



THE
POWER
TO KNOW.

SAS/STAT® 9.3 User's Guide

The MDS Procedure

(Chapter)



This document is an individual chapter from *SAS/STAT® 9.3 User's Guide*.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2011. *SAS/STAT® 9.3 User's Guide*. Cary, NC: SAS Institute Inc.

Copyright © 2011, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

For a Web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government Restricted Rights Notice: Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, July 2011

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at support.sas.com/publishing or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Chapter 55

The MDS Procedure

Contents

| | |
|---|-------------|
| Overview: MDS Procedure | 4482 |
| Getting Started: MDS Procedure | 4484 |
| Syntax: MDS Procedure | 4486 |
| PROC MDS Statement | 4487 |
| BY Statement | 4498 |
| ID Statement | 4498 |
| INVAR Statement | 4499 |
| MATRIX Statement | 4499 |
| VAR Statement | 4499 |
| WEIGHT Statement | 4500 |
| Details: MDS Procedure | 4500 |
| Formulas | 4500 |
| OUT= Data Set | 4502 |
| OUTFIT= Data Set | 4503 |
| OUTRES= Data Set | 4504 |
| INITIAL= Data Set | 4505 |
| Missing Values | 4505 |
| Normalization of the Estimates | 4505 |
| Comparison with Earlier Procedures | 4506 |
| Displayed Output | 4507 |
| ODS Table Names | 4508 |
| ODS Graphics | 4508 |
| Example: MDS Procedure | 4509 |
| Example 55.1: Jacobowitz Body Parts Data from Children and Adults | 4509 |
| References | 4518 |

Overview: MDS Procedure

Multidimensional scaling (MDS) refers to a class of methods. These methods estimate coordinates for a set of objects in a space of specified dimensionality. The input data are measurements of distances between pairs of objects. A variety of models can be used that include different ways of computing distances and various functions relating the distances to the actual data. The MDS procedure fits two- and three-way, metric and nonmetric multidimensional scaling models.

The data for the MDS procedure consist of one or more square symmetric or asymmetric matrices of similarities or dissimilarities between *objects* or *stimuli* (Kruskal and Wish 1978, pp. 7–11). Such data are also called *proximity* data. In psychometric applications, each matrix typically corresponds to a *subject*, and models that fit different parameters for each subject are called *individual difference* models.

Missing values are permitted. In particular, if the data are all missing except within some off-diagonal rectangle, the analysis is called *unfolding*. There are, however, many difficulties intrinsic to unfolding models (Heiser 1981). PROC MDS does not perform external unfolding; for analyses requiring external unfolding, use the TRANSREG procedure instead.

The MDS procedure estimates the following parameters by nonlinear least squares:

| | |
|---------------------------|--|
| configuration | the coordinates of each object in a Euclidean (Kruskal and Wish 1978, pp. 17–19) or weighted Euclidean space (Kruskal and Wish 1978, pp. 61–63) of one or more dimensions |
| dimension coefficients | for each data matrix, the coefficients that multiply each coordinate of the <i>common</i> or <i>group</i> weighted Euclidean space to yield the <i>individual</i> unweighted Euclidean space. These coefficients are the square roots of the <i>subject weights</i> (Kruskal and Wish 1978, pp. 61–63). A plot of the dimension coefficients is directly interpretable in that it shows how a unit square in the group space is transformed to a rectangle in each individual space. A plot of subject weights has no such simple interpretation. The weighted Euclidean model is related to the INDSCAL model (Carroll and Chang 1970). |
| transformation parameters | intercept, slope, or exponent in a linear, affine, or power transformation relating the distances to the data (Kruskal and Wish 1978, pp. 19–22). For a nonmetric analysis, monotone transformations involving no explicit parameters are used (Kruskal and Wish 1978, pp. 22–25). For a discussion of metric versus nonmetric transformations, see Kruskal and Wish (1978, pp. 76–78). |

Depending on the LEVEL= option, PROC MDS fits either a regression model of the form

$$fit(datum) = fit(trans(distance)) + error$$

or a measurement model of the form

$$fit(trans(datum)) = fit(distance) + error$$

where

| | |
|-----------------|--|
| <i>fit</i> | is a predetermined power or logarithmic transformation specified by the FIT= option. |
| <i>trans</i> | is an estimated (“optimal”) linear, affine, power, or monotone transformation specified by the LEVEL= option. |
| <i>datum</i> | is a measure of the similarity or dissimilarity of two objects or stimuli. |
| <i>distance</i> | is a distance computed from the estimated coordinates of the two objects and estimated dimension coefficients in a space of one or more dimensions. If there are no dimension coefficients (COEF=IDENTITY), this is an unweighted Euclidean distance. If dimension coefficients are used (COEF=DIAGONAL), this is a weighted Euclidean distance where the weights are the squares of the dimension coefficients; alternatively, you can multiply each dimension by its coefficient and compute an unweighted Euclidean distance. |
| <i>error</i> | is an error term assumed to have an approximately normal distribution and to be independently and identically distributed for all data. Under these assumptions, least-squares estimation is statistically appropriate. |

For an introduction to multidimensional scaling, see Kruskal and Wish (1978) and Arabie, Carroll, and DeSarbo (1987). A more advanced treatment is given by Young (1987). Many practical issues of data collection and analysis are discussed in Schiffman, Reynolds, and Young (1981). The fundamentals of psychological measurement, including both unidimensional and multidimensional scaling, are expounded by Torgerson (1958). Nonlinear least-squares estimation of PROC MDS models is discussed in Null and Sarle (1982).

Getting Started: MDS Procedure

The simplest application of PROC MDS is to reconstruct a map from a table of distances between points on the map (Kruskal and Wish 1978, pp. 7–9). A data set containing a table of flying mileages between 10 U.S. cities is available in the SasHELP library.

Since the flying mileages are very good approximations to Euclidean distance, no transformation is needed to convert distances from the model to data. The analysis can therefore be done at the absolute level of measurement, as displayed in the following PROC MDS step (LEVEL=ABSOLUTE). The following statements produce [Figure 55.1](#) and [Figure 55.2](#):

```
title 'Analysis of Flying Mileages between Ten U.S. Cities';

ods graphics on;

proc mds data=sasHELP.mileages level=absolute;
  id city;
run;
```

PROC MDS first displays the iteration history. In this example, only one iteration is required. The badness-of-fit criterion 0.001689 indicates that the data fit the model extremely well. You can also see that the fit is excellent in the fit plot in [Figure 55.2](#).

Figure 55.1 Iteration History from PROC MDS

| Analysis of Flying Mileages between Ten U.S. Cities | | | | |
|--|---------|---------------------------------|------------------------|------------------------|
| Multidimensional Scaling: Data=SASHELP.MILEAGES.DATA | | | | |
| Shape=TRIANGLE Condition=MATRIX Level=ABSOLUTE | | | | |
| Coef=IDENTITY Dimension=2 Formula=1 Fit=1 | | | | |
| Gconverge=0.01 Maxiter=100 Over=1 Ridge=0.0001 | | | | |
| Iteration | Type | Badness- of-Fit Criterion | Change in Criterion | Convergence Measure |
| 0 | Initial | 0.003273 | . | 0.8562 |
| 1 | Lev-Mar | 0.001689 | 0.001584 | 0.005128 |
| Convergence criterion is satisfied. | | | | |

While PROC MDS can recover the relative positions of the cities, it cannot determine absolute location or orientation. In this case, north is toward the bottom of the plot. (See the first plot in [Figure 55.2](#).)

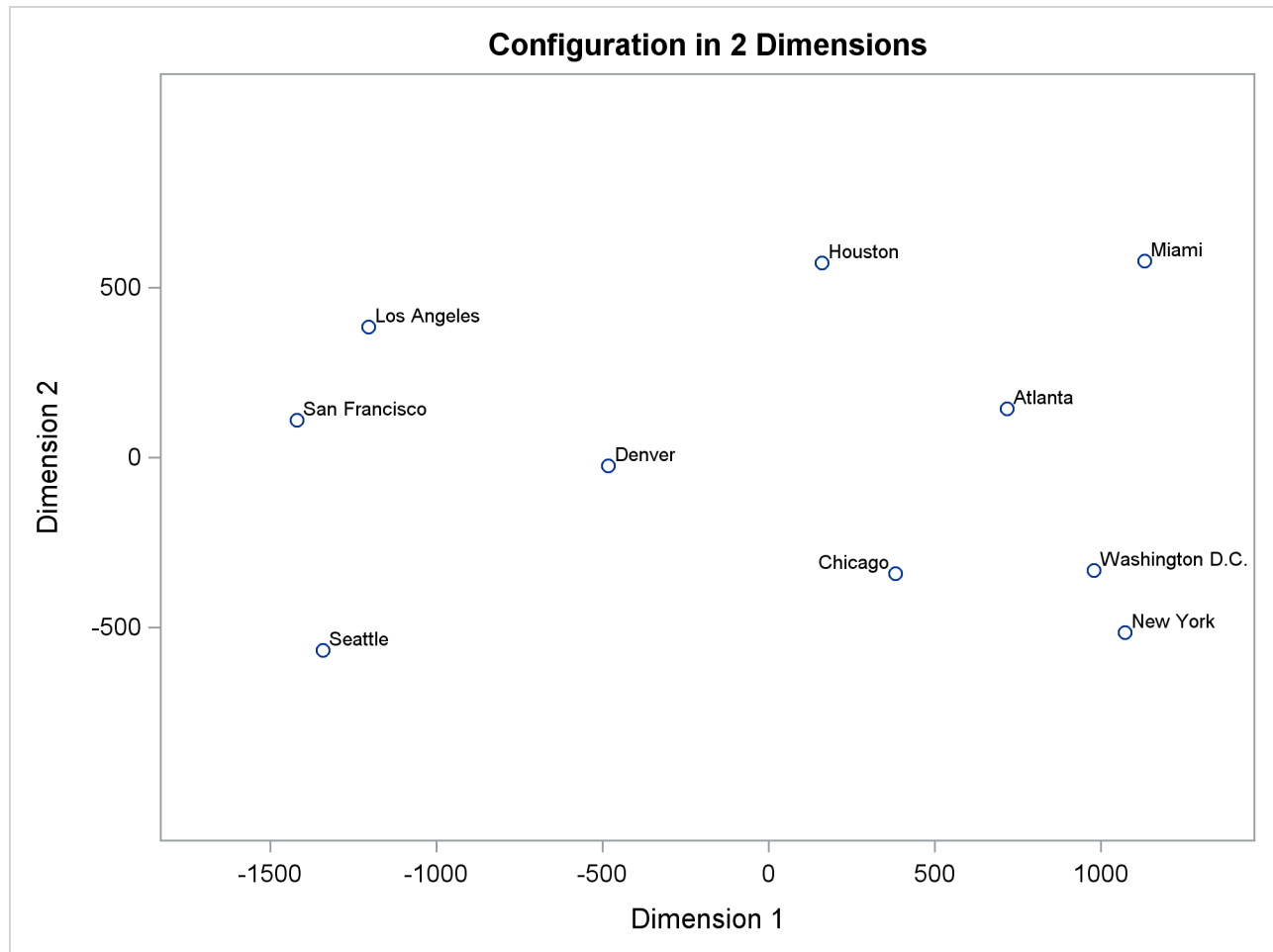
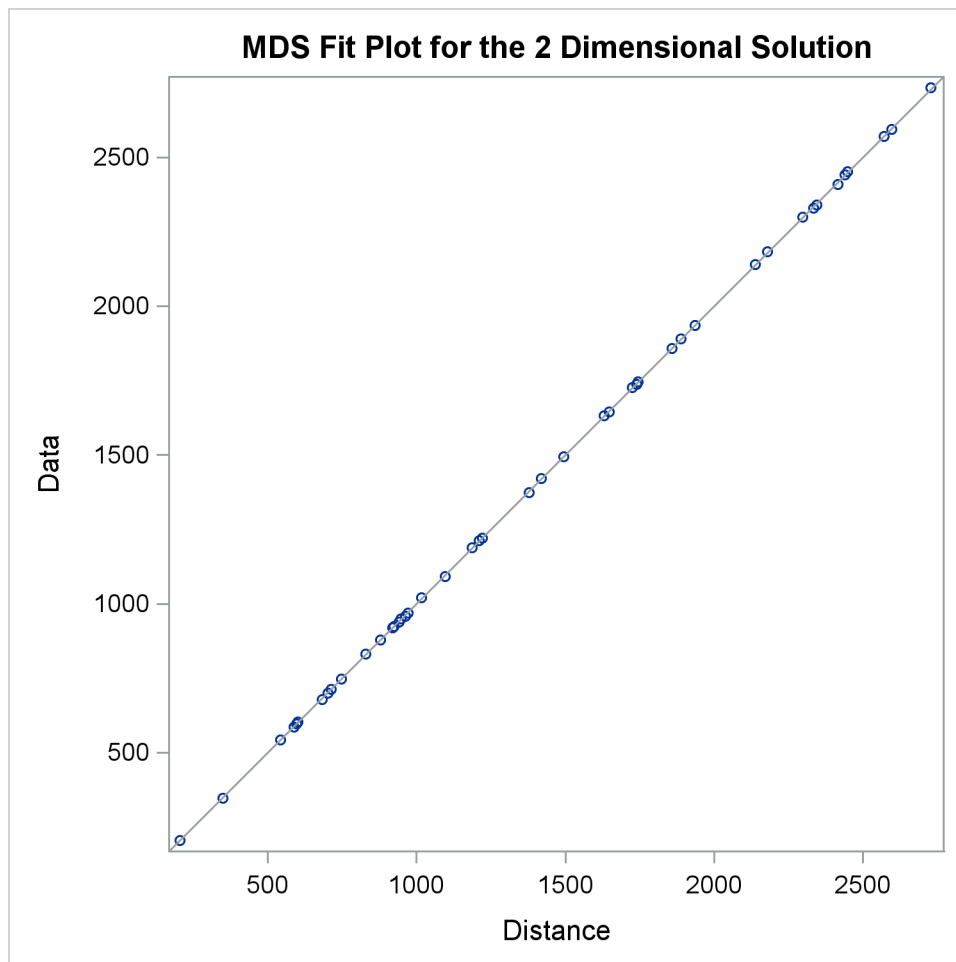
Figure 55.2 Plot of Estimated Configuration and Fit

Figure 55.2 continued



Syntax: MDS Procedure

You can specify the following statements with the MDS procedure:

```
PROC MDS < options > ;
  VAR variables ;
  INVAR variables ;
  ID | OBJECT variable ;
  MATRIX | SUBJECT variable ;
  WEIGHT variables ;
  BY variables ;
```

The PROC MDS statement is required. All other statements are optional.

