mca — Multiple and joint correspondence analysis

Description

mca performs multiple correspondence analysis (MCA) or joint correspondence analysis (JCA) on a series of categorical variables. MCA and JCA are two generalizations of correspondence analysis (CA) of a cross-tabulation of two variables (see [MV] **ca**) to the cross-tabulation of multiple variables.

mca performs an analysis of two or more integer-valued variables. Crossing (also called stacking) of integer-valued variables is also allowed.

Quick start

Multiple and joint correspondence analysis

Correspondence analysis of the multiway cross-tabulation of catvar1, catvar2, and catvar3

```
mca catvar1 catvar2 catvar3
```

Also produce the MCA coordinate plot

```
mca catvar1 catvar2 catvar3, plot
```

Use 3 dimensions instead of 2, and display the tables compactly

```
mca catvar1 catvar2 catvar3, dimensions(3) compact
```

Use principal normalization instead of the default standard normalization

```
mca catvar1 catvar2 catvar3, normalize(principal)
```

Joint correspondence analysis of catvar1, catvar2, and catvar3

```
mca catvar1 catvar2 catvar3, method(joint)
```

Multiple and joint correspondence analysis with crossed (stacked) variables

Multiple correspondence analysis of the multiway tabulation of catvar1, catvar2 and the crossed variables catvar3 and catvar4 with the crossed variables called z

```
mca catvar1 catvar2 (z: catvar3 catvar4)
```

Same as above, but joint correspondence analysis

```
mca catvar1 catvar2 (z: catvar3 catvar4), method(joint)
```

Menu

Statistics > Multivariate analysis > Correspondence analysis > Multiple correspondence analysis (MCA)

Syntax

Basic syntax for two or more categorical variables

mca *varlist* [*if*] [*in*] [*weight*] [, *options*]

Full syntax for use with two or more categorical or crossed (stacked) categorical variables

mca *speclist* [*if*] [*in*] [*weight*] [, *options*]

where

speclist = *spec* [*spec* . . .]

spec = *varlist* | (*newvar* : *varlist*)

options	Description
Model	
`supplementary(`*speclist*`)`	supplementary (passive) variables
`method(burt)`	use the Burt matrix approach to MCA; the default
`method(indicator)`	use the indicator matrix approach to MCA
`method(joint)`	perform a joint correspondence analysis (JCA)
`dimensions(#)`	number of dimensions (factors, axes); default is `dim(2)`
`normalize(standard)`	display standard coordinates; the default
`normalize(principal)`	display principal coordinates
`iterate(#)`	maximum number of `method(joint)` iterations; default is `iterate(250)`
`tolerance(#)`	tolerance for `method(joint)` convergence criterion; default is `tolerance(1e-5)`
`missing`	treat missing values as ordinary values
`noadjust`	suppress the adjustment of eigenvalues (`method(burt)` only)
Codes	
`report(variables)`	report coding of crossing variables
`report(crossed)`	report coding of crossed variables
`report(all)`	report coding of crossing and crossed variables
`length(min)`	use minimal length unique codes of crossing variables
`length(#)`	use # as coding length of crossing variables
Reporting	
`ddimensions(#)`	display # singular values; default is `ddim(.)` (all)
`points(`*varlist*`)`	display tables for listed variables; default is all variables
`compact`	display statistics table in a compact format
`log`	display the iteration log (`method(joint)` only)
`plot`	plot the coordinates (that is, `mcaplot`)
`maxlength(#)`	maximum number of characters for labels in plot; default is `maxlength(12)`

`bootstrap`, `by`, `collect`, `jackknife`, `rolling`, and `statsby` are allowed with `mca`; see **[U] 11.1.10 Prefix commands**. However, `bootstrap` and `jackknife` results should be interpreted with caution; identification of the `mca` parameters involves data-dependent restrictions, possibly leading to badly biased and overdispersed estimates (Milan and Whittaker 1995).

Weights are not allowed with the `bootstrap` prefix; see [R] **bootstrap**.

`fweight`s are allowed; see **[U] 11.1.6 weight**.

See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

`supplementary(`*speclist*`)` specifies that *speclist* are supplementary variables. Such variables do not affect the MCA solution, but their categories are mapped into the solution space. For `method(indicator)`, this mapping uses the first method described by Greenacre (2006). For `method(burt)` and `method(joint)`, the second and recommended method described by Greenacre (2006) is used, in which supplementary column principal coordinates are derived as a weighted average of the standard row coordinates, weighted by the supplementary profile. See the syntax diagram for the syntax of *speclist*.

`method(`*method*`)` specifies the method of MCA/JCA.

`method(burt)`, the default, specifies MCA, a categorical variables analogue to principal component analysis (see [MV] **pca**). The Burt method performs a CA of the Burt matrix, a matrix of the two-way cross-tabulations of all pairs of variables.

`method(indicator)` specifies MCA via a CA on the indicator matrix formed from the variables.

`method(joint)` specifies JCA, a categorical variables analogue to factor analysis (see [MV] **factor**). This method analyzes a variant of the Burt matrix, in which the diagonal blocks are iteratively adjusted for the poor diagonal fit of MCA.

`dimensions(#)` specifies the number of dimensions (= factors = axes) to be extracted. The default is `dimensions(2)`. If you specify `dimensions(1)`, the categories are placed on one dimension. The number of dimensions is no larger than the number of categories in the active variables (regular and crossed) minus the number of active variables, and it can be less. This excludes supplementary variables. Specifying a larger number than the dimensions available results in extracting all dimensions.

MCA is a hierarchical method so that extracting more dimensions does not affect the coordinates and decomposition of inertia of dimensions already included. The percentages of inertia accounting for the dimensions are in decreasing order as indicated by the singular values. The first dimension accounts for the most inertia, followed by the second dimension, and then the third dimension, etc.

`normalize(`*normalization*`)` specifies the normalization method, that is, how the row and column coordinates are obtained from the singular vectors and singular values of the matrix of standardized residuals.

`normalize(standard)` specifies that coordinates are returned in standard normalization (singular values divided by the square root of the mass). This is the default.

`normalize(principal)` specifies that coordinates are returned in principal normalization. Principal coordinates are standard coordinates multiplied by the square root of the corresponding principal inertia.

`iterate(#)` is a technical and rarely used option specifying the maximum number of iterations. `iterate()` is permitted only with `method(joint)`. The default is `iterate(250)`.

`tolerance(#)` is a technical and rarely used option specifying the tolerance for subsequent modification of the diagonal blocks of the Burt matrix. `tolerance()` is permitted only with `method(joint)`. The default is `tolerance(1e-5)`.

`missing` treats missing values as ordinary values to be included in the analysis. Observations with missing values are omitted from the analysis by default.

`noadjust` suppresses principal inertia adjustment and is allowed with `method(burt)` only. By default, the principal inertias (eigenvalues of the Burt matrix) are adjusted. The unmodified principal inertias present a pessimistic measure of fit because MCA fits the diagonal of **P** poorly (see Greenacre [1984]).

Codes

`report(`*opt*`)` displays coding information for the crossing variables, crossed variables, or both. `report()` is ignored if you do not specify at least one crossed variable.

`report(variables)` displays the coding schemes of the crossing variables, that is, the variables used to define the crossed variables.

`report(crossed)` displays a table explaining the value labels of the crossed variables.

