canon postestimation — Postestimation tools for canon

Description

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

Command	Description
`estat correlations`	show correlation matrices
`estat loadings`	show loading matrices
`estat rotate`	rotate raw coefficients, standard coefficients, or loading matrices
`estat rotatecompare`	compare rotated and unrotated coefficients or loadings
`screeplot`	plot canonical correlations

The following standard postestimation commands are also available:

Command	Description
`estat summarize`	summary statistics for the estimation sample
`estat vce`	variance–covariance matrix of the estimators (VCE)
`estimates`	cataloging estimation results
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of coefficients
`nlcom`	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
`predict`	predictions, residuals, influence statistics, and other diagnostic measures
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

Special-interest postestimation commands

`estat correlations` displays the correlation matrices calculated by `canon` for *varlist*$_1$ and *varlist*$_2$ and between the two lists.

`estat loadings` displays the canonical loadings computed by `canon`.

`estat rotate` performs orthogonal varimax rotation of the raw coefficients, standard coefficients, or canonical loadings. Rotation is calculated on the canonical loadings regardless of which coefficients or loadings are actually rotated.

`estat rotatecompare` displays the rotated and unrotated coefficients or loadings and the most recently rotated coefficients or loadings. This command may be used only if `estat rotate` has been performed first.

Syntax for predict

`predict` [*type*] *newvar* [*if*] [*in*] , *statistic** [`correlation(#)`]

*statistic**	Description
Main	
`u`	calculate linear combination of $varlist_1$
`v`	calculate linear combination of $varlist_2$
`stdu`	calculate standard error of the linear combination of $varlist_1$
`stdv`	calculate standard error of the linear combination of $varlist_2$

* There is no default statistic; you must specify one *statistic* from the list.

These statistics are available both in and out of sample; type `predict ... if e(sample) ...` if wanted only for the estimation sample.

Menu for predict

Statistics > Postestimation > Predictions, residuals, etc.

Options for predict

Main

`u` and `v` calculate the linear combinations of $varlist_1$ and $varlist_2$, respectively. For the first canonical correlation, `u` and `v` are the linear combinations having maximal correlation. For the second canonical correlation, specified in `predict` with the `correlation(2)` option, `u` and `v` have maximal correlation subject to the constraints that `u` is orthogonal to the `u` from the first canonical correlation, and `v` is orthogonal to the `v` from the first canonical correlation. The third and higher correlations are defined similarly. Canonical correlations may be chosen either with the `lc()` option to `canon` or by specifying the `correlation()` option to `predict`.

`stdu` and `stdv` calculate the standard errors of the respective linear combinations.

`correlation(#)` specifies the canonical correlation for which the requested statistic is to be computed. The default value for `correlation()` is 1. If the `lc()` option to `canon` was used to calculate a particular canonical correlation, then only this canonical correlation is in the estimation results. You can obtain estimates for it either by specifying `correlation(1)` or by omitting the `correlation()` option.

Syntax for estat

Display the correlation matrices

`estat correlations` [, `format(`*%fmt*`)`]

Display the canonical loadings

estat loadings [, format(%*fmt*)]

Perform orthogonal varimax rotation

estat rotate [, rawcoefs stdcoefs loadings format(%*fmt*)]

Display the rotated and unrotated coefficients or loadings

estat rotatecompare [, format(%*fmt*)]

Menu for estat

Statistics > Postestimation > Reports and statistics

Option for estat

format(%*fmt*) specifies the display format for numbers in matrices; see [D] **format**. format(%8.4f) is the default.

rawcoefs, an option for estat rotate, requests the rotation of raw coefficients. It is the default.

stdcoefs, an option for estat rotate, requests the rotation of standardized coefficients.

loadings, an option for estat rotate, requests the rotation of the canonical loadings.

Remarks and examples

stata.com

In addition to the coefficients presented by canon in computing canonical correlations, several other matrices may be of interest.

▷ Example 1: Predictions

Recall from canon the example of two scientists trying to describe how "big" a car is. One took physical measurements—the length, weight, headroom, and trunk space—whereas the second took mechanical measurements—engine displacement, mileage rating, gear ratio, and turning radius. We discovered that these two views are closely related, with the best linear combination of the two types of measurements, the largest canonical correlation, at 0.9476. We can prove that the first canonical correlation is correct by calculating the two linear combinations and then calculating the ordinary correlation.

```
. use http://www.stata-press.com/data/r13/auto
(1978 Automobile Data)
. quietly canon (length weight headroom trunk) (displ mpg gear_ratio turn)
. predict physical, u corr(1)
. predict mechanical, v corr(1)
```

```
. correlate mechanical physical
(obs=74)

             | mechan~l physical
-------------+------------------
  mechanical |   1.0000
    physical |   0.9476   1.0000

. drop mechanical physical
```

