J. R. Statist. Soc. A (2008)
171, *Part* 1, *pp.* 137–157

The analysis of change, Newton's law of gravity and association models

Mark de Rooij

Leiden University, The Netherlands

[Received September 2004. Final revision April 2007]

Summary. Newton's law of gravity states that the force between two objects in the universe is equal to the product of the masses of the two objects divided by the square of the distance between the two objects. In the first part of the paper it is shown that a model with a 'law-of-gravity' interpretation applies well to the analysis of longitudinal categorical data where the number of people changing their behaviour or choice from one category to another is a measure of force and the goal is to obtain estimates of mass for the two categories and an estimate of the distance between them. To provide a better description of the data dynamic masses and dynamic positions are introduced. It is shown that this generalized law of gravity is equivalent to Goodman's RC(M) association model. In the second part of the paper the model is generalized to two kinds of three-way data. The first case is when there are multiple two-way tables and in the second case we have change over three points of time. Conditional and partial association models are related to three-way distance models, like the INDSCAL model, and triadic distance models respectively.

Keywords: Categorical data; Euclidean distance; Gravity model; Longitudinal data; Square tables; Triadic distance

1. Introduction

This paper will be concerned with longitudinal categorical data, i.e. repeated measurements on a number of observational units with the same instrument. The main interest in studying longitudinal data is whether change occurred and, if so, what the nature of the change is. We shall confine ourselves to the case of categorical data. Our questions concern qualitative change, i.e. changes in attitude, opinion, behaviour or any other categorical variable. This is typically different from continuous data where it might be possible to describe change in terms of better or worse; for categorical data descriptions are in terms of 'different' or 'the same'.

Once longitudinal categorical data have been collected they can be represented in transition frequency tables, which are contingency tables where each way corresponds to the categories of a variable measured at a specific time point (we adopt the way mode terminology for the tables of Carroll and Arabie (1980)). The number of time points defines the number of ways of the transition frequency table. Having measured a group of people twice on a categorical variable, a square transition frequency table arises. If measurements are obtained at three time points the data can be gathered in a three-way contingency table, and so forth.

An example of such data is obtained from Upton and Särlvik (1981) who discussed changes in political voting in Sweden. The data are shown in Table 1. There are five political parties: the *Communists* COM; the *Social Democrats* SD; the *Centre Party* C; the *People's Party* P; the *Conservatives* CON. These are the anglicized names following Upton and Särlvik (1981). The rows correspond to the political parties in 1964 (in capital letters); the columns to the political parties in 1970 (lower-case letters).

Address for correspondence: Mark de Rooij, Methodology and Statistics Unit, Institute for Psychological Research, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands.
E-mail: rooijm@fsw.leidenuniv.nl

 0964–1998/08/171137

Table 1. Swedish voting data representing voting changes from 1964 (rows) to 1970 (columns)†

	com	sd	c	p	con
COM	(22)	27	4	1	0
SD	16	(861)	57	30	8
C	4	26	(248)	14	7
P	8	20	61	(201)	11
CON	0	4	31	32	(140)

†From Upton and Särlvik (1981).

The focus will be on change, i.e. on the off-diagonal entries. The values on the diagonal are within parentheses; for these cells special parameters (which are often called loyalty parameters) will be included in the models to be developed.

The question, once we have such change data, is not *whether* there is association but *what the pattern* of association looks like. We shall propose a model for these data based on Newton's law of gravity, which states that the force between any two objects in the universe is proportional to the masses of the two objects and inversely related to the squared distance between the two objects (Newton's law of gravity will be discussed in more detail in the next section). This deterministic model will be applied to the analysis of change where the objects in the universe are the categories of the variable. The force between two objects is measured by the number of respondents making a transition from one category to another. This number is not accurately measured, however, since a sample is obtained from a population. Therefore, the law of gravity is used as a probabilistic model assuming a multinomial sampling scheme (which is the usual sampling scheme for such data). The force will be modelled by the mass of the two categories and a function of the distance between the two objects.

The remainder of this paper is organized as follows. The next section discusses Newton's law of gravity in more detail. Section 3 describes the analysis of change in terms of Newton's law of gravity and introduces dynamic elements in the law to adapt for different data settings. After introducing the dynamic elements it will be shown that the model is a reparameterization of the RC(M) association model (Goodman, 1979, 1991). The usual identification constraints for this model, however, are not suited to the analysis of change. A new way of identifying the solution will be presented and finally the model will be applied to the data in Table 1. In Section 4 the model will be generalized to the case of multiple two-way tables. The gravity models that are developed are related to conditional association models (Clogg, 1982), but again the usual identification constraints are not suited to the analysis of change. Section 5 treats gravitational models for three time points. These models are related to partial association models (Clogg, 1982). Identification and an application to empirical data will be discussed. We shall conclude with discussion and reflection about the results obtained and show some limitations of the work presented.

2. Newton's law of gravity

One of the major laws of the natural sciences is Newton's law of gravity:

> 'All matter attracts all other matter with a force proportional to the product of their masses and inversely proportional to the square of the distance between them'.

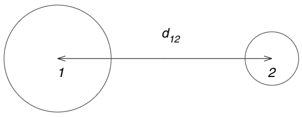

Fig. 1. Newton's law of gravity: the masses of objects 1 and 2 are represented by the area of the circles and d_{12} is the distance between the centres of the two objects

This law can be written in a formula as

$$\mathcal{F}_{ij} \propto \frac{m(i)\,m(j)}{d_{ij}^2}, \tag{1}$$

where $\mathcal{F}_{ij}$ denotes the force between objects i and j, $m(i)$ and $m(j)$ are the masses of the objects i and j, and d_{ij}^2 is the squared distance between the objects. A more general formulation of the law is

$$\mathcal{F}_{ij} \propto \frac{m(i)\,m(j)}{g(d_{ij})}, \tag{2}$$

where $g(\cdot)$ is $g(x)=x^2$, but may also be some other function. A graphical representation is shown in Fig. 1, where the masses are represented by the area of a circle. Newton explained a wide range of previously unrelated phenomena by using this law: the eccentric orbits of comets, the tides and their variations, the precession of the Earth's axis and motion of the Moon as perturbed by the gravity of the Sun. This work made Newton an international leader in scientific research.

In the next section we shall show that the law of gravity applies well to the analysis of social change. Therefore, first some other definitions of the function $g(\cdot)$ are provided and dynamic elements are introduced. The most general model that results is a reparameterization of a well-known model in statistics and social research, the RC(M) association model (Goodman, 1979, 1991). It should be noted, however, that by the time that we arrive at the RC(M) association model many properties of real forces as they are in the natural sciences have been lost. What is maintained is the interpretation in terms of mass and distance, and the analogy with Newton's law of gravity is meant more like a metaphor than reality.

3. The analysis of change

Where the task for Newton was to assess the force of the two objects on each other given their mass and their distance, we deal with the reverse problem. We assume that each object attracts people from other objects with some force. The resultant of these forces is flows of people between objects. These flows can be considered measurements of the attractional forces between objects, and thus (using the law of gravity) are the number of people going from one object to another proportional to the mass of the first object times the mass of the second object and inversely proportional to a function of the distance between the two objects.

In Table 1 it can be seen that there is a large number of people (57) who voted for the Social Democrats in 1964 and for the Centre Party in 1970, so there is a large force between these two categories. Similarly, the force between the Communists and the Conservatives is small (the frequency equals 0). Moreover, the force of one category on another is not equal to the reverse; for example, the force Communists–Social Democrats equals 27 and the force Social Democrats–Communists equals 16.

