Chapter 15

The Use of Covariates in Distance Association Models for the Analysis of Change

Mark de Rooij

Leiden University, Department of Psychology,
Leiden, The Netherlands

15.1 Introduction to Categorical Repeated Measurements

Many studies involve repeated measurements, i.e. measurements on a number of time points. This longitudinal research is at the heart of understanding human development, both at the individual level as well as at the group level. Statistical methods for longitudinal data can be divided in two types: Methods for numerical variables where change is described in better versus worse or methods for categorical variables where change is described in terms of different versus the same.

The central assumption in this paper is that people change step by step, making small moves. For example, when the length of a person is measured at repeated occasions the growth will be gradually and in (relatively) small steps. The same thing is true for learning to speak or to read, children tend to change gradually in their development of such abilities. As an example of see the learning to read study (Jansen and Bus, 1982) and the analysis of the data in Timmerman and Kiers (2002). This gradual development should be taken into account when modeling repeated measurements. For interval type data this is common, i.e. on a scale (in the first example, a length-scale) from small to long, people will make small steps from the left side (small) of this scale to the right side (long).

For categorical variables, however, this geometric framework of change is ignored in all common statistical methods for such data. Change is generally described by a number of transition probabilities, but no statistical method tries to recover the scale on which these changes occur. The terminology ‘scale’ is often thought of to be unidimensional. For categories of a variable like, for instance, political parties in the Netherlands, it is very plausible that they have a multidimensional nature (see Groenen, 2003). Therefore, the term scale is not very appropriate and ‘map’ will be used instead. Distance association models (De Rooij, 2001a, 2001b, 2002, 2005a, 2005b; De Rooij and Heiser, 2003, 2005) assume such a map is present and people change gradually over this map, i.e. categories that are close together on the map will have a high transition frequency, whereas categories that are far apart will have a low transition frequency. In other words, distance association models are the only models for longitudinal categorical data that take into account the geometric step by step aspect of change.

Distance association models (De Rooij, 2001a, 2001b, 2005a, 2005b; De Rooij and Heiser, 2003, 2005) are distance models for the analysis of contingency tables indexed by time. When there are measurements on two time points the table is square and the frequencies denote transitions between the categories of a variable. An example is political votes obtained from a sample of Dutch voters, where the categories are Dutch political parties, e.g., PvdA, CDA, VVD, D66, SP, etc.. The frequencies denote the number of voters that choose PvdA at the first time point and SP at the second. Similarly, when measurements are obtained at three time points the frequencies denote the number of people voting PvdA at the first time point, SP at the second, and D66 at the third time point. This frequency (in case of two time points) is dependent on two things: the mass of each of the parties at the two time points and the distance between the two parties. The distances are estimated from the data and from these distances a map can be constructed that provides insight into the field of change, in this case the political field. In De Rooij and Heiser (2003, 2005) the theory is developed for two-way transition data. In De Rooij (2005a) distance association models are generalized to three time points using triadic distance models (De Rooij and Heiser, 2001, De Rooij, 2002; Gower and De Rooij, 2003; De Rooij and Gower, 2003), that define distances between three points. These triadic distances are easily generalized to distances between more points (see De Rooij, 2001, chapter 7). The geometry of such generalizations will not change compared to the triadic case. Not only do distance association models take into account the geometric aspect of change (see above), since they are statistical models they can also be tested and statistical inference is possible. Other distance methods, often named multidimensional scaling, usually have an exploratory nature (see, for example, Borg and Groenen, 1997). The distance association models defined until this moment are suitable for change at the level of the whole group. It is, however, plausible that there are within group differences, for example between men and women.