PROC MDS Statement

PROC MDS <options> ;

PROC MDS produces an iteration history by default. Graphical displays are produced when ODS Graphics is enabled. Additional displayed output is controlled by the interaction of the PCONFIG, PCOEF, PTRANS, PFIT, and PFITROW options with the PININ, PINIT, PITER, and PFINAL options. The PCONFIG, PCOEF, PTRANS, PFIT, and PFITROW options specify *which* estimates and fit statistics are to be displayed. The PININ, PINIT, PITER, and PFINAL options specify *when* the estimates and fit statistics are to be displayed. If you specify at least one of the PCONFIG, PCOEF, PTRANS, PFIT, and PFITROW options but none of the PININ, PINIT, PITER, and PFINAL options, the final results (PFINAL) are displayed. If you specify at least one of the PININ, PINIT, PITER, and PFINAL options but none of the PCONFIG, PCOEF, PTRANS, PFIT, and PFITROW options, all estimates (PCONFIG, PCOEF, PTRANS) and the fit statistics for each matrix and for the entire sample (PFIT) are displayed. If you do not specify any of these nine options, no estimates or fit statistics are displayed (except the badness-of-fit criterion in the iteration history).

The types of estimates written to the OUT= data set are determined by the OCONFIG, OCOEF, OTRANS, and OCRIT options. If you do not specify any of these four options, the estimates of all the parameters of the PROC MDS model and the value of the badness-of-fit criterion appear in the OUT= data set. If you specify one or more of these options, only the information requested by the specified options appears in the OUT= data set. Also, the OITER option causes these statistics to be written to the OUT= data set after initialization and on each iteration, as well as after the iterations have terminated.

See [Table 55.2](#) for a list of options available in the PROC MDS statement.

Table 55.2 Summary of PROC MDS Statement Options

| Option | Description |
|-------------------------|--|
| Data Set Options | |
| DATA= | Specifies the input SAS data set |
| INITIAL= | Specifies the input SAS data set containing initial values |
| OUT= | Specifies the output data set |
| OUTFIT= | Specifies the output fit data set |
| OUTRES= | Specifies the output residual data set |
| Input Control | |
| CUTOFF= | Replaces data values with missing values |
| SHAPE= | Specifies the shape of the input data matrices |
| SIMILAR= | Specifies that the data are similarity measurements |
| Model | |
| COEF= | Specifies the type of matrix for the coefficients |
| CONDITION= | Specifies the conditionality of the data |
| DIMENSION= | Specifies the number of dimensions |
| LEVEL= | Specifies the measurement level |
| NEGATIVE | Permits slopes or powers to be negative |
| UNTIE | Permits tied data to be untied |

Table 55.2 *continued*

| Option | Description |
|---|--|
| Initialization | |
| INAV= | Affects the computation of initial coordinates |
| NOULB | Specifies the missing data initialization |
| RANDOM= | Specifies initial random coordinates |
| Estimation | |
| ALTERNATE= | Specifies the alternating-least-squares algorithm |
| CONVERGE= | Specifies the convergence criterion |
| EPSILON= | Specifies the amount added to squared distances |
| FIT= | Specifies a predetermined transformation |
| FORMULA= | Specifies the badness-of-fit formula |
| GCONVERGE= | Specifies the gradient convergence criterion |
| MAXITER= | Specifies the maximum number of iterations |
| MCONVERGE= | Specifies the monotone convergence criterion |
| MINCRIT= | Specifies the minimum badness-of-fit criterion |
| NONORM | Suppresses normalization of the initial and final estimates |
| OVER= | Specifies the maximum overrelaxation factor |
| RIDGE= | Specifies the initial ridge value |
| SINGULAR= | Specifies the singularity criterion |
| Control Output Data Set Contents | |
| OCOEF | Writes the dimension coefficients to the OUT= data set |
| OCONFIG | Writes the coordinates of the objects to the OUT= data set |
| OCRIT | Writes the badness-of-fit criterion to the OUT= data set |
| OITER | Writes current values after initialization and on every iteration |
| OTRANS | Writes the transformation parameter estimates to the OUT= data set |
| Control Displayed Output | |
| DECIMALS= | Specifies how many decimal places to use |
| NOPHIST | Suppresses the iteration history |
| PCOEF | Displays the estimated dimension coefficients |
| PCONFIG | Displays the estimated coordinates |
| PDATA | Displays each data matrix |
| PFINAL | Displays final estimates |
| PFIT | Displays the badness-of-fit criterion |
| PFITROW | Displays the badness-of-fit criterion for each row |
| PINAVDATA | Displays INAV= data set information |
| PINEIGVAL | Displays the initial eigenvalues |
| PINEIGVEC | Displays the initial eigenvectors |
| PININ | Displays values read from the INITIAL= data set |
| PINIT | Displays initial values |
| PITER | Displays estimates on each iteration |
| PLOTS= | Controls the graphical displays |
| PTRANS | Displays the estimated transformation parameters |

ALTERNATE | ALT=NONE | NO | N

ALTERNATE | ALT=MATRIX | MAT | M | SUBJECT | SUB | S

ALTERNATE | ALT=ROW | R < =n>

determines what form of alternating-least-squares algorithm is used. The default depends on the amount of memory available. The following ALTERNATE= options are listed in order of decreasing memory requirements:

| | |
|------------|--|
| ALT=NONE | causes all parameters to be adjusted simultaneously on each iteration. This option is usually best for a small number of subjects and objects. |
| ALT=MATRIX | adjusts all the parameters for the first subject, then all the parameters for the second subject, and so on, and finally adjusts all parameters that do not correspond to a subject, such as coordinates and unconditional transformations. This option usually works best for a large number of subjects with a small number of objects. |
| ALT=ROW | treats subject parameters the same way as the ALTERNATE=MATRIX option but also includes separate stages for unconditional parameters and for subsets of the objects. The ALT=ROW option usually works best for a large number of objects. Specifying ALT=ROW= n divides the objects into subsets of n objects each, except possibly for one subset when n does not divide the number of objects evenly. If you omit $=n$, the number of objects in the subsets is determined from the amount of memory available. The smaller the value of n , the less memory is required. |

When you specify the LEVEL=ORDINAL option, the monotone transformation is always computed in a separate stage and is listed as a separate iteration in the iteration history. In this case, estimation is done by iteratively reweighted least squares. The weights are recomputed according to the FORMULA= option on each monotone iteration; hence, it is possible for the badness-of-fit criterion to increase after a monotone iteration.

COEF=IDENTITY | IDEN | I

COEF=DIAGONAL | DIAG | D

specifies the type of matrix for the dimension coefficients.

| | |
|---------------|---|
| COEF=IDENTITY | is the default, which yields Euclidean distances. |
| COEF=DIAGONAL | produces weighted Euclidean distances, in which each subject can have different weights for the dimensions. The dimension coefficients that PROC MDS outputs are related to the square roots of what are called subject weights in PROC ALSCAL; the normalization in PROC MDS also differs from that in PROC ALSCAL. The weighted Euclidean model is related to the INDSCAL model (Carroll and Chang 1970). |

CONDITION | COND=UN | U**CONDITION | COND=MATRIX | MAT | M | SUBJECT | SUB | S****CONDITION | COND=ROW | R**

specifies the conditionality of the data (Young 1987, pp. 60–63). The data are divided into disjoint subsets called *partitions*. Within each partition, a separate transformation is applied, as specified by the **LEVEL=** option. The three types of conditionality are as follows:

| | |
|--------------------|---|
| COND=UN | (unconditional) puts all the data into a single partition. |
| COND=MATRIX | (matrix conditional) makes each data matrix a partition. |
| COND=ROW | (row conditional) makes each row of each data matrix a partition. |

The default is **CONDITION=MATRIX**. The **CONDITION=** option also determines the default value for the **SHAPE=** option. If you specify the **CONDITION=ROW** option and omit the **SHAPE=** option, each data matrix is stored as a square and possibly asymmetric matrix. If you specify the **CONDITION=UN** or **CONDITION=MATRIX** option and omit the **SHAPE=** option, only one triangle is stored. See the **SHAPE=** option for details.

CONVERGE | CONV=*p*

sets both the gradient convergence criterion and the monotone convergence criterion to p , where $0 \leq p \leq 1$. The default is **CONVERGE=0.01**; smaller values might greatly increase the number of iterations required. Values less than 0.0001 might be impossible to satisfy because of the limits of machine precision. (See the **GCONVERGE=** and **MCONVERGE=** options.)

CUTOFF=*n*

replaces data values less than n with missing values. The default is **CUTOFF=0**.

DATA=SAS-data-set

specifies the SAS data set containing one or more square matrices to be analyzed. In typical psychometric data, each matrix contains judgments from one subject, so there is a one-to-one correspondence between data matrices and subjects.

The data matrices contain similarity or dissimilarity measurements to be modeled and, optionally, weights for these data. The data are generally assumed to be dissimilarities unless you use the **SIMILAR** option. However, if there are nonmissing diagonal values and these values are predominantly larger than the off-diagonal values, the data are assumed to be similarities and are treated as if the **SIMILAR** option is specified. The diagonal elements are not otherwise used in fitting the model.

Each matrix must have exactly the same number of observations as the number of variables specified by the **VAR** statement or determined by defaults. This number is the number of objects or stimuli.

The first observation and variable are assumed to contain data for the first object, the second observation and variable are assumed to contain data for the second object, and so on.

When there are two or more matrices, the observations in each matrix must correspond to the same objects in the same order as in the first matrix.

The matrices can be symmetric or asymmetric, as specified by the **SHAPE=** option.

DECIMALS | DEC=*n*

specifies how many decimal places to use when displaying the parameter estimates and fit statistics. The default is DECIMALS=2, which is generally reasonable except in conjunction with the LEVEL=ABSOLUTE option and very large or very small data.

DIMENSION | DIMENS | DIM=*n* < TO *m* < BY=*i* >>

specifies the number of dimensions to use in the MDS model, where $1 \leq n, m < \text{number of objects}$. The parameter *i* can be either positive or negative but not zero. If you specify different values for *n* and *m*, a separate model is fitted for each requested dimension. If you specify only DIMENSION=*n*, then only *n* dimensions are fitted. The default is DIMENSION=2 if there are three or more objects; otherwise, DIMENSION=1 is the only valid specification. The analyses for each number of dimensions are done independently. For information about choosing the dimensionality, see Kruskal and Wish (1978, pp. 48–60).

EPSILON | EPS=*n*

specifies a number *n*, $0 < n < 1$, that determines the amount added to squared distances computed from the model to avoid numerical problems such as division by 0. This amount is computed as ϵ equal to *n* times the mean squared distance in the initial configuration. The distance in the MDS model is thus computed as

$$\text{distance} = \sqrt{\text{sqdist} + \epsilon}$$

where *sqdist* is the squared Euclidean distance or the weighted squared Euclidean distance.

The default is EPSILON=1E–12, which is small enough to have no practical effect on the estimates unless the FIT= value is nonpositive and there are dissimilarities that are very close to 0. Hence, when the FIT= value is nonpositive, dissimilarities less than *n* times 100 times the maximum dissimilarity are not permitted.

