

RXridge.LSP: “Shrinkage/Ridge Regression” ver. 2008

for use with...

XLisp-Stat release 3+

softRX freeware, c/o Bob Obenchain

5176 Upperwood Ct, Indianapolis IN 46250-1776, USA.

internet: wizbob@att.net <http://localcontrolstatistics.org>

Introduction

RXridge.LSP adds shrinkage regression calculation and graphical ridge “trace” display functionality to the XLisp-Stat release 3+ implementation of LISP-STAT, Tierney(1990). Shrinkage/ridge methods examine possible effects of ill-conditioning (numerical and/or statistical) among predictor variables on the relative magnitudes and numerical signs of fitted coefficients in multiple linear regression models. RXridge.LSP focuses on **maximum likelihood** approaches to statistical inference under normal distribution theory, Obenchain (1975, 1978, 1984, 1995, 1996a), concerning variance-bias trade-offs that can reduce overall mean-squared-error risk via shrinkage.

Other somewhat unique/innovative features of my softRX freeware™ systems for shrinkage/ridge regression include usage of (i) the MCAL = “multicollinearity allowance” measure for **extent-of-shrinkage** as the horizontal axis on all TRACE plots and (ii) a second ridge parameter, Q, that controls the **shape (or curvature) of the shrinkage path** through likelihood space. These two parameters, Q and m, are defined and implemented as follows:

The 2-parameter family of generalized ridge estimators implemented in RXridge.LSP can be written, Goldstein and Smith(1974), as

$$\beta^* = [X'X + k \cdot (X'X)^Q]^{-1} X'y,$$

where RXridge.LSP limits the Q-shape to integer and half-integer values within [-5, +5] and the range of k goes from $k = 0$ ($\beta^* = \text{OLS}$) to $k = +\infty$ ($\beta^* = 0$.) Writing the singular value decomposition of regressors as $X = H \Lambda^{1/2} G'$, where Λ is the diagonal matrix of eigenvalues of $X'X$, these estimators are of the general form...

$$\beta^* = G \Delta \Lambda^{-1/2} H' y = G \Delta c,$$

where the columns of G are the direction cosine vectors for the **principal axes** of X , Δ is the diagonal matrix of multiplicative **shrinkage factors**, $0 \leq \delta_i \leq 1$, and c is the column vector of **uncorrelated components** of the OLS solution. This shows that the shrinkage factor applied along the i-th principal axis is of the specific form

$$\delta_i = \frac{\lambda_i}{(\lambda_i + k \cdot \lambda_i^Q)} = \frac{1}{(1 + k \cdot \lambda_i^{Q-1})}.$$

Rather than use $k = \text{"additive eigenvalue inflation factor"}$ to index various extents of shrinkage along a Q-shape path, RXridge.LSP uses

$$MCAL = p - \delta_1 - \delta_2 - \dots - \delta_p = \text{rank}(X) - \text{trace}(\Delta).$$

Note that the range of MCAL is **finite**: $0 \leq MCAL \leq p$. Whatever may be your choice of Q-shape, the OLS solution always occurs at the beginning of the shrinkage path at $MCAL=0$ ($k=0$ and $\Delta = I$) and the terminus of the shrinkage path, where the fitted regression hyperplane becomes "horizontal" (slope=0 in all p-directions of X space) and $\hat{y} = \bar{y}$, always occurs at $MCAL=p$ ($k=+\infty$ and $\Delta = 0$). RXridge.LSP uses Newtonian descent methods to compute the numerical value of k corresponding to given values of MCAL and Q-shape.

In addition to having finite (rather than infinite) range, MCAL has a large number of other advantages over k when used as the scaling for the horizontal axis of ridge TRACE displays. For example, the first figure below illustrates that shrunken regression coefficients with **stable relative magnitudes** form STRAIGHT LINES when plotted versus MCAL. Similarly, the average value of all p shrinkage factors is $(p-MCAL)/p$, which is Theil's **proportion of Bayesian posterior precision** due to sample information rather than to prior information and which decreases linearly as MCAL increases. Perhaps most importantly, MCAL can frequently be interpreted as the **approximate deficiency in the rank of X**. For example, if a regressor $X'X$ matrix has only two relatively small eigenvalues, then the coefficient ridge trace of best Q-shape typically "stabilizes" at about $MCAL=2$. I.E., the coefficient trace then consists primarily of fairly

straight lines between $\text{MCAL}=2$ and $\text{MCAL}=p$. This last point is illustrated by the second plot below for the infamous Longley $p=6$ regressor dataset.

