly and Le Calvé (1995, p. 196), derived for the semi-perimeter model. As it turns out, their result holds for (14.4) and does not depend on the value of p.

Proposition 14.1. *Function $d^{(3)}_{ijl}$ satisfies* (14.4) *if and only if*

$$d^{(3)}_{ijl} = \left[d^{(3)}_{ij.} + d^{(3)}_{i.l} + d^{(3)}_{.jl}\right] - \left[d^{(3)}_{i..} + d^{(3)}_{.j.} + d^{(3)}_{..l}\right] + d^{(3)}_{...} \tag{14.7}$$

where

$$d^{(3)}_{ij.} = n^{-1}\sum_l d^{(3)}_{ijl}$$
$$d^{(3)}_{i..} = n^{-1}\sum_j d^{(3)}_{ij.}$$
$$\text{and} \qquad d^{(3)}_{...} = n^{-1}\sum_i d^{(3)}_{i..}.$$

Proof: Averaging over l, j, and i in (14.4) we obtain

$$pd^{(3)}_{ij.} = d_{ij} + d_{i.} + d_{j.}$$
$$pd^{(3)}_{i..} = 2d_{i.} + d_{..}$$
$$pd^{(3)}_{...} = 3d_{..}.$$

Expressing d_{ij} in terms of $d^{(3)}_{ij.}$, $d^{(3)}_{i..}$, and $d^{(3)}_{...}$, we obtain

$$d_{ij} = pd^{(3)}_{ij.} - \frac{p\left[d^{(3)}_{i..} + d^{(3)}_{.j.}\right]}{2} - \frac{pd^{(3)}_{...}}{3}. \tag{14.8}$$

Using (14.8) in (14.4) we obtain (14.7), which does not depend on p. □

Condition (14.7) for $d^{(3)}_{ijl}$ in (14.4) generalizes naturally to condition (14.9) for $d^{(4)}_{ijlh}$ in (14.5).

Proposition 14.2. *Function $d^{(4)}_{ijlh}$ satisfies* (14.5) *if and only if*

$$\begin{aligned} d^{(4)}_{ijlh} = &\left[d^{(4)}_{ijl.} + d^{(4)}_{ij.h} + d^{(4)}_{i.lh} + d^{(4)}_{.jlh}\right] - \left[d^{(4)}_{ij..} + d^{(4)}_{i.l.} + d^{(4)}_{i..h} + d^{(4)}_{.jl.} + d^{(4)}_{.j.h} + d^{(4)}_{..lh}\right] \\ &+ \left[d^{(4)}_{i...} + d^{(4)}_{.j..} + d^{(4)}_{..l.} + d^{(4)}_{...h}\right] - d^{(4)}_{....} \end{aligned} \tag{14.9}$$

where

$$d^{(4)}_{ijl.} = n^{-1}\sum_{l} d^{(4)}_{ijlh}$$
$$d^{(4)}_{ij..} = n^{-1}\sum_{k} d^{(4)}_{ijl.}$$
$$d^{(4)}_{i...} = n^{-1}\sum_{j} d^{(4)}_{ij..}$$
and $$d^{(4)}_{....} = n^{-1}\sum_{i} d^{(4)}_{i...}.$$

Proof: Averaging over h, l, j, and i in (14.5) we obtain

$$\begin{aligned} pd^{(4)}_{ijl.} &= d^{(3)}_{ijl} + d^{(3)}_{ij.} + d^{(3)}_{il.} + d^{(3)}_{jl.} \\ pd^{(4)}_{ij..} &= 2d^{(3)}_{ij.} + d^{(3)}_{i..} + d^{(3)}_{j..} \\ pd^{(4)}_{i...} &= 3d^{(3)}_{i..} + d^{(3)}_{...} \\ pd^{(4)}_{....} &= 4d^{(3)}_{...}. \end{aligned}$$

Expressing $d^{(3)}_{ijl}$ in terms of $d^{(4)}_{ijl.}$, $d^{(4)}_{ij..}$, $d^{(4)}_{i...}$, and $d^{(4)}_{....}$, we obtain

$$d^{(3)}_{ijl} = pd^{(4)}_{ijl.} - \frac{p\left[d^{(4)}_{ij..} + d^{(4)}_{i.l.} + d^{(4)}_{.jl.}\right]}{2} + \frac{p\left[d^{(4)}_{i...} + d^{(4)}_{.j..} + d^{(4)}_{..l.}\right]}{3} - \frac{pd^{(4)}_{....}}{4}. \tag{14.10}$$

Using (14.10) in (14.5) we obtain (14.9). □

We obtain a different generalization of (14.7) if $d^{(4)}_{ijlh}$ satisfies (14.6).