`report(all)` displays the codings of the crossing and crossed variables.

`length(`*opt*`)` specifies the coding length of crossing variables.

`length(min)` specifies that the minimal-length unique codes of crossing variables be used.

`length(#)` specifies that the coding length # of crossing variables be used, where # must be between 4 and 32.

Reporting

`ddimensions(#)` specifies the number of singular values to be displayed. If `ddimensions()` is greater than the number of singular values, all the singular values will be displayed. The default is `ddimensions(.)`, meaning all singular values.

`points(`*varlist*`)` indicates the variables to be included in the tables. By default, tables are displayed for all variables. Regular categorical variables, crossed variables, and supplementary variables may be specified in `points()`.

`compact` specifies that point statistics tables be displayed multiplied by 1,000, enabling the display of more columns without wrapping output. The compact tables can be displayed without wrapping for models with two dimensions at line size 79 and with three dimensions at line size 99.

`log` displays an iteration log. This option is permitted with `method(joint)` only.

`plot` displays a plot of the row and column coordinates in two dimensions. Use `mcaplot` directly to select different plotting points or for other graph refinements; see [MV] **mca postestimation plots**.

`maxlength(#)` specifies the maximum number of characters for labels in plots. The default is `maxlength(12)`. # must be less than 32.

Note: The reporting options may be specified during estimation or replay.

Remarks and examples

Remarks are presented under the following headings:

Introduction

Multiple correspondence analysis (MCA) and joint correspondence analysis (JCA) are methods for analyzing observations on categorical variables. MCA is usually viewed as an extension of simple correspondence analysis (CA) to more than two variables. CA analyzes a two-way contingency table; MCA and JCA analyze a multiway table.

MCA can be viewed as a generalization of principal component analysis where the variables to be analyzed are categorical, not continuous. MCA performs a CA on a Burt or indicator matrix; it explores the relationships within a set of variables, whereas CA has more focus on exploring relationships between two sets of variables. JCA is an extension of MCA that attempts to remedy discrepancies between CA and MCA.

For an introduction to MCA via CA, see Rencher and Christensen (2012) or Everitt and Dunn (2001). For an advanced introduction to MCA without previous discussion of CA, see Gower and Hand (1996). Greenacre (2006) approaches MCA from CA, then JCA from MCA, and gives a more advanced treatment. [MV] **ca** also introduces MCA concepts. Gower (1990) explores MCA history.

Three methods are implemented here. We will start with the simplest and most intuitive representation of unordered categorical data: the indicator matrix, usually denoted **Z**, a matrix of zeros and ones with columns for all categories of all variables and rows corresponding to observations. A value of one indicates that a category is observed; a zero indicates that it is not. MCA can be performed as a CA on the indicator matrix; an equivalent to this method is `method(indicator)` with `mca`.

Here is a manufactured indicator matrix **Z**, with 10 observations on three categorical variables, w, x, and y, each with two categories indicated by $w_1, w_2, x_1, x_2, y_1, y_2$:

$$\mathbf{Z} = \begin{array}{l|cccccc|} \text{obs} & w_1 & w_2 & x_1 & x_2 & y_1 & y_2 \\ 1. & 1 & 0 & 1 & 0 & 1 & 0 \\ 2. & 0 & 1 & 1 & 0 & 1 & 0 \\ 3. & 0 & 1 & 0 & 1 & 1 & 0 \\ 4. & 1 & 0 & 0 & 1 & 0 & 1 \\ 5. & 1 & 0 & 0 & 1 & 0 & 1 \\ 6. & 1 & 0 & 1 & 0 & 0 & 1 \\ 7. & 0 & 1 & 0 & 1 & 1 & 0 \\ 8. & 1 & 0 & 0 & 1 & 1 & 0 \\ 9. & 1 & 0 & 1 & 0 & 0 & 1 \\ 10. & 1 & 0 & 1 & 0 & 1 & 0 \end{array}$$

For large datasets with many variables and observations, the indicator matrix is burdensome and can exceed memory limitations. The Burt matrix, usually denoted as **B**, is a cross-tabulation of all categories of all variables. $\mathbf{B} = \mathbf{Z}'\mathbf{Z}$. The Burt matrix is smaller than the indicator matrix **Z**. Performing CA on the

Burt matrix and performing CA on the indicator matrix are related but not the same: they give equivalent standard coordinates, but the total principal inertias from the indicator matrix approach and Burt matrix approach are different. Performing a CA of the Burt matrix without further calculation is `method(burt)` along with the `noadjust` option.

The Burt matrix, **B**, corresponding to matrix **Z**, above:

$$\mathbf{B} = \begin{array}{c} \\ w_1 \\ w_2 \\ x_1 \\ x_2 \\ y_1 \\ y_2 \end{array} \begin{array}{c} \begin{array}{cccccc} w_1 & w_2 & x_1 & x_2 & y_1 & y_2 \end{array} \\ \begin{pmatrix} 7 & 0 & 4 & 3 & 3 & 4 \\ 0 & 3 & 1 & 2 & 3 & 0 \\ 4 & 1 & 5 & 0 & 3 & 2 \\ 3 & 2 & 0 & 5 & 3 & 2 \\ 3 & 3 & 3 & 3 & 6 & 0 \\ 4 & 0 & 2 & 2 & 0 & 4 \end{pmatrix} \end{array}$$

Applying CA to the Burt matrix or indicator matrix artificially inflates the χ^2 distances between profiles and the total inertia; see Gower (2006). This can be partially remedied after CA of the Burt matrix by scale readjustments of the MCA solution. Performing adjustments after a CA of the Burt matrix is the default, that is, if no method is specified `method(burt)` is assumed; unadjusted estimates may be obtained by using the option `noadjust`.

The third method, JCA, generalizes CA more naturally to more than two categorical variables. JCA attempts to account for the variation in the off-diagonal submatrices of the Burt matrix. It corrects the inflation of the total inertia. The solution can no longer be obtained by one application of the singular value decomposition (as in CA). JCA is obtained by specifying the option `method(joint)` to `mca`.

Compare MCA on two variables and CA

We illustrate MCA with a dataset from the International Social Survey Program on environment (ISSP 1993). This example is used in the MCA literature; see Greenacre (2006). We will look at the questions on attitudes toward science. We use data from the West German sample only and remove all observations containing missing data; 871 observations remain. The survey questions are

How much do you agree or disagree with each of these statements?

A. We believe too often in science, and not enough in feelings and faith.

B. Overall, modern science does more harm than good.

C. Any change humans cause in nature—no matter how scientific—is likely to make things worse.

D. Modern science will solve our environmental problems with little change to our way of life.

Each question has five possible response categories:

1. Agree strongly
2. Agree
3. Neither agree nor disagree
4. Disagree
5. Disagree strongly

Supplementary demographic information is also available as coded categorical variables:

Sex: male, female

Age: (six groups) 16–24, 25–34, 35–44, 45–54, 55–64, 65 and older

Education: (six groups) primary incomplete, primary completed, secondary incomplete, secondary completed, tertiary incomplete, tertiary completed.