FIT=DISTANCE | DIS | D**FIT=SQUARED | SQU | S****FIT=LOG | L****FIT=*n***

specifies a predetermined (not estimated) transformation to apply to both sides of the MDS model before the error term is added.

The default is FIT=DISTANCE or, equivalently, FIT=1, which fits data to distances.

The option FIT=SQUARED or FIT=2 fits squared data to squared distances. This gives greater importance to large data and distances and lesser importance to small data and distances in fitting the model.

The FIT=LOG or FIT=0 option fits log data to log distances. This gives lesser importance to large data and distances and greater importance to small data and distances in fitting the model.

In general, the FIT=*n* option fits *n*th-power data to *n*th-power distances. Values of *n* that are large in absolute value can cause floating-point overflows.

If the FIT= value is 0 or negative, the data must be strictly positive (see the EPSILON= option). Negative data might produce strange results with any value other than FIT=1.

FORMULA | FOR=0 | OLS | O**FORMULA | FOR=1 | USS | U****FORMULA | FOR=2 | CSS | C**

determines how the badness-of-fit criterion is standardized in correspondence with stress formulas 1 and 2 (Kruskal and Wish 1978, pp. 24–26). The default is FORMULA=1 unless you specify FIT=LOG, in which case the default is FORMULA=2. Data partitions are defined by the CONDITION= option.

FORMULA=0 fits a regression model by ordinary least squares (Null and Sarle 1982) without standardizing the partitions; this option cannot be used with the LEVEL=ORDINAL option. The badness-of-fit criterion is the square root of the error sum of squares.

FORMULA=1 standardizes each partition by the uncorrected sum of squares of the (possibly transformed) data; this option should not be used with the FIT=LOG option. With the FIT=DISTANCE and LEVEL=ORDINAL options, this is equivalent to Kruskal's stress formula 1 or an obvious generalization thereof. With the FIT=SQUARED and LEVEL=ORDINAL options, this is equivalent to Young's s-stress formula 1 or an obvious generalization thereof. The badness-of-fit criterion is analogous to $\sqrt{1 - R^2}$, where R is a multiple correlation about the origin.

FORMULA=2 standardizes each partition by the corrected sum of squares of the (possibly transformed) data; this option is the recommended method for unfolding. With the FIT=DISTANCE and LEVEL=ORDINAL options, this is equivalent to Kruskal's stress formula 2 or an obvious generalization thereof. With the FIT=SQUARED and LEVEL=ORDINAL options, this is equivalent to Young's s-stress formula 2 or an obvious generalization thereof. The badness-of-fit criterion is analogous to $\sqrt{1 - R^2}$, where R is a multiple correlation computed with a denominator corrected for the mean.

GCONVERGE | GCONV=p

sets the gradient convergence criterion to p , where $0 \leq p \leq 1$. The default is GCONVERGE=0.01; smaller values might greatly increase the number of iterations required. Values less than 0.0001 might be impossible to satisfy because of the limits of machine precision.

The gradient convergence measure is the multiple correlation of the Jacobian matrix with the residual vector, uncorrected for the mean. (See the **CONVERGE=** and **MCONVERGE=** options.)

INAV=DATA | D**INAV=SSCP | S**

affects the computation of initial coordinates. The default is INAV=DATA.

INAV=DATA computes a weighted average of the data matrices. Its value is estimated only if an element is missing from every data matrix. The weighted average of the data matrices with missing values filled in is then converted to a scalar products matrix (or what would be a scalar products matrix if the fit were perfect), from which the initial coordinates are computed.

INAV=SSCP estimates missing values in each data matrix and converts each data matrix to a scalar products matrix. The initial coordinates are computed from the unweighted average of the scalar products matrices.

INITIAL | IN=SAS-data-set

specifies a SAS data set containing initial values for some or all of the parameters of the MDS model. If the INITIAL= option is omitted, the initial values are computed from the data.

LEVEL=ABSOLUTE | ABS | A

LEVEL=RATIO | RAT | R

LEVEL=INTERVAL | INT | I

LEVEL=LOGINTERVAL | LOG | L

LEVEL=ORDINAL | ORD | O

specifies the measurement level of the data and hence the type of estimated (optimal) transformations applied to the data or distances (Young 1987, pp. 57–60; Krantz et al. 1971, pp. 9–12) within each partition as specified by the CONDITION= option. LEVEL=ORDINAL specifies a nonmetric analysis, while all other LEVEL= options specify metric analyses. The default is LEVEL=ORDINAL.

| | |
|-------------------|--|
| LEVEL=ABSOLUTE | permits no optimal transformations. Hence, the distinction between regression and measurement models is irrelevant. |
| LEVEL=RATIO | fits a regression model in which the distances are multiplied by a slope parameter in each partition (a linear transformation). In this case, the regression model is equivalent to the measurement model with the slope parameter reciprocated. |
| LEVEL=INTERVAL | fits a regression model in which the distances are multiplied by a slope parameter and added to an intercept parameter in each partition (an affine transformation). In this case, the regression and measurement models differ if there is more than one partition. |
| LEVEL=LOGINTERVAL | fits a regression model in which the distances are raised to a power and multiplied by a slope parameter in each partition (a power transformation). |
| LEVEL=ORDINAL | fits a measurement model in which a least-squares monotone increasing transformation is applied to the data in each partition. At the ordinal measurement level, the regression and measurement models differ. |

MAXITER | ITER=n

specifies the maximum number of iterations, where $n \geq 0$. The default is MAXITER=100.

MCONVERGE | MCONV=p

sets the monotone convergence criterion to p , where $0 \leq p \leq 1$, for use with the LEVEL=ORDINAL option. The default is MCONVERGE=0.01; if you want greater precision, MCONVERGE=0.001 is usually reasonable, but smaller values might greatly increase the number of iterations required.

The monotone convergence criterion is the Euclidean norm of the change in the optimally scaled data divided by the Euclidean norm of the optimally scaled data, averaged across partitions defined by the CONDITION= option. (See the [CONVERGE=](#) and [GCONVERGE=](#) options.)

MINCRIT | CRITMIN=*n*

causes iteration to terminate when the badness-of-fit criterion is less than or equal to *n*, where $n \geq 0$. The default is MINCRIT=1E-6.

NEGATIVE

permits slopes or powers to be negative with the LEVEL=RATIO, INTERVAL, or LOGINTERVAL option.

NONORM

suppresses normalization of the initial and final estimates.

NOPHIST | NOPRINT | NOP

suppresses the output of the iteration history.

NOULB

causes missing data to be estimated during initialization by the average nonmissing value, where the average is computed according to the FIT= option. Otherwise, missing data are estimated by interpolating between the Rabinowitz (1976) upper and lower bounds.

OCOEF

writes the dimension coefficients to the OUT= data set. See the OUT= option for interactions with other options.

OCONFIG

writes the coordinates of the objects to the OUT= data set. See the OUT= option for interactions with other options.

OCRIT

writes the badness-of-fit criterion to the OUT= data set. See the OUT= option for interactions with other options.

OITER | OUTITER

writes current values to the output data sets after initialization and on every iteration. Otherwise, only the final values are written to any output data sets. (See the OUT=, OUTFIT=, and OUTRES= options.)

OTRANS

writes the transformation parameter estimates to the OUT= data set if any such estimates are computed. There are no transformation parameters with the LEVEL=ORDINAL option. See the OUT= option for interactions with other options.

OUT=SAS-data-set

creates a SAS data set containing, by default, the estimates of all the parameters of the PROC MDS model and the value of the badness-of-fit criterion. However, if you specify one or more of the OCONFIG, OCOEF, OTRANS, and OCRIT options, only the information requested by the specified options appears in the OUT= data set. (See also the OITER option.)

OUTFIT=SAS-data-set

creates a SAS data set containing goodness-of-fit and badness-of-fit measures for each partition as well as for the entire data set. (See also the OITER option.)

OUTRES=SAS-data-set

creates a SAS data set containing one observation for each nonmissing data value from the DATA= data set. Each observation contains the original data value, the estimated distance computed from the MDS model, transformed data and distances, and the residual. (See also the OITER option.)

OVER=*n*

specifies the maximum overrelaxation factor, where $n \geq 1$. Values between 1 and 2 are generally reasonable. The default is OVER=2 with the LEVEL=ORDINAL, ALTERNATE=MATRIX, or ALTERNATE=ROW option; otherwise, the default is OVER=1. Use this option only if you have convergence problems.

PCOEF

produces the estimated dimension coefficients.

PCONFIG

produces the estimated coordinates of the objects in the configuration.

PDATA

displays each data matrix.

PFINAL

displays final estimates.

PFIT

displays the badness-of-fit criterion and various types of correlations between the data and fitted values for each data matrix, as well as for the entire sample.

PFITROW

displays the badness-of-fit criterion and various types of correlations between the data and fitted values for each row, as well as for each data matrix and for the entire sample. This option works only with the CONDITION=ROW option.

PINAVDATA

displays the sum of the weights and the weighted average of the data matrices computed during initialization with the INAV=DATA option.

PINEIGVAL

displays the eigenvalues computed during initialization.

PINEIGVEC

displays the eigenvectors computed during initialization.

PININ

displays values read from the INITIAL= data set. Since these values might be incomplete, the PFIT and PFITROW options do not apply.

PINIT

displays initial values.

PITER

displays estimates on each iteration.

PLOTS<(global-plot-option)> <= plot-request <(options)>>

PLOTS<(global-plot-options)> <= (plot-request <(options)> <... plot-request <(options)>>>

specifies options that control the details of the plots. When you specify only one plot request, you can omit the parentheses around the plot request.

ODS Graphics must be enabled before requesting plots. For example:

```
ods graphics on;

proc mds plots(flip);
run;

ods graphics off;
```

For more information about enabling and disabling ODS Graphics, see the section “[Enabling and Disabling ODS Graphics](#)” on page 609 in Chapter 21, “[Statistical Graphics Using ODS](#).”

The global plot option is as follows:

FLIP

flips or interchanges the X-axis and Y-axis dimensions for configuration and coefficient plots.

The plot requests include the following:

COEFFICIENTS(ONE)

combines the INDSCAL coefficients panel of plots into a single plot. By default, the display consists of a panel with two plots. The vectors are displayed in the left plot, and the labels are displayed in the right plot. The right plot provides a magnification of the region of the vector endpoints. In contrast, the single display, requested by COEFFICIENTS(ONE), is more compact, but there is less room for vector labels. It is often easier to identify the vectors in the default display.

NONE

suppresses all plots.

By default, a fit plot is produced. When more than one dimension is requested, plots of the configuration are also plotted. For individual differences models with more than one dimension, the subject weights or coefficients are plotted. When more than one value is specified for the DIMENSION= option, the badness-of-fit plot is produced.

PTRANS

displays the estimated transformation parameters if any are computed. There are no transformation parameters with the LEVEL=ORDINAL option.

RANDOM<=seed>

causes initial coordinate values to be pseudo-random numbers. In one dimension, the pseudo-random numbers are uniformly distributed on an interval. In two or more dimensions, the pseudo-random numbers are uniformly distributed on the circumference of a circle or the surface of a (hyper)sphere.

RIDGE=*n*

specifies the initial ridge value, where $n \geq 0$. The default is RIDGE=1E-4.

If you get a floating-point overflow in the first few iterations, specify a larger value such as RIDGE=0.01, RIDGE=1, or RIDGE=100.

If you know that the initial estimates are very good, using RIDGE=0 might speed convergence.

SHAPE=TRIANGULAR | TRIANGLE | TRI | T**SHAPE=SQUARE | SQU | S**

determines whether the entire data matrix for each subject or only one triangle of the matrix is stored and analyzed. If you specify the CONDITION=ROW option, the default is SHAPE=SQUARE. Otherwise, the default is SHAPE=TRIANGLE.

SHAPE=SQUARE causes the entire matrix to be stored and analyzed. The matrix can be asymmetric.

SHAPE=TRIANGLE causes only one triangle to be stored. However, PROC MDS reads both upper and lower triangles to look for nonmissing values and to symmetrize the data if needed. If corresponding elements in the upper and lower triangles both contain nonmissing values, only the average of the two values is stored and analyzed (Kruskal and Wish 1978, p. 74). Also, if an OUTRES= data set is requested, only the average of the two corresponding elements is output.

SIMILAR | SIM<=*max*>

causes the data to be treated as similarity measurements rather than dissimilarities. If =*max* is not specified, each data value is converted to a dissimilarity by subtracting it from the maximum value in the data set or BY group. If =*max* is specified, each data value is subtracted from the maximum of *max* and the data. The diagonal data are included in computing these maxima.

By default, the data are assumed to be dissimilarities unless there are nonmissing diagonal values and these values are predominantly larger than the off-diagonal values. In this case, the data are assumed to be similarities and are treated as if the SIMILAR option is specified.

SINGULAR=*p*

specifies the singularity criterion p , $0 \leq p \leq 1$. The default is SINGULAR=1E-8.