Proposition 14.3. *Function* $d^{(4)}_{ijlh}$ *satisfies* (14.6) *if and only if*

$$d^{(4)}_{ijlh} = \left[d^{(4)}_{ij..} + d^{(4)}_{i.l.} + d^{(4)}_{i..h} + d^{(4)}_{.jl.} + d^{(4)}_{.j.h} + d^{(4)}_{..lh}\right] - 2\left[d^{(4)}_{i...} + d^{(4)}_{.j..} + d^{(4)}_{..l.} + d^{(4)}_{...h}\right] + 3d^{(4)}_{....}. \tag{14.11}$$

Proof: Averaging over h, l, j, and i in (14.6) we obtain

$$\begin{aligned} pd^{(4)}_{ijl.} &= d_{ij} + d_{il} + d_{jl} + d_{i.} + d_{j.} + d_{l.} \\ pd^{(4)}_{ij..} &= d_{ij} + 2d_{i.} + 2d_{j.} + d_{..} \\ pd^{(4)}_{i...} &= 3d_{i.} + 3d_{..} \\ pd^{(4)}_{....} &= 6d_{..}. \end{aligned}$$

Expressing d_{ij} in terms of $d^{(4)}_{ij..}$, $d^{(4)}_{i...}$, and $d^{(4)}_{....}$, we obtain

$$d_{ij} = pd^{(4)}_{ij..} - \frac{2p\left[d^{(4)}_{i...} + d^{(4)}_{.j..}\right]}{3} + \frac{pd^{(4)}_{....}}{2}. \tag{14.12}$$

Using (14.12) in (14.6) yields (14.11). □

14.3 Metric properties

In this section we study metric properties of perimeter models (14.1) and (14.2). Consider metric inequalities

$$(k-1) \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}). \tag{14.13}$$

and

$$(k-2) \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}). \tag{14.14}$$

Inequality (14.13) implies inequality (14.14).

Proposition 14.4. *(i) Dissimilarity $d_n(x_{1,n})$ in* (14.2) *satisfies* (14.14). *(ii) Dissimilarity $d_n(x_{1,n})$ in* (14.2) *satisfies* (14.13) *if and only if $d(x_i, x_j)$ satisfies the triangle inequality.*
Proof (i): Using (14.2) in (14.14) we obtain

$$0 \leq (k-1) \sum_{i=1}^{k} d(x_i, x_{k+1})$$

which is true.
Proof (ii): Using (14.2) in (14.13) we obtain

$$\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} d(x_i, x_j) \leq (k-1) \sum_{i=1}^{k} d(x_i, x_{k+1}). \tag{14.15}$$

Applying (14.15) with the $(k+1)$-tuple $(x_1, x_2, x_3, ..., x_3)$ we obtain $d(x_1, x_2) \leq d(x_2, x_3) + d(x_1, x_3)$.

Conversely, inequality (14.15) follows from adding the k triangle inequalities formed by all pairs in the set $\{x_1, x_2, ..., x_k\}$ and x_{k+1}, for example, $d(x_1, x_2) \leq d(x_2, x_{k+1}) + d(x_1, x_{k+1})$. □

Consider metric inequalities

$$d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}) \tag{14.16}$$

and

$$u \times d_k(x_{1,k}) \leq \sum_{i=1}^{k} d_k(x_{1,k+1}^{-i}) \tag{14.17}$$

where u is a positive real number. Note that inequality (14.16) is implied by (14.13), (14.14) and (14.17).

Proposition 14.5. *(i) Dissimilarity $d_k(x_{1,k})$ in (14.1) satisfies (14.16). (ii) Dissimilarity $d_k(x_{1,k})$ in (14.1) satisfies (14.17) for $u > 1$ if $d_{k-1}(x_{1,k-1})$ satisfies*

$$(u-1) \times d_{k-1}(x_{1,k-1}) \leq \sum_{i=1}^{k-1} d_{k-1}(x_{1,k}^{-i}). \tag{14.18}$$

Proof (i): Using (14.1) in (14.16) we obtain

$$0 \leq (k-1) \sum_{i=1}^{k} d(x_i, x_{k+1})$$

which is true.
Proof (ii): Using (14.1) in (14.17) we obtain

$$(u-1) \sum_{i=1}^{k} d_{k-1}(x_{1,k}^{-i}) \leq 2\mathcal{S} \tag{14.19}$$

where $\mathcal{S}$ is the sum of the d_{k-1} dissimilarities that can be formed by all $(k-2)$-tuples in the set $\{x_1, x_2, ..., x_k\}$ and x_{k+1}. Inequality (14.19) follows from adding the k variants of (14.18) that can be formed by using each $(u-1) \times d_{k-1}(x_{1,k}^{-i})$ on the left-hand side of (14.19), on the left-hand side of each polyhedral inequality, and by summing the corresponding k dissimilarities from $\mathcal{S}$ on the right-hand side of the polyhedral inequality. □

14.4 Maximum distance

In the final section of this chapter on perimeter models we explore the multi-way extensions and properties of a somewhat different three-way function. For the three-way case, the maximum distance function is defined as

$$d_3(x_{1,3}) = \max\left[d(x_1, x_2), d(x_1, x_3), d(x_2, x_3)\right] \tag{14.20}$$

by both Heiser and Bennani (1997) and De Rooij and Gower (2003). Function (14.20) has two straightforward four-way generalizations, which are given by

$$d_4(x_{1,4}) = \max\left[d(x_1, x_2), d(x_1, x_3), d(x_1, x_4), d(x_2, x_3), d(x_2, x_4), d(x_3, x_4)\right] \tag{14.21}$$

and

$$d_4(x_{1,4}) = \max\left[d_3(x_{2,4}), d_3(x_{1,4}^{-2}), d_3(x_{1,4}^{-3}), d_3(x_{1,3})\right]$$

where $d_3(x_{1,3})$ is defined as in (14.20). Fortunately, the two formulations are equivalent.