3.1. A Gaussian link

In Newton's law of gravity the distance is defined by a three-dimensional Euclidean distance, i.e. our universe is three dimensional. For the analysis of change the dimensionality is not known in advance but will be denoted by P. For the data in Table 1, for example, it is likely that the parties are differentiated on the standard 'left–right' dimension that is often found in political systems. Furthermore, there might be another dimension that differentiates the five parties. In Section 5, for example, we find a 'rural–urban' dimension on which the political parties differ. Often the dimensions are interpreted after the solution has been found, on the basis of practical knowledge of the data at hand. The co-ordinates of object i in P-dimensional space are given by the vector $\mathbf{z}(i)=(z_{i1},\ldots,z_{iP})^{\mathrm{T}}$. The $\mathbf{z}(i)$s will, in turn, be collected in the $I \times P$ matrix $\mathbf{Z}=(\mathbf{z}(1),\mathbf{z}(2),\ldots,\mathbf{z}(I))^{\mathrm{T}}$. The squared Euclidean distance between objects i and j is given by

$$d_{ij}^2(\mathbf{Z}) = \sum_{p=1}^{P} (z_{ip}-z_{jp})^2. \tag{3}$$

Other distance measures might be used as well, e.g. any distance from the Minkowski family (see Borg and Groenen (2005), chapter 17). Where in the law of gravity $g(x)=x^2$ we shall employ $g(x)=\exp(x^2)$, the Gaussian transformation or Gaussian link function (de Rooij and Heiser, 2005; Nosofsky, 1985), which is a monotone function. Again, like for the distance formulation, other transformation functions might be used, but in Section 3.4 it will be shown that this function relates the law of gravity to a well-known model for the statistical analysis of contingency tables.

3.2. Dynamic masses

As discussed above the measured forces are not symmetric, i.e. the force from Communist on Social Democrats is measured to be 27, whereas the reverse force is 16. The law of gravity assumes symmetric forces and the asymmetry is a form of 'error'.

To deal with such asymmetries the model will be generalized in two ways. The first generalization is to make the masses of the objects dependent on time. So, we deal with *dynamic masses*. It is quite natural that masses change in the social sciences, i.e. an object might be popular at one time point and unpopular at another. For example, in brand switching data some brands come into fashion at one moment and go out of fashion another. When an object is popular it has a large mass; when it is unpopular it has a small mass. For our model this means that objects have a mass at each time point, and that mass will be denoted by $m_t(i)$, the mass of object i at time point t. In a graphical representation (like Fig. 1), each object would have two circles.

3.3. Dynamic positions

A second generalization is to make the positions time dependent. So, *dynamic positions* are introduced into the model. An interpretation of a dynamic position is that the content of an object has changed. For example, a political party might change its election programme after it has lost dramatically in the last election or when a loss is in prospect, and thereby change its relative position towards other parties. Each object has a position for each time point which is denoted by $\mathbf{z}_t(i)=(z_{it1},\ldots,z_{itP})^{\mathrm{T}}$ and the positions of all objects at time point t are gathered in a matrix $\mathbf{Z}_t=(\mathbf{z}_t(1),\mathbf{z}_t(2),\ldots,\mathbf{z}_t(I))^{\mathrm{T}}$. The one-mode Euclidean distance (3) is replaced by a two-mode Euclidean distance:

$$d_{ij}^2(\mathbf{Z}_1;\mathbf{Z}_2) = \sum_{p=1}^{P} (z_{i1p}-z_{j2p})^2. \tag{4}$$

In the graphical representation each object is shown twice: once for each time point.

3.4. Rewriting the model

The model with dynamic masses and dynamic positions is

$$F_{ij} \propto \frac{m_1(i)\, m_2(j)}{\exp\{d_{ij}^2(\mathbf{Z}_1;\mathbf{Z}_2)\}}. \tag{5}$$

This gravity model can be rewritten in a form that is well known in statistics and is often applied in sociological studies; the RC(M) association model (Goodman, 1979, 1991). By back-transforming the parameters of this model, estimates of the masses and co-ordinates of our gravity model are obtained. Furthermore, relationships of this well-known model to other models for contingency tables are well established, and are then also valid for our gravity model. However, the usual graphical displays for the RC(M) association model are susceptible to misinterpretation (for examples see de Rooij and Heiser (2005)), whereas our interpretation is more intuitive. The transformation from gravity to association model is as follows (de Rooij and Heiser, 2005):

$$\begin{aligned} F_{ij} &\propto \frac{m_1(i)\, m_2(j)}{\exp\left\{ \sum_{p=1}^{P} (z_{i1p}^2 + z_{j2p}^2 - 2 z_{i1p} z_{j2p}) \right\}} \\ &\propto \frac{m_1(i)\, m_2(j)}{\exp\left(\sum_{p=1}^{P} z_{i1p}^2 \right) \exp\left(\sum_{p=1}^{P} z_{j2p}^2 \right) \exp\left(\sum_{p=1}^{P} -2 z_{i1p} z_{j2p} \right)}. \end{aligned} \tag{6}$$

Defining $\alpha(i) = m_1(i)/\exp(\Sigma_{p=1}^{P} z_{i1p}^2)$ and $\beta(j) = m_2(j)/\exp(\Sigma_{p=1}^{P} z_{j2p}^2)$, we obtain

$$\begin{aligned} F_{ij} &\propto \frac{\alpha(i)\,\beta(j)}{\exp\left(\sum_{p=1}^{P} -2 z_{i1p} z_{j2p} \right)} \\ &\propto \alpha(i)\,\beta(j)\, \exp\left(\sum_{p=1}^{P} 2 z_{i1p} z_{j2p} \right) \\ &\propto \alpha(i)\,\beta(j)\, \exp\left(\sum_{p=1}^{P} \phi_p \mu_{ip} \nu_{jp} \right), \end{aligned} \tag{7}$$

with $z_{i1p} = \phi_p^{1/2} \mu_{ip}/\sqrt{2}$ and $z_{j2p} = \phi_p^{1/2} \nu_{jp}/\sqrt{2}$. The last line in expression (7) is Goodman's (1979, 1991) RC(M) association model. In summary, we started with (an adaptation of) Newton's law of gravity, introduced dynamic elements and ended up with this well-known model. The RC(M) association model is a reduced rank model for the association which equals the saturated model when the dimensionality equals $I-1$ and which equals the (quasi-) independence model when the dimensionality is 0. The model with stable positions is the *homogeneous* RC(M) association model and imposes a symmetry restriction on the association, and thus is a special case of the quasi-symmetry model (Caussinus, 1965). The model with stable masses and positions is a special case of the symmetry model.

Since our focus is on the off-diagonal entries we need parameters for the diagonal entries of the table. These are loyalty parameters for each class, i.e. the model becomes

$$F_{ij} \propto \frac{m_1(i)\, m_2(j)\, \exp(\delta_{ij} \lambda_i)}{\exp\{d_{ij}^2(\mathbf{Z}_1;\mathbf{Z}_2)\}}, \tag{8}$$

where δ_{ij} equals 1 if $i = j$ and 0 otherwise. The λ_i are object-specific loyalty parameters, of which there are I. The effect of these parameters is that the observations on the diagonal do not influence the masses and the distances, i.e. the gravity model pertains to change. Another effect of these parameters is that the expected frequencies equal the observed frequencies for these cells.

3.5. Distances, distances and inner products

The RC(M) association model is often thought of as a model for ordinal data (although, strictly, nowhere is an ordinal restriction imposed on the scale values) and the parameters μ_{ip} and ν_{jp} are often interpreted in terms of distances. This raises the question what is new about the distance interpretation in the gravity model. To answer this question we should distinguish between *within-set distances* and *between-set distances*. In the RC(M) association model distances within the set of row points can be interpreted such that, when $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_{i'}$, with $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{iP})^{\mathrm{T}}$, are (approximately) equal, categories i and i' have the same pattern of association to the column categories. To interpret the relationship between the row and column sets of categories in the RC(M) association model an inner product rule must be used, where the association equals the product of the length of the two vectors $\boldsymbol{\mu}_i$ and $\boldsymbol{\nu}_j$ multiplied by the cosine of the angle between these two vectors. The parameterization in terms of Newton's law of gravity provides a between-set distance interpretation, i.e. an interpretation of the distance between $\mathbf{z}_1(i)$ and $\mathbf{z}_2(j)$.