UNTIE

permits tied data to be assigned different optimally scaled values with the LEVEL=ORDINAL option. Otherwise, tied data are assigned equal optimally scaled values. The UNTIE option has no effect with values of the LEVEL= option other than LEVEL=ORDINAL.

BY Statement

BY *variables* ;

You can specify a BY statement with PROC MDS to obtain separate analyses on observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement for the MDS procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

If the INITIAL= data set contains the BY variables, the BY groups must appear in the same order as in the DATA= data set. If the BY variables are not in the INITIAL= data set, the entire data set is used to provide initial values for each BY group in the DATA= data set.

For more information about BY-group processing, see the discussion in *SAS Language Reference: Concepts*. For more information about the DATASETS procedure, see the discussion in the *Base SAS Procedures Guide*.

ID Statement

ID | **OBJECT** | **OBJ** *variable* ;

The ID statement specifies a variable in the DATA= data set that contains descriptive labels for the objects. The labels are used in the output and are copied to the OUT= data set. If there is more than one data matrix, only the ID values from the observations containing the first data matrix are used.

The ID variable is not used to establish any correspondence between observations and variables.

If the ID statement is omitted, the variable labels or names are used as object labels.

INVAR Statement

INVAR *variables* ;

The INVAR statement specifies the numeric variables in the INITIAL= data set that contain initial parameter estimates. The first variable corresponds to the first dimension, the second variable to the second dimension, and so on.

If the INVAR statement is omitted, the variables Dim1, . . . , Dim m are used, where m is the maximum number of dimensions.

MATRIX Statement

MATRIX | **MAT** | **SUBJECT** | **SUB** *variable* ;

The MATRIX statement specifies a variable in the DATA= data set that contains descriptive labels for the data matrices or subjects. The labels are used in the output and are copied to the OUT= and OUTRES= data sets. Only the first observation from each data matrix is used to obtain the label for that matrix.

If the MATRIX statement is omitted, the matrices are labeled 1, 2, 3, and so on.

VAR Statement

VAR *variables* ;

The VAR statement specifies the numeric variables in the DATA= data set that contain similarity or dissimilarity measurements on a set of objects or stimuli. Each variable corresponds to one object.

If the VAR statement is omitted, all numeric variables that are not specified in another statement are used.

To analyze a subset of the objects in a data set, you can specify the variable names corresponding to the columns in the subset, but you must also use a DATA step or a WHERE clause to specify the rows in the subset. PROC MDS expects to read one or more square matrices, and you must ensure that the rows in the data set correctly correspond to the columns in number and order.

WEIGHT Statement

WEIGHT *variables* ;

The WEIGHT statement specifies numeric variables in the DATA= data set that contain weights for each similarity or dissimilarity measurement. These weights are used to compute weighted least-squares estimates. The number of WEIGHT variables must be the same as the number of VAR variables, and the variables in the WEIGHT statement must be in the same order as the corresponding variables in the VAR statement.

If the WEIGHT statement is omitted, all data within a partition are assigned equal weights.

Data with zero or negative weights are ignored in fitting the model but are included in the OUTRES= data set and in monotone transformations.

Details: MDS Procedure

Formulas

The following notation is used:

| | |
|--------------|--|
| A_p | intercept for partition p |
| B_p | slope for partition p |
| C_p | power for partition p |
| D_{rcs} | distance computed from the model between objects r and c for subject s |
| F_{rcs} | data weight for objects r and c for subject s obtained from the c th WEIGHT variable, or 1 if there is no WEIGHT statement |
| f | value of the FIT= option |
| N | number of objects |
| O_{rcs} | observed dissimilarity between objects r and c for subject s |
| P_{rcs} | partition index for objects r and c for subject s |
| Q_{rcs} | dissimilarity after applying any applicable estimated transformation for objects r and c for subject s |
| R_{rcs} | residual for objects r and c for subject s |
| S_p | standardization factor for partition p |
| $T_p(\cdot)$ | estimated transformation for partition p |
| V_{sd} | coefficient for subject s on dimension d |

X_{nd} coordinate for object n on dimension d

Summations are taken over nonmissing values.

Distances are computed from the model as

$$\begin{aligned} D_{rcs} &= \sqrt{\sum_d (X_{rd} - X_{cd})^2} && \text{for COEF=IDENTITY:} \\ &&& \text{Euclidean distance} \\ &= \sqrt{\sum_d V_{sd}^2 (X_{rd} - X_{cd})^2} && \text{for COEF=DIAGONAL:} \\ &&& \text{weighted Euclidean distance} \end{aligned}$$

Partition indexes are

$$\begin{aligned} P_{rcs} &= 1 && \text{for CONDITION=UN} \\ &= s && \text{for CONDITION=MATRIX} \\ &= (s-1)N + r && \text{for CONDITION=ROW} \end{aligned}$$

The estimated transformation for each partition is

$$\begin{aligned} T_p(d) &= d && \text{for LEVEL=ABSOLUTE} \\ &= B_p d && \text{for LEVEL=RATIO} \\ &= A_p + B_p d && \text{for LEVEL=INTERVAL} \\ &= B_p d^{C_p} && \text{for LEVEL=LOGINTERVAL} \end{aligned}$$

For LEVEL=ORDINAL, $T_p(\cdot)$ is computed as a least-squares monotone transformation.

For LEVEL=ABSOLUTE, RATIO, or INTERVAL, the residuals are computed as

$$\begin{aligned} Q_{rcs} &= O_{rcs} \\ R_{rcs} &= Q_{rcs}^f - [T_{P_{rcs}}(D_{rcs})]^f \end{aligned}$$

For LEVEL=ORDINAL, the residuals are computed as

$$\begin{aligned} Q_{rcs} &= T_{P_{rcs}}(O_{rcs}) \\ R_{rcs} &= Q_{rcs}^f - D_{rcs}^f \end{aligned}$$

If f is 0, then natural logarithms are used in place of the f th powers.

For each partition, let

$$U_p = \frac{\sum_{r,c,s} F_{rcs}}{\sum_{r,c,s | P_{rcs}=p} F_{rcs}}$$

and

$$\overline{Q}_p = \frac{\sum_{r,c,s|P_{rcs}=p} Q_{rcs} F_{rcs}}{\sum_{r,c,s|P_{rcs}=p} F_{rcs}}$$

Then the standardization factor for each partition is

$$\begin{aligned} S_p &= 1 && \text{for FORMULA=0} \\ &= U_p \sum_{r,c,s|P_{rcs}=p} Q_{rcs}^2 F_{rcs} && \text{for FORMULA=1} \\ &= U_p \sum_{r,c,s|P_{rcs}=p} (Q_{rcs} - \overline{Q}_p)^2 F_{rcs} && \text{for FORMULA=2} \end{aligned}$$

The badness-of-fit criterion that the MDS procedure tries to minimize is

$$\sqrt{\sum_{r,c,s} \frac{R_{rcs}^2 F_{rcs}}{S_{P_{rcs}}}}$$

OUT= Data Set

The OUT= data set contains the following variables:

- BY variables, if any
- _ITER_ (if the OUTITER option is specified), a numeric variable containing the iteration number
- _DIMENS_, a numeric variable containing the number of dimensions
- _MATRIX_ or the variable in the MATRIX statement, identifying the data matrix or subject to which the observation pertains. This variable contains a missing value for observations that pertain to the data set as a whole and not to a particular matrix, such as the coordinates (_TYPE_='CONFIG').
- _TYPE_, a character variable of length 10 identifying the type of information in the observation

The values of _TYPE_ are as follows:

| | |
|-----------|---|
| CONFIG | the estimated coordinates of the configuration of objects |
| DIAGCOEF | the estimated dimension coefficients for COEF=DIAGONAL |
| INTERCEPT | the estimated intercept parameters |
| SLOPE | the estimated slope parameters |
| POWER | the estimated power parameters |
| CRITERION | the badness-of-fit criterion |

- _LABEL_ or the variable in the ID statement, containing the variable label or value of the ID variable of the object to which the observation pertains. This variable contains a missing value for observations that do not pertain to a particular object or dimension.

- `_NAME_`, a character variable containing the variable name of the object or dimension to which the observation pertains. This variable contains a missing value for observations that do not pertain to a particular object or dimension.
- `Dim1, ..., Dimm`, where m is the maximum number of dimensions

OUTFIT= Data Set

The OUTFIT= data set contains various measures of goodness and badness of fit. There is one observation for the entire sample plus one observation for each matrix. For the `CONDITION=ROW` option, there is also one observation for each row.

The OUTFIT= data set contains the following variables:

- BY variables, if any
- `_ITER_` (if the `OUTITER` option is specified), a numeric variable containing the iteration number
- `_DIMENS_`, a numeric variable containing the number of dimensions
- `_MATRIX_` or the variable in the `MATRIX` statement, identifying the data matrix or subject to which the observation pertains
- `_LABEL_` or the variable in the `ID` statement, containing the variable label or value of the ID variable of the object to which the observation pertains when `CONDITION=ROW`
- `_NAME_`, a character variable containing the variable name of the object or dimension to which the observation pertains when `CONDITION=ROW`
- `N`, the number of nonmissing data
- `WEIGHT`, the weight of the partition
- `CRITER`, the badness-of-fit criterion
- `DISCORR`, the correlation between the transformed data and the distances for `LEVEL=ORDINAL` or the correlation between the data and the transformed distances otherwise
- `UDISCORR`, the correlation uncorrected for the mean between the transformed data and the distances for `LEVEL=ORDINAL` or the correlation between the data and the transformed distances otherwise
- `FITCORR`, the correlation between the fit-transformed data and the fit-transformed distances
- `UFITCORR`, the correlation uncorrected for the mean between the fit-transformed data and the fit-transformed distances

OUTRES= Data Set

The OUTRES= data set has one observation for each nonmissing data value. It contains the following variables:

- BY variables, if any
- _ITER_ (if the OUTITER option is specified), a numeric variable containing the iteration number
- _DIMENS_, a numeric variable containing the number of dimensions
- _MATRIX_ or the variable in the MATRIX statement, identifying the data matrix or subject to which the observation pertains
- _ROW_, containing the variable label or value of the ID variable of the row to which the observation pertains
- _COL_, containing the variable label or value of the ID variable of the column to which the observation pertains
- DATA, the original data value
- TRANDATA, the optimally transformed data value when LEVEL=ORDINAL
- DISTANCE, the distance computed from the PROC MDS model
- TRANSDIST, the optimally transformed distance when the LEVEL= option is not ORDINAL or ABSOLUTE
- FITDATA, the data value further transformed according to the FIT= option
- FITDIST, the distance further transformed according to the FIT= option
- WEIGHT, the combined weight of the data value based on the WEIGHT variable(s), if any, and the standardization specified by the FORMULA= option
- RESIDUAL, FITDATA minus FITDIST

If you assign a nonmissing data value a weight of zero, PROC MDS will ignore it when the model is fit, but the value will still appear in the OUTRES= data set (see the section “[WEIGHT Statement](#)” on page 4500).

INITIAL= Data Set

The INITIAL= data set has the same structure as the OUT= data set but is not required to have all of the variables or observations that appear in the OUT= data set. You can use an OUT= data set previously created by PROC MDS (without the OUTITER option) as an INITIAL= data set in a subsequent invocation of the procedure.

The only variables that are required are Dim1, ..., Dim m (where m is the maximum number of dimensions) or equivalent variables specified in the INVAR statement. If these are the only variables, then all the observations are assumed to contain coordinates of the configuration; you cannot read dimension coefficients or transformation parameters.

To read initial values for the dimension coefficients or transformation parameters, the INITIAL= data set must contain the _TYPE_ variable and either the variable specified in the ID statement or, if no ID statement is used, the variable _NAME_. In addition, if there is more than one data matrix, either the variable specified in the MATRIX statement or, if no MATRIX statement is used, the variable _MATRIX_ or _MATNUM_ is required.

If the INITIAL= data set contains the variable _DIMENS_, initial values are obtained from observations with the corresponding number of dimensions. If there is no _DIMENS_ variable, the same observations are used for each number of dimensions analyzed.

If you want PROC MDS to read initial values from some but not all of the observations in the INITIAL= data set, use the WHERE= data set option to select the desired observations.

Missing Values

Missing data in the similarity or dissimilarity matrices are ignored in fitting the model and are omitted from the OUTRES= data set. Any matrix that is completely missing is omitted from the analysis.

Missing weights are treated as 0.

Missing values are also permitted in the INITIAL= data set, but a large number of missing values might yield a degenerate initial configuration.

Normalization of the Estimates

In multidimensional scaling models, the parameter estimates are not uniquely determined; the estimates can be transformed in various ways without changing their badness of fit. The initial and final estimates from PROC MDS are, therefore, normalized (unless you specify the NONORM option) to make it easier to compare results from different analyses.

The configuration always has a mean of 0 for each dimension.

With the COEF=IDENTITY option, the configuration is rotated to a principal-axis orientation. Unless you specify the LEVEL=ABSOLUTE option, the entire configuration is scaled so that the root-mean-square element is 1, and the transformations are adjusted to compensate.