The k-way formulation of (14.21) is given by

$$d_k(x_{1,k}) = \max\left[d(x_1, x_2), d(x_1, x_3), ..., d(x_{k-2}, x_k), d(x_{k-1}, x_k)\right]. \tag{14.22}$$

On the right-hand side of (14.22) we have the maximum dissimilarity that can be constructed from all pairs in the set $\{x_1, x_2, ..., x_k\}$. The multi-way function in (14.22) satisfies inequality (14.13) due to the following result.

Proposition 14.6. *Let $d(x_i, x_j) = 0$ if and only if $x_i = x_j$. Then $d_k(x_{1,k})$ in* (14.22) *satisfies* (14.13) *if $d(x_i, x_j)$ satisfies the triangle inequality.*
Proof for $k = 3$: It must be shown that

$$\begin{aligned}&2\max\left[d(x_1,x_2),d(x_1,x_3),d(x_2,x_3)\right]\leq\\&\quad\max\left[d(x_2,x_3),d(x_2,x_4),d(x_3,x_4)\right]+\max\left[d(x_1,x_3),d(x_1,x_4),d(x_3,x_4)\right]+\\&\quad\max\left[d(x_1,x_2),d(x_1,x_4),d(x_2,x_4)\right]\end{aligned} \tag{14.23}$$

holds. The proof is immediate if the maximum of the six dissimilarities is $d(x_i, x_4)$ for $i = 1, 2, 3$. For instance, if $d(x_1, x_4)$ is the largest, then (14.23) becomes

$$\begin{aligned}2\max\left[d(x_1,x_2),d(x_1,x_3),d(x_2,x_3)\right]\leq&2d(x_1,x_4)+\\&\max\left[d(x_2,x_3),d(x_2,x_4),d(x_3,x_4)\right]\end{aligned}$$

which is true, since $d(x_1, x_4) \geq \max\left[d(x_1, x_2), d(x_1, x_3), d(x_2, x_3)\right]$. Furthermore, suppose $d(x_1, x_2)$ is the maximum of the six values. Then (14.23) can be written as

$$\begin{aligned}d(x_1,x_2)\leq&\max\left[d(x_1,x_3),d(x_1,x_4),d(x_3,x_4)\right]+\\&\max\left[d(x_2,x_3),d(x_2,x_4),d(x_3,x_4)\right].\end{aligned} \tag{14.24}$$

Inequality (14.24) is true if the triangle inequality holds, which completes the proof for $k = 3$.
Proof for $k = 4$: It must be verified that

$$\begin{aligned}&3\max\left[d(x_1,x_2),d(x_1,x_3),d(x_1,x_4),d(x_2,x_3),d(x_2,x_4),d(x_3,x_4)\right]\leq\\&\quad\max\left[d(x_1,x_2),d(x_1,x_3),d(x_1,x_5),d(x_2,x_3),d(x_2,x_5),d(x_3,x_5)\right]+\\&\quad\max\left[d(x_1,x_2),d(x_1,x_4),d(x_1,x_5),d(x_2,x_4),d(x_2,x_5),d(x_4,x_5)\right]+\\&\quad\max\left[d(x_1,x_3),d(x_1,x_4),d(x_1,x_5),d(x_3,x_4),d(x_3,x_5),d(x_4,x_5)\right]+\\&\quad\max\left[d(x_2,x_3),d(x_2,x_4),d(x_2,x_5),d(x_3,x_4),d(x_3,x_5),d(x_4,x_5)\right].\end{aligned} \tag{14.25}$$

Again, the proof is immediate if the largest of the ten dissimilarities is $d(x_i, x_5)$ for $i = 1, ..., 4$. Suppose $d(x_1, x_2)$ is the maximum of the ten values. Then (14.25) can be written as

$$\begin{aligned}&d(x_1,x_2)\leq\\&\max\left[d(x_1,x_3),d(x_1,x_4),d(x_1,x_5),d(x_3,x_4),d(x_3,x_5),d(x_4,x_5)\right]+\\&\max\left[d(x_2,x_3),d(x_2,x_4),d(x_2,x_5),d(x_3,x_4),d(x_3,x_5),d(x_4,x_5)\right].\end{aligned} \tag{14.26}$$

