

cvplot — Plot cross-validation function after lasso

[Description](#)
[Options](#)
[Quick start](#)
[Remarks and examples](#)
[Menu](#)
[Also see](#)
[Syntax](#)

Description

`cvplot` graphs the cross-validation (CV) function after a lasso fit using `selection(cv)`, `selection(adaptive)`, `selection(bic)`, or `selection(none)`.

`cvplot` can be used after `lasso`, `elasticnet`, `sqrtlasso`, `telasso`, or any of the lasso inference commands.

Quick start

Graph the CV function after `lasso`, `sqrtlasso`, or `elasticnet`

```
cvplot
```

Same as above, and draw a reference line identifying the value selected by the one-standard-error rule

```
cvplot, seline
```

Graph the CV function after `elasticnet` for the $\alpha = 0.5$ lasso

```
cvplot, alpha(.5)
```

After any of the `ds` or `po` commands, graph the CV function for the dependent variable `y`

```
cvplot, for(y)
```

Same as above, and show standard error bands for the CV function

```
cvplot, for(y) se
```

After an `xpo` command without `resample`, graph the CV function for `x` in cross-fit fold 2

```
cvplot, for(x) xfold(2)
```

After an `xpo` command with `resample`, graph the CV function for `x` in cross-fit fold 2 for the first resample

```
cvplot, for(x) xfold(2) resample(1)
```

Same as above, but graph the CV function as a function of the ℓ_1 -norm of the standardized coefficient vector

```
cvplot, for(x) xfold(2) resample(1) xunits(l1norm)
```

After `telasso`, graph the CV function for the outcome variable `x` at treatment level 1

```
cvplot, for(y) tlevel(1)
```

Menu

Statistics > Postestimation

Syntax

After `lasso`, `sqrtlasso`, and `elasticnet`

```
cvplot [ , options ]
```

After `ds` and `po` commands

```
cvplot, for(varspec) [ options ]
```

After `xpo` commands without `resample`

```
cvplot, for(varspec) xfold(#) [ options ]
```

After `xpo` commands with `resample`

```
cvplot, for(varspec) xfold(#) resample(#) [ options ]
```

After `tlasso` for the outcome variable

```
cvplot, for(varspec) tlevel(#) [ options ]
```

After `tlasso` for the treatment variable

```
cvplot, for(varspec) [ options ]
```

After `tlasso` for the outcome variable with cross-fitting but without `resample`

```
cvplot, for(varspec) tlevel(#) xfold(#) [ options ]
```

After `tlasso` for the treatment variable with cross-fitting but without `resample`

```
cvplot, for(varspec) xfold(#) [ options ]
```

After `tlasso` for the outcome variable with cross-fitting and `resample`

```
cvplot, for(varspec) tlevel(#) xfold(#) resample(#) [ options ]
```

After `tlasso` for the treatment variable with cross-fitting and `resample`

```
cvplot, for(varspec) xfold(#) resample(#) [ options ]
```

varspec is *varname*, except after `poivregress` and `xpoivregress`, when it is either *varname* or `pred(varname)`.