With the COEF=DIAGONAL option, each dimension is scaled to a root-mean-square value of 1, and the dimension coefficients are adjusted to compensate. Unless you specify the LEVEL=ABSOLUTE option, the dimension coefficients are normalized as follows. If you specify the CONDITION=UN option, all of the dimension coefficients are scaled to a root-mean-square value of 1. For other values of the CONDITION= option, the dimension coefficients are scaled separately for each subject to a root-mean-square value of 1. In either case, the transformations are adjusted to compensate.

Each dimension is reflected to give a positive rank correlation with the order of the objects in the data set.

For the LEVEL=ORDINAL option, if the intercept, slope, or power parameters are fitted, the transformed data are normalized to eliminate these parameters if possible.

Comparison with Earlier Procedures

PROC MDS shares many of the features of the ALSCAL procedure (Young, Lewycky, and Takane 1986; Young 1982), as well as some features of the MLSCALE procedure (Ramsay 1986). Both PROC ALSCAL and PROC MLSCALE are no longer a part of SAS; however, they are described in the *SUGI Supplemental Library User's Guide, Version 5 Edition*. The MDS procedure generally produces results similar to those from the ALSCAL procedure (Young, Lewycky, and Takane 1986; Young 1982) if you use the following options in PROC MDS:

- FIT=SQUARED
- FORMULA=1 except for unfolding data, which require FORMULA=2
- PFINAL to get output similar to that from PROC ALSCAL

Running the MDS procedure with certain options generally produces results similar to those from using the MLSCALE procedure (Ramsay 1986) with other options. This is illustrated with the following statements:

```
proc mds fit=log level=loginterval ... ;

proc mlscale stvarnce=constant suvarnce=constant ... ;
```

Alternatively, using the FIT=DISTANCE option in the PROC MDS statement produces results similar to those from specifying the NORMAL option in the PROC MLSCALE statement.

Displayed Output

Unless you specify the NOPHIST option, PROC MDS displays the iteration history containing the following:

- Iteration number
- Type of iteration:

| | |
|----------|--------------------------|
| Initial | initial configuration |
| Monotone | monotone transformation |
| Gau-New | Gauss-Newton step |
| Lev-Mar | Levenberg-Marquardt step |
- Badness-of-Fit Criterion
- Change in Criterion
- Convergence Measures:

| | |
|----------|--|
| Monotone | the Euclidean norm of the change in the optimally scaled data divided by the Euclidean norm of the optimally scaled data, averaged across partitions |
| Gradient | the multiple correlation of the Jacobian matrix with the residual vector, uncorrected for the mean |

Depending on what options are specified, PROC MDS can also display the following tables:

- Data Matrix and possibly Weight Matrix for each subject
- Eigenvalues from the computation of the initial coordinates
- Sum of Data Weights and Pooled Data Matrix computed during initialization with INAV=DATA
- Configuration, the estimated coordinates of the objects
- Dimension Coefficients
- A table of transformation parameters, including one or more of the following:

| |
|-----------|
| Intercept |
| Slope |
| Power |
- A table of fit statistics for each matrix and possibly each row, including the following:

| |
|--|
| Number of Nonmissing Data |
| Weight of the matrix or row, permitting both observation weights and standardization factors |

Badness-of-Fit Criterion

Distance Correlation computed between the distances and data with optimal transformation

Uncorrected Distance Correlation not corrected for the mean

Fit Correlation computed after applying the FIT= transformation to both distances and data

Uncorrected Fit Correlation not corrected for the mean

ODS Table Names

PROC MDS assigns a name to each table it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in [Table 55.3](#). For more information about ODS, see Chapter 20, “[Using the Output Delivery System](#).”

Table 55.3 ODS Tables Produced by PROC MDS

| ODS Table Name | Description | Option |
|-------------------|------------------------------|---|
| ConvergenceStatus | Convergence status | default |
| DimensionCoef | Dimension coefficients | PCOEF with COEF= not IDENTITY |
| FitMeasures | Measures of fit | PFIT |
| IterHistory | Iteration history | default |
| PConfig | Configuration of coordinates | PCONFIG |
| PData | Data matrices | PDATA |
| PInAvData | INAV= data set information | PINAVDATA |
| PInEigval | Initial eigenvalues | PINEIGVAL |
| PInEigvec | Initial eigenvectors | PINEIGVEC |
| PInWeight | Initialization weights | PINWEIGHT |
| Transformations | Transformation parameters | PTRANS with LEVEL=RATIO, INTERVAL, LOGINTERVAL |

ODS Graphics

Statistical procedures use ODS Graphics to create graphs as part of their output. ODS Graphics is described in detail in Chapter 21, “[Statistical Graphics Using ODS](#).”

Before you create graphs, ODS Graphics must be enabled (for example, with the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section “[Enabling and Disabling ODS Graphics](#)” on page 609 in Chapter 21, “[Statistical Graphics Using ODS](#).”

The overall appearance of graphs is controlled by ODS styles. Styles and other aspects of using ODS Graphics are discussed in the section “[A Primer on ODS Statistical Graphics](#)” on page 608 in Chapter 21, “[Statistical Graphics Using ODS](#).”

All graphs are produced by default when they are appropriate. You can reference every graph produced through ODS Graphics with a name. The names of the graphs that PROC MDS generates are listed in Table 55.4, along with the required options.

Table 55.4 Graphs Produced by PROC MDS

| ODS Graph Name | Plot Description | Option |
|------------------|-------------------------|---|
| BadnessPlot | Badness of fit | DIMENSION= <i>range</i> |
| CoefficientsPlot | Individual coefficients | DIMENSION= <i>n</i> , <i>n</i> > 1 COEF=DIAGONAL |
| ConfigPlot | Configuration | DIMENSION= <i>n</i> , <i>n</i> > 1 |
| FitPlot | Fit | default |

Example: MDS Procedure

Example 55.1: Jacobowitz Body Parts Data from Children and Adults

Jacobowitz (1975) collected conditional rank-order data regarding perceived similarity of parts of the body from children of ages 6, 8, and 10 years and from college sophomores. The data set includes data from 15 children (6-year-olds) and 15 sophomores. The method of data collection and some results of an analysis are also described by Young (1987, pp. 4–10). The following statements create the input data set:

```
data body;
  title  'Jacobowitz Body Parts Data from 6-Year-Olds and Adults';
  title2 'First 15 Subjects (obs 1–225) Are Children';
  title3 'Second 15 Subjects (obs 226–450) Are Adults';
  input  (Cheek Face Mouth Head Ear Body Arm Elbow Hand
          Palm Finger Leg Knee Foot Toe) (2.);
  if _n_ <= 225 then Subject='C'; else subject='A';
  datalines;
0  2  1  3  4 10  5  9  6  7  8 11 12 13 14
2  0 12  1 13  3  8 10 11  9  7  4  5  6 14
3  2  0  1  4  9  5 11  6  7  8 10 13 12 14

... more lines ...

10 12 11 13  9 14  8  7  4  6  2  3  5  1  0
;
```

The data are analyzed as row conditional (CONDITION=ROW) at the ordinal level of measurement (LEVEL=ORDINAL) by using the weighted Euclidean model (COEF=DIAGONAL) in three dimensions (DIMENSION=3). The final estimates are displayed (PFINAL). The estimates (OUT=OUT) and fitted values (OUTRES=RES) are saved in output data sets. The following statements produce [Output 55.1.1](#):

```
ods graphics on;

proc mds data=body condition=row level=ordinal coef=diagonal
      dimension=3 pfinal out=out outres=res;
  subject subject;
  title5 'Nonmetric Weighted MDS';
run;
```

Output 55.1.1 Analysis of Body Parts Data

```
Jacobowitz Body Parts Data from 6-Year-Olds and Adults
  First 15 Subjects (obs 1-225) Are Children
  Second 15 Subjects (obs 226-450) Are Adults

Nonmetric Weighted MDS

Multidimensional Scaling: Data=WORK.BODY.DATA
  Shape=SQUARE Condition=ROW Level=ORDINAL
  Coef=DIAGONAL Dimension=3 Formula=1 Fit=1

Mconverge=0.01 Gconverge=0.01 Maxiter=100 Over=2 Ridge=0.0001 Alternate=MATRIX
```

| Iteration | Type | Badness- of-Fit Criterion | Change in Criterion | Convergence Measures | |
|-----------|----------|---------------------------------|------------------------|----------------------|----------|
| | | | | Monotone | Gradient |
| 0 | Initial | 0.5938 | . | . | . |
| 1 | Monotone | 0.2344 | 0.3594 | 0.4693 | 0.4028 |
| 2 | Gau-New | 0.2080 | 0.0264 | . | . |
| 3 | Monotone | 0.1963 | 0.0118 | 0.0556 | 0.2630 |
| 4 | Gau-New | 0.1927 | 0.003592 | . | . |
| 5 | Monotone | 0.1797 | 0.0130 | 0.0463 | 0.1544 |
| 6 | Gau-New | 0.1779 | 0.001809 | . | . |
| 7 | Monotone | 0.1744 | 0.003430 | 0.0225 | 0.1210 |
| 8 | Gau-New | 0.1736 | 0.000807 | . | . |
| 9 | Monotone | 0.1717 | 0.001929 | 0.0161 | 0.1128 |
| 10 | Gau-New | 0.1712 | 0.000474 | . | . |
| 11 | Monotone | 0.1698 | 0.001413 | 0.0135 | 0.1119 |
| 12 | Gau-New | 0.1696 | 0.000188 | . | . |
| 13 | Monotone | 0.1684 | 0.001261 | 0.0121 | 0.1121 |
| 14 | Gau-New | 0.1683 | 0.000117 | . | . |
| 15 | Monotone | 0.1672 | 0.001096 | 0.0111 | 0.1064 |
| 16 | Gau-New | 0.1670 | 0.000131 | . | . |
| 17 | Monotone | 0.1661 | 0.000902 | 0.0103 | 0.0965 |
| 18 | Gau-New | 0.1660 | 0.000160 | . | . |
| 19 | Monotone | 0.1652 | 0.000736 | 0.009740 | 0.0980 |
| 20 | Gau-New | 0.1651 | 0.000169 | . | 0.1062 |
| 21 | Gau-New | 0.1645 | 0.000542 | . | 0.0161 |
| 22 | Gau-New | 0.1645 | 4.2645E-6 | . | 0.009969 |

Convergence criteria are satisfied.