Inequality (14.26) is true if the triangle inequality holds, which completes the proof for $k = 4$.
General proof: From the proof for $k = 3$ and $k = 4$, the following pattern becomes apparent. After filling in (14.22) in (14.13), there are $k(k + 1)/2$ different two-way dissimilarities to consider. The proof is immediate if $d(x_i, x_{k+1})$ for $i = 1, 2, ..., k$ is the largest dissimilarity. This part of the proof does not require the triangle inequality. If any of the other dissimilarities is the largest, then (14.22) satisfies (14.13) if the triangle inequality holds. □

14.5 Epilogue

In this chapter multi-way generalizations of two three-way functions, the perimeter distance and the maximum distance, were presented. The extended perimeter distance is based on two-way dissimilarities or on $(k-1)$-way dissimilarities. The resulting multi-way perimeter models are different and posses different properties. We studied decompositions of the perimeter models for ordered tuples, not for tuples with distinct elements. The decomposition of the three-way perimeter model for triples with distinct elements can be found in Chepoi and Fichet (2007), Bennani-Dosse (1993) and Gower and De Rooij (2003). The case has not been studied here, but it may be noted that the decompositions of the two four-way perimeter models defined on tuples with distinct elements, provide similar and interesting formulas.

The maximum function may also be defined on two-way dissimilarities or on $(k-1)$-way dissimilarities; the different definitions are equivalent. Both the generalized perimeter distance and the maximum distance satisfy polyhedral inequality (12.4).

Validity of a multi-way metric inequality for $k \geq 3$ appears not to be important for methods used for multi-way dimensional scaling (Cox, Cox and Branco, 1991; Heiser and Bennani, 1997; Gower and De Rooij, 2003). The results in Section 14.3 therefore appear to be of theoretical interest only. From a theoretical point of view it is unfortunate that no well-established basic multi-way metric structure emerged from the study.

Perimeter models are simple functions with a straightforward interpretation. However, some empirical evidence suggests that using perimeter models is not the best approach to evaluating complex relationships between three or more objects at a time. Gower and De Rooij (2003) used the three-way perimeter model and compared multidimensional scaling of three-way distances to the scaling of two-way distances. These authors concluded that, when the three-way distances were linear transformations of the two-way information, the three-way analysis gained little or nothing over the conventional multidimensional scaling. De Rooij (2001, Chapter 5; 2002) noted that the problem seems to be that definitions of three-way distances in terms of two-way distances do not model true three-way interactions.

CHAPTER 15

Generalizations of Theorem 10.3

For the properties in this section we have a new use of the symbols a, b, c, and d already used for the 2×2 contingency table in Part I. With two-way dissimilarities, a function is called metric if it satisfies, among other things, the triangle inequality. Theorem 10.3 states that which states that if c is a positive constant and the two-way dissimilarity d satisfies the triangle inequality, then the function $d/(c+d)$ satisfies the triangle inequality. In this chapter generalizations of Theorem 10.3 for the triangle inequality are considered.

For the use in this chapter it suffices to define a multi-way metric on the number of dissimilarities involved. Multi-way dissimilarities can be used to measure the resemblance between two or more, say k, objects. Let d_i, $i = 1, 2, ..., n, n+1$ denote $n+1$ multi-way dissimilarities. A generalization of Theorem 10.3 is presented for the inequality

$$d_{n+1} \leq \sum_{i=1}^{n} d_i. \tag{15.1}$$

Furthermore, Conjecture 15.1 below is an attempt to generalize Theorem 10.3 to polyhedral inequality

$$(n-1) \times d_{n+1} \leq \sum_{i=1}^{n} d_i. \tag{15.2}$$

Inequality (15.2) portraits inequality (12.4) and (14.13) in the present simpler notation.

15.1 A generalization of Theorem 10.3.

Proposition 15.1 below is a first attempt to generalize Theorem 10.3, which states that if c is a positive constant and the two-way dissimilarity d satisfies the triangle inequality, then the function $d/(c+d)$ satisfies the triangle inequality. In Proposition 15.1 we consider the multi-way metrics that are characterized by (15.1). We first consider the proofs for $n = 2, 3, 4$. A general proof for Proposition 15.1 is straightforward after considering these proofs.

Proposition 15.1 *If the dissimilarities d_i for $i = 1, 2, ..., n, n+1$ satisfy n-way symmetry, then*

$$\frac{d_{n+1}}{c+d_{n+1}} \leq \sum_{i=1}^{n} \frac{d_i}{c+d_i} \quad \textit{if} \quad d_{n+1} \leq \sum_{i=1}^{n} d_i \quad \textit{holds.}$$

