Title

stata.com

mca postestimation — Postestimation tools for mca

Postestimation commands

The following postestimation commands are of special interest after `mca`:

Command	Description
`mcaplot`	plot of category coordinates
`mcaprojection`	MCA dimension projection plot
`estat coordinates`	display of category coordinates
`estat subinertia`	matrix of inertias of the active variables (after JCA only)
`estat summarize`	estimation sample summary
`screeplot`	plot principal inertias (eigenvalues)

The following standard postestimation commands are also available:

Command	Description
* `estimates`	cataloging estimation results
`predict`	row and category coordinates

*All `estimates` subcommands except `table` and `stats` are available; [R] **estimates**.

predict

Description for predict

`predict` creates a new variable containing predictions such as row scores and scores (coordinates) for the MCA variable.

Menu for predict

Statistics > Postestimation

Syntax for predict

`predict` [*type*] *newvar* [*if*] [*in*] [, *statistic* `normalize(`*norm*`)` `dimensions(#)`]

`predict` [*type*] {*stub*`*`|*newvarlist*} [*if*] [*in*] [, *statistic* `normalize(`*norm*`)` `dimensions(`*numlist*`)`]

statistic	Description
Main	
`rowscores`	row scores (coordinates); the default
`score(`*varname*`)`	scores (coordinates) for MCA variable *varname*

norm	Description
`standard`	use standard normalization
`principal`	use principal normalization

Options for predict

Main

`rowscores` specifies that row scores (row coordinates) be computed. The row scores returned are based on the indicator matrix approach to multiple correspondence analysis, even if another method was specified in the original `mca` estimation. The sample for which row scores are computed may exceed the estimation sample; for example, it may include supplementary rows (variables). `score()` and `rowscores` are mutually exclusive. `rowscores` is the default.

`score(`*varname*`)` specifies the name of a variable from the preceding MCA for which scores should be computed. The variable may be a regular categorical variable, a crossed variable, or a supplementary variable. `score()` and `rowscores` are mutually exclusive.

Options

`normalize(`*norm*`)` specifies the normalization of the scores (coordinates). `normalize(standard)` returns coordinates in standard normalization. `normalize(principal)` returns principal scores. The default is the normalization method specified with `mca` during estimation, or `normalize(standard)` if no method was specified.

`dimensions(#)` or `dimensions(`*numlist*`)` specifies the dimensions for which scores (coordinates) are computed. The number of dimensions specified should equal the number of variables in *newvarlist*. If `dimensions()` is not specified, scores for dimensions 1, ..., k are returned, where k is the number of variables in *newvarlist*. The number of variables in *newvarlist* should not exceed the number of dimensions extracted during estimation.

estat

Description for estat

`estat coordinates` displays the category coordinates, optionally with column statistics.

`estat subinertia` displays the matrix of inertias of the active variables (after JCA only).

`estat summarize` displays summary information of MCA variables over the estimation sample.

Menu for estat

Statistics > Postestimation

Syntax for estat

Display of category coordinates

`estat coordinates` [*varlist*] [, *coordinates_options*]

Matrix of inertias of the active variables (after JCA only)

`estat subinertia`

Estimation sample summary

`estat summarize` [, *summarize_options*]

Note: Variables in *varlist* must be from the preceding `mca` and may refer to either a regular categorical variable or a crossed variable. The variables in *varlist* may also be chosen from the supplementary variables.

coordinates_options	Description
`normalize(standard)`	standard coordinates
`normalize(principal)`	principal coordinates
`stats`	include mass, distance, and inertia
`format(%`*fmt*`)`	display format; default is `format(%9.4f)`

summarize_options	Description
Main	
`crossed`	summarize crossed and uncrossed variables as used
`labels`	display variable labels
`noheader`	suppress the header
`noweights`	ignore weights

Options for estat

Options for `estat` are presented under the following headings:

Options for estat coordinates

`normalize(`*norm*`)` specifies the normalization of the scores (coordinates). `normalize(standard)` returns coordinates in standard normalization. `normalize(principal)` returns principal scores. The default is the normalization method specified with `mca` during estimation, or `normalize(standard)` if no method was specified.