3.6. Identification constraints

The RC(M) association model is not identified; it needs location, scaling and cross-dimensional constraints on the row and column scores. Usually the scores are centred, the sum of squares is set equal to 1 and the dimensions are made orthogonal. For the analysis of change, however, these standard identification constraints prevent substantial conclusions.

Let us denote the centred row scores by $\tilde{\mathbf{z}}_1(i)$. It is possible to transform these centred row scores linearly by $\mathbf{z}_1(i) = \mathbf{T}\,\tilde{\mathbf{z}}_1(i) + \mathbf{a}$ for diagonal $\mathbf{T}$ and a vector $\mathbf{a}$, and to adapt accordingly $\mathbf{z}_2(j) = \mathbf{T}^{-1}\,\tilde{\mathbf{z}}_2(j) - \mathbf{a}$ and the estimates of the masses without changing the expected frequencies. The vector $\mathbf{a}$ changes the mean position of the row points on each dimension whereas the diagonal matrix $\mathbf{T}$ changes the spread of the row points on each dimension. These transformations, however, do not change the order or relative spacings between row points; they are a dimensionwise linear transformation of the row points. The centred column points are transformed by using the inverse of this linear transformation. To obtain optimal location ($\mathbf{a}$) and scalings ($\mathbf{T}$) the correlation between squared distances ($d_{ij}^2(\mathbf{Z}_1;\mathbf{Z}_2)$) and $\hat{F}_{ij}/\hat{\alpha}_i\hat{\beta}_j$ is minimized. This can be done by using the procedures that are described in de Rooij (2007).

It is important that the total mass is equal over the time points, i.e. that the total mass at the first time point is equal to the total mass at the second time point. Therefore the identification constraints on the masses were chosen such that this restriction is true. The mass will be represented by the area of the circle. To draw the circles we shall therefore use a radius equal to $r(i) = \sqrt{\{m(i)/\pi\}}$.

3.7. Estimation

Several researchers have discussed estimation of the RC(M) association model (Goodman,

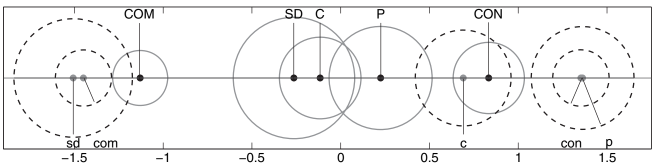

Fig. 2. Graphical representation of transitions between Swedish political parties: ○, COM, SD, C, P, CON, mass and positions of the parties in 1964; ◌, sd, com, c, con, p, mass and positions of the parties in 1970

1979; Becker, 1990; Haberman, 1995). By back-transforming using equations (7) and (6) we can obtain estimates of our model (5). The program *LEM* (Vermunt, 1997) will be used to fit the models; the transformation to a distance model and the way of identifying the solution are performed in MATLAB (Mathworks, 2006). The model with stable masses but dynamic positions cannot be written as an association model and thus cannot be estimated with available software for association models.

3.8. Application to Swedish politics data

The data in Table 1 were analysed by using the gravity models proposed. First some benchmark models were fitted. The quasi-independence model, the symmetry model and the quasi-symmetry model do not fit these data (X^2-values respectively 103.13, 78.03 and 22.86; G^2 101.12, 83.13 and 23.37, with 11, 10 and 6 degrees of freedom df). Since the quasi-symmetry model does not fit the data we expect that the models with stable positions do not fit either, which is indeed the case. With one dimension $X^2=28.97$, $G^2=27.29$ and $\mathrm{df}=7$ and with two dimensions the fit barely increased: $X^2=22.88$, $G^2=23.40$ and $\mathrm{df}=4$. With dynamic positions a good fit was obtained in a single dimension, $X^2=4.96$, $G^2=5.92$ and $\mathrm{df}=4$. The solution is shown in Fig. 2. We see several positional changes there: in 1964 the positions of the five parties are as expected on the left–right dimension, and also as described by Upton and Särlvik (1981). Ordered from left to right the Communists, the Social Democrats, the Centre Party, the People's Party and the Conservatives.

The positional changes from 1964 to 1970 can be summarized as follows: the Communists and Social Democrats grouped together on the left wing whereas the Conservatives and People's Party grouped on the right wing. Especially the Social Democrats made a big change to the left. The Centre Party moved from the centre to a more right-wing position. It seems that some polarization happened that distinguishes the two left-wing parties from the three other parties. Such a grouping was also found in Upton and Särlvik (1981). Although it may seem strange that the Social Democrats are more leftist than the Communists this has also been observed at several points in time by Lewin *et al.* (1972), page 226.

Concerning the masses it can be seen that the masses of the Communists, the People's Party and the Social Democrats stay the same, the Centre Party is the winner and the Conservative Party is the political party that lost mass.

4. Multiple two-way tables and multiple universes

Up to this point discussion has been confined to two-way tables. In the remainder of this paper we shall generalize the models to three-way tables. This section treats the case of multiple

two-way tables, e.g. transition data that are obtained in different countries, or different groups or at different time points.

4.1. The model

In this section we shall develop models for multiple two-way tables. Each of these tables can be modelled by the gravity models of the previous section, resulting in a universe with objects and masses for each layer ($k=1,\ldots,K$) of the table. The most general model is

$$F_{ijk} \propto \frac{m_{1k}(i)\,m_{2k}(j)}{\exp\{d_{ij}^2(\mathbf{Z}_{1k};\mathbf{Z}_{2k})\}}, \tag{9}$$

where $m_{tk}(i)$ is the mass of object i at time point t in layer k and the vector $\mathbf{z}_{tk}(i)$ gives the position of object i at time point t in layer k, i.e. each layer is represented by the gravity model of the previous section.

Restrictions can be imposed to relate the different universes. For example, the masses of different layers or the co-ordinates of different layers can be constrained to be equal or equal up to a scaling constant. The most natural choice is to restrict the co-ordinates ($\mathbf{Z}_{tk}$). Examples of restrictions are

$$\mathbf{Z}_{tk}=\mathbf{Z}_t\mathbf{W}_k, \tag{10}$$

$$\mathbf{Z}_{tk}=\mathbf{Z}\mathbf{W}_k, \tag{11}$$

$$\mathbf{Z}_{tk}=\mathbf{Z}_t, \tag{12}$$

$$\mathbf{Z}_{tk}=\mathbf{Z}, \tag{13}$$

where $\mathbf{W}_k$ is a diagonal matrix, specifying positive weights that stretch or shrink the dimensions of each layer, and $\mathbf{Z}$ is a matrix with co-ordinates of the points, which can be dependent on time ($\mathbf{Z}_t$) or not ($\mathbf{Z}$). When $w_{ppk}>1$, dimension p for layer k is stretched, meaning that for layer k the objects are more differentiated on this dimension. When $w_{ppk}<1$ the dimension shrinks, i.e. for layer k the objects are less differentiated on dimension p. The restriction in equation (11) defines the well-known INDSCAL-type of three-way distance model (Carroll and Chang, 1970), with stable positions of the categories for each layer of the table. The first restriction, equation (10), defines a two-mode version of the INDSCAL model, i.e. an INDSCAL distance model with dynamic positions. The third restriction, equation (12), defines two-mode distances which are equal across the layers and the fourth restriction, equation (13), defines one-mode distances that are equal across the layers. An example of a model with stable positions within each layer that is stretched or shrunk in the different layers and with equal masses over the layers can be found in de Rooij (2001).