Table 15.1. Cross-classification of opinions on Clarence Thomas in September and October (Favorable (F); Not Favorable (U); NO clear Opinion (N) differentiated by Political orientation (Liberal (L); Moderate (M); Conservative (C). From Bergsma and Croon (2005)

			O	
P	*S*	F	U	N
L	f	18	4	7
L	u	4	43	7
L	n	43	32	54
M	f	77	7	31
M	u	12	36	12
M	n	108	41	127
C	f	99	3	21
C	u	8	11	4
C	n	55	24	59

This paper aims at extending distance association models for square contingency tables to include covariates, variables that represent differences between people.

To facilitate the discussion we will use a small data set from Bergsma and Croon (2005). Bergsma and Croon (2005, p.84) describe this data set as follows: “These data were obtained in a longitudinal survey in which a panel of respondents was interviewed in September and October of the same year on their opinion on the Supreme Court candidate Clarence Thomas (CT) who was nominated by President George Bush Senior. This nomination stimulated some public debate because of CT’s Afro-American background and his allegedly extremely conservative views. Moreover, after the first survey was held on September 3-5, a charge for sexual harassment was brought against CT by his former aide Anita Hill on September 25. The second survey was then conducted on October 9. [...]. The variables *S* and *O* refer to the respondents’ opinion in September and October, respectively. [...]. The variable *P* refers to Political orientation. We have grouped the original responses to this variable in three categories: Liberal, Moderate and Conservative.” The data is reproduced in Table 15.1, using codes ‘F, ‘U’, ‘N’ for Favorable, Un-Favorable, No clear Opinion. September opinions are in lower case, while October opinions are in Upper Case. For Political orientation the codes Liberal (L), Moderate (M), and Conservative (C) are used.

The paper is organized as follows. In the next section we briefly outline the distance association models plus an analysis of the pooled data (Table 1). The third section discusses two ways of dealing with covariates: 1) Using design matrices for a two-way table; 2) Using space transforming matrices. We will describe both methods and show relationships to log-linear models. In the fourth Section we will analyze Table 1 and we will conclude with some discussion and possible extensions.

15.2 Distance Association Model

15.2.1 The Model

Distance association models for the analysis of two-way contingency tables, $\mathbf{F} = \{f_{ij}\}$, with $i = 1, \ldots, I$ and $j = 1, \ldots, J$, were proposed by De Rooij and Heiser (2005). For longitudinal data the i indicate categories at the first time point, and j for the second, and typically $I = J$. The ingredients of distance association models are masses of the categories at the first and second time point and distances between categories in a Euclidean space. This Euclidean space provides the map on which the transitions occur, i.e. the multidimensional scale comparable to the one-dimensional length scale in the introductory section. This map is of utmost importance to understand the change process. The most general distance association model for a two-way table is the two-mode distance association model, defined as

$$\pi_{ij} = \alpha_i \beta_j \exp\left(-d^2(\mathbf{x}_i, \mathbf{y}_j)\right) \tag{1}$$

where $d^2(\mathbf{x}_i, \mathbf{y}_j)$ is the squared Euclidean distance between two points defined by the p-dimensional vectors $\mathbf{x}_i = [x_{i1}, \ldots, x_{ip}]^T$ and $\mathbf{y}_j = [y_{j1}, \ldots, y_{jp}]^T$. The dimensionality cannot exceed $\min(I-1, J-1)$. When $p = \min(I-1, J-1)$ the model equals the saturated log-linear model for a two-way table; when $p = 0$ the model equals the independence model for such a table.

In some cases the model where the row vectors and column vectors for a specific category are constrained to be equal, i.e. $\mathbf{x}_i = \mathbf{y}_i$, is of special interest, leading to a symmetric association pattern. The model constrained in that way is named the one-mode distance association model, defined by

$$\pi_{ij} = \alpha_i \beta_j \exp\left(-d^2(\mathbf{x}_i, \mathbf{x}_j)\right). \tag{2}$$

Again, the maximum dimensionality is $I - 1$, in which case the model equals the quasi-symmetry model (Caussinus, 1965).