▷ Example 1

Here is a summary of these data.

```
. use https://www.stata-press.com/data/r19/issp93
(Selection from ISSP (1993))

. describe

Contains data from https://www.stata-press.com/data/r19/issp93.dta
 Observations:           871                  Selection from ISSP (1993)
    Variables:             8                  17 May 2024 09:36
                                              (_dta has notes)
────────────────────────────────────────────────────────────────────────────────
Variable      Storage   Display    Value
    name         type    format    label      Variable label
────────────────────────────────────────────────────────────────────────────────
id              int     %9.0g                 Respondent identifier
A               byte    %26.0g     agree5     Too much science, not enough
                                                feelings & faith
B               byte    %26.0g     agree5     Science does more harm than good
C               byte    %26.0g     agree5     Any change makes nature worse
D               byte    %26.0g     agree5     Science will solve environmental
                                                problems
sex             byte    %9.0g      sex        Sex
age             byte    %9.0g      age        Age (6 categories)
edu             byte    %20.0g     edu        Education (6 categories)
────────────────────────────────────────────────────────────────────────────────
Sorted by:
```

We begin by comparing MCA with two variables to CA with two variables. The default MCA analysis is a CA of the Burt matrix, performed with adjustment to the principal inertias. The unadjusted results from MCA will give different principal inertias from the CA, although the standard coordinates would be the same. With adjustment, results are nearly identical. For a detailed discussion of the output of ca and mca, see [MV] **ca**.

```
. mca A B

Multiple/joint correspondence analysis         Number of obs     =       871
                                               Total inertia     = .23775345
    Method: Burt/adjusted inertias             Number of axes    =         2

                 |   Principal                Cumul.
      Dimension  |     inertia    Percent    percent
    -------------+-----------------------------------
          Dim 1  |    .1686131      70.92      70.92
          Dim 2  |    .0586785      24.68      95.60
          Dim 3  |     .010444       4.39      99.99
          Dim 4  |    .0000178       0.01     100.00
    -------------+-----------------------------------
          Total  |    .2377535     100.00

Statistics for column categories in standard normalization

                 |           Overall          |         Dimension_1
      Categories |  Mass   Quality    %inert  |   Coord    Sqcorr   Contrib
    -------------+----------------------------+-----------------------------
    A            |                            |
    Agree stro~y |  0.068    1.000     0.100  |   1.017     0.502     0.071
           Agree |  0.185    0.997     0.042  |   0.560     0.982     0.058
    Neither ag~e |  0.117    0.930     0.062  |   0.248     0.083     0.007
        Disagree |  0.102    0.922     0.123  |  -1.239     0.907     0.157
    Disagree s~y |  0.028    0.954     0.174  |  -2.741     0.845     0.207
    -------------+----------------------------+-----------------------------
    B            |                            |
    Agree stro~y |  0.041    0.982     0.146  |   1.571     0.490     0.101
           Agree |  0.100    0.962     0.034  |   0.667     0.932     0.044
    Neither ag~e |  0.118    0.840     0.043  |   0.606     0.716     0.043
        Disagree |  0.161    0.769     0.043  |  -0.293     0.228     0.014
    Disagree s~y |  0.080    0.994     0.235  |  -1.926     0.900     0.298
    -------------------------------------------------------------------------

                 |         Dimension_2
      Categories |  Coord    Sqcorr   Contrib
    -------------+----------------------------
    A            |
    Agree stro~y |  1.718     0.498     0.202
           Agree |  0.116     0.015     0.002
    Neither ag~e | -1.344     0.847     0.212
        Disagree | -0.268     0.015     0.007
    Disagree s~y |  1.672     0.109     0.077
    -------------+----------------------------
    B            |
    Agree stro~y |  2.671     0.493     0.291
           Agree | -0.201     0.029     0.004
    Neither ag~e | -0.427     0.124     0.022
        Disagree | -0.764     0.541     0.094
    Disagree s~y |  1.055     0.094     0.089
    ------------------------------------------
```

With ca we use the norm(standard) option to obtain results in the same normalization as the preceding mca. [MV] **ca** discusses the normalize() option; mca has some of the normalize() options of ca.

The top table in the output for both ca and mca reports the principal inertias. The principal inertias reported by ca are simply the squares of the singular values. Two pieces of information reported by ca that are not reported by mca are the χ^2 value in the table and the explained inertia in the header. The χ^2 value is taken as a measure of dissimilarity between the row and column profiles of the two-way table. χ^2 distance makes sense for a two-way table but has less justification when applied to the rows and columns of either the indicator matrix or the Burt matrix (Greenacre 2006). The explained inertia is the value from the cumulative percent column in the top table corresponding to the number of dimensions selected; it is simply not reported in the mca header.

The coordinates reported in the second table are the same. The mass, inertia, and contributions are half as large in mca as they are in ca; in ca the row and column each sum to the principal inertia, but in mca there are only columns, which sum to the principal inertia.

```
. ca A B, norm(standard)
Correspondence analysis                       Number of obs     =        871
                                              Pearson chi2(16)  =     207.08
                                              Prob > chi2       =     0.0000
                                              Total inertia     =     0.2378
    5 active rows                             Number of dim.    =          2
    5 active columns                          Expl. inertia (%) =      95.60
                |   Singular   Principal                              Cumul.
      Dimension |      value     inertia          chi2    Percent   percent
    ------------+-----------------------------------------------------------
          Dim 1 |   .4106252    .1686131        146.86      70.92     70.92
          Dim 2 |   .2422364    .0586785         51.11      24.68     95.60
          Dim 3 |   .1021961     .010444          9.10       4.39     99.99
          Dim 4 |   .0042238    .0000178          0.02       0.01    100.00
    ------------+-----------------------------------------------------------
          Total |               .2377535        207.08        100
Statistics for row and column categories in standard normalization
                 |          Overall          |        Dimension_1
      Categories |    Mass  Quality   %inert |   Coord   Sqcorr  Contrib
    -------------+---------------------------+---------------------------
    A            |                           |
    Agree stro~y |   0.137    1.000    0.200 |   1.017    0.502    0.141
           Agree |   0.370    0.997    0.084 |   0.560    0.982    0.116
    Neither ag~e |   0.234    0.930    0.123 |   0.248    0.083    0.014
        Disagree |   0.204    0.922    0.245 |  -1.239    0.907    0.314
    Disagree s~y |   0.055    0.954    0.348 |  -2.741    0.845    0.414
    -------------+---------------------------+---------------------------
    B            |                           |
    Agree stro~y |   0.082    0.982    0.291 |   1.571    0.490    0.201
           Agree |   0.200    0.962    0.068 |   0.667    0.932    0.089
    Neither ag~e |   0.235    0.840    0.086 |   0.606    0.716    0.087
        Disagree |   0.323    0.769    0.086 |  -0.293    0.228    0.028
    Disagree s~y |   0.161    0.994    0.470 |  -1.926    0.900    0.596
    -------------------------------------------------------------------------
```

```
                 |        Dimension_2
      Categories |   Coord   Sqcorr  Contrib
-----------------+---------------------------
A                |
Agree stro~y     |   1.718    0.498    0.403
           Agree |   0.116    0.015    0.005
Neither ag~e     |  -1.344    0.847    0.423
        Disagree |  -0.268    0.015    0.015
Disagree s~y     |   1.672    0.109    0.154
-----------------+---------------------------
B                |
Agree stro~y     |   2.671    0.493    0.582
           Agree |  -0.201    0.029    0.008
Neither ag~e     |  -0.427    0.124    0.043
        Disagree |  -0.764    0.541    0.188
Disagree s~y     |   1.055    0.094    0.179
-----------------------------------------------
```

◁

MCA on four variables

Now we will take a look at MCA with more than two variables and at the different methods of performing MCA or JCA.