4.2. Rewriting the model

Like the model for two-way tables this model can be written as an association model. In this case we deal with the conditional association models as proposed by Clogg (1982) and Becker and Clogg (1989). As in the two-way case, the link between the two types of models makes available software for fitting our gravity models and helps in understanding relationships between our gravity models and other models for contingency tables. The formulae for transforming one model into the other are similar to expressions (6) and (7). For example, with the restriction in equation (10) we have

$$
\begin{aligned}
F_{ijk} &\propto \frac{m_{1k}(i)\, m_{2k}(j)}{\exp\{d_{ij}^2(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{W}_k)\}} \\
&\propto \frac{m_{1k}(i)\, m_{2k}(j)}{\exp\{\sum_p w_{kp}^2 (z_{i1p} - z_{j2p})^2\}} \\
&\propto \frac{m_{1k}(i)\, m_{2k}(j)}{\exp\big(\sum_p w_{kp}^2 z_{i1p}^2 + w_{kp}^2 z_{j2p}^2 - 2 w_{kp}^2 z_{i1p} z_{j2p}\big)} \\
&\propto \frac{m_{1k}(i)}{\exp\big(\sum_p w_{kp}^2 z_{i1p}^2\big)} \frac{m_{2k}(j)}{\exp\big(\sum_p w_{kp}^2 z_{j2p}^2\big)} \exp\big(\sum_p 2 w_{kp}^2 z_{i1p} z_{j2p}\big) \\
&\propto \alpha_{ik} \beta_{jk} \exp\big(\sum_p \phi_{kp} \mu_{ip} \nu_{jp}\big),
\end{aligned}
\tag{14}
$$

where the difference between z_{i1p} and μ_{ip} is a scaling factor $\sqrt{2}$ and $\phi_{kp} = w_{kp}^2$. The last line in expression (14) is the conditional association model (Clogg, 1982; Becker and Clogg, 1989).

Model (9) represents a reduced rank association model for each two-way table. By using restrictions (10)–(13) the co-ordinates of the different layers are functions of each other. With restrictions (11) and (13) symmetry restrictions on the association are imposed, whereas restrictions (12) and (13) result in models without three-way association.

4.3. Identification

The conditional association model needs location and scaling constraints but no cross-dimensional constraints (Wong (2001), page 207). Similar to the situation that was discussed in the previous section, new locations (**a**) and scalings (**T**) are found by minimizing the correlation between the elements e_{ij}, defined as

$$e_{ij} = \sum_k \frac{\hat{F}_{ijk}}{\hat{\alpha}_{ik}\hat{\beta}_{jk}},$$

and the squared unweighted two-mode distances. The one-mode distance models ((11) and (13)) are identified.

The degrees of freedom for the conditional association models were discussed in Wong (2001), pages 205–207. For our models these numbers should be adapted for the loyalty parameters for the non-movers, i.e. the cell frequencies represented within parentheses in Table 2. The model that is given in expression (9) is multiplied by the term $\exp(\delta_{ij|k}\lambda_{ik})$ where $\delta_{ij|k}$ equals 1 if $i = j$ and 0 otherwise. There are IK of these parameters.

4.4. Estimation

Conditional association models can be estimated with LEM (Vermunt, 1997). From the expected frequencies we obtain an identified solution with a MATLAB procedure.

A cautionary note is in order here: when the model with constraint (10) or (11) is estimated by using the conditional association model to obtain estimates, although not regularly encountered, negative association parameters may occur for some layer. In that case there is not a distance representation of the conditional association model. A traditional (i.e. for the association model) graphical display of the association must also reflect the dimension for the row or column points for the specific layer. If such a negative association coefficient occurs it is

Table 2. Mobility data from three countries: USA, UK and Japan†

Country	*Father*	*Son*				
		un	ln	um	lm	fa
US	UN	(1276)	364	274	272	17
	LN	1055	(597)	394	443	31
	UM	1043	587	(1045)	951	47
	LM	1159	791	1323	(2046)	52
	FA	666	496	1031	1632	(646)
UK	UN	(474)	129	87	124	11
	LN	300	(218)	171	220	8
	UM	438	254	(669)	703	16
	LM	601	388	932	(1789)	37
	FA	76	56	125	295	(191)
Japan	UN	(127)	101	24	30	12
	LN	86	(207)	64	61	13
	UM	43	73	(122)	60	13
	LM	35	51	62	(66)	11
	FA	109	206	184	253	(325)

†From Yamaguchi (1987): UN, upper non-manual; LN, lower non-manual; UM, upper manual; LM, lower manual; FA, farm.

reasonable to assume that the model does not provide a good description of the association for that specific layer.

4.5. Application

To illustrate, the gravity model will be applied to data from Yamaguchi (1987) (see also Caussinus and Thelot (1976)), where occupational mobility is given for three countries: the USA, the UK and Japan. The data are reproduced in Table 2. Each occupational mobility table has five occupational categories: *upper non-manual* UN; *lower non-manual* LN; *upper manual* UM; *lower manual* LM; *farmer* FA. Again the focus is on change and for all cells within parentheses loyalty parameters are included in the models to be discussed.

Two benchmark models for these data are the conditional quasi-independence model and the no-three-way interaction model. The conditional quasi-independence model of father and son given country has $X^2 = 1409.76$ and $G^2 = 1336.20$ with df $= 33$ and the no-three-way association model has $X^2 = 36.24$ and $G^2 = 36.21$ with df $= 22$. The latter model with symmetry restrictions on the father–son association term has $X^2 = 125.24$ and $G^2 = 106.67$ with df $= 28$ (the degrees of freedom in the latter two models are computed by adding the number of boundary parameters to the usual degrees of freedom). The last model shows that the assumption of a symmetric association pattern is not tenable, i.e. a two-mode distance model will be needed.

The one-dimensional model with dynamic masses and dynamic positions constrained to be equal for the three layers (the restriction which is defined by equation (12)) fits the data marginally ($X^2 = 37.75$, $G^2 = 37.72$ and df $= 26$), which is reasonable considering the large sample size. The solution is shown in Fig. 3.

The major positional change is that of the farmer category, which is for the fathers at the right but for the sons in the middle, closer to the non-manual categories. When looking at the masses it can be seen that the USA and the UK have a similar pattern, whereas the pattern in Japan is typically different. In the USA and the UK the non-manual classes gained mass, whereas the lower manual and farmer classes lost mass. In the USA the upper manual class gained mass

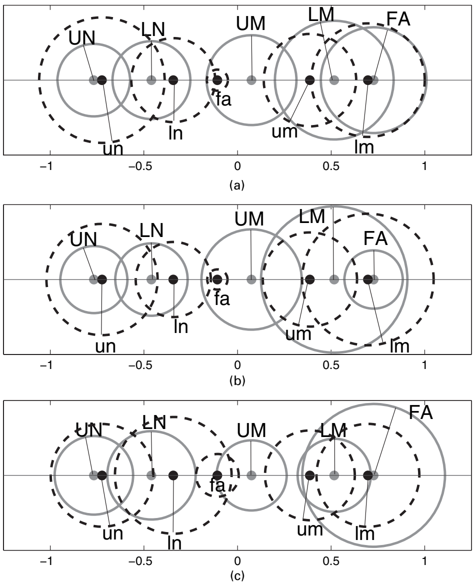

Fig. 3. Graphical representation of occupational mobility data from Yamaguchi (1987) (○, UN, LN, UM, LM, FA, mass and positions of the categories for the fathers; ◌, un, ln, fa, um, lm, mass and positions of the categories for the sons): (a) USA; (b) UK; (c) Japan

whereas in the UK this category lost mass. In Japan the farmers lost mass whereas all other categories gained mass.