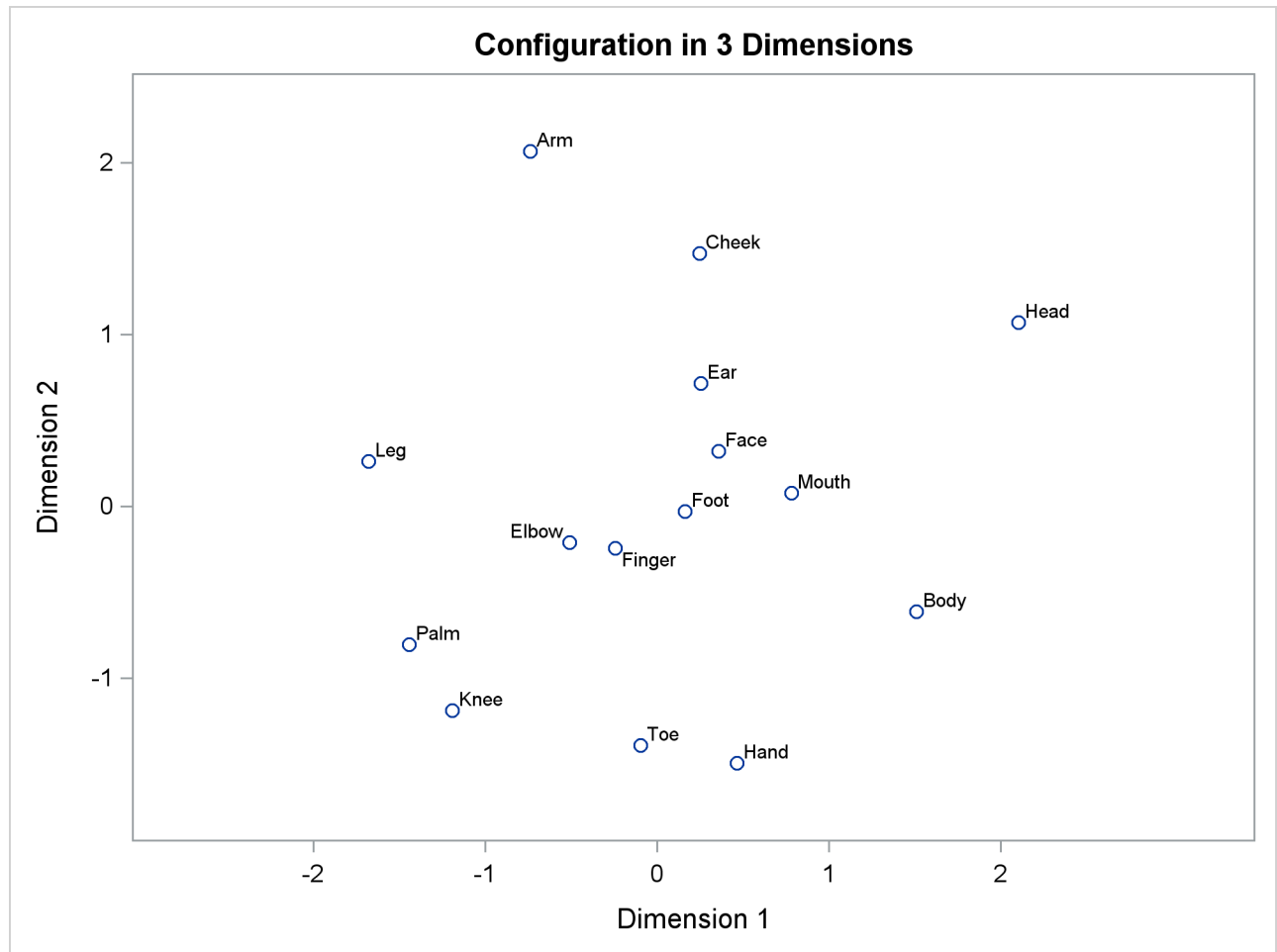
Output 55.1.1 continued

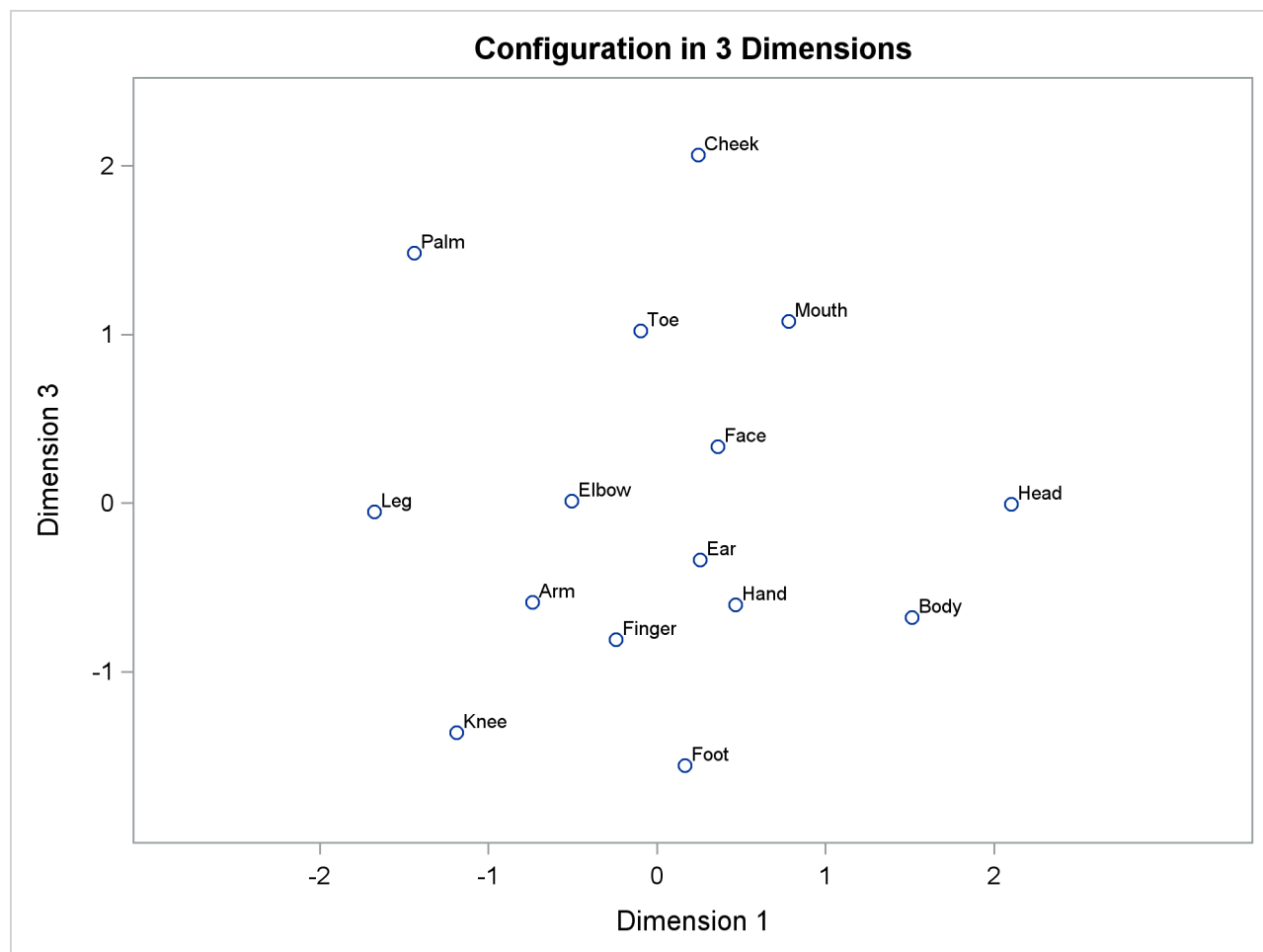
| Configuration | | | |
|------------------------|-------|-------|-------|
| | Dim1 | Dim2 | Dim3 |
| Cheek | 0.25 | 1.47 | 2.06 |
| Face | 0.36 | 0.32 | 0.33 |
| Mouth | 0.78 | 0.08 | 1.08 |
| Head | 2.10 | 1.07 | -0.01 |
| Ear | 0.26 | 0.72 | -0.34 |
| Body | 1.51 | -0.61 | -0.68 |
| Arm | -0.74 | 2.07 | -0.59 |
| Elbow | -0.51 | -0.21 | 0.01 |
| Hand | 0.46 | -1.50 | -0.60 |
| Palm | -1.44 | -0.81 | 1.48 |
| Finger | -0.24 | -0.24 | -0.81 |
| Leg | -1.68 | 0.26 | -0.05 |
| Knee | -1.19 | -1.19 | -1.36 |
| Foot | 0.16 | -0.03 | -1.56 |
| Toe | -0.10 | -1.39 | 1.02 |
| Dimension Coefficients | | | |
| Subject | 1 | 2 | 3 |
| C | 1.00 | 1.12 | 0.86 |
| C | 0.96 | 1.02 | 1.01 |
| C | 0.98 | 1.05 | 0.98 |
| C | 1.02 | 1.08 | 0.89 |
| C | 0.95 | 1.04 | 1.01 |
| C | 0.99 | 1.12 | 0.89 |
| C | 1.07 | 1.00 | 0.93 |
| C | 1.04 | 1.02 | 0.94 |
| C | 0.99 | 1.15 | 0.83 |
| C | 0.89 | 1.11 | 0.99 |
| C | 1.04 | 1.03 | 0.92 |
| C | 1.06 | 1.01 | 0.93 |
| C | 0.92 | 1.24 | 0.78 |
| C | 0.97 | 0.98 | 1.05 |
| C | 1.03 | 1.00 | 0.97 |
| A | 0.93 | 1.17 | 0.88 |
| A | 0.89 | 1.12 | 0.97 |
| A | 0.88 | 1.17 | 0.94 |
| A | 0.81 | 1.14 | 1.02 |
| A | 0.90 | 1.11 | 0.98 |
| A | 0.90 | 1.17 | 0.91 |
| A | 0.92 | 1.17 | 0.88 |
| A | 0.97 | 1.19 | 0.80 |
| A | 0.95 | 1.16 | 0.87 |
| A | 1.08 | 1.07 | 0.83 |
| A | 0.95 | 1.20 | 0.81 |
| A | 1.00 | 0.97 | 1.02 |
| A | 0.89 | 1.18 | 0.91 |
| A | 0.97 | 1.15 | 0.86 |
| A | 0.93 | 1.21 | 0.82 |

Output 55.1.1 *continued*

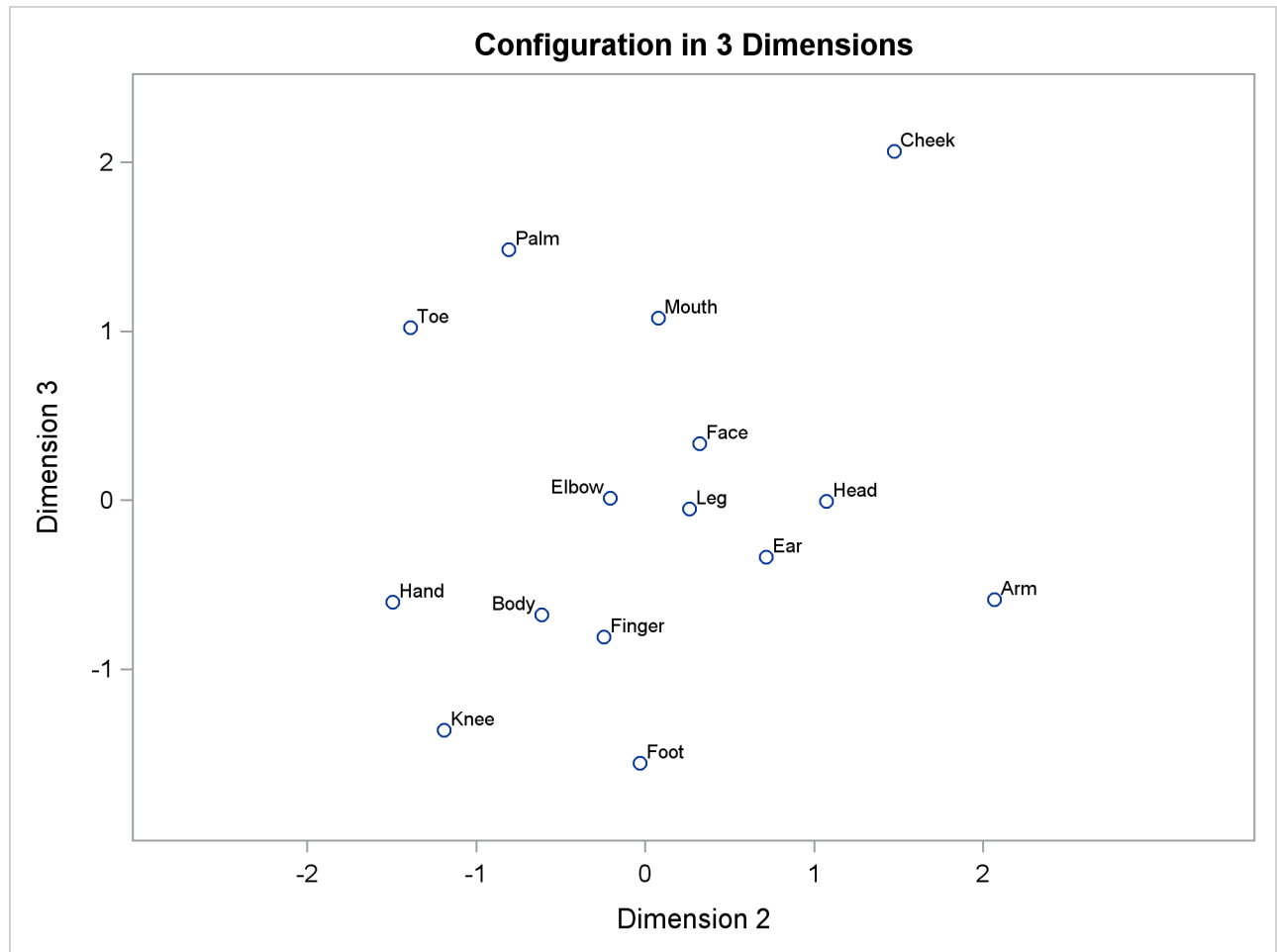
| Subject | Number of Nonmissing Data | Weight | Badness-of- Fit Criterion | Distance Correlation | Uncorrected Distance Correlation |
|---------|---------------------------------|--------|---------------------------------|-------------------------|--|
| C | 160 | 0.03 | 0.15 | 0.86 | 0.99 |
| C | 163 | 0.03 | 0.19 | 0.78 | 0.98 |
| C | 166 | 0.03 | 0.20 | 0.79 | 0.98 |
| C | 158 | 0.03 | 0.16 | 0.84 | 0.99 |
| C | 173 | 0.03 | 0.18 | 0.83 | 0.98 |
| C | 164 | 0.03 | 0.14 | 0.90 | 0.99 |
| C | 158 | 0.03 | 0.20 | 0.77 | 0.98 |
| C | 170 | 0.03 | 0.18 | 0.83 | 0.98 |
| C | 156 | 0.03 | 0.15 | 0.88 | 0.99 |
| C | 165 | 0.03 | 0.18 | 0.79 | 0.98 |
| C | 153 | 0.03 | 0.19 | 0.79 | 0.98 |
| C | 162 | 0.03 | 0.17 | 0.83 | 0.98 |
| C | 161 | 0.03 | 0.14 | 0.90 | 0.99 |
| C | 164 | 0.03 | 0.17 | 0.83 | 0.99 |
| C | 161 | 0.03 | 0.18 | 0.81 | 0.98 |
| A | 163 | 0.03 | 0.15 | 0.87 | 0.99 |
| A | 174 | 0.04 | 0.17 | 0.85 | 0.99 |
| A | 172 | 0.03 | 0.15 | 0.89 | 0.99 |
| A | 175 | 0.04 | 0.17 | 0.85 | 0.98 |
| A | 171 | 0.03 | 0.15 | 0.87 | 0.99 |
| A | 163 | 0.03 | 0.16 | 0.86 | 0.99 |
| A | 173 | 0.03 | 0.14 | 0.90 | 0.99 |
| A | 160 | 0.03 | 0.14 | 0.89 | 0.99 |
| A | 164 | 0.03 | 0.14 | 0.90 | 0.99 |
| A | 158 | 0.03 | 0.16 | 0.86 | 0.99 |
| A | 165 | 0.03 | 0.16 | 0.87 | 0.99 |
| A | 168 | 0.03 | 0.18 | 0.82 | 0.98 |
| A | 175 | 0.04 | 0.15 | 0.89 | 0.99 |
| A | 172 | 0.03 | 0.16 | 0.88 | 0.99 |
| A | 175 | 0.04 | 0.15 | 0.90 | 0.99 |
| - All - | 4962 | 1.00 | 0.16 | 0.85 | 0.99 |