<i>options</i>	Description
Main	
<code>xunits(<i>x_unit_spec</i>)</code>	<i>x</i> -axis units (scale); default is <code>xunits(rlnlambda)</code> , where <code>rlnlambda</code> denotes λ on a reverse logarithmic scale
<code>minmax</code>	add labels for the minimum and maximum <i>x</i> -axis units
<code>*for(<i>varspec</i>)</code>	lasso for <i>varspec</i> ; <code>telasso</code> , <code>ds</code> , <code>po</code> , and <code>xpo</code> commands only
<code>*xfold(#)</code>	lasso for the <i>#</i> th cross-fit fold; <code>xpo</code> commands and <code>telasso</code> with <code>xfolds</code> only
<code>*resample(#)</code>	lasso for the <i>#</i> th resample; <code>xpo</code> commands and <code>telasso</code> with <code>resample</code> only
<code>*tlevel(#)</code>	lasso for the outcome model with the treatment level <i>#</i> ; <code>telasso</code> only
<code>alpha(#)</code>	graph CV function for $\alpha = \#$; default is the selected value α^* ; allowed after <code>elasticnet</code> only
<code>lineopts(<i>cline_options</i>)</code>	affect rendition of the plotted lines
S.E. plot	
<code>se</code>	show standard error bands for the CV function
<code>seopts(<i>rcap_options</i>)</code>	affect rendition of the standard error bands
Reference lines	
<code>cvlineopts(<i>cline_options</i>)</code>	affect rendition of reference line identifying the minimum of the CV function or other stopping rule
<code>nocvline</code>	suppress reference line identifying the minimum of the CV function or other stopping rule
<code>lslineopts(<i>cline_options</i>)</code>	affect rendition of reference line identifying the value selected using <code>lassoselect</code>
<code>nolsline</code>	suppress reference line identifying the value selected using <code>lassoselect</code>
<code>selineopts(<i>cline_options</i>)</code>	affect rendition of reference line identifying the value selected by the one-standard-error rule
<code>[no]seline</code>	draw or suppress reference line identifying the value selected by the one-standard-error rule; shown by default for <code>selection(cv, serule)</code>
<code>hrefline</code>	add horizontal reference lines that intersect the vertical reference lines
<code>rlabelopts(<i>r_label_opts</i>)</code>	change look of labels for reference line
Data	
<code>data(<i>filename</i> [, <i>replace</i>])</code>	save plot data to <i>filename</i>
Y axis, X axis, Titles, Legend, Overall	<i>twoway_options</i> any options other than <code>by()</code> documented in [G-3] <i>twoway_options</i>

`*for(varspec)` is required for all `ds`, `po`, and `xpo` commands and for `telasso`.

`xfold(#)` is required for all `xpo` commands and for `telasso` when the option `xfolds(#)` was specified.

`resample(#)` is required for `xpo` and for `telasso` when the option `resample(#)` was specified.

`tlevel(#)` is required for the outcome model in `telasso`.

<i>x_unit_spec</i>	Description
<code>rlnlambda</code>	λ on a reverse logarithmic scale; the default
<code>lnlambda</code>	λ on a logarithmic scale
<code>l1norm</code>	ℓ_1 -norm of standardized coefficient vector
<code>l1normraw</code>	ℓ_1 -norm of unstandardized coefficient vector

<i>r_label_opts</i>	Description
<code>labgap(<i>size</i>)</code>	margin between tick and label
<code>labstyle(<i>textstyle</i>)</code>	overall style of label
<code>labsize(<i>textsizestyle</i>)</code>	size of label
<code>labcolor(<i>colorstyle</i>)</code>	color and opacity of label

Options

Main

`xunits(x_unit_spec)` specifies the x -axis units used for graphing the CV function. The following *x_unit_specs* are available:

`rlnlambda` specifies x -axis units λ on a reverse logarithmic scale. This is the default.

`lnlambda` specifies x -axis units λ on a logarithmic scale.

`l1norm` specifies x -axis units ℓ_1 -norm of the standardized coefficient vector.

`l1normraw` specifies x -axis units ℓ_1 -norm of the unstandardized coefficient vector.

`minmax` adds labels for the minimum and maximum x -axis units to the graph of the CV function.

`for(varspec)` specifies a particular lasso after `telasso` or after a `ds`, `po`, or `xpo` estimation command fit using the option `selection(cv)`, `selection(adaptive)`, or `selection(bic)`. For all commands except `poivregress` and `xpoivregress`, *varspec* is always *varname*.

For the `ds`, `po`, and `xpo` commands except `poivregress` and `xpoivregress`, *varspec* is either *depvar*, the dependent variable, or one of *varsofinterest* for which inference is done.

For `poivregress` and `xpoivregress`, *varspec* is either *varname* or `pred(varname)`. The lasso for *depvar* is specified with its *varname*. Each of the endogenous variables have two lassos, specified by *varname* and `pred(varname)`. The exogenous variables of interest each have only one lasso, and it is specified by `pred(varname)`.

For `telasso`, *varspec* is either the outcome variable or the treatment variable.

This option is required after `telasso` and after the `ds`, `po`, and `xpo` commands.

`xfold(#)` specifies a particular lasso after an `xpo` estimation command or after `telasso` when the option `xfolds(#)` was specified. For each variable to be fit with a lasso, K lassos are done, one for each cross-fit fold, where K is the number of folds. This option specifies which fold, where $\# = 1, 2, \dots, K$. `xfold(#)` is required after an `xpo` command and after `telasso` when the option `xfolds(#)` was specified.