5. Generalizations to change over three time points

5.1. *The model*

Above we treated models for two time points. Often, however, data are gathered at more time points. For three time points a gravity model can be built by using triadic distance models (de Rooij and Gower, 2003; Gower and de Rooij, 2003; de Rooij and Heiser, 2000, Heiser and Benanni, 1997; Daws, 1996; Joly and Le Calvé, 1995; Cox *et al.*, 1991; Pan and Harris, 1991), models that define a distance between three points simultaneously. An extension of model (5) is

$$F_{ijk} \propto \frac{m_1(i)\, m_2(j)\, m_3(k)}{\exp\{d^2_{ijk}(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{Z}_3)\}}. \tag{15}$$

Different forms of $d^2_{ijk}(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{Z}_3)$ can be considered. de Rooij and Gower (2003) provided an extensive description of possibilities plus the geometry of the options. A natural choice in the current framework is to consider the generalized Euclidean model, in which case d^2_{ijk} is defined as

$$d^2_{ijk}(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{Z}_3) = d^2_{ij}(\mathbf{Z}_1;\mathbf{Z}_2) + d^2_{ik}(\mathbf{Z}_1;\mathbf{Z}_3) + d^2_{jk}(\mathbf{Z}_2;\mathbf{Z}_3), \tag{16}$$

where each dyadic distance is defined as in equation (4). The interpretation of a triadic distance is facilitated when the isocontours are known, which are the lines with constant triadic distance with two fixed points. The isocontours for the generalized Euclidean model are circular around the centre of the two fixed points (de Rooij and Gower (2003), Fig. 3) just like the isocontours for a regular Euclidean distance. The distance that is defined in equation (16) is a triadic three-mode distance, where categories have dynamic positions as before in the two-mode distance. The positions can be constrained to be stable; then the triadic one-mode distance is obtained, in which case $\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}_3 = \mathbf{Z}$.

5.2. *Rewriting the model*

Model (15) can be rewritten as a partial association model (Clogg, 1982) as follows:

$$\begin{aligned}
F_{ijk} &\propto \frac{m_1(i)\, m_2(j)\, m_3(k)}{\exp\{d^2_{ijk}(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{Z}_3)\}} \\
&\propto \frac{m_1(i)\, m_2(j)\, m_3(k)}{\exp\{d^2_{ij}(\mathbf{Z}_1;\mathbf{Z}_2) + d^2_{ik}(\mathbf{Z}_1;\mathbf{Z}_3) + d^2_{jk}(\mathbf{Z}_2;\mathbf{Z}_3)\}} \\
&\propto \frac{m_1(i)\, m_2(j)\, m_3(k)}{\exp\big(\sum_p z^2_{i1p} + z^2_{j2p} - 2z_{i1p}z_{j2p} + z^2_{i1p} + z^2_{k3p} - 2z_{i1p}z_{k3p} + z^2_{j2p} + z^2_{k3p} - 2z_{j2p}z_{k3p}\big)} \\
&\propto \frac{m_1(i)\, m_2(j)\, m_3(k)}{\exp\big(\sum_p 2z^2_{i1p} + 2z^2_{j2p} + 2z^2_{k3p} - 2z_{i1p}z_{j2p} - 2z_{i1p}z_{k3p} - 2z_{j2p}z_{k3p}\big)}.
\end{aligned} \tag{17}$$

Defining

$$\alpha(i) = m_1(i) / \exp\Big(\sum_{p=1}^{P} 2z^2_{i1p}\Big),$$

$$\beta(j) = m_2(j) / \exp\Big(\sum_{p=1}^{P} 2z^2_{j2p}\Big)$$

and

$$\gamma(k) = m_3(k) / \exp\Big(\sum_{p=1}^{P} 2z^2_{k3p}\Big)$$

we obtain

$$F_{ijk} \propto \alpha(i)\, \beta(j)\, \gamma(k)\, \exp\Big(\sum_p 2z_{i1p}z_{j2p} + 2z_{i1p}z_{k3p} + 2z_{j2p}z_{k3p}\Big). \tag{18}$$

The partial association model with restricted row, column or layer terms such that the row scores are equal in the association with the columns and with the layers, and similarly for the column scores and layer scores, is

$$F_{ijk} \propto \alpha(i)\,\beta(j)\,\gamma(k)\,\exp\Big(\sum_p \phi_{1p}\mu_{ip}\nu_{jp} + \phi_{2p}\mu_{ip}\kappa_{kp} + \phi_{3p}\nu_{jp}\kappa_{kp}\Big). \tag{19}$$

Model (18) can be obtained from this by defining

$$z_{i1p} = \frac{a_p\mu_{ip}}{\sqrt{2}},$$

$$z_{j2p} = \frac{b_p\nu_{jp}}{\sqrt{2}},$$

$$z_{k3p} = \frac{c_p\kappa_{kp}}{\sqrt{2}}$$

where $a_p = \sqrt{(\phi_{1p}\phi_{2p}/\phi_{3p})}$, $b_p = \phi_{1p}/a_p$ and $c_p = \phi_{2p}/a_p$.

Again, the link between the gravity and association model makes software available for fitting the gravity model and provides insight into the relationships between our gravity model and other models for contingency tables.

5.3. Identification

For the partial association models location constraints are necessary for each variable, whereas scaling and cross-dimensional constraints are necessary for only one of the three variables (Anderson and Vermunt (2000), page 95, and Wong (2001), page 204). In the triadic three-mode model we can find new locations for the variable at the first time point $\mathbf{a}$, for the second time point $\mathbf{b}$ and for the third time point $\mathbf{c} = -(\mathbf{a}+\mathbf{b})$ by minimizing the correlation between $\hat{F}_{ijk}/\hat{\alpha}_i\hat{\beta}_j\hat{\gamma}_k$ and the squared triadic distance $d^2_{ijk}(\mathbf{Z}_1;\mathbf{Z}_2;\mathbf{Z}_3)$. The triadic one-mode model, i.e. the model with stable positions, is identified.

As before we focus on change by including loyalty parameters in the model for the people who made the same choice on all three occasions. In other words, the model defined in expression (15) is multiplied by the term $\exp(\delta_{ijk}\lambda_i)$, where δ_{ijk} equals 1 if $i = j = k$, and 0 otherwise. The λ_i are loyalty parameters, of which there are I.

5.4. Estimation

In LEM it is not possible to estimate the partial association models in more than two dimensions. In most situations in which we want to represent a model graphically this will be enough. However, for comparing against higher dimensional alternatives it is not satisfactory.

For the triadic one-mode distance function with stable positions ($\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}_3 = \mathbf{Z}$) the association model should be fitted with equality restrictions on the row, column or layer scores such that $\mu_{ip} = \nu_{ip} = \kappa_{ip}$, but also a restriction on the association parameters $\phi_{1p} = \phi_{2p} = \phi_{3p}$.

5.5. Application

To illustrate, model (15) will be applied to data obtained from Upton (1978), page 128, where a sample of 1651 Swedish people were asked for their votes at three consecutive elections (Table 3). There are four political parties, the Social Democrats SD, the Centre Party C, the People's Party P and the Conservatives CON. Table 3 gives the measurements of forces between the four political parties. For example, there are low forces between the Social Democrats in 1964, the Centre Party in 1968 and the People's Party in 1970 (the force equals 6) and between the People's Party in 1964, the Centre Party in 1968 and the Social Democrats in 1970 (the force equals 1).