De Rooij and Heiser (2005) show that the two-mode distance association model and the *RC(M)*-association model (Goodman, 1985) are equivalent in the expected frequencies. The interpretation of both models, however, is quite different: The distance association model is interpreted in terms of distances where the *RC(M)*-association models should be interpreted using inner products. For a detailed discussion of the different interpretations and possible complications see De Rooij and Heiser (2005) and De Rooij (2005a, 2005b).

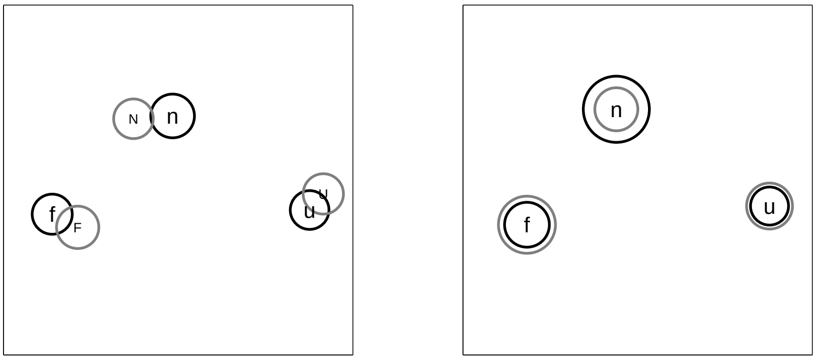

Figure 15.1. Distance displays for the 3×3 table, ignoring political orientation. Left frame is the two-dimensional two-mode distance association model. Lower case letters and black circles represent categories at September; Upper case letters and grey circles the categories at October. Right frame shows the two-dimensional one-mode distance association model, where black circles represent masses at September and grey circles masses at October.

15.2.2 Application to CT data with neglecting the covariate

For sake of illustration of the distance association model, suppose we had not observed the respondents political orientation. In that case we have one table and one might be interested in whether the opinion about CT changed or not. The 3×3-table obtained by summing over Political orientation, is therefore analyzed. The two-dimensional two-mode distance association model is equal to the saturated log-linear model. Looking at its graphical display does give much insight into the structure (see left frame of Figure 15.1). The masses (main effects) are represented by the area of circles, like is discussed in De Rooij (2005a).

We see that there is a large continuity in the opinion of CT, i.e. the favorable categories at September and October are close together, as well as for the other two categories, unfavorable and no clear opinion. Furthermore, we see that the distance from unfavorable at September to favorable in October ($d(u, F)$) is smaller than the other way around ($d(f, U)$), indicating that given the masses there more transitions from unfavorable to favorable then reverse. Some more minor asymmetries are detectable from the left frame, but one might ask whether these asymmetries are due to sampling error or not. Therefore, the two-dimensional one-mode model is fitted, and this model indeed fits the data well ($X^2 = 2.84$; $G^2 = 2.89$; $df = 1$), and the solution is shown in the right frame of Figure 15.1. There we see that the masses of favorable and unfavorable increase while the mass of no clear opinion declined. Furthermore we see that the distance between no clear opinion and favorable is smallest, indicating relatively many transitions, while the distance between favorable and unfavorable is large, indicating relatively few transitions.

15.3 Covariates in Distance Association Models

In many cases of research involving longitudinal categorical data covariates are present and of utmost importance. Consider cases like pre-test post-test designs, where in between the two measurements subgroups get different treatments. Another case is where the subjects can be divided into natural groups, like boys/girls. In the CT example we have an intermediate case, where the covariate is not an experimental condition or a natural grouping, but distinguishes the respondents on a presumed important characteristic.

In this Section we will discuss two approaches to incorporate such covariates into our distance association model. The first way is to reorganize the table as a two-way table and use the two-mode distance association model on that table. Design matrices for the row coordinates might be used to obtain simpler models. The second way is to define a three-way table where the third way corresponds to the covariate-mode and using three-way distance models. We will discuss both approaches and link them to hierarchical log-linear models like the saturated model and the no three-factor association model (see, for example, Agresti, 1990, p. 148).

