A SPECIAL JACKKNIFE FOR MULTIDIMENSIONAL SCALING

Jan de Leeuw

Jacqueline Meulman

Department of Data Theory FSW/RUL

Middelstegracht 4

2312 TW Leiden

Netherlands

Abstract

In this paper we develop a version of the Jackknife which seems especially suited for Multidimensional Scaling. It deletes one stimulus at the time, and combines the resulting solutions by a least squares matching method. The results can be used for stability analysis, and for purposes of cross validation.

Keywords

Multidimensional Scaling, Jackknife, Cross Validation, Stability.

1 INTRODUCTION

The results of a data analysis are not complete without some form of information about the stability of the solution. There are many types of stability that can be distinguished (Gifi, 1981), and consequently there are many forms of stability analysis (or sensitivity analysis). Most of these techniques have in common that they study the effect of a small perturbation of the data on the solution, or aspects of the solution. A method is stable, in a particular application, if small changes in the data produce only small changes in the solution. It is clear that statistical analysis, in the form of computation of standard errors or confidence intervals for example, is a particular kind of sensitivity analysis, which studies the perturbations induced by random sampling.

In most applications of Multidimensional Scaling (MDS) the stability of the solutions is not studied at all. There are exceptions, however. The method of Maximum Likelihood (ML) has been applied to MDS by Ramsay, Takane, Carroll, de Soete, Zinnes, and others. Compare Ramsay (1982) for a review. At least in some cases standard large sample theory for ML estimation can be applied to obtain information about statistical stability. It is true, however, that the optimality properties of ML estimation often do not apply directly to MDS situations.

This is because of non-standard aspects of such situations, which are riddled with incidence parameters, dependencies, lack of replications, and so on. Compare the discussion of Ramsay (1982).

If there are independent replications in a MDS experiment, then statistical information can be obtained even without assuming parametric distributional models or specific response functions. This is indicated in Stoop, Heiser, and De Leeuw (1981). One can use, in this case, the standard delta method results. One could also apply the currently popular resampling techniques such as the Jackknife or the Bootstrap (Efron, 1982). Heiser and Meulman (1983a, 1983b) review the use of resampling in MDS.

If there are no replications, matters become more complicated. We need a parametric model to derive statistical information on stability, but very often the prior knowledge required to make the necessary assumptions with some confidence is not available. And there are other situations in which the whole idea of random sampling does not make much sense, and in which statistical stability is consequently not interesting, in fact not even defined. Nevertheless it is still possible, in such cases, to perform other forms of stability analysis. The idea of a small change in the data can still be defined in various ways, and the effects of such small perturbations can still be studied. This was already indicated,

in the MDS context, by Kruskal and Wish (1978, page 58-60). Our contribution makes their suggestions explicit, and implements and applies them.

A first suggestion, and a rather natural one, is to make a small change in one dissimilarity in a metric MDS problem. The implicit function theorem can be used to compute the effect of such a small change on the loss function and the configuration computed by the MDS algorithm. For classical metric scaling (also known as Torgerson-Gower scaling or Young-Householder scaling) some of the necessary results have already been given by Sibson (1979). We can use these results to define a type of influence function for multidimensional scaling (Hampel, 1974). Under random sampling assumptions such an influence function is a sort of 'infinitesimal' Jackknife (Jaeckel, 1972). Influence functions for the closely related principal components problem have been computed recently by Critchley (1984), and influence functions for various forms of correspondence analysis and canonical analysis problems can be deduced from Gifi (1981), De Leeuw (1984). The more familiar 'discrete' version of the Jackknife, reviewed by Gray and Schucany (1972) and Miller (1974), corresponds with deleting one dissimilarity and treating it as missing. We then reanalyze the data with this dissimilarity omitted, and combine the $\binom{n}{2}$ different solutions obtained by omitting each dissimilarity successively. This scheme can also be

applied to nonmetric MDS, and may be quite useful. For a large number of objects, however, we have to perform many MDS analyses, and the perturbations by leaving out one dissimilarity will be very small indeed.