Table 3. Swedish voting data representing voting changes from 1964 to 1968 to 1970

1964 party	*1968 party*	*1970 party*			
		SD	C	P	CON
SD	SD	(812)	27	16	5
	C	5	20	6	0
	P	2	3	4	0
	CON	3	3	4	2
C	SD	21	6	1	0
	C	3	(216)	6	2
	P	0	3	7	0
	CON	0	9	0	4
P	SD	15	2	8	0
	C	1	37	8	0
	P	1	17	(157)	4
	CON	0	2	12	6
CON	SD	2	0	0	1
	C	0	13	1	4
	P	0	3	17	1
	CON	0	12	11	(126)

As in the two-way tables the force from a to b was not equal to the force from b to a; in the three-way table the force abc is not equal to the forces acb, bac, bca, cab and cba. The focus is on the movers, meaning that subjects who made the same choice at all three time points are excluded from the analysis with the gravity model, by including loyalty parameters for the cells within parentheses.

Some benchmark models are the no-three-way association model, which has $X^2 = 27.13$ and $G^2 = 29.00$ with df $= 23$ with a Bayesian information criterion (BIC) statistic of -141.40. With symmetry restrictions on the association terms we obtain $X^2 = 47.25$ and $G^2 = 49.49$ with df $= 32$ with a BIC statistic of -187.60. The latter model just does not fit: $p = 0.04$ by using the X^2-statistic. Another benchmark model is the first-order Markov model; it has $X^2 = 427.04$ and $G^2 = 207.33$ with df $= 36$. Its BIC statistic equals -59.39.

Application of the gravity models with stable positions (triadic one-mode distance) gives with one dimension $X^2 = 166.97$ and $G^2 = 180.06$ with df $= 47$ and $X^2 = 138.16$ and $G^2 = 138.07$ with df $= 45$ with two dimensions. Using dynamic positions $X^2 = 116.00$ and $G^2 = 131.25$ with df $= 41$ in the one-dimensional solution and $X^2 = 74.36$ and $G^2 = 80.86$ with df $= 33$ in the two-dimensional solution.

Looking at BIC statistics we obtain the following. For the model with stable positions in one dimension the BIC statistic equals -168.17, whereas in two dimensions it is -195.34. With dynamic positions in one dimension the BIC statistic is -172.52, and in two dimensions it is -163.64.

Fig. 4 shows the model with the lowest BIC statistic, the two-dimensional model with dynamic masses but stable positions. This is a model with a three-way symmetric association pattern, i.e. all asymmetry in the data is captured by the masses. We see that the People's Party and the Conservatives are close in space ($d^2 = 0.04$), whereas the Social Democrats are far from all other parties ($d^2 = 1.78$, $d^2 = 1.53$ and $d^2 = 2.06$ to the Centre Party, People's Party and Conservatives respectively). The Centre Party is closer to the People's Party ($d^2 = 0.85$) and the Conservatives ($d^2 = 0.96$) than to the Social Democrats ($d^2 = 1.78$). The largest triadic distance

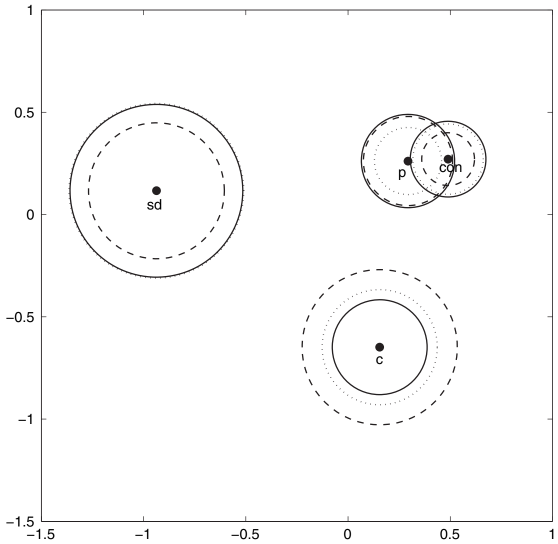

Fig. 4. Graphical representation of transitions between Swedish political parties from 1964 to 1968 to 1970 (the horizontal dimension can be interpreted as the traditional left–right dimension, whereas the vertical dimension can be interpreted as a rural–urban dimension): ○, mass at 1964; ○, mass at 1968; ○, mass at 1970

is between the Social Democrats, the Centre Party and the Conservatives ($d^2_{ijk} = 4.79$), whereas the smallest is between the Centre Party, the People's Party and the Conservatives ($d^2_{ijk} = 1.84$). The other two triadic distances are $d^2_{ijk} = 4.16$ for the combination Social Democrats, Centre Party and People's Party and $d^2_{ijk} = 3.63$ for the combination Social Democrats, People's Party and Conservatives. Also triples with a recurring party have a triadic distance, which is in the triadic one-mode distance the square root of twice the squared dyadic distance. Since these are often smaller than the triadic distances between three different parties, the pattern is such that more people transit between two than between three parties.

Concerning the masses we see that the Social Democrats first stay stable but then lose mass, the Centre Party gains mass twice, the People's Party first loses and then regains mass, and finally the Conservatives lose mass twice.

The horizontal dimension is the traditional left–right dimension, where again the Centre Party, the People's Party and the Conservatives group on the right-hand side. The vertical dimension differentiates the Centre Party from the other parties and can be understood as a rural–urban dimension since the Centre Party used to be the Agrarian Party, attracting many farmers and people from the small villages (Lewin *et al.* (1972), page 221).

For the interpretation as a model of change consider the odds of choosing the Social Democrats *versus* the Centre Party at the third time point given the People's Party at the first and Conservatives at the second time point. The odds are

$$\begin{aligned}\frac{\pi_{\text{P,CON,SD}}}{\pi_{\text{P,CON,C}}} &= \frac{m_1(\text{P})\, m_2(\text{CON})\, m_3(\text{SD})\, \exp(-d^2_{\text{P,CON,SD}})}{m_1(\text{P})\, m_2(\text{CON})\, m_3(\text{C})\, \exp(-d^2_{\text{P,CON,C}})}\\ &= \frac{m_3(\text{SD})}{m_3(\text{C})}\, \frac{\exp(-d^2_{\text{CON,SD}})}{\exp(-d^2_{\text{CON,C}})}\, \frac{\exp(-d^2_{\text{P,SD}})}{\exp(-d^2_{\text{P,C}})}.\end{aligned}$$

So, the knowledge that at the first time point the choice was for the People's Party changed the odds by a factor $\exp(-d^2_{\text{P,SD}})/\exp(-d^2_{\text{P,C}}) = 0.50$.

As another example consider the odds of the People's Party *versus* the Conservatives given twice the Social Democrats:

$$\begin{aligned}\frac{\pi_{\text{SD,SD,P}}}{\pi_{\text{SD,SD,CON}}} &= \frac{m_1(\text{SD})\, m_2(\text{SD})\, m_3(\text{P})\, \exp(-d^2_{\text{SD,SD,P}})}{m_1(\text{SD})\, m_2(\text{SD})\, m_3(\text{CON})\, \exp(-d^2_{\text{SD,SD,CON}})}\\ &= \frac{m_3(\text{P})}{m_3(\text{CON})}\, \frac{\exp(-d^2_{\text{SD,P}})}{\exp(-d^2_{\text{SD,CON}})}\, \frac{\exp(-d^2_{\text{SD,P}})}{\exp(-d^2_{\text{SD,CON}})}\\ &= \frac{0.14}{0.05}\, \frac{\exp(-1.53)}{\exp(-2.06)}\, \frac{\exp(-1.53)}{\exp(-2.06)}\\ &= \frac{0.14}{0.05}\, \frac{0.22}{0.13}\, \frac{0.22}{0.13}\\ &= 2.8 \times 1.69 \times 1.69\\ &= 7.98.\end{aligned}$$

So, after twice choosing the Social Democrats the odds are largely in favour of the People's Party instead of the Conservatives, which can be judged from the larger mass of the People's Party and the smaller distance from the Social Democrats to the People's Party compared with the distance Social Democrats–Conservatives.