`stats` includes the column mass, the distance of the columns to the centroid, and the column inertias in the table.

`format(`*%fmt*`)` specifies the display format for the matrix, for example, `format(%8.3f)`. The default is `format(%9.4f)`.

Options for estat summarize

Main

`crossed` specifies summarizing the crossed variables if crossed variables are used in the MCA, rather than the crossing variables from which they are formed. The default is to summarize the crossing variables and single categorical variables used in the MCA.

`labels` displays variable labels.

`noheader` suppresses the header.

`noweights` ignores the weights, if any. The default when weights are present is to perform a weighted summarize on all variables except the weight variable itself. An unweighted summarize is performed on the weight variable.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Postestimation statistics

We continue to examine the ISSP (1993) dataset on the environment. We perform joint correspondence analysis.

Example 1: Category coordinates and subinertias

```
. use http://www.stata-press.com/data/r15/issp93
(Selection from ISSP (1993))
. mca A-D, method(joint)
Multiple/Joint correspondence analysis          Number of obs     =        871
                                                Total inertia     =  .18242477
    Method: Joint (JCA)                         Number of axes    =          2
```

Dimension	principal inertia	percent	cumul percent
dim 1	.099091	54.32	54.32
dim 2	.0650329	35.65	89.97
Total	.1824248	100.00	

Statistics for column categories in standard normalization

Categories	overall mass	overall quality	overall %inert	dimension_1 coord	dimension_1 sqcorr	dimension_1 contrib
A						
agree stro~y	0.034	0.964	0.052	-1.456	0.759	0.072
agree	0.092	0.774	0.020	-0.536	0.733	0.027
neither ag~e	0.059	0.884	0.030	0.154	0.025	0.001
disagree	0.051	0.887	0.046	1.157	0.810	0.068
disagree s~y	0.014	0.899	0.060	2.258	0.636	0.070
B						
agree stro~y	0.020	0.957	0.093	-2.490	0.735	0.126
agree	0.050	0.851	0.031	-0.850	0.636	0.036
neither ag~e	0.059	0.953	0.033	-0.639	0.393	0.024
disagree	0.081	0.705	0.029	0.618	0.579	0.031
disagree s~y	0.040	0.977	0.149	2.014	0.594	0.163
C						
agree stro~y	0.044	0.983	0.149	-2.104	0.704	0.193
agree	0.091	0.665	0.020	-0.327	0.257	0.010
neither ag~e	0.057	0.839	0.047	0.539	0.188	0.016
disagree	0.044	0.907	0.054	1.429	0.899	0.090
disagree s~y	0.015	0.944	0.065	1.864	0.431	0.052
D						
agree stro~y	0.017	0.850	0.034	-0.674	0.124	0.008
agree	0.067	0.145	0.008	0.165	0.120	0.002
neither ag~e	0.058	0.769	0.038	0.181	0.027	0.002
disagree	0.065	0.178	0.010	0.116	0.047	0.001
disagree s~y	0.043	0.929	0.030	-0.402	0.125	0.007

Categories	dimension_2 coord	sqcorr	contrib
A			
agree stro~y	-0.934	0.205	0.030
agree	0.158	0.042	0.002
neither ag~e	1.103	0.859	0.071
disagree	-0.440	0.077	0.010
disagree s~y	-1.796	0.264	0.044
B			
agree stro~y	-1.690	0.222	0.058
agree	0.611	0.216	0.019
neither ag~e	0.942	0.560	0.052
disagree	0.356	0.126	0.010
disagree s~y	-1.995	0.383	0.160
C			
agree stro~y	-1.634	0.279	0.116
agree	0.508	0.407	0.023
neither ag~e	1.237	0.651	0.087
disagree	-0.166	0.008	0.001
disagree s~y	-2.509	0.513	0.094
D			
agree stro~y	-2.018	0.727	0.070
agree	0.092	0.024	0.001
neither ag~e	1.166	0.741	0.079
disagree	0.239	0.131	0.004
disagree s~y	-1.256	0.804	0.068

If we wish to see the coordinates in the principal normalization, we do not need to rerun our analysis. We can use `estat coordinates` to display them.

```
. estat coordinates, norm(principal)
Column principal coordinates
```