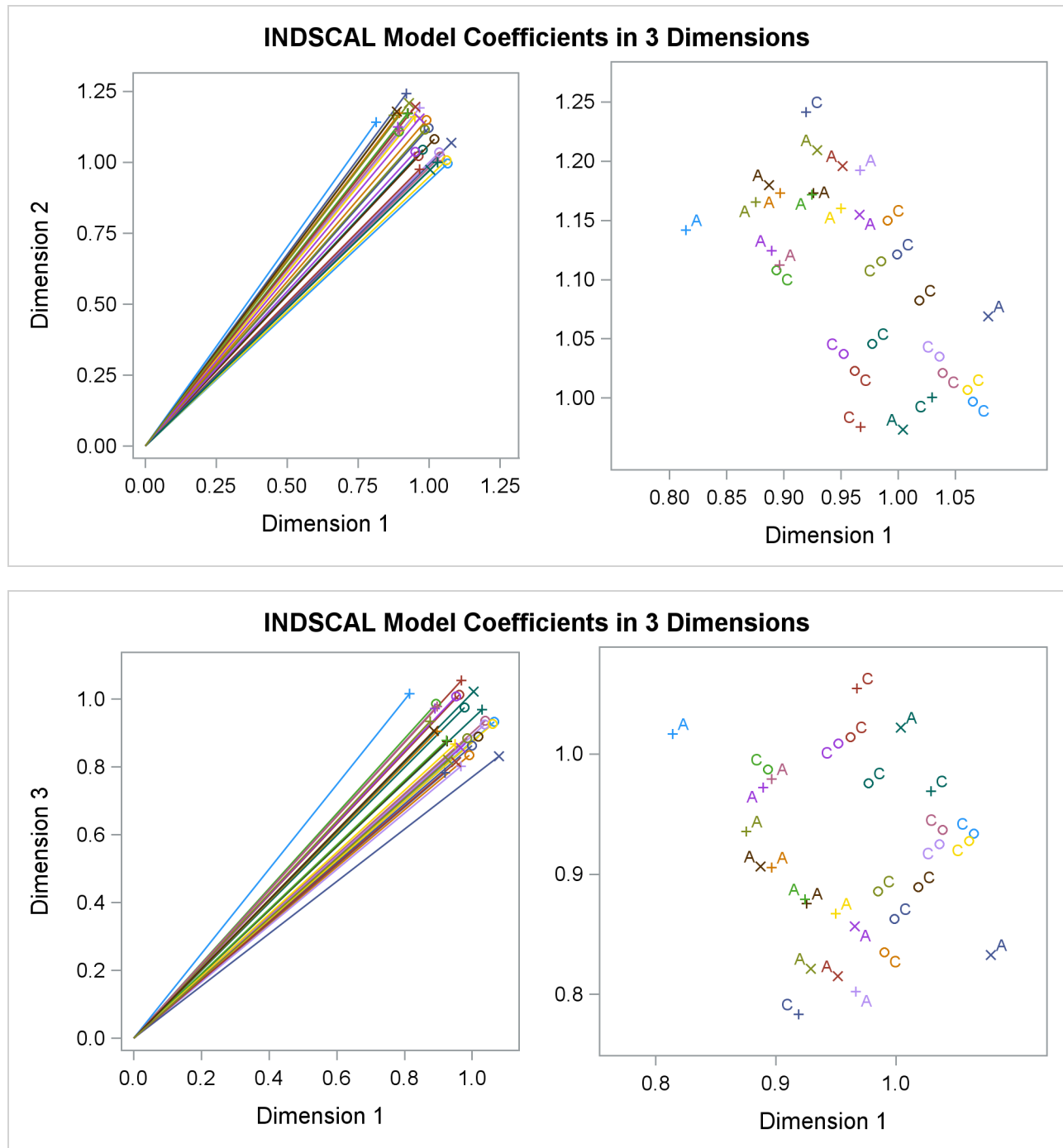
Output 55.1.1 *continued*



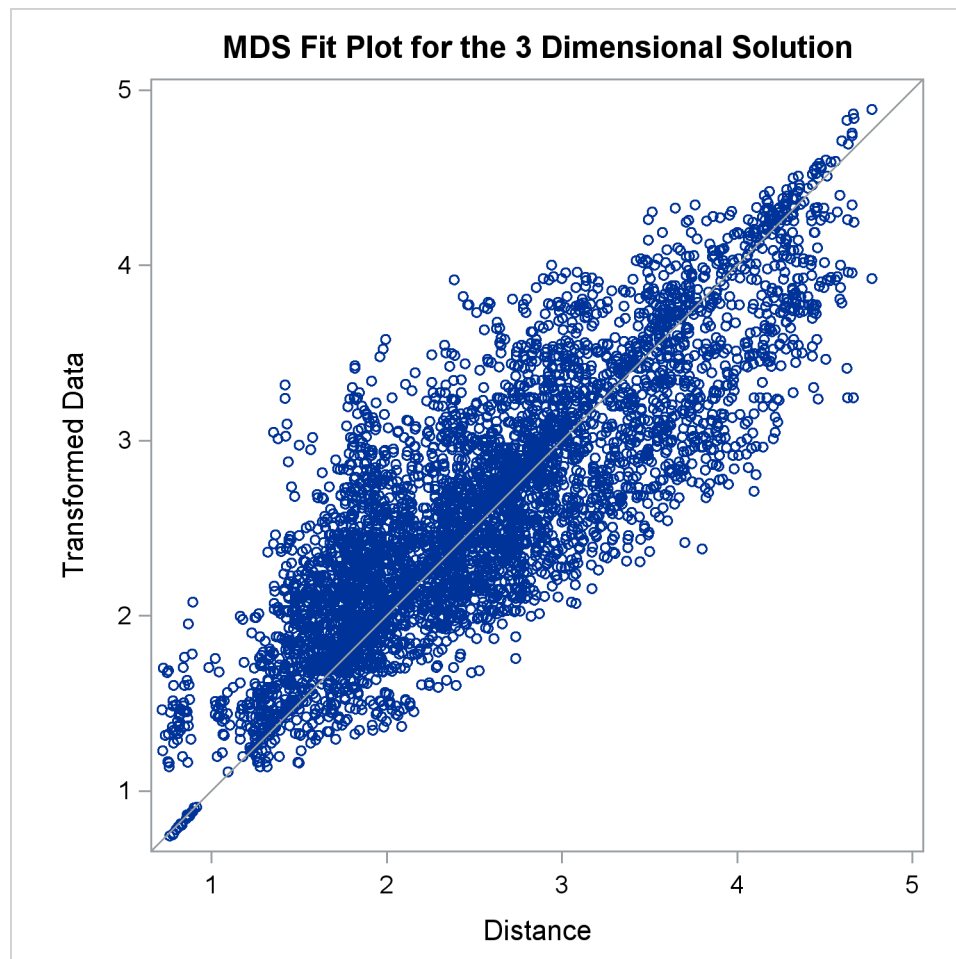
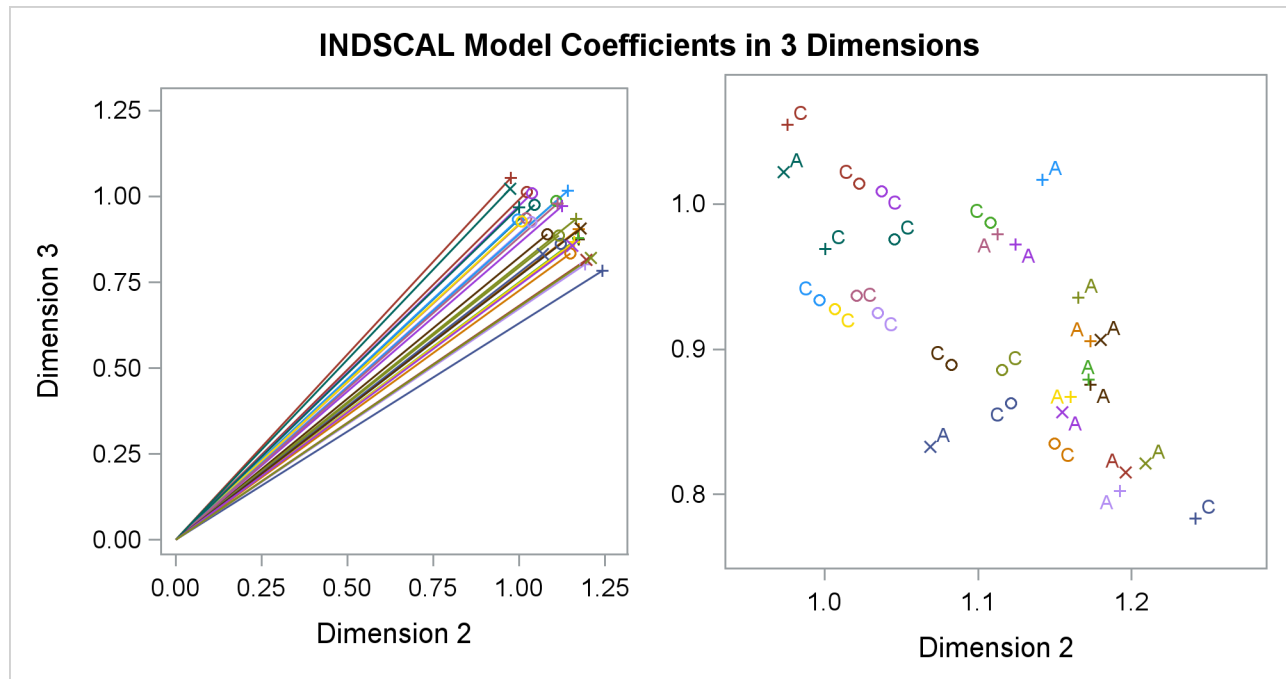
Output 55.1.1 *continued*

Output 55.1.1 *continued*



Output 55.1.1 *continued*

Output 55.1.1 continued



Often the output of greatest interest in an MDS analysis is the graphical output. The first plots show two-dimensional view of the three-dimensional configuration. Next, the coefficients are plotted. The last plot is the fit plot.

In the fit plot, the transformed data are plotted on the vertical axis, and the distances from the model are plotted on the horizontal axis. If the model fits perfectly, all points lie on a diagonal line from lower left to upper right. The vertical departure of each point from this diagonal line represents the residual of the corresponding observation.

The configuration has a tripodal shape with Body at the apex. The three legs of the tripod can be distinguished in the plot of dimension 2 by dimension 1, which shows three distinct clusters with Body in the center. Dimension 1 separates head parts from arm and leg parts. Dimension 2 separates arm parts from leg parts. The plot of dimension 3 by dimension 1 shows the tripod from the side. Dimension 3 distinguishes the more inclusive body parts (at the top) from the less inclusive body parts (at the bottom).

The plots of dimension coefficients show that children differ from adults primarily in the emphasis given to dimension 2. Children give about the same weight (approximately 1) to each dimension. Adults are much more variable than children, but all have coefficients less than 1.0 for dimension 2, with an average of about 0.7. Referring back to the configuration plot, you can see that adults consider arm parts to be more similar to leg parts than children do. Many adults also give a high weight to dimension 1, indicating that they consider head parts to be more dissimilar from arm and leg parts than children do. Dimension 3 shows considerable variability for both children and adults.