In this paper we study another data analytical version of the Jackknife, which investigates the effect of somewhat larger perturbations of the data. We emphasize, from the start, that our results are not supposed to have inferential applications. We do not use them to assess statistical stability in some sense or another. Our purposes are purely data analytical. The starting point is that we ask ourselves what happens to our multidimensional scaling solution if we delete all information on one object. 'Likewise stimuli could be eliminated from the data matrix, and solutions determined for the remaining stimuli using the "Jackknife" idea of J.W. Tukey.' (Kruskal and Wish (1978), page 59). Instead of analyzing the dissimilarities between n stimuli once, we analyze, in addition, n times the dissimilarities between n-1 stimuli, by deleting each one in turn. The stability question associated with this scheme is interesting in its own right, and the results of the n+1 analyses can be portrayed quite nicely in a single plot (cf below). Thus our technique is well suited for data analysis. It is clear that it can be applied to both metric and nonmetric MDS. We shall indicate further that it can also be combined with cross validation types of assessment, along the lines indicated

by Stone (1974), although a great deal of further study seems required to find out what its properties are in that context. We shall investigate, tentatively, if our leave-one-out method can be used to determine a 'correct' or 'optimal' dimensionality. Previous procedures for choosing dimensionality are based on large scale Monte Carlo studies, and were reviewed recently by Spence (1983). Monte Carlo studies of robustness of MDS and of dimensionestimation differ from our scheme, because Monte Carlo methods are used in situations in which the 'correct' answer to the question that is studied is known beforehand.

2 TECHNIQUE

The basic idea behind our Jackknife is simple. We perform n+1 multidimensional scaling analyses. The solution to the original problem is X_0, the solutions to the n additional problems are X_i (i=1,....,n), where X_i is the solution to the problem without object i. Thus X_0 is a n x p configuration matrix, while the X_i only have information on the location of n-1 points. For convenience we also locate them in n x p matrices, where row i of X_i is always zero. Of course these zeroes are 'structural' zeroes, which must be distinguished from other zeroes which may occur in the matrix. All X_i can be computed with an ordinary multidimensional scaling program, both in metric and nonmetric cases. Of course if we intend to use this form of the Jackknife in a routine matter, then it will be convenient to build appropriate modifications into the software of the programs.

The major problem in comparing the X_i is, of course, that information on the location of one point is missing from each of them. In addition comparisons are complicated because of the familiar translational and rotational indeterminacy of scaling solutions. We propose to do the comparison by performing matching of the configurations X_i (i=1,....,n).

Define

$$Y_i = a_i X_i K_i + e_i b_i' + e c_i',$$

with K_i an orthogonal rotation, c_i a translation vector, b_i the location of the missing point, and a_i a scalar for uniform stretching or shrinking. Vector e_i is the i^{th} unit vector, and vector e has all elements equal to unity. In metric scaling it is often better to set $a_i = 1$ for all i, but in nonmetric scaling we usually want to compute the optimal a_i. We perform matching by minimizing the least squares loss function.

$$\sigma(a_i, b_i, c_i, K_i, Y_0) = \sum_{i=1}^{n} \text{tr} \ (Y_0 - Y_i)'(Y_0 - Y_i) \qquad (1)$$

over all unknowns. The n x p matrix Y_0 is the <u>comparison matrix</u>, it is also unknown. We discuss some normalization conventions. In the first place the data matrices X_i are are supposed to be centered columnwise. If we solve for the a_i we suppose in addition that tr $X_i'X_i = 1$. This merely implies a particular scaling of the data, it has no influence on the solution. For identification purposes we also require that the columns of Y_0 sum to zero. Again this causes no loss of generality. If we solve for the optimal a_i we need an additional constraint to prevent degeneracy. We use $\Sigma a_i^2 = 1$.

We minimize (1) in three steps. In the first step we minimize over b_i and c_i, for fixed a_i, K_i, and Y_0. The optimal c_i and b_i are

$$\hat{c}_i = -y_i^0/(n - 1), \qquad (2a)$$

$$\hat{b}_i = n y_i^0/(n - 1). \qquad (2b)$$

Here y_i^0 is row i of Y_0. If we substitute this in (1) we find the 'partial minimum'

$$\sigma(a_i,*,*,K_i,Y_0) = \sum_{i=1}^{n} \text{tr}(Y_0-a_iX_iK_i)'(Y_0-a_iX_iK_i) - \frac{n}{n-1}\text{tr}Y_0'Y_0. \quad (3)$$