▷ Example 2

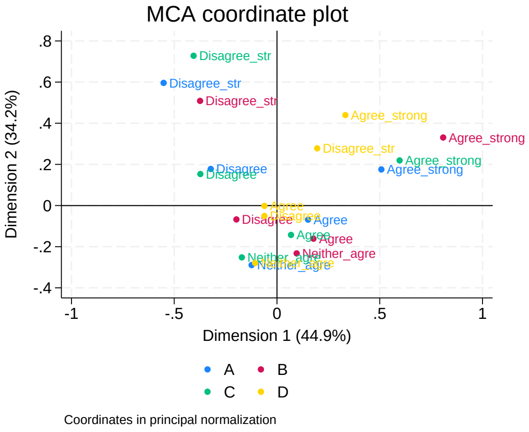
We continue to use the ISSP (1993) dataset, looking at all four questions on attitudes toward science. We use the default method of MCA, which is a CA of the Burt matrix for the data, followed by simple scale adjustments. We choose the principal normalization `normalize(principal)`, which scales the coordinates by the principal inertias. MCA with the Burt matrix and adjustments explains at least 79.1% of the total inertia in the first two dimensions.

```
. mca A-D, normalize(principal)

Multiple/joint correspondence analysis         Number of obs    =        871
                                               Total inertia    =  .17024552
    Method: Burt/adjusted inertias             Number of axes   =          2

                 |   Principal               Cumul.
       Dimension |     inertia    Percent    percent
-----------------+-----------------------------------
           Dim 1 |    .0764553     44.91      44.91
           Dim 2 |    .0582198     34.20      79.11
           Dim 3 |     .009197      5.40      84.51
           Dim 4 |    .0056697      3.33      87.84
           Dim 5 |    .0011719      0.69      88.53
           Dim 6 |    6.61e-06      0.00      88.53
-----------------+-----------------------------------
           Total |    .1702455    100.00
```

```
Statistics for column categories in principal normalization

                 |           Overall            |        Dimension_1
      Categories |    Mass  Quality  %inert  |   Coord   Sqcorr  Contrib
    -------------+-----------------------------+---------------------------
    A            |                              |
    Agree stro~y |   0.034    0.963    0.060  |   0.508    0.860    0.115
           Agree |   0.092    0.659    0.023  |   0.151    0.546    0.028
    Neither ag~e |   0.059    0.929    0.037  |  -0.124    0.143    0.012
        Disagree |   0.051    0.798    0.051  |  -0.322    0.612    0.069
    Disagree s~y |   0.014    0.799    0.067  |  -0.552    0.369    0.055
    -------------+-----------------------------+---------------------------
    B            |                              |
    Agree stro~y |   0.020    0.911    0.100  |   0.809    0.781    0.174
           Agree |   0.050    0.631    0.027  |   0.177    0.346    0.021
    Neither ag~e |   0.059    0.806    0.027  |   0.096    0.117    0.007
        Disagree |   0.081    0.620    0.033  |  -0.197    0.555    0.041
    Disagree s~y |   0.040    0.810    0.116  |  -0.374    0.285    0.074
    -------------+-----------------------------+---------------------------
    C            |                              |
    Agree stro~y |   0.044    0.847    0.122  |   0.597    0.746    0.203
           Agree |   0.091    0.545    0.024  |   0.068    0.101    0.006
    Neither ag~e |   0.057    0.691    0.045  |  -0.171    0.218    0.022
        Disagree |   0.044    0.788    0.054  |  -0.373    0.674    0.080
    Disagree s~y |   0.015    0.852    0.071  |  -0.406    0.202    0.032
    -------------+-----------------------------+---------------------------
    D            |                              |
    Agree stro~y |   0.017    0.782    0.039  |   0.333    0.285    0.025
           Agree |   0.067    0.126    0.012  |  -0.061    0.126    0.003
    Neither ag~e |   0.058    0.688    0.044  |  -0.106    0.087    0.009
        Disagree |   0.065    0.174    0.014  |  -0.061    0.103    0.003
    Disagree s~y |   0.043    0.869    0.034  |   0.196    0.288    0.022
    ---------------------------------------------------------------------------
```

```
                       Dimension_2
 Categories       Coord   Sqcorr  Contrib

A
Agree stro~y      0.176    0.103    0.018
       Agree     -0.069    0.113    0.007
Neither ag~e     -0.289    0.786    0.084
    Disagree      0.178    0.186    0.028
Disagree s~y      0.596    0.430    0.084

B
Agree stro~y      0.331    0.131    0.038
       Agree     -0.161    0.285    0.022
Neither ag~e     -0.233    0.690    0.055
    Disagree     -0.068    0.065    0.006
Disagree s~y      0.509    0.526    0.179

C
Agree stro~y      0.219    0.101    0.036
       Agree     -0.143    0.444    0.032
Neither ag~e     -0.252    0.473    0.062
    Disagree      0.153    0.114    0.018
Disagree s~y      0.728    0.650    0.136

D
Agree stro~y      0.440    0.497    0.057
       Agree     -0.002    0.000    0.000
Neither ag~e     -0.280    0.601    0.078
    Disagree     -0.051    0.071    0.003
Disagree s~y      0.278    0.581    0.057
```

```
. mcaplot, overlay origin
```

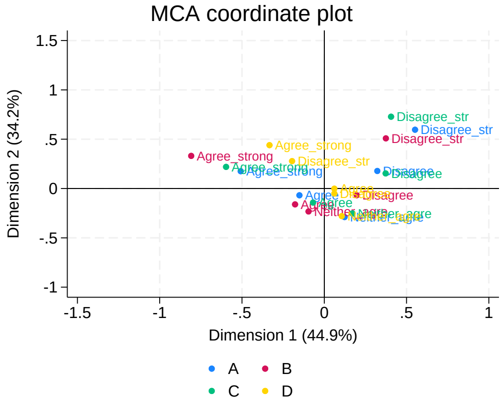


We displayed the origin axes in the plot with option `origin`. This plot makes it easier to see data associations. For more information on `mcaplot`, see [MV] **mca postestimation plots**.

A clear pattern is seen in the plot. Results from questions A, B, and C are clustered together, whereas question D forms a pattern of its own. Question D is formulated differently from A, B, and C, and the plot shows its incompatibility with the others.

Greenacre (2006, 70) produces this same plot. To obtain equivalent results, we reflect over the y axis with `xnegate` and adjust the scale of the graph with `xlabel()` and `ylabel()`.

```
. mcaplot, overlay xnegate origin ylabel(-1(.5)1.5) xlabel(-1.5(.5)1)
```



Coordinates in principal normalization

◁

❑ Technical note

The percentage of inertia in the top table of the `mca` output does not add up to 100%, although all singular values are displayed. Why? The percentages are lower-bound estimates in the Burt method with adjustments.

❑

❑ Technical note

Readers who are familiar with the discussion of the `normalize()` option in [MV] **ca** might be wondering why we are doing analysis of coordinates in the principal normalization. Principal normalization in `ca` does not allow good associations between the row and column categories, and so the symmetric normalization is the default used for `ca`. Principal normalization does allow studying the row categories or column categories separately from each other in `ca`. In `mca` there are only column categories. Consequently, the principal normalization is often preferred.

