absence–absence match. Then

$$d_{rs} = \frac{\sum_f I^f_{rs} d^f_{rs}}{\sum_f I^f_{rs}}. \tag{9.2}$$

The classical or metric method

In the classical or metric method of multidimensional scaling, often known as *principal coordinate analysis* (Gower, 1966) but going back to Schoenberg (1935), Young & Householder (1938) and Torgerson (1952, 1958), we assume that the dissimilarities were derived as Euclidean distances between n points in p dimensions, for unknown p. Given the distances, we obviously cannot recover the observations themselves, since the distances are invariant to rigid motions (translations, rotation and reflections) of $\mathbb{R}^p$. It transpires that this is the only freedom allowed.

Proposition 9.5 *For any symmetric matrix T, define the matrix*

$$T' = -\frac{1}{2}\left[T - \frac{(T\mathbf{1})\mathbf{1}^T}{n} - \frac{\mathbf{1}(T\mathbf{1})^T}{n} + \frac{\mathbf{1}^T T\mathbf{1}}{n^2}\right]$$

by subtracting row and column means and adding back the overall mean, or, equivalently, by removing row means then column means.

(a) *Given any configuration X of n points in $\mathbb{R}^p$, the matrix $T = (d^2_{rs} = \|\mathbf{x}_r - \mathbf{x}_s\|^2)$ gives a non-negative definite $T' = XX^T$. Such a set of distances is called* Euclidean.

(b) *Given a symmetric $n \times n$ matrix T with non-negative definite T', we can find a configuration of points in $\mathbb{R}^{(n-1)}$ such that $T = (d^2_{rs})$.*

(c) *A necessary and sufficient condition for an $n \times n$ matrix T to be a squared distance matrix is that $\mathbf{w}^T T\mathbf{w} \leqslant 0$ for all $\mathbf{w}$ with $\mathbf{w}^T\mathbf{1} = 0$.*

(d) *Any two configurations of n points with the same (d^2_{rs}) differ only by a shift and a rigid motion of $\mathbb{R}^p$, so lie in (shifted) subspaces of the same minimal dimension, the rank of T'.*

Proof: (a) Without loss of generality, centre the data so every column of X has zero mean. Then $T = (\|\mathbf{x}_r - \mathbf{x}_s\|^2) = (\|\mathbf{x}_r\|^2 + \|\mathbf{x}_s\|^2 - 2\mathbf{x}_r^T\mathbf{x}_s) = E\mathbf{1}^T + \mathbf{1}E^T - 2XX^T$ where $E = (\|\mathbf{x}_r\|^2)$. Let $e = E^T\mathbf{1}$ so $T\mathbf{1} = nE + e\mathbf{1}$ and $\mathbf{1}^T E\mathbf{1} = 2ne$. Thus

$$\begin{aligned} -2T' &= E\mathbf{1}^T + \mathbf{1}E^T - 2XX^T - E\mathbf{1}^T - e\mathbf{1}\mathbf{1}^T/n \\ &\qquad - \mathbf{1}E^T - e\mathbf{1}\mathbf{1}^T/n + 2ne\mathbf{1}\mathbf{1}^T/n^2 \\ &= -2XX^T \end{aligned}$$