Figure 1: Coefficients with perfectly stable relative magnitudes, path Q-shape=+1.

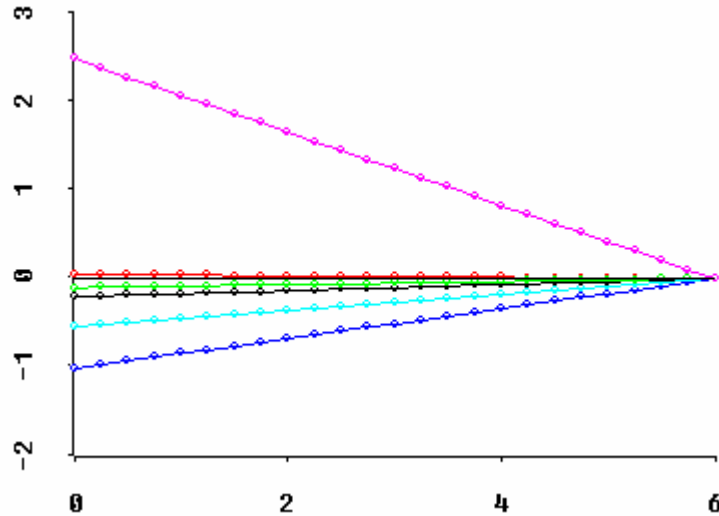
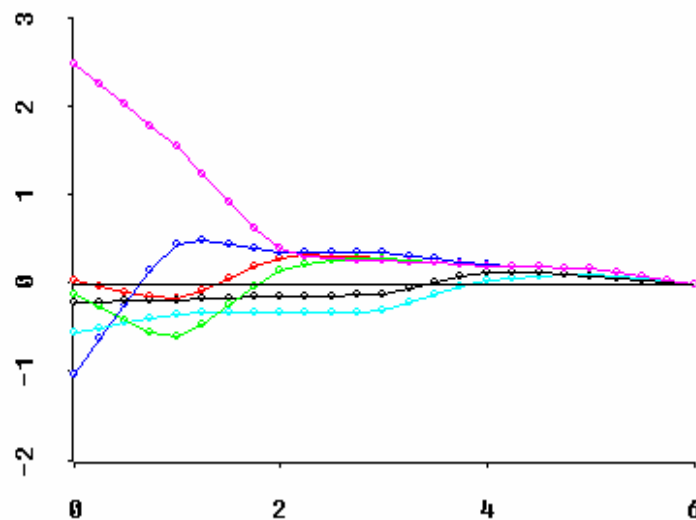


Figure 2: Coefficients that approximately stabilize at and beyond roughly $\text{MCAL} = 2$ for the Longley Data when the path has Q-shape = -1.5.



To locate the 2-parameter shrinkage/ridge estimator **most likely** to minimize overall MSE risk, Obenchain (1996a) provides both (i) a closed form expression for the optimal value of k (and thus $m = \text{MCAL}$) given Q and (ii) a simple expression [the “curlicue” function, $\text{CRL}(Q)$] that, when maximized via numerical search, identifies the shrinkage path of optimal Q -shape. `RXridge.LSP` implements these as well as a number of other

sound, objective, data analytical methods for picking an appropriate form and extent of shrinkage.

Displaying RXridge Menus

There are three somewhat different ways to invoke the RXridge.LSP code to analyze a dataset. Personally, I prefer method “C” ...described on page 7.

Possibility A:

You start by **loading RXridge.LSP** either from XLisp-Stat’s “**File**” menu item or else by entering a XLisp expression like...