❑

CA of the indicator matrix

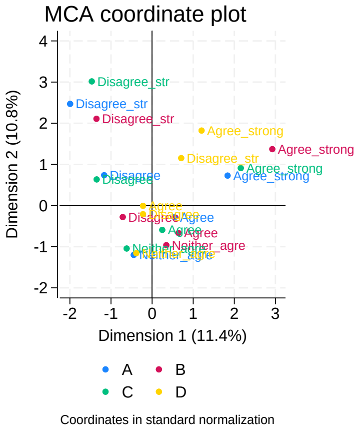
▷ Example 3

We compare the previous result with that obtained using the `method(indicator)` option to perform an equivalent analysis to CA on the indicator matrix for these data. The first two dimensions explain only 22.2% of the total principal inertia.

```
. mca A-D, method(indicator)

Multiple/joint correspondence analysis          Number of obs     =        871
                                                Total inertia     =          4
    Method: Indicator matrix                    Number of axes    =          2

                |   Principal               Cumul.
      Dimension |     inertia    Percent   percent
    ------------+---------------------------------
          Dim 1 |    .4573792      11.43     11.43
          Dim 2 |    .4309658      10.77     22.21
          Dim 3 |    .3219257       8.05     30.26
          Dim 4 |    .3064732       7.66     37.92
          Dim 5 |    .2756747       6.89     44.81
          Dim 6 |     .251928       6.30     51.11
          Dim 7 |    .2425591       6.06     57.17
          Dim 8 |    .2349506       5.87     63.05
          Dim 9 |     .225468       5.64     68.68
         Dim 10 |    .2206291       5.52     74.20
         Dim 11 |    .2098376       5.25     79.44
         Dim 12 |    .1971485       4.93     84.37
         Dim 13 |    .1778833       4.45     88.82
         Dim 14 |    .1691119       4.23     93.05
         Dim 15 |    .1528191       3.82     96.87
         Dim 16 |    .1252462       3.13    100.00
    ------------+---------------------------------
          Total |           4     100.00

Statistics for column categories in standard normalization

                 |           Overall          |        Dimension_1
      Categories |   Mass  Quality   %inert   |   Coord   Sqcorr  Contrib
    -------------+----------------------------+---------------------------
    A            |                            |
    Agree stro~y |  0.034    0.280    0.054   |   1.837    0.244    0.078
           Agree |  0.092    0.100    0.039   |   0.546    0.080    0.019
    Neither ag~e |  0.059    0.218    0.048   |  -0.447    0.028    0.008
        Disagree |  0.051    0.220    0.050   |  -1.166    0.160    0.047
    Disagree s~y |  0.014    0.260    0.059   |  -1.995    0.106    0.037
    -------------+----------------------------+---------------------------
    B            |                            |
    Agree stro~y |  0.020    0.419    0.057   |   2.924    0.347    0.118
           Agree |  0.050    0.095    0.050   |   0.642    0.047    0.014
    Neither ag~e |  0.059    0.140    0.048   |   0.346    0.017    0.005
        Disagree |  0.081    0.127    0.042   |  -0.714    0.111    0.028
    Disagree s~y |  0.040    0.527    0.052   |  -1.354    0.161    0.050
    -------------+----------------------------+---------------------------
    C            |                            |
    Agree stro~y |  0.044    0.525    0.052   |   2.158    0.450    0.137
           Agree |  0.091    0.102    0.040   |   0.247    0.016    0.004
    Neither ag~e |  0.057    0.189    0.048   |  -0.619    0.051    0.015
        Disagree |  0.044    0.216    0.051   |  -1.349    0.179    0.054
    Disagree s~y |  0.015    0.312    0.059   |  -1.468    0.063    0.022
    -------------+----------------------------+---------------------------
    D            |                            |
    Agree stro~y |  0.017    0.155    0.058   |   1.204    0.049    0.017
           Agree |  0.067    0.008    0.046   |  -0.221    0.008    0.002
    Neither ag~e |  0.058    0.195    0.048   |  -0.385    0.020    0.006
        Disagree |  0.065    0.015    0.046   |  -0.222    0.008    0.002
    Disagree s~y |  0.043    0.168    0.052   |   0.708    0.048    0.015
    --------------------------------------------------------------------------
```

```
                │        Dimension_2
     Categories │   Coord   Sqcorr  Contrib
 ───────────────┼──────────────────────────
 A              │
 Agree stro~y   │   0.727    0.036    0.012
          Agree │  -0.284    0.020    0.005
 Neither ag~e   │  -1.199    0.190    0.055
       Disagree │   0.737    0.060    0.018
 Disagree s~y   │   2.470    0.153    0.055
 ───────────────┼──────────────────────────
 B              │
 Agree stro~y   │   1.370    0.072    0.025
          Agree │  -0.667    0.048    0.015
 Neither ag~e   │  -0.964    0.123    0.036
       Disagree │  -0.280    0.016    0.004
 Disagree s~y   │   2.108    0.367    0.117
 ───────────────┼──────────────────────────
 C              │
 Agree stro~y   │   0.909    0.075    0.024
          Agree │  -0.592    0.086    0.021
 Neither ag~e   │  -1.044    0.137    0.040
       Disagree │   0.635    0.037    0.012
 Disagree s~y   │   3.017    0.249    0.089
 ───────────────┼──────────────────────────
 D              │
 Agree stro~y   │   1.822    0.106    0.038
          Agree │  -0.007    0.000    0.000
 Neither ag~e   │  -1.159    0.175    0.051
       Disagree │  -0.211    0.007    0.002
 Disagree s~y   │   1.152    0.120    0.038
 ───────────────┴──────────────────────────
```

```
. mcaplot, overlay origin
```



Even though the first two dimensions explain only 22.2% of the total inertia, and difficulties exist in justifying the full χ^2 geometry for the CA of the indicator matrix, the same clear pattern we saw in the previous plot is seen here. Variables A, B, and C are related, but variable D does not belong with the rest.