Proof for $n = 2$: It must be shown that the quantity a given by

$$\begin{aligned} a &= (c+d_1)(c+d_2)(c+d_3)\left[\frac{d_1}{c+d_1} + \frac{d_2}{c+d_2} - \frac{d_3}{c+d_3}\right] \\ &= d_1(c+d_2)(c+d_3) + d_2(c+d_1)(c+d_3) - d_3(c+d_1)(c+d_2) \\ &= c^2(d_1+d_2-d_3) + 2cd_1d_2 + d_1d_2d_3. \end{aligned}$$

is positive. Since $d_3 \leq d_1 + d_2$ under the conditions of the assertion, the quantity a is positive, which completes the proof for $n = 2$.
Proof for $n = 3$: It must be shown that the quantity a given by

$$\begin{aligned} a =& (c+d_1)(c+d_2)(c+d_3)(c+d_4)\left[\frac{d_1}{c+d_1} + \frac{d_2}{c+d_2} + \frac{d_3}{c+d_3} - \frac{d_4}{c+d_4}\right] \\ =& d_1(c+d_2)(c+d_3)(c+d_4) + \\ & d_2(c+d_1)(c+d_3)(c+d_4) + \\ & d_3(c+d_1)(c+d_2)(c+d_4) - d_4(c+d_1)(c+d_2)(c+d_3) \end{aligned}$$

is positive. Expanding the equation in polynomial form we obtain

$$\begin{aligned} a =& c^3(d_1+d_2+d_3-d_4) + 2c^2(d_1d_2 + 2d_1d_3 + 2d_2d_3) + \\ & c(3d_1d_2d_3 + d_1d_2d_4 + d_1d_3d_4 + d_2d_3d_4) + 2d_1d_2d_3d_4. \end{aligned}$$

Only the coefficient of c^3 needs to be checked since all other coefficients are positive. The coefficient of c^3 is positive if $d_4 \leq d_1 + d_2 + d_3$ (the condition of the assertion). This completes the proof for $n = 3$.

Proof for $n = 4$: It must be shown that the quantity a given by

$$\begin{aligned}
a =& \prod_{i=1}^{5}(c+d_i)\left[\sum_{i=1}^{4}\frac{d_i}{c+d_i} - \frac{d_5}{c+d_5}\right] \\
=& d_1(c+d_2)(c+d_3)(c+d_4)(c+d_5)+ \\
& d_2(c+d_1)(c+d_3)(c+d_4)(c+d_5)+ \\
& d_3(c+d_1)(c+d_2)(c+d_4)(c+d_5)+ \\
& d_4(c+d_1)(c+d_2)(c+d_3)(c+d_5) - d_5(c+d_1)(c+d_2)(c+d_3)(c+d_4)
\end{aligned}$$

is positive. Expanding the equation in polynomial form we obtain

$$\begin{aligned}
a =& c^4(d_1+d_2+d_3+d_4-d_4) \\
&+ 2c^3(d_1d_2+d_1d_3+d_1d_4+d_2d_3+d_2d_4+d_3d_4) \\
&+ 3c^2(d_1d_2d_3+d_1d_2d_4+d_1d_3d_4+d_2d_3d_4) \\
&+ 2c^2d_5(d_1d_2+d_1d_3+d_1d_4+d_2d_3+d_2d_4+d_3d_4) \\
&+ 4cd_1d_2d_3d_4 + 2d_5(d_1d_2d_3+d_1d_2d_4+d_1d_3d_4+d_2d_3d_4) \\
&+ 3d_1d_2d_3d_4d_5.
\end{aligned}$$

Only the coefficient of c^4 needs to be checked since all other coefficients are positive. The coefficient of c^4 is positive under the conditions of the assertion. This completes the proof for $n = 4$.

Outline general proof: It must be shown that the quantity

$$a = \prod_{i=1}^{n+1}(c+d_i) \times \left[\sum_{i=1}^{n}\frac{d_i}{c+d_i} - \frac{d_{n+1}}{c+d_{n+1}}\right]$$

is positive. After expanding the equation in polynomial form only the coefficient of c^n needs to be checked. This coefficient is positive under the conditions of the assertion. □

Conjecture 15.1 in Section 15.3 is a (potentially) stronger result compared to Proposition 15.1. With Conjecture 15.1 we attempt to prove Proposition 15.1 not for inequality (15.1), but for inequality (15.2). Before presenting this attempt, the next section is first used to present some auxiliary results.

15.2 Auxiliary results

We first repeat Proposition 12.1 in Proposition 15.2, using the more convenient notation.

Proposition 15.2. *If the dissimilarities d_i for $i = 1, 2, ..., n, n+1$ satisfy n-way symmetry, then (for $n \geq 3$)* (15.2) *implies*

$$(n-2)d_n \leq \sum_{i=1}^{n-1} d_i.$$

Proof: Interchanging the roles of d_n and d_{n+1} and dividing by $n-1$ in (15.2), we may obtain the inequalities

$$(n-1)d_n \leq d_{n+1} + \sum_{i=1}^{n-1} d_i$$

and

$$d_{n+1} \leq \left[\frac{1}{n-1}\right] \sum_{i=1}^{n} d_i.$$

Adding the two inequalities and multiplying by $(n-1)/n$ gives the required inequality. □

The inequality in Proposition 15.4 below concerns one of the inequalities required in Conjecture 15.1 below. First, we present a stronger result, which is then used in the proof of Proposition 15.4.