Arguments over which we have minimized are replaced by an asterix. In the second step we minimize (3) over Y_0, for fixed a_i and K_i. The solution is

$$\hat{Y}_0 = \frac{n-1}{n(n-2)} \sum_{i=1}^{n} a_iX_iK_i. \quad (4)$$

The new partial minimum is

$$\sigma(a_i,*,*,K_i,*) = \sum_{i=1}^{n} a_i^2 \text{tr}(X_i'X_i) - \frac{n-1}{n(n-2)} \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j \text{tr}(K_i'X_i'X_jK_j). \quad (5$$

If the a_i are fixed, as in metric scaling, then all that remains to be done is maximization of $\Sigma\Sigma \text{tr}(K_i'X_i'X_jK_j)$ over the K_i. If the a_i are free, then we must maximize $\Sigma\Sigma a_ia_j \text{tr}(K_i'X_i'X_jK_j)$ over the K_i and over the a_i. This is exactly the problem discussed by Gower (1975) and Ten Berge (1977a, 1977b), who discuss simple alternating least squares algorithms to solve this problem and find the optimal K_i (and a_i). Together with (2) and (4) this defines $\hat{Y}_i$, the optimally matched and completed configuration. Finally we rotate X_0 in such a way that X_0 and Y_0 are matched optimally, which again facilitates comparisons between our different solutions.

3 HOW TO USE THE RESULTS

After matching we have n+2 solutions that can be compared: we have our original X_0, we have the n matched solutions Y_i (i=1,....,n), and their average Y_0. Our first assessment of stability is obtained by comparing the Y_i. This can be done graphically, by plotting all Y_i as the endpoints of stars with centers given by Y_0. This will be illustrated in section 4. In addition to the graphical comparisons, we can compute dispersions of the rows of Y_i around their centroids Y_0, or around the original solution X_0. In the illustrations in section 4 we will use the following measure:

$$\text{STAB} : 1 - \frac{\sum_{i=1}^{n} \left\| Y_i - Y_0 \right\|^2}{\sum_{i=1}^{n} \left\| Y_i \right\|^2}$$

This stability measure can be interpreted as the ratio of Between to Total variance (cf. Heiser & Meulman,1983b).

For cross validation purposes we can investigate if the "predicted" position of object i, which is row i of Y_i, corresponds with the "actual" position of object i, which is row i of X_0. It follows from (2a) and (2b) that row i of Y_i is equal to row i of Y_0. Thus each row of Y_0 comes from a different Multidimensional Scaling analysis, although all rows are only known after a single matching analysis. Similarity between Y_0 and X_0, established

either numerically or graphically, can consequently also be used for cross validation. It is possible that, in some sense, the difference between these two configurations can be used to correct for bias (as in the ordinary Jackknife). We do not have any theory yet which would indicate how such a combination can be made. To measure cross validity we will use

$$\text{CROSS} : 1 - \frac{n \left\| X_0 - Y_0 \right\|^2}{\sum_{i=1}^{n} \left\| Y_i \right\|^2}$$

By normalizing STAB and CROSS in this manner, solutions are comparable across methods, and the dispersion around the original solution X_0 (DISP) can easily be derived. Since

$$\frac{1}{n} \sum_{i=1}^{n} \left\| Y_i - X_0 \right\|^2 = \frac{1}{n} \sum_{i=1}^{n} \left\| Y_i - Y_0 \right\|^2 + \left\| X_0 - Y_0 \right\|^2$$

it follows that DISP = 2 - (STAB + CROSS).

4 ILLUSTRATIONS

In the following paragraphs the Jackknife will be illustrated with two different examples. Various MDS methods have been applied, which we summarize below. To avoid ambiguous terminology we will denote them with acronyms in terms of the loss function that is minimized by them.

Insert table 1 about here

Table 1 uses the following notation:

δ_{ij} : dissimilarity between object i and object j

X : configuration in p-dimensional space

d_{ij} : distance between object i and object j in configuration X

B : $-\frac{1}{2}$ double centered matrix of squared dissimilarities

α : additive constant

$\hat{d}$: optimally transformed dissimilarities by a monotone function

LSSP is least squares on the scalar products; it is a metric method and also known as classical, or Torgerson-Gower MDS;

LSD is least squares on the distances; it is a metric method;
LSDA is like LSD, with in addition an additive constant estimated;
LSDN is least squares on the distances, nonmetric, and also known as Kruskal-Shepard MDS. A special version is called LSDR, which denotes that the monotone regression is performed across the separate rows of the data matrix.
LSSD is least squares on the squared distances. The metric version has been applied;
LSLD is least squares on the logarithms of the distances.
LSD, LSDA, LSDN, and LSDR have been minimized by means of the SMACOF algorithm (see e.g. De Leeuw & Heiser, 1980). For LSSD the ALSCAL program has been used (see e.g. Takane, Young & De Leeuw, 1977), and for LSLD MULTISCALE has been applied (see e.g. Ramsay, 1977). LSSP is performed by the computation of the standard initial configuration in the SMACOF program.