◁

CA of the Burt matrix

▷ Example 4

Now we will look at results with `method(burt)` and the `noadjust` option. This performs a CA on the Burt matrix without adjusting the principal inertias. This model does slightly better than the indicator matrix approach, explaining 35.0% of the principal inertia in the first two dimensions. We display column points only for variable A to reduce the output.

```
. mca A-D, method(burt) noadjust points(A)

Multiple/joint correspondence analysis         Number of obs     =        871
                                               Total inertia     =  1.1276841
    Method: Burt/unadjusted inertias           Number of axes    =          2

                |   Principal               Cumul.
      Dimension |     inertia   Percent    percent
    ------------+---------------------------------
                |
          Dim 1 |    .2091957     18.55      18.55
          Dim 2 |    .1857315     16.47      35.02
          Dim 3 |    .1036362      9.19      44.21
          Dim 4 |    .0939258      8.33      52.54
          Dim 5 |    .0759966      6.74      59.28
          Dim 6 |    .0634677      5.63      64.91
          Dim 7 |    .0588349      5.22      70.12
          Dim 8 |    .0552018      4.90      75.02
          Dim 9 |    .0508358      4.51      79.53
         Dim 10 |    .0486772      4.32      83.84
         Dim 11 |    .0440318      3.90      87.75
         Dim 12 |    .0388675      3.45      91.20
         Dim 13 |    .0316425      2.81      94.00
         Dim 14 |    .0285988      2.54      96.54
         Dim 15 |    .0233537      2.07      98.61
         Dim 16 |    .0156866      1.39     100.00
    ------------+---------------------------------
                |
          Total |    1.127684    100.00

Statistics for column categories in standard normalization

                 |          Overall          |        Dimension_1
      Categories |    Mass  Quality   %inert |   Coord   Sqcorr  Contrib
    -------------+---------------------------+---------------------------
                 |                           |
    A            |                           |
    Agree stro~y |   0.034    0.445    0.055 |   1.837    0.391    0.115
           Agree |   0.092    0.169    0.038 |   0.546    0.136    0.028
    Neither ag~e |   0.059    0.344    0.047 |  -0.447    0.047    0.012
        Disagree |   0.051    0.350    0.050 |  -1.166    0.258    0.069
    Disagree s~y |   0.014    0.401    0.060 |  -1.995    0.170    0.055
    ---------------------------------------------------------------------

                 |        Dimension_2
      Categories |   Coord   Sqcorr  Contrib
    -------------+---------------------------
                 |
    A            |
    Agree stro~y |   0.727    0.054    0.018
           Agree |  -0.284    0.033    0.007
    Neither ag~e |  -1.199    0.298    0.084
        Disagree |   0.737    0.092    0.028
    Disagree s~y |   2.470    0.231    0.084
    -----------------------------------------
```

We do not provide a plot for this mca example; it would be the same as the previous one. MCA via the indicator matrix or the Burt matrix produces the same standard coordinates, although they produce different principal inertias. If normalize(principal) is used, different coordinates will be produced. Principal normalization relies on the principal inertias for normalization.

For more information on normalization, see either *Methods and formulas* in mca or the discussion of normalization in [MV] **ca**.

◁

Joint correspondence analysis

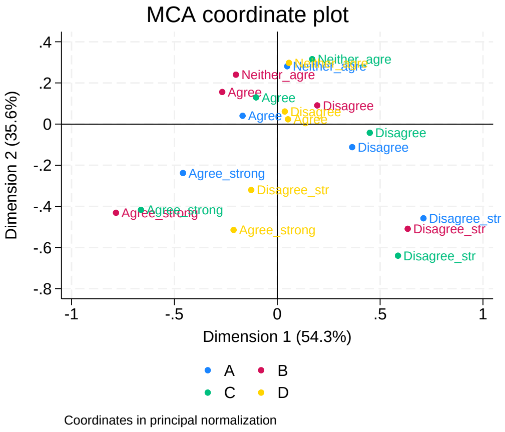
▷ Example 5

JCA attempts to remedy inflation of the total inertia by the block diagonal submatrices of the Burt matrix and is implemented as method(joint). Results using JCA explain 90% of the principal inertia in the first two dimensions. With JCA, we must specify the dimensions we want in advance. The output contains principal inertias only for the first two dimensions (the default). For other methods, principal inertias for all dimensions are displayed. More dimensions can be requested with the dim() option. JCA is an iterative method, with repeated CAs and adjustments taking place until convergence is achieved.

We ask for coordinates in the principal normalization, norm(principal), and ask for coordinates of only variable A to be displayed.

```
. mca A-D, method(joint) norm(principal) points(A)
Multiple/joint correspondence analysis         Number of obs     =        871
                                               Total inertia     =  .18242477
    Method: Joint (JCA)                        Number of axes    =          2

             |   Principal               Cumul.
   Dimension |     inertia    Percent   percent
  -----------+-----------------------------------
       Dim 1 |     .099091      54.32     54.32
       Dim 2 |    .0650329      35.65     89.97
  -----------+-----------------------------------
       Total |    .1824248     100.00
Statistics for column categories in principal normalization

              |           Overall           |        Dimension_1
   Categories |  Mass  Quality   %inert     |  Coord   Sqcorr  Contrib
  ------------+-----------------------------+---------------------------
  A           |                             |
  Agree stro~y | 0.034    0.964    0.052    | -0.458    0.759    0.072
         Agree | 0.092    0.774    0.020    | -0.169    0.733    0.027
  Neither ag~e | 0.059    0.884    0.030    |  0.048    0.025    0.001
      Disagree | 0.051    0.887    0.046    |  0.364    0.810    0.068
  Disagree s~y | 0.014    0.899    0.060    |  0.711    0.636    0.070
  ------------+-----------------------------+---------------------------

              |        Dimension_2
   Categories |  Coord   Sqcorr  Contrib
  ------------+---------------------------
  A           |
  Agree stro~y | -0.238    0.205    0.030
         Agree |  0.040    0.042    0.002
  Neither ag~e |  0.281    0.859    0.071
      Disagree | -0.112    0.077    0.010
  Disagree s~y | -0.458    0.264    0.044
  ------------+---------------------------
```

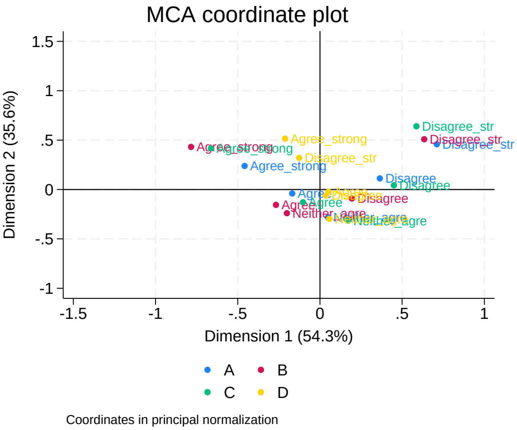
```
. mcaplot, overlay origin
```



The plot shows the same relationships among the variables that we saw in the one after the indicator matrix approach to MCA. The main difference between the plots is the change in scale, partially because of the normalization and largely because of the change of method.

These same data are analyzed and plotted in Greenacre (2006, 66). To obtain an equivalent plot, we perform a reflection on the data with ynegate, add in the origin axes with origin, and use the same scale with options xlabel() and ylabel(),

```
. mcaplot, overlay ynegate origin ylabel(-1(.5)1.5) xlabel(-1.5(.5)1)
```



Note the similarities between this plot and the one obtained through the default approach to MCA via the Burt matrix and adjustments.

