Lasso
Variable selection, prediction, inference

Cross-validation plot
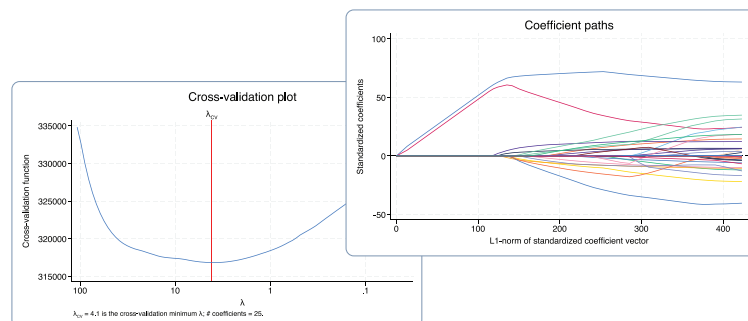
Coefficient paths

- ▪ All the tools you expect for lasso machine learning
  - ▪ Lasso, square-root lasso, and elastic net
  - ▪ Cross-validation
  - ▪ Adaptive lasso
  - ▪ Knot analysis
  - ▪ Coefficient paths
  - ▪ Adjustment for clustered data
- ▪ Alongside cutting-edge inferential methods
  - ▪ Robust to mistakes in variable selection
  - ▪ Proper inference for coefficients of interest
  - ▪ Double selection
  - ▪ Partialing out
  - ▪ Cross-fit partialing out
  - ▪ Double machine learning
  - ▪ Treatment-effects estimation
  - ▪ Inference for clustered data

Select predictors for continuous, binary, count, and survival-time outcomes

Lasso with selection via cross-validation

```
. lasso linear   y x1-x1000
. lasso logit    y x1-x1000
. lasso probit   y x1-x1000
. lasso poisson  y x1-x1000
. lasso cox        x1-x1000
```

Adaptive lasso

```
. lasso linear y x1-x1000, selection(adaptive)
```

Selection via BIC

```
. lasso linear y x1-x1000, selection(bic)
```

Selection via plugin method

```
. lasso linear y x1-x1000, selection(plugin)
```

Elastic net with selection via cross-validation

```
. elasticnet linear   y x1-x1000
. elasticnet logit    y x1-x1000
. elasticnet probit   y x1-x1000
. elasticnet poisson  y x1-x1000
. elasticnet cox        x1-x1000
```

Square-root lasso

```
. sqrtlasso y x1-x1000
```

Examine the results

View selected variables

```
. lassoknots
. lassoinfo
. lassocoef
```

Plot cross-validation function

```
. cvplot
```

Plot coefficient path

```
. coefpath
```

Obtain predictions

```
. use newdata
. predict yhat
```

Evaluate fit

```
. lassogof
```

Viewer - view lasso1.smcl

view lasso1.smcl

Dialog ▾   Also see ▾   Jump to ▾

```
. lasso linear y x1-x10

Lasso linear model                      No. of obs        =        69
                                        No. of covariates =        10
Selection: Cross-validation             No. of CV folds   =        10
```

ID	Description	lambda	No. of nonzero coef.	Out-of-sample R-squared	CV mean prediction error
1	first lambda	4.69114	0	-0.0090	34.22157
17	lambda before	1.0588	3	0.5641	14.78393
* 18	selected lambda	.9647388	3	0.5648	14.76141
19	lambda after	.8790341	3	0.5645	14.77163
22	last lambda	.664957	5	0.5613	14.87944

```
* lambda selected by cross-validation.

. estimates store cv

. lassocoef cv adaptive
```

	cv	adaptive
x5	x	x
x6	x	x
x7	x	
_cons	x	x

Lasso for inference

With lasso inferential methods, you can estimate coefficients, standard errors, test statistics, and confidence intervals for variables of interest while using lassos to select from a potentially large number of control variables.

Double-selection method; estimate coefficients for **x1** and categorical **x2**; selection of controls via plugin

```
. dsregress y x1 i.x2, controls(c1-c1000)
```

Logit model for binary outcome; estimate odds ratios for **x1** and **x2**

```
. dslogit y x1 i.x2, controls(c1-c1000)
```

Poisson model for count outcome; estimate incidence-rate ratios for **x1** and **x2**

```
. dspoisson y x1 i.x2, controls(c1-c1000)
```

Selection of controls via cross-validation

```
. dsregress y x1 i.x2, controls(c1-c1000)
           selection(cv)
```

Partialing-out method

```
. poregress y x1 i.x2, controls(c1-c1000)
```

Cross-fit partialing-out method (double machine learning)

```
. xporegress y x1 i.x2, controls(c1-c1000)
```