Example 2: Canonical loadings

Researchers are often interested in the canonical loadings, the correlations between the original variable lists and their canonical variates. The canonical loadings are used to interpret the canonical variates. However, as shown in the technical note later in this entry, Rencher (1988; 1992; 1998, sec. 8.6.3) and Rencher and Christensen (2012, 397) have shown that there is no information in these correlations about how one variable list contributes jointly to canonical correlation with the other. Loadings are still often discussed, and `estat loadings` reports these as well as the cross-loadings or correlations between *varlist*$_1$ and the canonical variates for *varlist*$_2$ and the correlations between *varlist*$_2$ and the canonical variates for *varlist*$_1$. The loadings and cross-loadings are all computed by `canon`.

```
. estat loadings
Canonical loadings for variable list 1
```

	1	2	3	4
length	0.9664	0.2481	0.0361	-0.0566
weight	0.9972	-0.0606	-0.0367	0.0235
headroom	0.5140	-0.1295	0.7134	-0.4583
trunk	0.6941	0.0644	-0.0209	-0.7167

```
Canonical loadings for variable list 2
```

	1	2	3	4
displacement	0.9404	-0.3091	0.1050	0.0947
mpg	-0.8569	-0.1213	0.1741	0.4697
gear_ratio	-0.7945	0.3511	0.4474	-0.2129
turn	0.9142	0.3286	-0.0345	0.2345

```
Correlation between variable list 1 and canonical variates from list 2
```

	1	2	3	4
length	0.9158	0.0844	0.0023	-0.0025
weight	0.9449	-0.0206	-0.0023	0.0011
headroom	0.4871	-0.0440	0.0452	-0.0205
trunk	0.6577	0.0219	-0.0013	-0.0320

```
Correlation between variable list 2 and canonical variates from list 1
```

	1	2	3	4
displacement	0.8912	-0.1051	0.0067	0.0042
mpg	-0.8120	-0.0413	0.0110	0.0210
gear_ratio	-0.7529	0.1194	0.0284	-0.0095
turn	0.8663	0.1117	-0.0022	0.0105

```
. mat load2 = r(canload22)
```

▷ Example 3: Predictions and correlation matrices

In example 2, we saved the loading matrix for *varlist*$_2$, containing the mechanical variables, and we wish to verify that it is correct. We predict the canonical variates for *varlist*$_2$ and then find the canonical correlations between the canonical variates and the original mechanical variables as a means of getting the correlation matrices, which we then display using `estat correlations`. The mixed correlation matrix is the same as the loading matrix that we saved.

```
. predict mechanical1, v corr(1)
. predict mechanical2, v corr(2)
. predict mechanical3, v corr(3)
. predict mechanical4, v corr(4)
. quietly canon (mechanical1-mechanical4) (displ mpg gear_ratio turn)
. estat correlation
Correlations for variable list 1
```

	mechan~1	mechan~2	mechan~3	mechan~4
mechanical1	1.0000			
mechanical2	-0.0000	1.0000		
mechanical3	-0.0000	0.0000	1.0000	
mechanical4	-0.0000	-0.0000	-0.0000	1.0000

```
Correlations for variable list 2
```

	displa~t	mpg	gear_r~o	turn
displacement	1.0000			
mpg	-0.7056	1.0000		
gear_ratio	-0.8289	0.6162	1.0000	
turn	0.7768	-0.7192	-0.6763	1.0000

```
Correlations between variable lists 1 and 2
```

	mechan~1	mechan~2	mechan~3	mechan~4
displacement	0.9404	-0.3091	0.1050	0.0947
mpg	-0.8569	-0.1213	0.1741	0.4697
gear_ratio	-0.7945	0.3511	0.4474	-0.2129
turn	0.9142	0.3286	-0.0345	0.2345

```
. matlist load2, format(%8.4f) border(bottom)
```

	1	2	3	4
displacement	0.9404	-0.3091	0.1050	0.0947
mpg	-0.8569	-0.1213	0.1741	0.4697
gear_ratio	-0.7945	0.3511	0.4474	-0.2129
turn	0.9142	0.3286	-0.0345	0.2345

◁

▷ Example 4: Rotated canonical loadings

Here we observe the results of rotation of the canonical loadings, via the Kaiser varimax method outlined in Cliff and Krus (1976). This observation is often done for interpretation of the results; however, rotation destroys several fundamental properties of canonical correlation.

```
. quietly canon (length weight headroom trunk) (displ mpg gear_ratio turn)
. estat rotate, loadings

    Criterion          varimax
    Rotation class     orthogonal
    Normalization      none

Rotated canonical loadings
```

	1	2	3	4
length	0.3796	0.7603	0.4579	0.2613
weight	0.6540	0.5991	0.3764	0.2677
headroom	0.0390	0.1442	0.3225	0.9347
trunk	0.1787	0.2052	0.8918	0.3614
displacement	0.7638	0.4424	0.2049	0.4230
mpg	-0.3543	-0.4244	-0.8109	-0.1918
gear_ratio	-0.9156	-0.3060	-0.2292	0.1248
turn	0.3966	0.8846	0.2310	0.0832

```
Rotation matrix
```