Before we start the discussion some notation is needed: The measurements at the two time-points will be denoted by T_1 and T_2 and the covariate by Z. The categories of the variable at T_1 will be indexed by $i = 1, \ldots, I$, those at T_2 with $j = 1, \ldots, J$ and the categories of the covariate with $k = 1, \ldots, K$.

15.3.1 Approach I: Design Matrices

The way to take covariates into account discussed in this section is based on methods proposed by Takane (1998). Basically, the table is re-ordered to be a two-way table and the two-mode distance association can be applied to this table. Since the table is in fact three-way, we will label the model as follows:

$$\pi_{ijk} = \alpha_{ik}\beta_j \exp\left(-d^2(\mathbf{x}_{ik}, \mathbf{y}_j)\right). \tag{3}$$

For the CT example, the data is re-ordered to be of size 9×3. Without constraints on the row coordinates and in two dimensions the distance association model equals the saturated model for such a table, and thus also for the original three-way table. By imposing a dimensionality constraint (i.e. a one-dimensional solution), the number of parameters is reduced and we obtain a testable model.

If we write $\mathbf{X} = [\mathbf{x}_{11}^T, \ldots, \mathbf{x}_{ik}^T, \ldots, \mathbf{x}_{IK}^T]^T$ then restrictions on these coordinates can be imposed by using design matrices $\mathbf{A}$, i.e. $\mathbf{X} = \mathbf{A}\mathbf{X}^*$. These design matrices are equal to the design matrices used in regression analysis with categorical predictors. Without restrictions the matrix $\mathbf{A}$ might be written as

$$\mathbf{A} = [\mathbf{a}_1, \mathbf{A}_{M_1}, \mathbf{A}_{M_2}, \mathbf{A}_A] = \left(\begin{array}{c|cc|cc|cccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & -1 & -1 & 0 & 0 & -1 & -1 \\ 1 & -1 & -1 & 1 & 0 & -1 & 0 & -1 & 0 \\ 1 & -1 & -1 & 0 & 1 & 0 & -1 & 0 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{array}\right).$$

The vertical lines correspond to different parts, the constant, the main effects and the association effect all of the association: $\mathbf{A} = [\mathbf{a}_1, \mathbf{A}_{M_1}, \mathbf{A}_{M_2}, \mathbf{A}_A]$. The 'main effect of the association' indicates a two variable association, where the 'association of the association' indicates a three variable association. Since the coordinates for the rows and columns are usually centered, the constant term ($\mathbf{a}_1$) is not necessary for fitting. However, to obtain the correct number of degrees of freedom it is necessary. To obtain simpler models we could, for example, leave the $\mathbf{A}_A$ out of $\mathbf{A}$. In that case, the distances represent an association between the covariate(s) and the measurement on the second time point and an association between the two time points, i.e. in that case the model in maximum dimensionality equals the no three-factor association log-linear model ($[ZT_1][ZT_2][T_1T_2]$).

We could also apply the matrix $\mathbf{A}$ to the main effect vector $\boldsymbol{\alpha}$, with $\boldsymbol{\alpha} = \{\alpha_{ik}\}$. Then instead of an interaction between the covariates and the measurement at the first time point, the $\boldsymbol{\alpha}$ represents the two main effects. The model using the design matrix on both the main effect as well as the row coordinates in full dimensionality equals the log-linear model $[ZT_2][T_1T_2]$.

It should be noted that in this modeling approach the one-mode distance association model is not possible, or put otherwise, is equal to the model for the 3×3-table with symmetry restrictions.