Treatment-effects estimation; estimate the ATE of **treat**, controlling for **x1-x1000** in the outcome model and **w1-w1000** in the treatment model

```
. telasso (y x1-x1000) (treat w1-w1000)
```

```
Viewer - view lasso2.smcl                                    —    □    ×
view lasso2.smcl   ×
+                                              Dialog ▾  Also see ▾  Jump to ▾
. dsregress y x1 i.x2, controls(c1-c8)

Double-selection linear model        Number of obs             =          69
                                     Number of controls        =           8
                                     Number of selected controls =         6
                                     Wald chi2(5)              =        6.24
                                     Prob > chi2               =      0.2835

                            Robust
         y   Coefficient  std. err.      z    P>|z|     [95% conf. interval]

        x1     .1272712    .256027    0.50   0.619    -.3745326    .629075

        x2
         2     .2792513   1.270518    0.22   0.826    -2.210918    2.76942
         3    -.2613078   1.358118   -0.19   0.847     -2.92317    2.400554
         4     .7492284   1.427334    0.52   0.600    -2.048295    3.546752
         5     4.082883   1.905783    2.14   0.032      .3476163    7.81815

                                                          CAP  NUM  INS
```
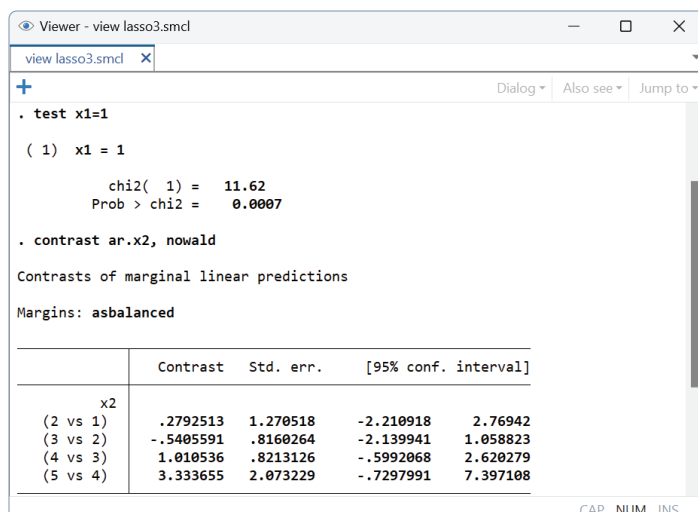
Evaluate results using Stata's standard tools

Perform tests on coefficients

```
. test x1=1
```

Estimate contrasts such as differences across levels

```
. contrast ar.x2
```

```
Viewer - view lasso3.smcl                          —    □    ×
view lasso3.smcl   ×
+                                       Dialog ▾  Also see ▾  Jump to ▾
. test x1=1

 ( 1)  x1 = 1

        chi2( 1) =    11.62
        Prob > chi2 =    0.0007

. contrast ar.x2, nowald

Contrasts of marginal linear predictions

Margins: asbalanced

                 Contrast   Std. err.    [95% conf. interval]

          x2
   (2 vs 1)      .2792513   1.270518    -2.210918    2.76942
   (3 vs 2)     -.5405591   .8160264    -2.139941    1.058823
   (4 vs 3)      1.010536   .8213126    -.5992068    2.620279
   (5 vs 4)      3.333655   2.073229    -.7297991    7.397108

                                                  CAP  NUM  INS
```
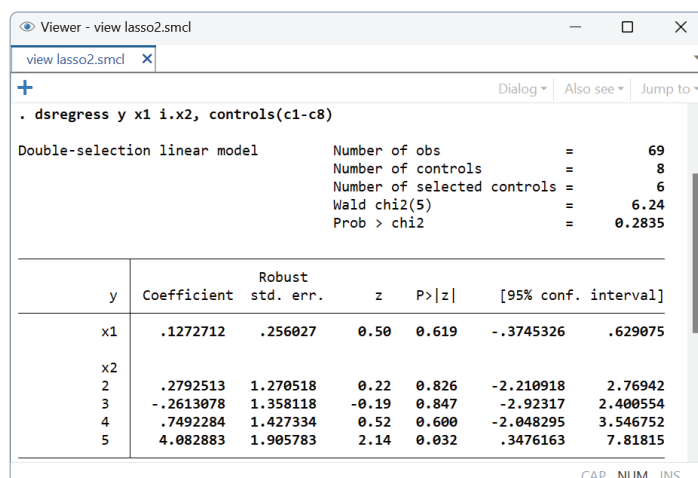
Explore underlying lassos

View the selected controls in the lasso for **y**

```
. lassocoef (.,for(y))
```

Plot coefficient paths in the lasso for **y**

```
. coefpath, for(y)
```


Coefficient paths