References

- Arabie, P., Carroll, J. D., and DeSarbo, W. S. (1987), *Three-Way Scaling and Clustering*, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-065, Beverly Hills and London: Sage Publications.
- Carroll, J. D. and Chang, J. J. (1970), "Analysis of Individual Differences in Multidimensional Scaling via an N-way Generalization of the 'Eckart-Young' Decomposition," *Psychometrika*, 35, 283–319.
- Heiser, W. J. (1981), *Unfolding Analysis of Proximity Data*, Leiden: Department of Data Theory, University of Leiden.
- Jacobowitz, D. (1975), *The Acquisition of Semantic Structures*, Ph.D. dissertation, University of North Carolina at Chapel Hill.
- Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971), *Foundations of Measurement*, New York: Academic Press.
- Kruskal, J. B. and Wish, M. (1978), *Multidimensional Scaling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-011, Beverly Hills and London: Sage Publications.
- Null, C. H. and Sarle, W. S. (1982), "Multidimensional Scaling by Least Squares," in *Proceedings of the Seventh Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc.
- Rabinowitz, G. (1976), "A Procedure for Ordering Object Pairs Consistent with the Multidimensional Unfolding Model," *Psychometrika*, 41, 349–373.

- Ramsay, J. O. (1986), "The MLSCALE Procedure," in *SUGI Supplemental Library User's Guide, Version 5 Edition*, Cary, NC: SAS Institute Inc.
- Schiffman, S. S., Reynolds, M. L., and Young, F. W. (1981), *Introduction to Multidimensional Scaling*, New York: Academic Press.
- Torgerson, W. S. (1958), *Theory and Methods of Scaling*, New York: John Wiley & Sons.
- Young, F. W. (1982), "Enhancements in ALSCAL-82," in *Proceedings of the Seventh Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc.
- Young, F. W. (1987), *Multidimensional Scaling: History, Theory, and Applications*, ed. R. M. Hamer, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Young, F. W., Lewyckyj, R., and Takane, Y. (1986), "The ALSCAL Procedure," in *SUGI Supplemental Library User's Guide, Version 5 Edition*, Cary, NC: SAS Institute Inc.

Subject Index

A

alternating least squares
MDS procedure, 4489
asymmetric
data (MDS), 4497

B

badness of fit
MDS procedure, 4492, 4494, 4495, 4502, 4503

C

conditional data
MDS procedure, 4490
configuration
MDS procedure, 4482
convergence criterion
MDS procedure, 4490, 4492, 4493

D

dimension coefficients
MDS procedure, 4482, 4483, 4489, 4494, 4495, 4501, 4502
dissimilarity data
MDS procedure, 4482, 4490, 4497
distance
data (MDS), 4482
distance data
MDS procedure, 4490, 4497

E

Euclidean distances
MDS procedure, 4483, 4489, 4501
external unfolding
MDS procedure, 4482

I

individual difference models
MDS procedure, 4482
INDSCAL model
MDS procedure, 4482, 4489
initial values
MDS procedure, 4492–4496, 4505

L

level of measurement
MDS procedure, 4483, 4493

M

MDS procedure
alternating least squares, 4489
asymmetric data, 4497
badness of fit, 4492, 4494, 4495, 4502, 4503
conditional data, 4490
configuration, 4482, 4494, 4495, 4502
convergence criterion, 4490, 4492, 4493
coordinates, 4494, 4495, 4501, 4502
data weights, 4500
dimension coefficients, 4482, 4483, 4489, 4494, 4495, 4501, 4502
dissimilarity data, 4482, 4490, 4497
distance data, 4482, 4490, 4497
Euclidean distances, 4483, 4489, 4501
external unfolding, 4482
individual difference models, 4482
INDSCAL model, 4482, 4489
initial values, 4492–4496, 4505
measurement level, 4483, 4493
metric multidimensional scaling, 4482
missing values, 4505
multidimensional scaling, 4482
nonmetric multidimensional scaling, 4482
normalization of the estimates, 4505
ODS Graph names, 4509
optimal transformations, 4482, 4483, 4493
output table names, 4508
partitions, 4490, 4501
plot of configuration, 4518
plot of dimension coefficients, 4518
plot of linear fit, 4518
proximity data, 4482, 4490, 4497
residuals, 4495, 4500, 4501, 4504, 4518
similarity data, 4482, 4490, 4497
stress formula, 4492, 4502
subject weights, 4482, 4489
three-way multidimensional scaling, 4482
ties, 4497
transformations, 4482, 4483, 4493, 4494, 4496, 4500–4502
transformed data, 4504

- transformed distances, 4504
- unfolding, 4482
- weighted Euclidean distance, 4483, 4489, 4501
- weighted Euclidean model, 4482, 4489
- weighted least squares, 4500
- measurement level
 - MDS procedure, 4483, 4493
- metric multidimensional scaling
 - MDS procedure, 4482
- missing values
 - MDS procedure, 4505
- multidimensional scaling
 - MDS procedure, 4482
 - metric (MDS), 4482
 - nonmetric (MDS), 4482
 - three-way (MDS), 4482

N

- nonmetric multidimensional scaling
 - MDS procedure, 4482
- normalization of the estimates
 - MDS procedure, 4505

O

- ODS Graph names
 - MDS procedure, 4509
- optimal
 - transformations (MDS), 4482, 4483, 4493
- output table names
 - MDS procedure, 4508

P

- partitions
 - MDS procedure, 4490, 4501
- plots
 - of configuration (MDS), 4518
 - of dimension coefficients (MDS), 4518
 - of linear fit (MDS), 4518
- proximity data
 - MDS procedure, 4482, 4490, 4497

R

- residuals
 - MDS procedure, 4495, 4500, 4501, 4504, 4518

S

- similarity data
 - MDS procedure, 4482, 4490, 4497
- stress formula

- MDS procedure, 4492, 4502
- subject weights
 - MDS procedure, 4482, 4489

T

- three-way multidimensional scaling
 - MDS procedure, 4482
- ties
 - MDS procedure, 4497
- transformations
 - MDS procedure, 4482, 4483, 4493, 4494, 4496, 4500–4502
- transformed data
 - MDS procedure, 4504
- transformed distances
 - MDS procedure, 4504

U

- unfolding
 - MDS procedure, 4482

W

- weighted Euclidean distance
 - MDS procedure, 4483, 4489, 4501
- weighted Euclidean model
 - MDS procedure, 4482, 4489
- weighted least squares
 - MDS procedure, 4500

Syntax Index

A

ALTERNATE= option
PROC MDS statement, [4489](#)

B

BY statement
MDS procedure, [4498](#)

C

COEF= option
PROC MDS statement, [4489](#)
CONDITION= option
PROC MDS statement, [4490](#)
CONVERGE= option
PROC MDS statement, [4490](#)
CRITMIN= option
PROC MDS statement, [4494](#)
CUTOFF= option
PROC MDS statement, [4490](#)

D

DATA= option
PROC MDS statement, [4490](#)
DECIMALS= option
PROC MDS statement, [4491](#)
DIMENSION= option
PROC MDS statement, [4491](#)

E

EPSILON= option
PROC MDS statement, [4491](#)

F

FIT= option
PROC MDS statement, [4491](#)
FORMULA= option
PROC MDS statement, [4492](#)

G

GCONVERGE= option
PROC MDS statement, [4492](#)

I

ID statement
MDS procedure, [4498](#)
INAV= option
PROC MDS statement, [4492](#)
INITIAL= option
PROC MDS statement, [4493](#)
INVAR statement, MDS procedure, [4499](#)
ITER= option
PROC MDS statement, [4493](#)

L

LEVEL= option
PROC MDS statement, [4493](#)

M

MATRIX statement
MDS procedure, [4499](#)
MAXITER= option
PROC MDS statement, [4493](#)
MCONVERGE= option
PROC MDS statement, [4493](#)
MDS procedure
syntax, [4486](#)
MDS procedure, BY statement, [4498](#)
MDS procedure, ID statement, [4498](#)
MDS procedure, INVAR statement, [4499](#)
MDS procedure, MATRIX statement, [4499](#)
MDS procedure, PROC MDS statement, [4487](#)
ALTERNATE= option, [4489](#)
COEF= option, [4489](#)
CONDITION= option, [4490](#)
CONVERGE= option, [4490](#)
CRITMIN= option, [4494](#)
CUTOFF= option, [4490](#)
DATA= option, [4490](#)
DECIMALS= option, [4491](#)
DIMENSION= option, [4491](#)
EPSILON= option, [4491](#)
FIT= option, [4491](#)
FORMULA= option, [4492](#)
GCONVERGE= option, [4492](#)
INAV= option, [4492](#)
INITIAL= option, [4493](#)
ITER= option, [4493](#)

LEVEL= option, 4493
MAXITER= option, 4493
MCONVERGE= option, 4493
MINCRIT= option, 4494
NEGATIVE option, 4494
NONORM option, 4494
NOPHIST option, 4494
NOPRINT option, 4494
NOULB option, 4494
OCOEF option, 4494
OCONFIG option, 4494
OCRIT option, 4494
OTRANS option, 4494
OUT= option, 4494
OUTFIT= option, 4494
OUTITER option, 4494
OUTRES= option, 4495
OVER= option, 4495
PCOEF option, 4495
PCONFIG option, 4495
PDATA option, 4495
PFINAL option, 4495
PFIT option, 4495
PFITROW option, 4495
PINAVDATA option, 4495
PINEIGVAL option, 4495
PINEIGVEC option, 4495
PININ option, 4495
PINIT option, 4495
PITER option, 4495
PTRANS option, 4496
RANDOM= option, 4496
RIDGE= option, 4497
SHAPE= option, 4497
SIMILAR= option, 4497
SINGULAR= option, 4497
UNTIE option, 4497

MDS procedure, VAR statement, 4499

MDS procedure, WEIGHT statement, 4500

MINCRIT= option
PROC MDS statement, 4494

N

NEGATIVE option
PROC MDS statement, 4494
NONORM option
PROC MDS statement, 4494
NOPHIST option
PROC MDS statement, 4494
NOPRINT option
PROC MDS statement, 4494
NOULB option
PROC MDS statement, 4494

O

OCOEF option
PROC MDS statement, 4494
OCONFIG option
PROC MDS statement, 4494
OCRIT option
PROC MDS statement, 4494
OTRANS option
PROC MDS statement, 4494
OUT= option
PROC MDS statement, 4494
OUTFIT= option
PROC MDS statement, 4494
OUTITER option
PROC MDS statement, 4494
OUTRES= option
PROC MDS statement, 4495
OVER= option
PROC MDS statement, 4495

P

PCOEF option
PROC MDS statement, 4495
PCONFIG option
PROC MDS statement, 4495
PDATA option
PROC MDS statement, 4495
PFINAL option
PROC MDS statement, 4495
PFIT option
PROC MDS statement, 4495
PFITROW option
PROC MDS statement, 4495
PINAVDATA option
PROC MDS statement, 4495
PINEIGVAL option
PROC MDS statement, 4495
PINEIGVEC option
PROC MDS statement, 4495
PININ option
PROC MDS statement, 4495
PINIT option
PROC MDS statement, 4495
PITER option
PROC MDS statement, 4495
PROC MDS statement, *see* MDS procedure
PTRANS option
PROC MDS statement, 4496

R

RANDOM= option

PROC MDS statement, [4496](#)
RIDGE= option
PROC MDS statement, [4497](#)

S

SHAPE= option
PROC MDS statement, [4497](#)
SIMILAR= option
PROC MDS statement, [4497](#)
SINGULAR= option
PROC MDS statement, [4497](#)

U

UNTIE option
PROC MDS statement, [4497](#)

V

VAR statement
MDS procedure, [4499](#)

W

WEIGHT statement
MDS procedure, [4500](#)

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **`yourturn@sas.com`**. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to **`suggest@sas.com`**.

SAS® Publishing Delivers!

Whether you are new to the work force or an experienced professional, you need to distinguish yourself in this rapidly changing and competitive job market. SAS® Publishing provides you with a wide range of resources to help you set yourself apart. Visit us online at support.sas.com/bookstore.

SAS® Press

Need to learn the basics? Struggling with a programming problem? You'll find the expert answers that you need in example-rich books from SAS Press. Written by experienced SAS professionals from around the world, SAS Press books deliver real-world insights on a broad range of topics for all skill levels.

support.sas.com/saspress

SAS® Documentation

To successfully implement applications using SAS software, companies in every industry and on every continent all turn to the one source for accurate, timely, and reliable information: SAS documentation. We currently produce the following types of reference documentation to improve your work experience:

- Online help that is built into the software.
- Tutorials that are integrated into the product.
- Reference documentation delivered in HTML and PDF – **free** on the Web.
- Hard-copy books.

support.sas.com/publishing

SAS® Publishing News

Subscribe to SAS Publishing News to receive up-to-date information about all new SAS titles, author podcasts, and new Web site features via e-mail. Complete instructions on how to subscribe, as well as access to past issues, are available at our Web site.

support.sas.com/spn



**THE
POWER
TO KNOW®**