15.3.2 Approach II: Space Transforming Matrices

The second way of treating covariates has its roots in papers of Carroll and Chang (1970) and Carroll and Wish (1974) where they discuss individual difference models in multidimensional scaling. The table to be analyzed in these proposals is three-way where two of its ways represent the same mode. The above mentioned papers only discuss one-mode models, but the ideas can easily be generalized to two-mode models. The idea of this approach is that for each $k = 1, \ldots, K$ a different distance model can be fitted, but that these dis-

Table 15.2. Fit statistics for distance association models using design matrices. The numbers refer to X^2/G^2 (df).

	dimensionality	
model	$p = 1$	$p = 2$
$\alpha_{ik}\beta_j \exp\left(-d^2(x_{ik}, y_j)\right)$	21.19/20.99 (7)	0/0 (0)
$\alpha_{ik}\beta_j \exp\left(-d^2(x_{i+k}, y_j)\right)$	33.93/32.47 (11)	7.67/7.80 (8)
$\alpha_{i+k}\beta_j \exp\left(-d^2(x_{i+k}, y_j)\right)$	67.39/68.12 (15)	41.73/42.41 (12)
$\alpha_{ik}\beta_j \exp\left(-d^2(x_i, y_j)\right)$	57.45/49.46 (13)	28.15/27.15 (12)

tance models can be related to each other in a functional way. The function is then defined by a space transforming matrix ($\mathbf{B}_k$). Without the functional constraints the model is

$$\pi_{ijk} = \alpha_{ik}\beta_{jk} \exp\left(-d^2(\mathbf{x}_{ik}, \mathbf{y}_{jk})\right) \tag{4}$$

Let us define the matrices $\mathbf{X}_k$ and $\mathbf{Y}_k$ as $\mathbf{X}_k = [\mathbf{x}_{1k}, \ldots, \mathbf{x}_{ik}, \ldots, \mathbf{x}_{Ik}]^T$, and $\mathbf{Y}_k = [\mathbf{y}_{1k}, \ldots, \mathbf{y}_{jk}, \ldots, \mathbf{y}_{Jk}]^T$. Now, in order to obtain simpler models the $\mathbf{X}_k$'s ($\mathbf{Y}_k$'s) can be constrained for $k = 1, \ldots, K$ to be related to each other in a number of ways, where always the following equality holds: $\mathbf{X}_k = \mathbf{X}\mathbf{B}_k$. The different possibilities are that $\mathbf{B}_k$ is non-singular, is diagonal, or is equal to the identity matrix. Furthermore, one-mode restrictions, $\mathbf{X}_k = \mathbf{Y}_k$, are possible.

The model defined in (4) in its most general form is equal to the saturated log-linear model ($[ZT_1T_2]$); without the distance part the model equals the log-linear model ($[ZT_1][ZT_2]$). So, the distances represent the two-way association between T_1 and T_2, but also the three-way association ZT_1T_2. When the matrices $\mathbf{B}_k$ do not depend on the index k, no three-way association is represented by the distances.

15.4 Approach II: Application CT to Data

In this section we will apply the described approaches to the data in Table 1. The two approaches will be dealt with in the two subsections.

15.4.1 Approach I

In Table 15.2 the results for some models are summarized, where we use the subscript $i + k$ to indicate that a design matrix with $\mathbf{A}_A$ removed was used, i.e., $\mathbf{A} = [\mathbf{a}_1, \mathbf{A}_{M_1}, \mathbf{A}_{M_2}]$, and the subscript ik for the unrestricted case.