Windows> (load “drive:\\path\\rxridge.lsp”)

or

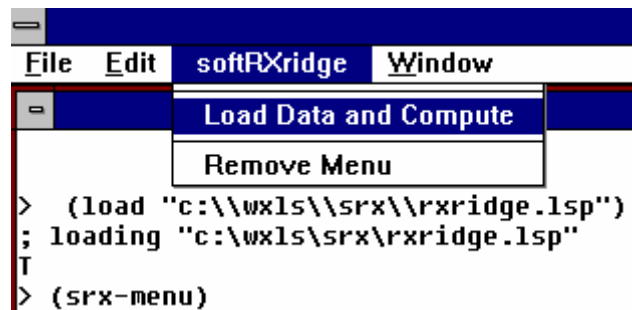
Mac> (load “path.to.rxridge.lsp”)

You will probably find loading of RXridge.LSP using XLisp-Stat’s “**File**” menu much easier than entering a (load “path-etc.”) expression whenever your copy of the **RXridge.LSP** file is not in XLisp-Stat’s main **xlslib** MS-DOS sub-directory or Mac folder.

Next, enter the Xlisp expression...

(srx-menu)

This second action adds a new main entry to the XLisp-Stat menu bar named 'softRXridge' with two menu items:



This is an easy way to use RXridge.LSP to perform shrinkage-regression on a dataset you have used before in XLisp-Stat.

Possibility B:

First, **load** RXridge.LSP as in option 'A' above. Next, invoke (RXridge-load-compute) by entering...

(RXridge-load-compute)

or

(def my-model (RXridge-load-compute))

The latter tactic saves your regression-model object in the XLisp-Stat variable MY-MODEL. This has the advantage that you can send messages to your model object such as...

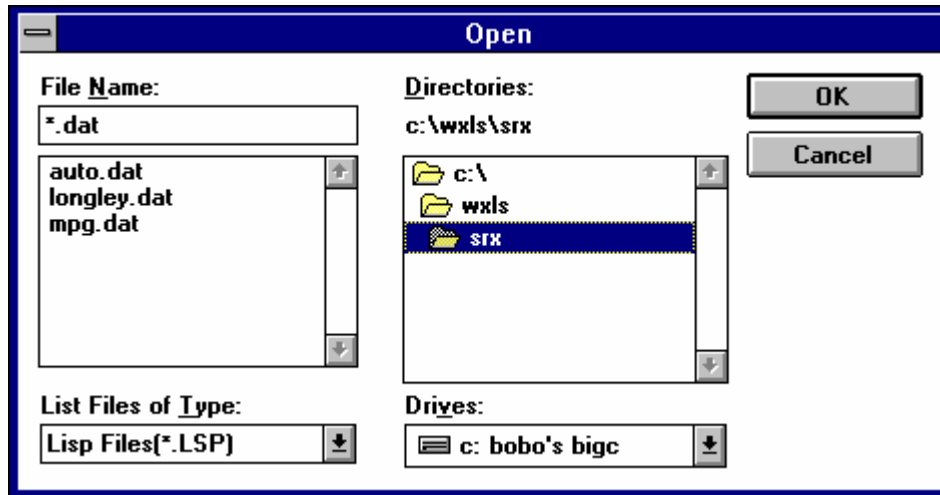
(send my-model :RXridge-qp)

(send my-model :RXridge-seed #\$(1 # (214748 920333 169369 773360)))

to display the current setting of the shrinkage path Q-shape parameter and to reset the saved random number generator seed/state, respectively. (You will probably need to copy and paste from the XLisp-Stat Listener window to reset the seed/state as shown above. Valid seeds can consist of 4 ten-digit numbers rather than just the six-digits shown in the above example.) To learn about all possible 'RXridge-' messages, you would need to examine the RXridge.LSP code file.