Proposition 15.3. *Dissimilarities d_i for $i = 1, 2, ..., n, n+1$ satisfy*

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n+1} d_i d_j \geq \left[\frac{n^2-n-1}{2(n-1)}\right] \sum_{i=1}^{n+1} d_i^2$$

if (15.2) *holds.*
Proof: Inequality (15.2) can be written as

$$d_1 \geq (n-1)d_{n+1} - \sum_{i=2}^{n} d_i. \tag{15.3}$$

Squaring both sides of (15.3) we obtain

$$d_1^2 \geq (n-1)^2 d_{n+1}^2 + \sum_{i=2}^{n} d_i^2 - 2(n-1)d_{n+1} \sum_{i=2}^{n} d_i + 2 \sum_{i=2}^{n} \sum_{j=i+1}^{n+1} d_i d_j \tag{15.4}$$

(for $n = 2$ the last term of the inequality equals zero).

There are $(n+1)$ variants of d_i^2 in (15.4) and $n(n+1)/2$ variants of $d_i d_j$. The number of variants of the inequality is given by the smallest common multiple of $(n+1)$ and $n(n+1)/2$. Instead, consider the multiple $n(n+1)^2/2$. Adding up all $n(n+1)^2/2$ variants of (15.4) we obtain

$$\begin{aligned}\frac{n(n+1)}{2}\sum_{i=1}^{n+1} d_i^2 \geq & \frac{(n-1)^2 n(n+1)}{2}\sum_{i=1}^{n+1} d_i^2 \\ & + \frac{(n-1)n(n+1)}{2}\sum_{i=1}^{n+1} d_i^2 \\ & - 2(n-1)^2(n+1)\sum_{i=1}^{n}\sum_{j=i+1}^{n+1} d_i d_j \\ & + (n-1)(n-2)(n+1)\sum_{i=1}^{n}\sum_{j=i+1}^{n+1} d_i d_j\end{aligned}$$

which equals the required inequality. This completes the proof. □

The inequality in Proposition 15.4 is one of the inequalities required in Conjecture 15.1 in Section 15.3. The proof of this inequality makes use of the stronger result in Proposition 15.3.

Proposition 15.4. *Dissimilarities d_i for $i = 1, 2, ..., n, n+1$ satisfy*

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_i d_j \geq \left[\frac{n-2}{2}\right] d_{n+1}\sum_{i=1}^{n} d_i$$

if (15.2) *holds.*
Proof: Using the equality

$$\left[\sum_{i=1}^{n} d_i\right]^2 - \sum_{i=1}^{n} d_i^2 = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_i d_j$$

the quantity a given by

$$a = 2(n-1)\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_i d_j - (n-1)(n-2)d_{n+1}\sum_{i=1}^{n} d_i$$

can be written as $a = b_1 + b_2$, where

$$b_1 = (n-2)\left[\sum_{i=1}^{n} d_i\right]\left[\sum_{i=1}^{n} d_i - (n-1)d_{n+1}\right]$$

and

$$b_2 = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_i d_j - (n-2)\sum_{i=1}^{n} d_i^2.$$

The assertion follows if it can be shown that the quantity a is positive. Under the condition of the proposition (the last term of) quantity b_1 is positive. Furthermore, in the light of Proposition 15.3, quantity b_2 is positive since

$$\frac{n^2-n-1}{2(n-1)} \geq \frac{n-2}{2}.$$

Hence quantity a is positive, which completes the proof. □

15.3 A stronger generalization of Theorem 10.3

Conjecture 15.1 below is an attempt to generalize Theorem 10.3, which states that if c is a positive constant and the two-way dissimilarity d satisfies the triangle inequality, then the function $d/(c+d)$ satisfies the triangle inequality. Below, proofs for small n are presented, but no proof is offered for any n. With respect to Conjecture 15.1, it is assumed that the multi-way dissimilarities satisfy n-way symmetry, which makes the use of Proposition 15.2 possible. Note that also for $n = 2$, Theorem 10.3 is a special case of Conjecture 15.1.

Conjecture 15.1 *If the dissimilarities d_i for $i = 1, 2, ..., n, n+1$ satisfy n-way symmetry, then*

$$\frac{(n-1)d_{n+1}}{c+d_{n+1}} \leq \sum_{i=1}^{n} \frac{d_i}{c+d_i}$$

if (15.2) *holds.*

Proof for $n = 2$: It must be shown that the quantity a given by

$$\begin{aligned} a &= (c+d_1)(c+d_2)(c+d_3)\left[\frac{d_1}{c+d_1} + \frac{d_2}{c+d_2} - \frac{d_3}{c+d_3}\right] \\ &= d_1(c+d_2)(c+d_3) + d_2(c+d_1)(c+d_3) - d_3(c+d_1)(c+d_2) \\ &= c^2(d_1+d_2-d_3) + 2cd_1d_2 + d_1d_2d_3 \end{aligned}$$

is positive. Since $d_3 \leq d_1 + d_2$ by Proposition 15.2, the quantity a is positive, which completes the proof for $n = 2$.