4.1 Illustration 1: Color data

The data for illustration 1 have been taken from Torgerson (1958), and concern dissimilarities between 9 Munsell colors all of the same red hue but differing from each other in brightness and saturation. The judgments were obtained from 38 subjects. These data have been used in two different versions: an asymmetric data matrix, and

a symmetric one with an additive constant eliminated (Torgerson, 1958, table 2 on page 285, and table 7 on page 287 respectively). The symmetric data have been analysed by LSSP, LSD, and LSDN, while LSDR has been applied to the asymmetric tabel.

The Jackknife has been performed for this example to investigate whether Torgerson's conclusions about the dimensionality of the data could be sustained. Table 2 summarizes results for STAB and CROSS for the different methods with varying dimensionality.

Insert table 2 about here

Within the same dimensionality results for STAB and CROSS are almost perfectly consistent, while across dimensions the performances of the various methods differ slightly. It is clear that the one-dimensional solutions are special and discredited by STAB and CROSS results. When we inspect the significant decimals for the two- and three-dimensional solutions, ignoring the overall stability, it is clear that the two-dimensional solutions have to be preferred. STAB and CROSS confirm Torgerson's conclusion, which he had drawn using a completely different criterion while performing LSSP. For none of the dimensionalities

it is true that STAB and CROSS are monotone increasing with the values of stress for the different solutions (which are ordered, by definition, from LSDR upto LSSP).

4.2 Illustration 2: Nation data

The second illustration concerns data from Kruskal & Wish (1978). The degree of overall similarity between 12 nations was rated by 18 students on a scale from 1 (very different) to 9 (very similar) and was averaged to obtain a mean similarity matrix. Since MDS methods approximate dissimilarities rather than similarities, the data were transformed by substracting the mean rating from 9, the largest value of the scale. This resulted in a dissimilarity matrix labelled D^{+}. This procedure immediately suggests the first application of the Jackknife, which is to investigate whether we are dealing with an additive constant problem contained in the transformed data. Various metric analyses of D^{+} have been performed. In addition they have been applied to D^{-}, a dissimilarity matrix with an additive constant eliminated. This estimation was done by applying the Messick and Abelson procedure, as described in Torgerson, 1958. To complete the results D^{+} has been analyzed by minimizing LSDA, i.e. iterative estimation of the additive constant.

Insert table 3 about here

It is clear from table 3 that LSD suffers most from the additive constant problem: both STAB and CROSS increase when the additive constant is eliminated, either by analyzing D^- or by analyzing D^+ and estimating the additive constant iteratively. LSSP, which involves the double-centering of the squared dissimilarity matrix, is not much affected by the additive constant.

The second application of the Jackknife for the nation data again concerns the dimensionality. Kruskal & Wish remark that the data might be four or even five dimensional (Kruskal & Wish, 1978, page 56). To see if this is really the case, LSSP and LSD have been applied to D^-, and LSDA and LSDN to D^+, with varying dimensionality. Results are given in table 4.

Insert table 4 about here

The most important conclusion that can be drawn from table 4 is that the three-dimensional solution has to be preferred, according to both STAB and CROSS for all methods. Furthermore, LSSP is most stable in three dimensions, compared with the other methods, but this is not true for the four-dimensional solution, where LSD performs best. These findings are consistent with the results in table 2 for the color data in two and three dimensions, and might very well be explained by the way the dissimilarities are approximated. By LSSP they are approximated from below, while by the other methods the approximation is simultaneously from below and from above.

Our final application concerns the comparison of LSD with its most important rivals, which are LSSD and LSLD. Our hypothesis is that LSSD, because of its relation with LSSP (cf. De Leeuw & Heiser, 1982), will give similar results, thus LSSD will be more stable than LSD.

Insert table 5 about here

This hypothesis is confirmed by both STAB and CROSS. According to both criteria LSLD gives less stable results compared with other methods.