`resample(#)` specifies a particular lasso after an `xpo` estimation command or after `telasso` fit using the option `resample(#)`. For each variable to be fit with a lasso, $R \times K$ lassos are done, where R is the number of resamples and K is the number of cross-fitting folds. This option specifies which resample, where $\# = 1, 2, \dots, R$. `resample(#)`, along with `xfold(#)`, is required after an `xpo` command and after `telasso` with resampling.

`tlevel(#)` specifies the lasso for the outcome variable at the specified treatment level after `telasso`.

This option is required to refer to the outcome model after `telasso`.

`alpha(#)` graphs the CV function for $\alpha = \#$. The default is `alpha(α^*)`, where α^* is the selected α . `alpha(#)` may only be specified after `elasticnet`.

`lineopts(cline_options)` affects the rendition of the plotted line. See [G-3] [cline_options](#).

S.E. plot

`se` shows standard error bands for the CV function.

`seopts(rcap_options)` affects the rendition of the standard error bands. See [G-3] [rcap_options](#).

Reference lines

`cvlineopts(cline_options)` affects the rendition of the reference line identifying the minimum CV value, the value selected when the stopping tolerance is reached, or the grid-minimum value. See [G-3] [cline_options](#).

`nocvline` suppresses the reference line identifying the minimum CV value, the value selected when the stopping tolerance is reached, or the grid-minimum value.

`lslineopts(cline_options)` affects the rendition of the reference line identifying the value selected using `lassoselect`. See [G-3] [cline_options](#).

`noslline` suppresses the reference line identifying the value selected using `lassoselect`.

`selineopts(cline_options)` affects the rendition of the reference line identifying the value selected by the one-standard-error rule. See [G-3] [cline_options](#).

`[no]seline` draws or suppresses a reference line identifying the value selected by the one-standard-error rule. By default, the line is shown when `selection(cv, serule)` was the selection method for the lasso. For other selection methods, the line is not shown by default.

`hrefline` adds horizontal reference lines that intersect the vertical reference lines.

`rlabelopts(r_label_opts)` changes the look of labels for the reference line. The label options `labgap(relativesize)`, `labstyle(textstyle)`, `labsize(textsizestyle)`, and `labcolor(colorstyle)` specify details about how the labels are presented. See [G-4] [size](#), [G-4] [textstyle](#), [G-4] [textsizestyle](#), and [G-4] [colorstyle](#).

Data

`data(filename [, replace])` saves the plot data to a Stata data file.

Y axis, X axis, Titles, Legend, Overall

`twoway_options` are any of the options documented in [G-3] [twoway_options](#), excluding `by()`. These include options for titling the graph (see [G-3] [title_options](#)) and options for saving the graph to disk (see [G-3] [saving_option](#)).

Remarks and examples

stata.com

CV plots graph the CV function over the search grid for the lasso penalty parameter λ . For linear models, the CV function is the mean squared error of the predictions in the CV samples. For logit and Poisson models, the CV function is the mean deviance in the CV samples.

The search grid can be shown as the log of the lasso penalty parameter λ , `xunits(lnlambda)`; the reverse of that scale, `xunits(rlnlambda)`; the ℓ_1 -norm of the standardized coefficients, `xunits(l1norm)`; or the ℓ_1 -norm of the unstandardized coefficients, `xunits(l1normraw)`. The reverse log of lambda is the default because it represents the CV search path over λ , with the first λ tried on the left and the last λ tried on the right.

CV plots can be drawn after any command that directly searches over a grid of λ 's—that is, after any command that used the option `selection(cv)`, `selection(adaptive)`, or `selection(none)`. They can be drawn after commands `lasso`, `elasticnet`, `sqrtlasso`, `telasso`, or any of the 11 lasso inference commands.

Examples that demonstrate how to use `cvplot` after the `lasso` command can be found in *The CV function* in [LASSO] `lasso`.

Examples after `elasticnet` can be found starting in *example 2* of [LASSO] `elasticnet`.

Also see

[LASSO] `lasso postestimation` — Postestimation tools for lasso for prediction

[LASSO] `lasso inference postestimation` — Postestimation tools for lasso inferential models

[CAUSAL] `telasso postestimation` — Postestimation tools for `telasso`

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2023 StataCorp LLC, College Station, TX, USA. All rights reserved.