6. Swedish politics revisited

The two examples that were discussed dealt with votes from Sweden in the period 1964–1970 (see Sections 3.8 and 5.5). The first example included the Communists for which we have no data in the second example. The conclusion from the first example (1964–1970) was that there have been major changes in the positions, whereas the conclusion from the second example (1964–1968–1970) is that the positions are unchanged. These solutions cannot really be compared, since the first shows marginal association whereas the second shows conditional association. It is well known that these two differ in many cases (see Agresti (2002), chapter 2).

To compare the positions further we analysed four tables: each of the two-way marginal tables of the 1964–1968–1970 data (Table 3) and the 1964–1970 data (Table 1) but without the Communists. For these 4×4 tables, the one-dimensional model with dynamic positions has 0 degrees of freedom, i.e. it is a saturated model. To make comparisons we changed from object-specific loyalty parameters to a single overall loyalty parameter (i.e. $\lambda_1 = \lambda_2 = \ldots = \lambda_I$), the effect of which is that the masses and distances now also contribute to the fit of the diagonal elements in the data matrix. Moreover, these diagonal elements are not necessarily fitted perfectly as is the case with object-specific loyalty parameters. With these settings the model with dynamic positions has 3 degrees of freedom, whereas the model with stable position has 5.

For all data sets, except for the 1968–1970 data, dynamic positions are needed. For the 1968–1970 data the model with stable positions provides an adequate fit. The results are shown in Fig. 5. Figs 5(a)–5(c) pertain to the marginal tables that were obtained from Table 3, whereas Fig. 5(d) pertains to the data from Table 1 but without the Communists. From Fig. 5 it can be concluded that the positions of the parties in 1964 and 1970 are in all cases roughly the same. Again, as in Fig. 2, the Centre Party, the People's Party and the Conservatives seem to cluster together over time, which is mostly due to the period 1964–1968, since this effect is visible in both the analysis of the 1964–1968 and the 1964–1970 data. Concerning the 1968 positions in the analysis of 1964–1968 data there is a reversal of the People's Party and the Conservatives at the right-hand side of the scale, which cannot be found back in the 1968–1970 data. Note

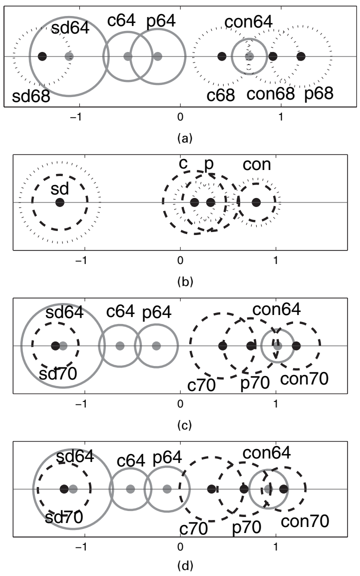

Fig. 5. Comparison of results from four two-way tables (○, masses for 1964; ○, masses for 1968; ◌, masses for 1970): (a) 1964–1968 marginal table (obtained from the three-way table); (b) 1968–1970 marginal table (obtained from the three-way table); (c) 1964–1970 marginal table (obtained from the three-way table); (d) 1964–1970 result for Table 1 without the Communists

again that this reversal is not due to the new identification constraints. The same reversal in 1968 was found by Lewin *et al.* (1972), page 220, and can possibly be explained by 'the conservative party in the sixties is our best example of a party in the throes of a crisis of identity' (Lewin *et al.* (1972), page 285), which is also apparent by their change of name from (literal translation) 'The Right Party' to 'The Moderates' (see Upton and Särlvik (1981)). When looking at the 1968–1970 data, however, this reversal of the 1968 positions is not preserved. Finally, note that the analyses of the two data sets pertaining to changes from 1964 to 1970 give (almost) identical results.

7. Discussion

The analysis of change was discussed in terms of Newton's law of gravity. It was shown that a well-known and often-applied model for the analysis of contingency tables, the RC(M) association model, can be interpreted in terms of mass and distance, and thus has an interpretation that is similar to the law of gravity. Both masses and positions can be stable as well as dynamic. These dynamic elements were discussed extensively; dynamic masses relate to popularity of objects which might change whereas dynamic positions relate to content changes of objects (in the case of stable masses no change in content took place). The RC(M) association model needs location and scaling constraints for identification. The usual constraints, however, are troublesome in the analysis of change, and therefore a new way of identifying the solution was discussed. This new way of identifying the solution makes it possible to interpret the solution in terms of polarization, as we did in the application. However, if everything (i.e. all voters and all parties) makes the same shift in one direction, our method will not find this shift in location since all relative positions remain the same. An example of such a situation, as a referee pointed out, is that as a consequence of world events (e.g. global warming) the nation becomes more left or right wing (i.e. 'green'). All parties will feel this shift and will adapt their stances as a consequence. This common shift will not be noted by our gravity model. If, however, some parties shift more than others then we will see that the relative positions change.

The new interpretation in terms of mass and distance of the RC(M) association model is simple since both mass and distance are fairly well-understood concepts, at least better than main effects and inner products (projection). In the examples that were shown in Section 3 (and 4) a one-dimensional solution was obtained in which the interpretation of the graphical display in terms of distances is much easier than the product of lengths of vectors (as in the inner product parameterization). So, a new interpretation to a well-known model was provided, which might be of great value, since the new interpretation has roots in the natural sciences and is well understood by many people.

The gravity models that were proposed can be considered a generalization of the loyalty–distance models that were proposed by Upton and Särlvik (1981). Compared with their model our model is not dependent on an *a priori* ordering of the objects; our model can be used for multidimensional solutions; and our model allows for changing positions of the objects. It can be assumed that the solutions that are obtained with our unidimensional model with stable positions and the loyalty–distance model of Upton and Särlvik are approximately the same.

After the case of square contingency tables we looked at the case where there are multiple tables. Bridges between conditional association models and weighted Euclidean distances (the INDSCAL model) were shown, but also other solutions (further or less restricted) were discussed for such data. As for the square table case, we developed a new interpretation in terms of mass and distance. As for the standard RC(M) association model the identification constraints had to be adapted; we developed a manner to do so.

The generalization to tables for three time points needed a further expansion of the law of gravity. Distances between three points, triadic distances, were discussed and their relationship to partial association models was shown. Again the triadic distance is easier to understand than the sum of three inner products, as modelled in the partial association models. Although this representation is moving away from Newton's law of gravity, the basic ingredients are still mass and distance. The models that were discussed can theoretically be generalized to more time points by using tetradic (or polyadic) distances, which can be defined as sums of all (squared) dyadic distances (as in equation (16)). However, the relationship between partial association models and triadic distance models as shown in equations (18) and (19) does not generalize to more time points (variables). In that case the gravity models become more restricted than the association models.

The triadic distance models are good in showing and representing change. They lack a formal change mechanism, however, as in for example (latent) Markov models. However, the triadic gravity models can be conceived as second-order Markov models with restrictions, and when interpreted by using the correct temporal ordering there is an influence of the first time point on the second, and an influence of the first and second time points on the third. We provided some examples of interpretation at the end of Section 5.5. For triadic three-mode distances similar statements about conditional odds can be obtained.

The triadic gravity models are not collapsible, i.e. the change from time point 1 to 2 in triadic distance models is different from the change that is obtained when the table was collapsed over the third time point. A model is collapsible when the conditional association equals the marginal association, and this is generally so for conditional independence models (Bishop *et al.*, 1975). For the example that was discussed in Section 5 this would be the non-fitting first-order Markov model. A possible reason for the failure of this model is subject heterogeneity, which is due to ignoring relevant covariates in the analysis (Agresti (2002), page 478). In other words, the group of people who make the same change from time point 1 to 2 do not form a homogeneous group. By using the choices that are made at the third time point we obtain more reliable change estimates.