	1	2	3	4
1	0.5960	0.6359	0.3948	0.2908
2	-0.6821	0.6593	0.1663	-0.2692
3	-0.3213	0.1113	-0.3400	0.8768
4	0.2761	0.3856	-0.8372	-0.2724

```
. estat rotatecompare
Rotated canonical loadings — orthogonal varimax
```

	1	2	3	4
length	0.3796	0.7603	0.4579	0.2613
weight	0.6540	0.5991	0.3764	0.2677
headroom	0.0390	0.1442	0.3225	0.9347
trunk	0.1787	0.2052	0.8918	0.3614
displacement	0.7638	0.4424	0.2049	0.4230
mpg	-0.3543	-0.4244	-0.8109	-0.1918
gear_ratio	-0.9156	-0.3060	-0.2292	0.1248
turn	0.3966	0.8846	0.2310	0.0832

```
Unrotated canonical loadings
```

	1	2	3	4
length	0.9664	0.2481	0.0361	-0.0566
weight	0.9972	-0.0606	-0.0367	0.0235
headroom	0.5140	-0.1295	0.7134	-0.4583
trunk	0.6941	0.0644	-0.0209	-0.7167
displacement	0.9404	-0.3091	0.1050	0.0947
mpg	-0.8569	-0.1213	0.1741	0.4697
gear_ratio	-0.7945	0.3511	0.4474	-0.2129
turn	0.9142	0.3286	-0.0345	0.2345

◁

❑ Technical note

`estat loadings` reports the canonical loadings or correlations between a *varlist* and its corresponding canonical variates. It is widely claimed that the loadings provide a more valid interpretation of the canonical variates. Rencher (1988; 1992; 1998, sec. 8.6.3) and Rencher and Christensen (2012, 397) has shown that a weighted sum of the correlations between an $x_j \in varlist_1$ and the canonical variates from $varlist_1$ is equal to the squared multiple correlation between x_j and the variables in $varlist_2$. The correlations do not give new information on the importance of a given variable in the context of the others. Rencher and Christensen (2012, 397) notes, "The researcher who uses these correlations for interpretation is unknowingly reducing the multivariate setting to a univariate one."

❑

Stored results

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

Matrices

`r(corr_var1)`	correlations for $varlist_1$
`r(corr_var2)`	correlations for $varlist_2$
`r(corr_mixed)`	correlations between $varlist_1$ and $varlist_2$

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

Matrices

`r(canload11)`	canonical loadings for $varlist_1$
`r(canload22)`	canonical loadings for $varlist_2$
`r(canload21)`	correlations between $varlist_2$ and the canonical variates for $varlist_1$
`r(canload12)`	correlations between $varlist_1$ and the canonical variates for $varlist_2$

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

Macros

`r(coefficients)`	coefficients rotated
`r(class)`	rotation classification
`r(criterion)`	rotation criterion

Matrices

`r(AT)`	rotated coefficient matrix
`r(T)`	rotation matrix

Methods and formulas

Cliff and Krus (1976) state that they use the Kaiser varimax method with normalization for rotation. The loading matrix, the correlation matrix between the original variables and their canonical variates, is already normalized. Consequently, normalization is not required, nor is it offered as an option.

Rotation after canonical correlation is a subject fraught with controversy. Although some researchers wish to rotate coefficients and loadings for greater interpretability, and Cliff and Krus (1976) have shown that some properties of canonical correlations are preserved by orthogonal rotation, rotation does destroy some of the fundamental properties of canonical correlation. Rencher (1992), Rencher and Christensen (2012), and Thompson (1996) contribute on the topic. Rencher speaks starkly against rotation. Thompson explains why rotation is desired as well as why it is at odds with the principles of canonical correlation analysis.

The researcher is encouraged to consider carefully his or her goals in canonical correlation analysis and these references when evaluating whether rotation is an appropriate tool to use.

Harris (2001) gives an amusing critique on the misuse of canonical loadings in the interpretation of canonical correlation analysis results. As mentioned, Rencher (1988; 1992; 1998, sec. 8.6.3) and Rencher and Christensen (2012, 397) critique the use of canonical loadings.

References

Cliff, N., and D. J. Krus. 1976. Interpretation of canonical analysis: Rotated vs. unrotated solutions. *Psychometrika* 41: 35–42.

Harris, R. J. 2001. *A Primer of Multivariate Statistics*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum.

Rencher, A. C. 1988. On the use of correlations to interpret canonical functions. *Biometrika* 75: 363–365.

——. 1992. Interpretation of canonical discriminant functions, canonical variates, and principal components. *American Statistician* 46: 217–225.

——. 1998. *Multivariate Statistical Inference and Applications*. New York: Wiley.

Rencher, A. C., and W. F. Christensen. 2012. *Methods of Multivariate Analysis*. 3rd ed. Hoboken, NJ: Wiley.

Thompson, B. 1996. *Canonical Correlation Analysis: Uses and Interpretation*. Thousand Oaks, CA: Sage.

Also see

[MV] **canon** — Canonical correlations

[MV] **rotatemat** — Orthogonal and oblique rotations of a Stata matrix

[MV] **screeplot** — Scree plot