◁

Stored results

mca stores the following in e():

Scalars

e(N)	number of observations
e(f)	number of dimensions
e(inertia)	total inertia
e(ev_unique)	1 if all eigenvalues are distinct, 0 otherwise
e(adjust)	1 if eigenvalues are adjusted, 0 otherwise (method(burt) only)
e(converged)	1 if successful convergence, 0 otherwise (method(joint) only)
e(iter)	number of iterations (method(joint) only)
e(inertia_od)	proportion of off-diagonal inertia explained by extracted dimensions (method(joint) only)

Macros

e(cmd)	mca
e(cmdline)	command as typed
e(names)	names of MCA variables (crossed or actual)
e(supp)	names of supplementary variables
e(defs)	per crossed variable: crossing variables separated by “\”
e(missing)	missing, if specified
e(crossed)	1 if there are crossed variables, 0 otherwise
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in output
e(method)	burt, indicator, or joint
e(norm)	standard or principal
e(properties)	nob noV eigen
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins

Matrices

e(Coding#)	row vector with coding of variable #
e(A)	standard coordinates for column categories
e(F)	principal coordinates for column categories
e(cMass)	column mass
e(cDist)	distance column to centroid
e(cInertia)	column inertia
e(cGS)	general statistics of column categories [.,1] column mass [.,2] overall quality [.,3] inertia/sum(inertia) [.,3*f+1] dim f: coordinate in e(norm) normalization [.,3*f+2] dim f: contribution of the profiles to principal axes [.,3*f+3] dim f: contribution of principal axes to profiles (= squared correlation of profile and axes)
e(rSCW)	weight matrix for row standard coordinates
e(Ev)	principal inertias/eigenvalues
e(inertia_e)	explained inertia (percent)
e(Bmod)	modified Burt matrix of active variables (method(joint) only)
e(inertia_sub)	variable-by-variable inertias (method(joint) only)

Functions

e(sample)	marks estimation sample

Methods and formulas

Methods and formulas are presented under the following headings:

Notation

We use notation that is fairly standard in the literature on correspondence analysis (for example, Greenacre [2006]). Let $x_1, \dots, x_q$ be categorical variables on N observations that are active in the analysis. To simplify notation, but without loss of generality, we assume that x_j is coded with consecutive integers $1, \dots, n_j$. Let $\mathbf{Z}^{(j)}$ be the $N \times n_j$ binary indicator matrix associated with x_j, $\mathbf{Z}^{(j)}_{ih} = 1$ if and only if $x_{ij} = h$. Let

$$\mathbf{Z} = \left(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(q)}\right)$$

be the $N \times J$ indicator matrix of the set of active x-variables, where $J = n_1 + \cdots + n_q$.

We will be consistent in letting i index observations $1, \dots, N$, j index variables $1, \dots, q$, and h index categories $1, \dots, n_j$, or $1, \dots, J$.

The $J \times J$ Burt matrix is defined as $\mathbf{B} = \mathbf{Z}'\mathbf{Z}$, or $\mathbf{B} = \mathbf{Z}'\mathbf{D}(w)\mathbf{Z}$, where w is the weight for the analysis and $\mathbf{D}(w)$ is a $J \times J$ square matrix with the weights on the diagonal and 0 off diagonal. The diagonal block of $\mathbf{B}$ associated with variable x_j is a diagonal matrix with the frequencies of x_j on the diagonal. The off-diagonal block of $\mathbf{B}$ associated with variables x_j and x_k is the two-way cross-tabulation of x_j and x_k.

In an analogous way, we define $\mathbf{B}^*$, the Burt matrix with more rows containing cross-tabulation from the supplementary variables. $\mathbf{B}^* = \mathbf{Z}^{*\prime}\mathbf{Z}$, where $\mathbf{Z}^*$ is the indicator matrix with more columns for the supplementary variables.

$\mathbf{D}(v)$, in general, represents a diagonal matrix with the elements of vector v on the diagonal and 0 off diagonal; $\mathbf{1}$ is a column vector of ones where length is defined by the context.

Using ca to compute MCA

The indicator approach to MCA involves an equivalent technique to performing a standard CA on the indicator matrix $\mathbf{Z}$; see Greenacre (2006). We refer to *Methods and formulas* in [MV] **ca** for details. The indicator approach could be handled by forming $\mathbf{Z}$ and invoking `camat` on $\mathbf{Z}$. If you had k categorical variables named `v1`, ..., `vk`, you could perform CA on the indicator matrix by taking the steps in the following code fragment:

```
tab v1, gen(V1_)
...
tab vk, gen(Vk_)
mkmat V1_* ... Vk_*, matrix(VALL)
camat VALL
```

CA of an indicator or Burt matrix

$\mathbf{Z}$ may be a rather large matrix: the number of rows is equal to the number of observations, and memory limitations would pose a limit for applications. Moreover, the interest in MCA is mostly on the column space, and so much computation would be wasted. Therefore, the implementation in `mca` exploits the close analogy between the indicator and Burt matrix approach for efficiency (see Greenacre [1984, chap. 5; 2006, app. A.3]). The consequence is that the CA results for the rows of $\mathbf{Z}$ are not computed. You may use the above scheme if you need row information.

Much literature (for example, Everitt and Dunn [2001]; Rencher and Christensen [2012]) identifies MCA with the application of CA on the Burt matrix $\mathbf{B} = \mathbf{Z}'\mathbf{Z}$.

This discussion is related to that of *Methods and formulas* in [MV] **ca**. We define

$$\begin{aligned}
\mathbf{B}_{++} &= \sum_{k=1}^{J}\sum_{h=1}^{J}\mathbf{B}_{kh} \\
\mathbf{P} &= \mathbf{B}/\mathbf{B}_{++} \\
\mathbf{c} &= \sum_{k=1}^{J}\mathbf{P}_{k*} = \mathbf{P}_{+*} = \mathbf{P}'\mathbf{1} \qquad \text{here } * \text{ represents all possible values} \\
\mathbf{S} &= \mathbf{D}(\mathbf{c})^{-1/2}(\mathbf{P}-\mathbf{cc}')\mathbf{D}(\mathbf{c})^{-1/2}
\end{aligned}$$

$\mathbf{c}$ is called the column mass. $\mathbf{D}(\mathbf{c})$ is the diagonal matrix with diagonal $\mathbf{c}$; $\mathbf{D}(\mathbf{c})^{-1/2}$ is therefore the diagonal matrix with elements $1/\sqrt{c_t}$, where c_t is an element of $\mathbf{c}$. $\mathbf{B}$ and $\mathbf{S}$ are symmetric. Thus the singular value decomposition commonly associated with CA is equivalent to the spectral or eigen decomposition of $\mathbf{S}$.

$$\mathbf{S} = \mathbf{V}\boldsymbol{\Phi}\mathbf{V}', \qquad \phi_1 \ge \phi_2 \ge \dots$$

For identification purposes and to facilitate resampling methods, the eigenvectors of $\mathbf{V}$ are directed so that for h, $\mathbf{V}_{kh} > 0$ for the first k for which $\mathbf{V}_{kh} \neq 0$. The standard column coordinates $\mathbf{A}$ are the same for the indicator and the Burt approach and are defined as

$$\mathbf{A} = \mathbf{D}(\mathbf{c})^{-1}\mathbf{V}$$

$\mathbf{D}(\mathbf{c})^{-1}$ is the diagonal matrix with elements $1/c_t$, where c_t is an element of $\mathbf{c}$.

In the indicator approach to MCA, the tth principal inertia is defined as $\lambda_t = \phi_t$, the total inertia as $\sum_t \lambda_t$. The inertia of column j or variable j, $\mathbf{In}^{(j)}$, is computed elementwise as

$$\text{In}_h^{(j)} = \sum_{i=1}^{N} w_i \frac{(Z_{ih}^{(j)} - qc_h^{(j)})^2}{q^2 c_h^{(j)} w_+}$$

where $Z_{ih}^{(j)}$ is the i,hth element of the indicator matrix for variable j, w_i is the weight for observation i, q is the number of active variables, $c_h^{(j)}$ is the column mass of variable j for category h, and w_+ is the sum of the weights over the observations.