In this paper all models were estimated with LEM. This yielded some problems; for example, for square tables the model with stable masses but dynamic positions cannot be fitted by using software for association models. Furthermore, the models for observations on three time points could not be estimated in more than two dimensions, and for the model for multiple two-way tables sometimes a negative association coefficient occurs for a particular layer which under a distance model is not possible. To deal with such problems special software should be written.

We tried to build models for the whole data set. Another way of analysing square tables is to decompose the table into a symmetric part and a skew symmetric part, and these are then analysed separately. Often this is done by using least squares techniques. The best known of such procedures is that due to Gower (1977) and discussed also in Constantine and Gower (1978). Bilinear forms for skew symmetry were discussed in van der Heijden and Mooijaart (1995). For a treatment of skew symmetry in the three-way case see de Rooij and Heiser (2000).

The use of gravity models to explain social phenomena is not new. Tobler (1976), for example, used a social gravity model for migration data. Gravitational models are also often used in economic and transportational studies (see, for example, Sen and Smith (1995)). In these cases, however, the distances are often known in advance, i.e. they are real distances, or the distances are measured by several variables. In our case, however, the distances must be estimated from the data. The fact that Newton's gravity model is used more often in other areas is due to its simple and understandable nature: mass and distance are easily grasped concepts. In this paper

these concepts are used for the analysis of change. For change from one time point to another these models give a very natural description of the change process.

Acknowledgements

This project was finished while the author was sponsored by the Netherlands Organisation for Scientific Research, innovational grant 452-06-002. I thank Paul Eilers, Willem Heiser, Carolyn Anderson and the referees for discussion and comments on earlier versions of the manuscript.

Appendix A: Software note

We used the program LEM to obtain fitted frequencies. The scaling and locations were found by using MATLAB. A MATLAB shield was built around LEM such that no manual copying is needed in performing the analysis. On request the MATLAB and LEM files can be obtained from the author.

References

Agresti, A. (2002) *Categorical Data Analysis*, 2nd edn. New York: Wiley.
Anderson, C. J. and Vermunt, J. K. (2000) Log-multiplicative association models as latent variable models for nominal and/or ordinal data. *Sociol. Methodol.*, **30**, 81–121.
Becker, M. P. (1990) Algorithm AS 253: Maximum likelihood estimation of the RC(M) association model. *Appl. Statist.*, **39**, 152–167.
Becker, M. P. and Clogg, C. C. (1989) Analysis of sets of two-way contingency tables using association models. *J. Am. Statist. Ass.*, **84**, 142–151.
Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Massachusetts: Massachusetts Institute of Technology Press.
Borg, I. and Groenen, P. J. F. (2005) *Modern Multidimensional Scaling; Theory and Applications*, 2nd edn. New York: Springer.
Carroll, J. D. and Arabie, P. (1980) Multidimensional scaling. *A. Rev. Psychol.*, **31**, 607–649.
Carroll, J. D. and Chang, J. J. (1970) Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition. *Psychometrika*, **35**, 283–319.
Caussinus, H. (1965) Contribution à l'analyse statistique des tableaux de corrélation [Contributions to the statistical analysis of correlation matrices]. *Ann. Fac. Sci. Univ. Toul.*, **29**, 715–720.
Caussinus, H. and Thelot, C. (1976) Note complémentair sur l'analyse statistique des migrations [Complementary note on the statistical analysis of migrations]. *Ann. INSEE*, **22–23**, 83–100.
Clogg, C. C. (1982) Some models for the analysis of association in multiway cross-classifications having ordered categories. *J. Am. Statist. Ass.*, **77**, 803–815.
Constantine, A. G. and Gower, J. C. (1978) Graphical representation of asymmetric matrices. *Appl. Statist.*, **27**, 297–304.
Cox, T. F., Cox, M. A. and Branco, J. A. (1991) Multidimensional scaling for *n*-tuples. *Br. J. Math. Statist. Psychol.*, **44**, 195–206.
Daws, J. T. (1996) The analysis of free-sorting data: beyond pairwise comparison. *J. Classificn*, **13**, 57–80.
Goodman, L. A. (1979) Simple models for the analysis of association in cross classifications having ordered categories. *J. Am. Statist. Ass.*, **74**, 537–552.
Goodman, L. A. (1991) Measures, models, and graphical displays in the analysis of cross-classified data. *J. Am. Statist. Ass.*, **86**, 1085–1111.
Gower, J. C. (1977) The analysis of asymmetry and orthogonality. In *Recent Developments in Statistics* (eds J. Barra, F. Brodeau, G. Romer and B. van Cutsem), pp. 109–123. Amsterdam: North-Holland.
Gower, J. C. and de Rooij, M. (2003) A comparison of the multidimensional scaling of triadic and dyadic distances. *J. Classificn*, **20**, 115–136.
Haberman, S. J. (1995) Computation of maximum likelihood estimates in association models. *J. Am. Statist. Ass.*, **90**, 1438–1446.
van der Heijden, P. G. M. and Mooijaart, A. (1995) Some new log-bilinear models for the analysis of asymmetry in square contingency tables. *Sociol. Meth. Res.*, **24**, 7–29.
Heiser, W. J. and Bennani, M. (1997) Triadic distance models: axiomatization and least squares representation. *J. Math. Psychol.*, **41**, 189–206.
Joly, S. and Le Calvé, G. (1995) Three-way distances. *J. Classificn*, **12**, 191–205.
Lewin, L., Jansson, B. and Sörbom, D. (1972) *The Swedish Electorate 1887-1968*. Stockholm: Almqvist and Wiksell.

Mathworks (2006) *MATLAB: the Language of Technical Computing*. Natick: Mathworks.
Nosofsky, R. M. (1985) Overall similarity and the identification of separable-dimension stimuli: a choice model analysis. *Perceptn Psychphys.*, **38**, 415–432.
Pan, G. and Harris, D. P. (1991) A new multidimensional scaling technique based upon association of triple objects Pijk and its application to the analysis of geochemical data. *Math. Geol.*, **23**, 861–886.
de Rooij, M. (2001) Distance association models for the analysis of repeated transition frequency data. *Statist. Neerland.*, **55**, 157–181.
de Rooij, M. (2007) The distance perspective of generalized biadditive models: scalings and transformations. *J. Computnl Graph. Statist.*, **16**, 210–227.
de Rooij, M. and Gower, J. C. (2003) The geometry of triadic distance models. *J. Classificn*, **20**, 181–220.
de Rooij, M. and Heiser, W. J. (2000) Triadic distance models for the analysis of asymmetric three-way proximity data. *Br. J. Math. Statist. Psychol.*, **53**, 99–119.
de Rooij, M. and Heiser, W. J. (2005) Graphical representations and odds ratios in a distance-association model for the analysis of cross-classified data. *Psychometrika*, **70**, 99–122.
Sen, A. and Smith, T. E. (1995) *Gravity Models of Spatial Interaction Behavior*. Heidelberg: Springer.
Tobler, W. (1976) Spatial interaction patterns. *J. Environ. Syst.*, **6**, 271–301.
Upton, G. J. G. (1978) *The Analysis of Cross-tabulated Data*. Chichester: Wiley.
Upton, G. J. G. and Särlvik, B. (1981) A loyalty–distance model for voting change. *J. R. Statist. Soc.* A, **144**, 247–259.
Vermunt, J. K. (1997) *lEM*: a general program for the analysis of categorical data. *Technical Report*. Tilburg University, Tilburg.
Wong, R. S. K. (2001) Multidimensional association models: a multilinear approach. *Sociol. Meth. Res.*, **30**, 197–240.
Yamaguchi, K. (1987) Models for comparing mobility tables: towards parsimony and substance. *Am. Sociol. Rev.*, **52**, 482–494.