Proof for $n = 3$: It must be shown that the quantity a given by

$$\begin{aligned}a =&(c+d_1)(c+d_2)(c+d_3)(c+d_4)\left[\frac{d_1}{c+d_1}+\frac{d_2}{c+d_2}+\frac{d_3}{c+d_3}-\frac{2d_4}{c+d_4}\right]\\ =&d_1(c+d_2)(c+d_3)(c+d_4)+\\ &d_2(c+d_1)(c+d_3)(c+d_4)+\\ &d_3(c+d_1)(c+d_2)(c+d_4)-2d_4(c+d_1)(c+d_2)(c+d_3)\end{aligned}$$

is positive. Expanding the equation in polynomial form we obtain

$$\begin{aligned}a =&c^3(d_1+d_2+d_3-2d_4)+\\ &c^2(2d_1d_2+2d_1d_3+2d_2d_3-d_1d_4-d_2d_4-d_3d_4)+\\ &3cd_1d_2d_3+d_1d_2d_3d_4.\end{aligned}$$

The coefficient of c^3 is positive if $2d_4 \le d_1+d_2+d_3$. The coefficient of c^2 is positive if $d_3 \le d_1+d_2$, since it can be written as

$$d_1(d_2+d_3-d_4)+d_2(d_1+d_3-d_4)+d_3(d_1+d_2-d_4).$$

Thus, the quantity a is positive by Proposition 15.2, which completes the proof for $n = 3$.
Proof for $n = 4$: It must be shown that the quantity a given by

$$\begin{aligned}a =&\prod_{i=1}^{5}(c+d_i)\left[\sum_{i=1}^{4}\frac{d_i}{c+d_i}-\frac{3d_5}{c+d_5}\right]\\ =&d_1(c+d_2)(c+d_3)(c+d_4)(c+d_5)+\\ &d_2(c+d_1)(c+d_3)(c+d_4)(c+d_5)+\\ &d_3(c+d_1)(c+d_2)(c+d_4)(c+d_5)+\\ &d_4(c+d_1)(c+d_2)(c+d_3)(c+d_5)-3d_5(c+d_1)(c+d_2)(c+d_3)(c+d_4)\end{aligned}$$

is positive. Expanding the equation in polynomial form we obtain

$$\begin{aligned}a =\ &c^4(d_1+d_2+d_3+d_4-3d_4)\\ &+2c^3(d_1d_2+d_1d_3+d_1d_4+d_2d_3+d_2d_4+d_3d_4-d_1d_5-d_2d_5-d_3d_4-d_4d_5)\\ &+3c^2(d_1d_2d_3+d_1d_2d_4+d_1d_3d_4+d_2d_3d_4)\\ &-c^2d_5(d_1d_2+d_1d_3+d_1d_4+d_2d_3+d_2d_4+d_3d_4)\\ &+4cd_1d_2d_3d_4+d_1d_2d_3d_4d_5.\end{aligned}$$

The coefficient of c^4 is positive if $3d_5 \le d_1+d_2+d_3+d_4$. The coefficient of c^3 is positive if $2d_4 \le d_1+d_2+d_3$, since it can be written as

$$\begin{aligned}&d_1(d_2+d_3+d_4-2d_5)+d_2(d_1+d_3+d_4-2d_5)+\\ &d_3(d_1+d_2+d_4-2d_5)+d_4(d_1+d_2+d_3-3d_5).\end{aligned}$$

Alternatively, the coefficient of c^3 is positive by Proposition 15.4.

The coefficient of c^2 is positive if $d_3 \leq d_1 + d_2$, since it can be written as

$$d_1 d_2(d_3 + d_4 - d_5) + d_1 d_3(d_2 + d_4 - d_5) + d_1 d_4(d_2 + d_3 - d_5)+$$
$$d_2 d_3(d_1 + d_4 - d_5) + d_2 d_4(d_1 + d_3 - d_5) + d_3 d_4(d_1 + d_2 - d_5).$$

Thus, the quantity a is positive by Proposition 15.2, which completes the proof for $n = 4$.
Proof for $n = 5$*:* It must be shown that the quantity a given by

$$a = \prod_{i=1}^{6}(c + d_i) \left[\sum_{i=1}^{5} \frac{d_i}{c + d_i} - \frac{4d_6}{c + d_6} \right]$$

is positive. Quantity a can be written as

$$\begin{aligned} a = {} & d_1(c + d_2)(c + d_3)(c + d_4)(c + d_5)(c + d_6) \\ & + d_2(c + d_1)(c + d_3)(c + d_4)(c + d_5)(c + d_6) \\ & + d_3(c + d_1)(c + d_2)(c + d_4)(c + d_5)(c + d_6) \\ & + d_4(c + d_1)(c + d_2)(c + d_3)(c + d_5)(c + d_5) \\ & + d_5(c + d_1)(c + d_2)(c + d_3)(c + d_4)(c + d_5) \\ & - 4d_6(c + d_1)(c + d_2)(c + d_3)(c + d_4)(c + d_5). \end{aligned}$$

