Description

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

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

Quick start

Graph the CV function after lasso, sqrtlasso, or elasticnet

cvplot

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)

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)

As above, but graph the CV function as a function of the \( \ell_1 \)-norm of the standardized coefficient vector

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

Menu

Statistics \( \rightarrow \) Postestimation
Syntax

After lasso, sqrtlasso, and elasticnet

\[ \text{cvplot} \left[ , \text{options} \right] \]

After ds and po commands

\[ \text{cvplot, for} (\text{varspec}) \left[ \text{options} \right] \]

After xpo commands without resample

\[ \text{cvplot, for} (\text{varspec}) \text{ xfold(\#)} \left[ \text{options} \right] \]

After xpo commands with resample

\[ \text{cvplot, for} (\text{varspec}) \text{ xfold(\#)} \text{ resample(\#)} \left[ \text{options} \right] \]

\texttt{varspec} is a \texttt{varname}, except after poivregress and xpoivregress, when it is either a \texttt{varname} or \texttt{pred(varname)}. 
### Options

<table>
<thead>
<tr>
<th>Description</th>
<th>&lt;br&gt;options</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$-axis units (scale); default is $xunits(rlnlambda)$, where $rlnlambda$ denotes $\lambda$ on a reverse logarithmic scale</td>
<td>$xunits(x_unit_spec)$</td>
</tr>
<tr>
<td>add labels for the minimum and maximum $x$-axis units</td>
<td>minmax</td>
</tr>
<tr>
<td>lasso for $varspec$; $ds$, $po$, and $xpo$ commands only</td>
<td>$\ast for(varspec)$</td>
</tr>
<tr>
<td>lasso for the $#$th cross-fit fold; $xpo$ commands only</td>
<td>$\ast xfold(#)$</td>
</tr>
<tr>
<td>lasso for the $#$th resample; $xpo$ commands with resample only</td>
<td>$\ast resample(#)$</td>
</tr>
<tr>
<td>graph CV function for $\alpha = #$; default is the selected value $\alpha^*$; allowed after elasticnet only</td>
<td>$alpha(#)$</td>
</tr>
<tr>
<td>affect rendition of the plotted lines</td>
<td>lineopts($cline_options$)</td>
</tr>
<tr>
<td>show standard-error bands for the CV function</td>
<td>se</td>
</tr>
<tr>
<td>affect rendition of the standard-error bands</td>
<td>seopts($rcap_options$)</td>
</tr>
<tr>
<td>affect rendition of reference line identifying the minimum of the CV function or other stopping rule</td>
<td>cvlineopts($cline_options$)</td>
</tr>
<tr>
<td>suppress reference line identifying the minimum of the CV function or other stopping rule</td>
<td>nocvline</td>
</tr>
<tr>
<td>affect rendition of reference line identifying the value selected using lassoselect</td>
<td>lslineopts($cline_options$)</td>
</tr>
<tr>
<td>suppress reference line identifying the value selected using lassoselect</td>
<td>nolsline</td>
</tr>
<tr>
<td>affect rendition of reference line identifying the value selected by the one-standard-error rule</td>
<td>selineopts($cline_options$)</td>
</tr>
<tr>
<td>draw or suppress reference line identifying the value selected by the one-standard-error rule; shown by default for selection(cv, serule)</td>
<td>[no]seline</td>
</tr>
<tr>
<td>add horizontal reference lines that intersect the vertical reference lines</td>
<td>hrefline</td>
</tr>
<tr>
<td>change look of labels for reference line</td>
<td>rlabelopts($r_label_opts$)</td>
</tr>
<tr>
<td>save plot data to $filename$</td>
<td>data($filename$, replace)</td>
</tr>
</tbody>
</table>

*for($varspec$) is required for all $ds$, $po$, and $xpo$ commands. $xfold(#)$ is required for all $xpo$ commands. resample(#) is required for $xpo$ when the option resample(#) was specified.

### $x\_unit\_spec$ Description

| $rlnlambda$ | $\lambda$ on a reverse logarithmic scale; the default |
| $lnlambda$ | $\lambda$ on a logarithmic scale |
| $l1norm$ | $\ell_1$-norm of standardized coefficient vector |
| $l1normraw$ | $\ell_1$-norm of unstandardized coefficient vector |
cvplot — Plot cross-validation function after lasso

### r_label_opts

<table>
<thead>
<tr>
<th>Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>labgap(size)</td>
<td>margin between tick and label</td>
</tr>
<tr>
<td>labstyle(textstyle)</td>
<td>overall style of label</td>
</tr>
<tr>
<td>labsize(textsize)</td>
<td>size of label</td>
</tr>
<tr>
<td>labcolor(colorstyle)</td>
<td>color and opacity of label</td>
</tr>
</tbody>
</table>

### Options

- **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 $\lambda$ on a reverse logarithmic scale. This is the default.
  - lnlambda specifies $x$-axis units $\lambda$ on a logarithmic scale.
  - l1norm specifies $x$-axis units $\ell_1$-norm of the standardized coefficient vector.
  - l1normraw specifies $x$-axis units $\ell_1$-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 a ds, a po, or an xpo estimation command fit using the option selection(cv) or selection(adaptive). For all commands except poivregress and xpoivregress, varspec is always a varname; it 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).
  This option is required after ds, po, and xpo commands.
- **xfold(#)** specifies a particular lasso after an xpo estimation command. 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 $# = 1, 2, \ldots, K$. It is required after an xpo command.
- **resample(#)** specifies a particular lasso after an xpo estimation command 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 $# = 1, 2, \ldots, R$. This option, along with xfold(#), is required after an xpo command with resampling.
- **alpha(#)** graphs the CV function for $\alpha = #$. The default is alpha($\alpha^*$), where $\alpha^*$ is the selected $\alpha$. 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.

cvl ine 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.

nolsline 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.

[noseline] 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.


Data

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

Remarks and examples

CV plots graph the CV function over the search grid for the lasso penalty parameter \(\lambda\). 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 either the log of the lasso penalty parameter \(\lambda\), xunits(lnlambda); the reverse of that scale, xunits(rlnlambda); the \(\ell_1\)-norm of the standardized coefficients, xunits(l1norm); or the \(\ell_1\)-norm of the unstandardized coefficients, xunits(l1normraw). The reverse log of lambda is the default because it represents the CV search path over \(\lambda\), with the first \(\lambda\) tried on the left and the last \(\lambda\) tried on the right.
CV plots can be drawn after any command that directly searches over a grid of $\lambda$'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, 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