Insert Figure 1 about here

Insert Figure 2 about here

Insert Figure 3 about here

Insert Figure 4 about here

To complete our second illustration we present parts of the Jackknife results in a graphical way. We have chosen the solutions for LSSP, LSSD, LSD and LSLD, which can be compared in this manner. By looking at the separate stars in a configuration specific stability information can be deduced for each object. Because of the large overall stability, the endpoints of the stars have not all been labelled, since this would yield very smudgy plots. When performing an actual data analysis, however, it might be very worthwile to enlarge the scale of the figure and look for special patterns. To give some idea, we have labelled the endpoint of each star indicating the position of the object when object 1 (Brazil) was left out of the analysis. We see e.g. that Egypt and Cuba move away from each other; the same is true for the U.S.A. and Israel, while Russia, Yugoslavia and Cuba retain their interrelations. This pattern is consistent across the four methods.

It is beyond the scope of this paper to scrutinize all aspects of the various solutions. We want to draw attention, however, to the close resemblance between the solutions for LSSP and LSSD. It seems that apart from details, like the position of Russia and Yugoslavia, the ALSCAL solutions hardly move away from their initial configuration, which is obtained by the Torgerson scaling method. Apart from the fact that LSD is more stable than LSLD, the SMACOF and MULTISCALE configurations don't show major differences. They are more similar to each other

than to the ALSCAL solution.
We do not present graphical results concerning the cross validation aspects of our Jackknife. As can be deduced from tables 3-5 the "actual" position and the "predicted" position in the three dimensional solutions are almost identical. This state of affairs, ofcourse, credits the analyses we have performed. Multidimensional Scaling again proves to be a very robust technique, even when large perturbations of the data are applied.

5 DISCUSSION

In this paper a special Jackknife is presented that gives information about the stability of MDS solutions, especially for the case when there are no replications. This procedure gives very useful results, that are however obtained at the cost of a lot additional computation. Each analysis has to be performed n + 1 times, where n is the number of objects in the data matrix. This might be prohibitive when the available CPU-time is limited. Moreover, a lot of separate manipulations are needed as long as the procedures like the Jackknife are not built into standard MDS programs.
Finally we want to report some general experiences in our stability study. Experimentation by comparing the FORTRAN SMACOF program with a program in APL indicated that in

some cases the APL results, which are more precise, were less stable than the FORTRAN results. This might be related to our finding that incomplete convergence of the algorithms tends to give overoptimistic stability stability results, i.e. when the (stable) LSSP solution is used as the initial configuration. If we would have started with random initial configurations, we expect the reverse to be true: incomplete convergence will then lead to less stable solutions.

REFERENCES

Critchley, F. (1984), Influence in principal components analysis. To appear in Biometrika.

De Leeuw, J. (2984), Statistical properties of multiple correspondence analysis. Submitted for publication.

De Leeuw, J. & Heiser, W.J. (1980), Multidimensional scaling with restrictions on the configuration. In P.R. Krishnaiah (ed.), Multivariate analysis, Vol. 5 Amsterdam: North Holland.

De Leeuw, J. & Heiser, W.J. (1982), Theory of multidimensional scaling. In P.R. Krisnaiah & L. Kanal (Eds.), Handbook of statistics, Vol. 2, Amsterdam: North Holland

Efron, B (1982), The Jackknife, the Bootstrap and other resampling plans. CBMS NSF Regional Conference Series in Applied Mathematics, No. 38, Philadelphia: SIAM.

Gifi, A. (1981), Nonlinear multivariate analysis. Leiden: Dept. of Data Theory.

Gower, J.C. (1975), Generalized Procrustes analyses. Psychometrika, 40, 33-51.

Gray, H.L. & Schucany, W.R. (1972), The generalized Jackknife. New York: Marcel Dekker, Inc.:

Hampel, F.R. (1974), The influence curve and its role in robust estimation. Journal of the American Statistical Association, 69, 383-393.

Heiser, W.J. & Meulman, J. (1983a), Analyzing rectangular tables by joint and constrained multidimensional scaling. Journal of Econometrics, 22, 139-167.

Heiser, W.J. & Meulman, J. (1983b), Constrained multidimensional scaling, including confirmation, Applied Psychological Measurement, 7, 381-404.

Jaeckel, L. (1972), The infinitesimal Jackknife, Memorandum MM 72-1215-11. Murray Hill: Bell Laboratories.