In Table 15.2 we see that all one-dimensional models give a bad fit. For the two-dimensional models we see that only the saturated (i.e. without con-

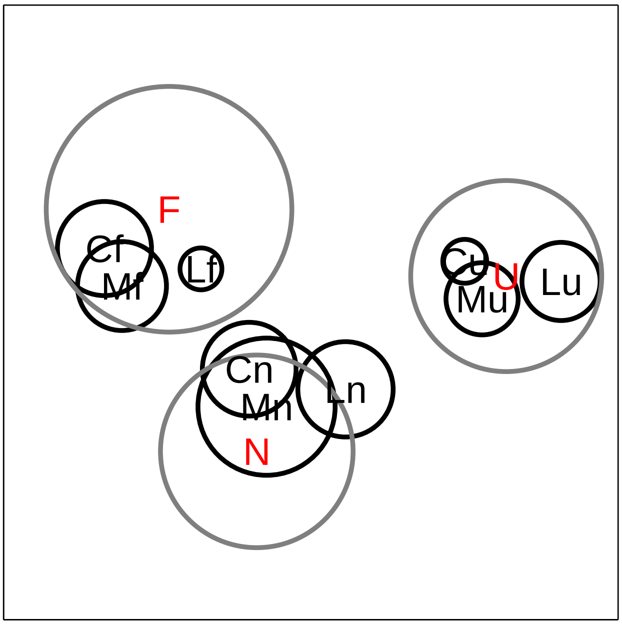

Figure 15.2. Distance display in two dimension with a design matrix on the row categories for the 9×3 *table. Double letters with black circles refer to combinations of political orientation and opinion about CT at September. Upper case letters with grey circles refer to opinion at October.*

straints) and the model with restrictions on the row coordinates fit well. The latter model equals the model with all pairwise associations.

The solution is shown in Figure 15.2. In this graphical representation the masses of row points are difficult to compare to the masses of the column points. We can however, easily compare them within the row or column categories. Concerning the distances, like in the analysis without covariate, a clear consistency in opinion is visible, that is, the favorable categories at September are close to the favorable category of October, and similarly for the other opinion categories. Furthermore, there are little differences between political orientations: The liberals are generally somewhat closer to the unfavorable opinion; The conservatives somewhat closer to the favorable condition; The moderates are closer to the no clear opinion category. Are these differences important? On the last line of Table 15.2 the fit statistics of the model in which these points are constraint to be equal shows that they are indeed important, i.e. that model does not fit the data.

15.4.2 Approach II

In this Section we discuss the second approach to dealing with covariates for the data represented in Table 1. Fit statistics are given in Table 15.3. Again, one-dimensional models do not fit the data adequately. From the two-dimensional models the most simple one, the two-dimensional one-mode with restriction $\mathbf{B}_k = \mathbf{I}$ fits the data well and is shown in Figure 15.3. The association is the same for each group of subjects and is symmetric. Thus for each of the tables we have a quasi-symmetric form, where the margins may differ but the association structure is symmetric. In Figure 15.3 the masses are represented by the radius of the circle, where black circles are used for T_1 and grey circles for T_2. We see that the frames of Figure 15.3, corresponding to the three groups of political orientation, hardly differ. Some minor noticeable differences are that for the conservative group (right frame) the masses for unfavorable are relatively small, while for the liberal group (left frame) the masses of the favorable group are relatively small. Moreover, it seems that in the moderate group (middle frame) the masses of no clear opinion is relatively large. The pattern of masses is the same in all three groups: favorable and unfavorable gain mass where no clear opinion loses mass.

Table 15.3. Fit statistics for distance association models using space transforming matrices. The numbers refer to X^2/G^2 (df).

	dimensionality	
Distance defined by	$p = 1$	$p = 2$
$\mathbf{X}_k$, $\mathbf{Y}_k$	18.66/18.98 (3)	0/0 (0)
$\mathbf{XB}_k$, $\mathbf{YB}_k$ with $\mathbf{B}_k$ diagonal	31.11/28.96 (7)	0.47/0.47 (2)
$\mathbf{X}$, $\mathbf{Y}$	31.35/29.53 (9)	7.67/7.80 (8)
$\mathbf{X}_k$	21.97/22.19 (6)	5.10/5.34 (3)
$\mathbf{XB}_k$ with $\mathbf{B}_k$ diagonal	32.73/30.76 (8)	5.13/5.37 (4)
$\mathbf{X}$	33.30/31.42 (10)	10.16/10.23 (9)