Categories	dim1	dim2
A		
agree_stro~y	-.4582629	-.2381115
agree	-.1686314	.0402091
neither_ag~e	.0484366	.2811716
disagree	.3642677	-.1123168
disagree_s~y	.7106647	-.4578886
B		
agree_stro~y	-.783911	-.4310436
agree	-.2674646	.1558017
neither_ag~e	-.2010783	.2402487
disagree	.1944504	.0906804
disagree_s~y	.6341215	-.5088398
C		
agree_stro~y	-.6623101	-.4166016
agree	-.1029922	.1295649
neither_ag~e	.169804	.3155628
disagree	.4496893	-.0423339
disagree_s~y	.5867913	-.6397215
D		
agree_stro~y	-.2123187	-.5145647
agree	.05208	.0233723
neither_ag~e	.0569168	.297262
disagree	.0365233	.0609881
disagree_s~y	-.1264563	-.3203889

We may also be interested in the contributions of each submatrix of the Burt matrix to the total inertia. This can be obtained by `estat subinertia`. Compare with Greenacre (2006, table A.12).

```
. estat subinertia
Subinertias: decomposition of total inertia
```

Variable	A	B	C	D
A	.0074502			
B	.0148596	.022442		
C	.012149	.0185838	.0210336	
D	.0032898	.0053016	.0096583	.0038148

Predicting new variables

Coordinates (scores) can be predicted after `mca`. You can predict either the column coordinates or the row coordinates.

Example 2: Predictions

We will calculate row scores by hand and via `predict, rowscore`. Row scores are calculated off the indicator method approach. This calculation would not work with JCA, because the standard coordinates and the principal inertias are different for JCA. The principal inertias are also a problem with the Burt method, with and without adjustment. We use the `points(A)` option to restrict the output from `mca`. Four variables were used in the original `mca`. `predict` without the `dimensions()` option predicts the first dimension.

```
. mca A-D, method(indicator) points(A)
Multiple/Joint correspondence analysis         Number of obs     =        871
                                               Total inertia     =          4
    Method: Indicator matrix                   Number of axes    =          2
```

Dimension	principal inertia	percent	cumul 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
```

Categories	overall mass	overall quality	overall %inert	dimension_1 coord	dimension_1 sqcorr	dimension_1 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

Categories	dimension_2 coord	dimension_2 sqcorr	dimension_2 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

```
. predict double a1, score(A)
. predict double b1, score(B)
. predict double c1, score(C)
```

```
. predict double d1, score(D)
. predict double r1, rowscore
. mat Ev = e(Ev)
. scalar phi1 = Ev[1,1]
. generate double rc = (a1+b1+c1+d1)/(4*sqrt(phi1))
. assert reldif(rc, r1) < 1e-14
```

In the indicator method approach, we can also find Cronbach's alpha either via the `alpha` command (see [MV] **alpha**) or by hand.

```
. alpha a1 b1 c1 d1
Test scale = mean(unstandardized items)
Average interitem covariance:     .2768234
Number of items in the scale:            4
Scale reliability coefficient:      0.6045
. scalar alpha = (4/(4-1))*(1-1/(4*phi1))
. display alpha
.60454338
```

◁

Stored results

`estat summarize` stores the following in `r()`:

Matrices
- `r(stats)` — $k \times 4$ matrix of means, standard deviations, minimums, and maximums

`estat coordinates` stores the following in `r()`:

Macros
- `r(norm)` — normalization method of the coordinates

Matrices
- `r(Coord)` — column coordinates
- `r(Stats)` — column statistics: mass, distance, and inertia (option `stats` only)

`estat subinertia` stores the following in `r()`:

Matrices
- `r(inertia_sub)` — variable-by-variable inertias

Methods and formulas

See *Methods and formulas* in [MV] **mca**.

References

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

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

Also see

[MV] **mca** — Multiple and joint correspondence analysis

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

[MV] **screeplot** — Scree plot

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

[MV] **ca postestimation** — Postestimation tools for ca and camat

[U] **20 Estimation and postestimation commands**