Expanding the equation in polynomial form we obtain

$$\begin{aligned} a = {} & c^5 \left[\sum_{i=1}^{5} d_i - 4d_6 \right] \\ & + c^4 \left[2 \sum_{i=1}^{4} \sum_{j=i+1}^{5} d_i d_j - 3d_6 \sum_{i=1}^{5} d_i \right] \\ & + c^3 \left[3 \sum_{i=1}^{3} \sum_{j=i+1}^{4} \sum_{r=j+1}^{5} d_i d_j d_r - 2d_6 \sum_{i=1}^{4} \sum_{j=i+1}^{5} d_i d_j \right] \\ & + c^2 \left[4 \sum_{i=1}^{2} \sum_{j=i+1}^{3} \sum_{r=j+1}^{4} \sum_{s=r+1}^{5} d_i d_j d_r d_s - d_6 \sum_{i=1}^{3} \sum_{j=i+1}^{4} \sum_{l=j+1}^{5} d_i d_j d_l \right] \\ & + 5c \prod_{i=1}^{5} d_i + \prod_{i=1}^{6} d_i. \end{aligned}$$

The coefficient of c^5 is positive if $4d_6 \leq \sum_{i=1}^{5} d_i$. The coefficient of c^4 is positive if $3d_5 \leq \sum_{i=1}^{4} d_i$, since it can be written as

$$d_1(d_2 + d_3 + d_4 + d_5 - 3d_6) + d_2(d_1 + d_3 + d_4 + d_5 - 3d_6)+$$
$$d_3(d_1 + d_2 + d_4 + d_5 - 3d_6) + d_4(d_1 + d_2 + d_3 + d_5 - 3d_6)+$$
$$d_5(d_1 + d_2 + d_3 + d_4 - 3d_6).$$

Alternatively, the coefficient of c^4 is positive by Proposition 15.4.

The coefficient of c^3 is positive if $2d_4 \leq \sum_{i=1}^{3} d_i$, since it can be written as

$$\begin{aligned}
&d_1d_2(d_3 + d_4 + d_5 - 2d_6) + d_1d_3(d_2 + d_4 + d_5 - 2d_6)+\\
&d_1d_4(d_2 + d_3 + d_5 - 2d_6) + d_1d_5(d_2 + d_3 + d_4 - 2d_6)+\\
&d_2d_3(d_1 + d_4 + d_5 - 2d_6) + d_2d_4(d_1 + d_3 + d_5 - 2d_6)+\\
&d_2d_5(d_1 + d_3 + d_4 - 2d_6) + d_3d_4(d_1 + d_2 + d_5 - 2d_6)+\\
&d_3d_5(d_1 + d_2 + d_4 - 2d_6) + d_4d_5(d_1 + d_2 + d_3 - 2d_6).
\end{aligned}$$

The coefficient of c^2 is positive if $d_3 \leq d_1 + d_2$, since it can be written as

$$\begin{aligned}
&d_1d_2d_3(d_4 + d_5 - d_6) + d_1d_2d_4(d_3 + d_5 - d_6)+\\
&d_1d_2d_5(d_3 + d_4 - d_6) + d_1d_3d_4(d_2 + d_5 - d_6)+\\
&d_1d_3d_5(d_2 + d_4 - d_6) + d_1d_4d_5(d_2 + d_3 - d_6)+\\
&d_2d_3d_4(d_1 + d_5 - d_6) + d_2d_3d_5(d_1 + d_4 - d_6)+\\
&d_2d_4d_5(d_1 + d_3 - d_6) + d_3d_4d_5(d_1 + d_2 - d_6).
\end{aligned}$$

Hence, a is positive, which completes the proof $n = 5$.
Outline general proof: It must be shown that the quantity

$$a = \prod_{i=1}^{n+1}(c + d_i) \times \left[\sum_{i=1}^{n} \frac{d_i}{c + d_i} - \frac{(n-1)d_{n+1}}{c + d_{n+1}}\right]$$

is positive. Due to Proposition 15.2 each metric inequality also implies all weaker metric inequalities. The quantity a can be written as a polynomial function of $c^n, c^{n-1}, ..., c^2, c$ and a constant $\prod_{i=1}^{n+1} d_i$. The coefficient belonging to the linear part c and the constant $\prod_{i=1}^{n+1} d_i$ are always positive. It must be shown that the remaining $(n-1)$ coefficients are also positive. The coefficient corresponding to c^n appears to be positive if the metric inequality $(n-1)d_{n+1} \leq \sum_{i=1}^{n} d_i$ holds.

15.4 Epilogue

Theorem 10.3, which states that if two-way dissimilarity d satisfies the triangle inequality, then so does the function $d/(c + d)$, was generalized to the multi-way case in this chapter. In the first generalization, Proposition 15.1, multi-way metrics were considered that are characterized by inequality $d_{n+1} \leq \sum_{i=1}^{n} d_i$. In the second attempt, Conjecture 15.1, we tried to proof the generalization for the stronger class of multi-way metrics characterized by $(n-1)d_{n+1} \leq \sum_{i=1}^{n} d_i$. The proof of Proposition 15.1 turned out to be straightforward, especially in contrast to the proof of Conjecture 15.1.