NOTE: The only real difference between options 'A' and 'B' is that no 'softRXridge' menu will appear in approach 'B.'
--

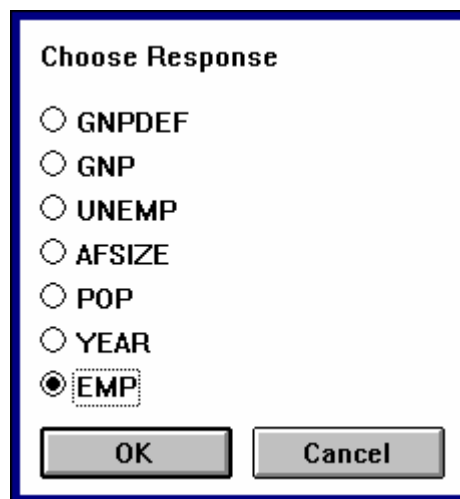
Upon invoking (RXridge-load-compute) via either option 'A' or 'B', a system **FILE locator dialog** window appears, allowing you to select a data file with a name of the form *.dat. Unfortunately, the XLisp-Stat default file mask is designed to locate only '*.lsp' files. You may, of course, edit the '*.lsp' mask ...changing it to '*.dat' ...and then press the enter key. This changes the default mask and allows you to locate all *.dat files.



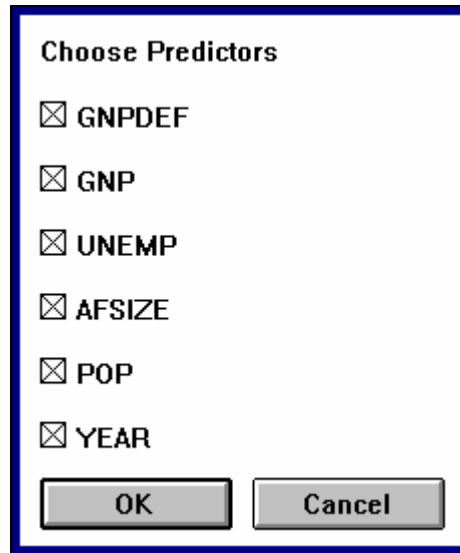
On the other hand, you could edit the '*.lsp' mask ...changing it to the full 'filename.dat' for the specific file you wish to read ...then press the enter key.

Once you select a specific *.dat file, the code searches for a file with the same initial filename, but extension '.nam' instead of '.dat'. You may list NAMES for variables (in column order) in a *.nam file. But names are optional in the sense that, when this '.nam' file cannot be found, variables will be named 'Var1', 'Var2', ..., etc.

Next, a **dialog box** appears that allows you to select the response variable by clicking on its name.



After that, another dialog box pops up with all the remaining variables pre-selected. If you want to exclude a variable from computation, deselect it here so that it will not be included as a regressor. Be certain you deselect any **character variables** that represent observation labels!



When your regression-model is first computed, XLisp-Stat's "built-in" OLS statistics are written to the Listener. Then a new entry will appear on the XLisp-Stat menu bar named by the filename for your dataset. This is the RXridge Main-Menu for use ONLY with the current dataset and regression-model object.

Possibility C:

Instead of loading RXridge.LSP directly, you can load a specially prepared "example" file, such as "mpg.lsp", that contains the (**require "RXridge"**) directive, reads in your data, names the variables, names the regression-model object, and creates an RXridge Menu for analysis of the given dataset. The "mpg.lsp" example file (contained in the softRX distribution archive) for the Hocking(1976) miles-per-gallon dataset illustrates Lisp coding details for this option, 'C.' Simply use your favorite text editor or **LispEdit** to mimic/modify the contents of 'mpg.lsp' or one of the other example files to create/save an initialization file for your own dataset.

Example: Contents of a Type 'C' *.LSP Specification File

```
(require "RXridge")
(def mpg-dat (read-data-columns "mpg.dat" 6)) ; last column contains
labels
(def cyl (select mpg-dat 0)) ; get single variables
(def cub (select mpg-dat 1)) ; out of read data
(def hp (select mpg-dat 2))
(def wght (select mpg-dat 3))
(def mpg (select mpg-dat 4))
(def mpg-reg (regression-model (list cyl cub hp wght) mpg
:predictor-names '("CYL" "CUB" "HP" "WGHT" )
```

```

      :response-name "MPG"
    )
  )
  (format t "(The RXridge regression model object is 'mpg-reg') ~%")
  (send mpg-reg :RXridge-menu "MPG")
  (send mpg-reg :RXridge-compute-summary)

```

Option 'C' is, perhaps, best for users who want to keep on file a permanent record of exactly which datafile was read-in and how variables and objects were named. This sort of user would probably also want to turn “Dribble” on (via the Windows “File” menu or Mac “Command” menu) to capture everything that RXridge.LSP writes to the Listener in a permanent text file.