15.5 Stability

To obtain stability estimates of parameters data reuse methods can be used, like the bootstrap and the jackknife. Here the jackknife procedure will be exploited since for contingency tables it is computationally very cheap (Clogg and Shidadeh, 1994, pp. 34-38; Dayton, 1998, pp. 22-23). Let ξ be any parameter of interest and $\hat{\xi}$ be its estimator. For a data set with n observations an estimate of ξ is obtained for n samples where the i-th observation is deleted, $i = 1, \ldots, n$. Let $\hat{\xi}_i$ be the estimate with the i-th observation deleted. The estimated standard

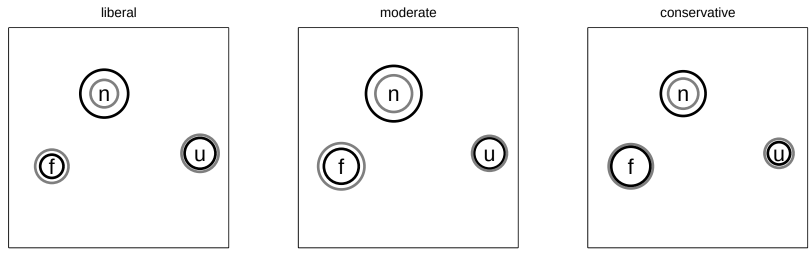

Figure 15.3. The two-dimensional one-mode solution with restriction $\mathbf{B}_k = \mathbf{I}$. Left plot for the 'liberals', middle plot for the 'moderates', and right plot for the 'conservatives'. Masses are represented by the area of the circles: Black circles for September; Grey for October.

error using the jackknife is defined as (see Dayton, 1998, pp. 22-23)

$$SE_J(\hat{\xi}) = \sqrt{\frac{n}{n-1}\sum_{i=1}^{n}(\hat{\xi}_i - \hat{\xi})^2}. \qquad (5)$$

There is some evidence that the standard errors obtained from the jackknife procedure are somewhat too large (see Dayton, 1998, p. 37). In general the number of computations to be done is equal to the sample size. However, for contingency tables each observation in a given cell produces the same calculation as other observations in this cell. Therefore the number of calculations reduces from the sample size (n) to the number of cells of the contingency ta-

Table 15.4. Coordinates and Standard errors for Solution shown in Figure 15.2.

Category	Coordinate-1	SE	Coordinate-2	SE
Lf	-0.4289	0.1056	0.1697	0.0953
Lu	0.9777	0.0676	0.1202	0.0587
Ln	0.1361	0.0779	-0.3013	0.0921
Mf	-0.7373	0.0922	0.1022	0.0692
Mu	0.6693	0.0740	0.0527	0.0726
Mn	-0.1723	0.0788	-0.3689	0.0635
Cf	-0.8053	0.0804	0.2487	0.0633
Cu	0.6012	0.0941	0.1991	0.0823
Cn	-0.2403	0.0948	-0.2224	0.0843
F	-0.5529	0.0510	0.4021	0.0467
U	0.7634	0.0250	0.1414	0.0464
N	-0.2105	0.0679	-0.5435	0.0252

ble. Confidence intervals can be obtained from these estimated standard errors by $\hat{\xi} \pm 1.96 \times SE_J(\hat{\xi})$.

We will not show the standard errors of all solutions, but use the solution shown in Figure 15.2 as an exemplar. Table 15.4 gives the coordinates of all categories on the two dimensions plus the standard errors. It can be seen that the solution is quite stable[1]. This is probably due to the fact that the data contain few small frequencies.

15.6 Conclusions and Discussion

Distance association models were discussed for longitudinal categorical data. These models provide a map or scale on which the transitions occur. The incorporation of categorical covariate variables was introduced. Two approaches to incorporate covariates into the distance association models were treated: The first approach has the advantage that the final model can be represented in a single display, whereas two displays are needed for the representation of the model in the second approach; The second approach has the advantage that the strength of the association between the measurements of the two time points in different groups can be compared: the larger the distances the stronger the association.

