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

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).

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).

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'.

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

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 $r(i) = \surd\{m(i)/\pi\}$.

3.7. Estimation

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

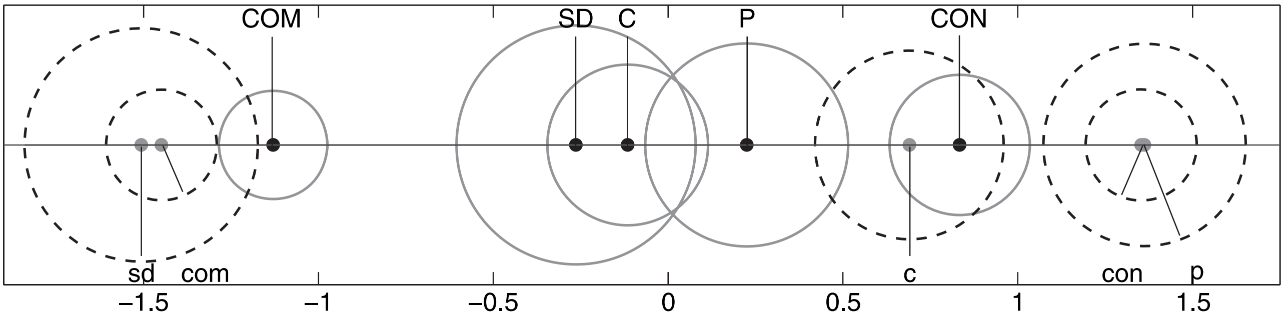

Fig. 2. Graphical representation of transitions between Swedish political parties: ○, COM, SD, C, P, CON, mass and positions of the parties in 1964; ◌, sd, com, c, con, p, mass and positions of the parties in 1970

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 df = 7 and with two dimensions the fit barely increased: $X^2 = 22.88$, $G^2 = 23.40$ and df = 4. With dynamic positions a good fit was obtained in a single dimension, $X^2 = 4.96$, $G^2 = 5.92$ and 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