Kruskal, J.B. & Wish, M. (1978), Multidimensional Scaling. Beverley Hills and London: Sage Publications.

Miller, R.G. (1974), The Jackknife - a review, Biometrika, 61, 1-17.

Ramsay, J.O. (1977), Maximum Likelihood estimation in multidimensional scaling, Psychometrika, 42, 241-266.

Ramsay, J.O. (1982), Some statistical approaches to multidimensional scaling data, Journal of the Royal Statistical Association, Series A, 145, 285-312.

Sibson, R. (1979), Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling. Journal of the Royal Statistical Society Series B, 41, 217-229.

Spence, I. (1983), Monte Carlo simulation studies. Applied Psychological Measurement, 7, 405-425.

Stone, M. (1974), Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society, Series B, 36, 111-147.

Stoop, I, Heiser, W.J. & De Leeuw, J. (1981), How to use SMACOF-IA. Leiden: Dept. of Data Theory.

Takane, Y., Young, F.W. & De Leeuw, J. (1977), Nonmetric individual differences scaling: An alternating least squares method with optimal scaling features. Psychometrika, 42, 1-67.

Ten Berge, J.M.F. (1977a), Optimizing factorial invariance. Doctoral Dissertation, University of Groningen.

Ten Berge, J.M.F. (1977b), Orthogonal Procrustes rotation for two or more matrices. Psychometrika, 42, 267-276.

Torgerson, W.S. (1958), Theory and methods of scaling. New York: Wiley.

Table 1. MDS methods that have been applied.

1. LSSP	$\sigma^2(X)$	$= \text{tr}\ (BXX')^2$
2. LSD	$\sigma^2(X)$	$= \sum_{ij}\sum (\delta_{ij}-d_{ij}(X))^2$
3. LSDA	$\sigma^2(X,\alpha)$	$= \sum_{ij}\sum ((\delta_{ij}+\alpha)-d_{ij}(X))^2$
4. LSDN	$\sigma^2(X,\hat{d})$	$= \sum_{ij}\sum (\hat{d}_{ij}-d_{ij}(X))^2$
5. LSSD	$\sigma^2(X)$	$= \sum_{ij}\sum (\delta^2_{ij}-d^2_{ij}(X))^2$
6. LSLD	$\sigma^2(X)$	$= \sum_{ij}\sum (\log \delta_{ij}-\log d_{ij}(X))^2$

Table 2. Jacckknife results for the color data.

METHOD	STAB			CROSS		
	1-dim	2-dim	3-dim	1-dim	2-dim	3-dim
LSSP	.8596	.9994	.9951	.9801	1.000	.9994
LSD	.8423	.9994	.9974	.9678	1.000	.9999
LSDN	.8600	.9988	.9979	.9829	.9998	.9997
LSDR	.7721	.9991	.9960	.9400	.9999	.9992

Table 3. Jackknife results for nation data: additive constant problem.

METHOD	STAB 2-dim	CROSS 2-dim
LSSP (D^{+}_{-})	.9814	.9993
LSSP (D^{-}_{+})	.9812	.9993
LSD (D^{+}_{-})	.9633	.9926
LSD (D^{-}_{+})	.9721	.9959
LSDA (D^{+})	.9728	.9933

Table 4. Jackknife results for the nation data: dimensionality problem.

METHOD	STAB			CROSS		
	2-dim	3-dim	4-dim	2-dim	3-dim	4-dim
LSSP (D^-)	.9812	.9891	.9630	.9993	.9997	.9974
LSD (D^-)	.9721	.9880	.9809	.9959	.9996	.9995
LSDA (D^+)	.9728	.9808	.9640	.9933	.9980	.9970
LSDN (D^+)	.9564	.9585	.9533	.9892	.9905	.9870

Table 5. Jackknife results for the nation data: three major loss functions

METHOD	STAB	CROSS
	3-dim	3-dim
LSSD	.9887	.9998
LSD	.9880	.9996
LSLD	.9697	.9956

Figure captions

Figure 1. Stability results for LSSP: first 2 dimensions of the three-dimensional solution for the nation data.

Figure 2. Stability results for LSSD: first 2 dimensions of the three-dimensional solution for the nation data.

Figure 3. Stability results for LSD: first 2 dimensions of the three-dimensional solution for the nation data.

Figure 4. Stability results for LSLD: first 2 dimensions of the three-dimensional solution for the nation data.