In the Burt approach to MCA, the unadjusted principal inertia is defined as $\lambda_t^{\text{unadj}} = \phi_t^2$, the total unadjusted inertia as $\sum_t \lambda_t^{\text{unadj}}$, and the unadjusted column inertias as $\mathbf{1}'\mathbf{S}\odot\mathbf{S}$, with $\odot$ the Hadamard or elementwise matrix product. The adjusted principal inertia, λ_t^{adj}, is defined as

$$\lambda_t^{\text{adj}} = \left(\frac{q}{q-1}\right)^2\left(\phi_t^2 - \frac{1}{q}\right)^2 \qquad \text{provided } q\phi_t > 1$$

The total inertia is defined as

$$\text{total inertia } = \left(\frac{q}{q-1}\right)\sum \phi_t^2 - \frac{(J-q)}{q^2}$$

The standard coordinates are independent of the principal inertia; with or without adjustment, these are defined as before

$$\mathbf{A} = \mathbf{D}(\mathbf{c})^{-1}\mathbf{V}$$

The principal coordinates **F** are defined as

$$\mathbf{F} = \mathbf{A}\mathbf{D}(\boldsymbol{\Lambda})^{1/2}$$

where $\boldsymbol{\Lambda}$ is a vector of adjusted or unadjusted principal inertias and $\mathbf{D}(\boldsymbol{\Lambda})^{1/2}$ is the diagonal matrix with elements $\lambda_t^{1/2}$ on the diagonals.

JCA

The implementation of JCA uses the alternating least-squares method proposed by Greenacre (1988, 2006). This algorithm forms a modification of the Burt matrix, changing the diagonal blocks associated with the variables, keeping the off-diagonal blocks unchanged. In the first iteration, $\mathbf{B}_0 = \mathbf{B}$. In iteration m, the blocks are replaced by the *f*-dimensional MCA approximation of the adjusted Burt matrix $\mathbf{B}_{m-1}$. Iteration continues until the change in the elements of $\mathbf{B}_m$ and $\mathbf{B}_{m-1}$ falls below a convergence tolerance. The JCA coordinates and inertias are computed from the converged solution, $\mathbf{B}_\infty$, analogous to the (adjusted) Burt method. The total inertia of the modified Burt matrix is defined as the sum of the inertias of the off-diagonal blocks.

To compute the *f*-dimensional MCA approximation of the adjusted Burt matrix $\mathbf{B}_{m-1}$, we perform MCA on $\mathbf{B}_{m-1}$ and then reconstruct the approximation of the data from the solution

$$\widehat{\mathbf{B}}_{hk} = \mathbf{B}_{++}c_h c_k \left(1 + \sum_{t=1}^{f} \phi_t^2 A_{ht} A_{kt}\right)$$

where A_{ht} is an element of the standard coordinate matrix $\mathbf{A}$, c_h and c_k are the column masses, and ϕ_t are the eigenvalues as in the computation of the CA of the Burt matrix. We then update the main diagonal submatrices of $\mathbf{B}_{m-1}$ with the corresponding entries of $\widehat{\mathbf{B}}$ to obtain $\mathbf{B}_m$.

Supplementary variables

The coordinates of supplementary variables are computed as weighted averages of the column coordinates by using the so-called CA transition formulas. As outlined by Greenacre (2006), standard coordinates may be used for averaging, with the profiles of the indicator representation of supplementary columns as weights. Supplementary principal column coordinates are computed as weighted averages of the standard active column coordinates, and then supplementary standard coordinates are computed by division by the principal inertias.

To compute, we add the supplementary variables to the Burt matrix as more rows; if **B** is the Burt matrix of the active variables then $\mathbf{B}^*$ is the Burt matrix with the additional cross-tabulation from the supplementary variables. Define **P** as above and $\mathbf{P}^*$ analogously with the supplementary variables added. MCA is performed on **B** as before, and information from this solution is then applied to $\mathbf{B}^*$. Let p^* represent the elements of $\mathbf{P}^*$. Let k index categories of supplementary variables, h index categories of active variables, and t index dimensions of the solution. Let **A** be the standard coordinates of the active variables, as computed previously. Then the principal coordinate for category k and dimension s is computed as:

$$g_{kt} = \sum_{h=1}^{J} \frac{p^*_{hk}}{p^*_{+k}} A_{ht}$$

Coordinates in standard coordinates are obtained by division by the square root of the corresponding principal inertia as described above.

predict

`predict` after `mca` produces variables with the MCA coordinates as displayed by `mca` for both active and supplementary variables. Formulas are shown above. `predict` can also compute row coordinates also known as row scores. Row coordinates computed are always based on the indicator method. The standard row coordinate for the tth dimension for the ith observation with indicator matrix elements Z_{ih} is computed as

$$R_{it} = \sum_{h=1}^{J} \frac{Z_{ih} A_{ht}}{q\sqrt{\phi_t}}$$

where **A** is the matrix of standard coordinates, q is the number of active variables in the analysis, and ϕ_t is an eigenvalue of the CA on the Burt matrix. To get the row coordinate in principal normalization, one multiplies by the square root of the corresponding principal inertia.

References

Everitt, B. S., and G. Dunn. 2001. *Applied Multivariate Data Analysis.* 2nd ed. London: Arnold.

Gower, J. C. 1990. Fisher's optimal scores and multiple correspondence analysis. *Biometrics* 46: 947–961. https://doi.org/10.2307/2532440.

———. 2006. "Divided by a common language—analyzing and visualizing two-way arrays". In *Multiple Correspondence Analysis and Related Methods*, edited by M. J. Greenacre and J. Blasius. Boca Raton, FL: Chapman and Hall/CRC.

Gower, J. C., and D. J. Hand. 1996. *Biplots.* London: Chapman and Hall.

Greenacre, M. J. 1984. *Theory and Applications of Correspondence Analysis.* London: Academic Press.

———. 1988. Correspondence analysis of multivariate categorical data by weighted least-squares. *Biometrika* 75: 457–467. https://doi.org/10.2307/2336595.

———. 2006. "From simple to multiple correspondence analysis". In *Multiple Correspondence Analysis and Related Methods*, edited by M. J. Greenacre and J. Blasius. Boca Raton, FL: Chapman and Hall.

Greenacre, M. J., and J. Blasius, eds. 1994. *Correspondence Analysis in the Social Sciences.* London: Academic Press.

———. 2006. *Multiple Correspondence Analysis and Related Methods.* Boca Raton, FL: Chapman and Hall.

ISSP. 1993. International Social Survey Programme. https://www.issp.org.

Milan, L., and J. C. Whittaker. 1995. Application of the parametric bootstrap to models that incorporate a singular value decomposition. *Journal of the Royal Statistical Society*, C ser., 44: 31–49. https://doi.org/10.2307/2986193.

Rencher, A. C., and W. F. Christensen. 2012. *Methods of Multivariate Analysis.* 3rd ed. Hoboken, NJ: Wiley. https://doi.org/10.1002/9781118391686.

Also see

[MV] **mca postestimation** — Postestimation tools for mca

[MV] **mca postestimation plots** — Postestimation plots for mca

[MV] **ca** — Simple correspondence analysis

[MV] **canon** — Canonical correlations

[MV] **factor** — Factor analysis

[MV] **pca** — Principal component analysis

[U] 20 Estimation and postestimation commands

For suggested citations, see the FAQ on citing Stata documentation.