We only discussed models for a single covariate. In practical situations it is most likely that more information about the subjects is available. In that case both approaches can still be used. In the first approach the design matrices just become larger with more main effects and more interaction effects. In the second approach more displays are obtained, but the space transforming matrices itself might be subjected to design matrices itself.

In the case of measurements at three time points *triadic distance models* might be used. Triadic distance association models were discussed in De Rooij (2005a). The incorporation of covariates in such a framework can follow the lines suggested here. However, with increasing time points and increasing number of covariates the number of cells in the table also increases rapidly and one has to be careful not to obtain very sparse tables. Moreover, in that case there is also risk of loosing subjects over time. How to deal with such patterns with missing data in the current framework is still a subject of research.

References

Agresti, A. (1990). *Categorical data analysis.* New York: John Wiley and Sons.

[1]This is also true for the other solutions shown in this chapter

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling; theory and applications.* New York: Springer.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart- Young' decomposition. *Psychometrika*, *35*, 283-319.

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krant, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (p. 57-105). San Francisco: Freeman.

Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation [contributions to the statistical analysis of correlation matrices]. *Annals of the faculty of Science, University of Toulouse*, *29*, 715-720.

Clogg, C. C., & Shidadeh, E. S. (1994). *Statistical models for ordinal variables.* Thousand oaks: Sage Publications.

Dayton, C. M. (1998). *Latent class scaling analysis.* Thousand Oaks, CA: Sage.

De Rooij, M. (2001a). Distance association models for the analysis of repeated transition frequency data. *Statistica Neerlandica*, *55*, 157–181.

De Rooij, M. (2001b). *Distance models for the analysis of transition frequencies.* Unpublished doctoral dissertation, Leiden University.

De Rooij, M. (2002). Distance models for three-way tables and three-way association. *Journal of Classification*, *19*, 161–178.

De Rooij, M. (2005a). The analysis of change, newton's law of gravity, and association models. *submitted*.

De Rooij, M. (2005b). The geometry of generalized biadditive models. *submitted*.

De Rooij, M., & Gower, J. C. (2003). The geometry of triadic distance models. *Journal of Classification*, *20*.

De Rooij, M., & Heiser, W. J. (2000). Triadic distance models for the analysis of asymmetric three-way proximity data. *British Journal of Mathematical and Statistical Psychology*, *53*, 99–119.

De Rooij, M., & Heiser, W. J. (2003). A distance representation of the quasi-symmetry model and related distance models. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J. J. Meulman (Eds.), *New developments on psychometrics: proceedings of the international meeting of the psychometric society* (p. 487-494). Tokyo: Springer-Verlag.

De Rooij, M., & 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.

Goodman, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetric models for contingency tables with or without missing entries. *The Annals of Statistics, 13*, 10-69.

Gower, J. C., & De Rooij, M. (2003). A comparison of the multidimensional scaling of triadic and dyadic distances. *Journal of Classification, 20*, 115–136.

Groenen, P. J. F. (2003). *Dynamische meerdimenionele schaling: statistiek op de kaart.* Rotterdam, The Netherlands: Inaugural address Erasmus Research Institute of Management.

Jansen, G. G. H., & Bus, A. G. (1982). *An application of a logistic bio-assay model to achievement test data* (Tech. Rep.). Department of Education: Goningen University.

Takane, Y. (1998). Visualization in ideal point discriminant analysis. In J. Blasius & M. J. Greenacre (Eds.), *Visualization of categorical data* (pp. 441–459). New York: Academic Press.

Timmerman, M. E., & Kiers, H. A. L. (2002). Three-mode principal component analysis with smoothness constraints. *Computational Statistics & Data Analysis, 40*, 447-470.