Using an RXridge “Dataset Specific” Menu

An RXridge “Dataset Specific” Menu is a **main-menu entry** named after your dataset. There can be several such main-menu-items at any one time, each dealing only with its own dataset and model objects. This strategy would allow a user to visually compare results from different models ...possibly for the same data, e.g. different regressor subsets and/or x-axes (re-)scalings.

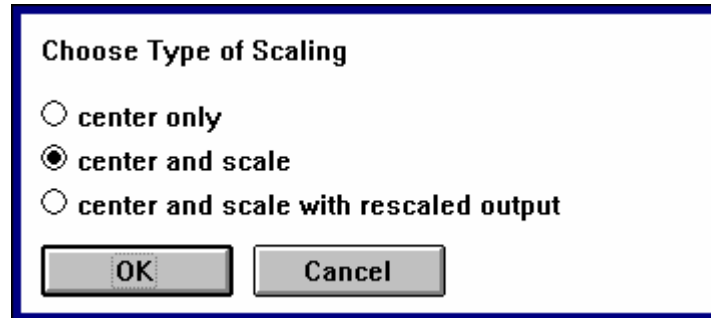
XLISP-STAT
C:\W\XLS\SRX\LONGLEY Window
Choice of Regressor-Response Scaling
Compute Principal Axis Summary
Identify Most Likely Path Q-shape Set Shrinkage Path Q-shape Set Number of STEPS per Shrinkage Unit Identify Most Likely Shrinkage M-extent
List Shrinkage Trace Details Display Shrinkage TRACE Plots
Visual Re-Regression and Influence Plot Component Size vs. Significance
Specify True Values of Parameters List True Shrinkage Details Display Expected Traces, True MSE Risks List Shrinkage Simulation Details Display Simulated Traces, True SE Losses
Remove this MENU

Note that the RXridge main-menu contains 16 items arranged into 6 groups. The natural order for invocation of items is generally from-top-to-bottom. But you can skip over any

item that is of little interest to you ...say, because that item controls a parameter setting whose default value you currently consider satisfactory.

Group One:

Choice of Regressor-Response Scaling



Choose Type of Scaling

☐ center only

☒ center and scale

☐ center and scale with rescaled output

OK Cancel

Note the default here is to not only “center” the response and regressor variables at zero (by subtracting off their sample means) but also to “scale” each variable by dividing its centered values by their sample standard deviation. The resulting sum-of-squares of centered and rescaled observations on each variable will then be (n-1), the number of observations minus one. This choice of preliminary scaling eliminates what Marquardt(1980) called “non-essential” ill-conditioning **before** computing summary statistics for the principal axis rotation of regressor variables.

Note that the last option is to perform initial centering and scaling, estimate all effects using these standardized coordinates, and then (at the very end) re-express results back in terms of the original, given X and y coordinates. Thus, although principal axis decompositions are not scale invariant, this option does lead to predictions, $\hat{y} = X\beta^*$, which are scale invariant by their very definition!

Compute Principal Axis Summary

You should generally invoke this item after any change in regressor/response scaling. But, if you don’t click here first, the corresponding code will always still execute when needed prior to computations spawned by other RXridge menu items.

You may also use this item simply to refresh your memory about summary statistics that characterize the form and extent of ill-conditioning ...i.e. rather than scrolling back through RXridge output previously written to the XLisp-Stat Listener

Example: Principal Axis Summary Statistics for the Longley Data

***** Shrinkage/Ridge Regression *****

softRX freeware (c) by Bob Obenchain for XLisp-Stat 2.1 Release 3+.
Version 9602...Vast majority of Lisp code by Bernhard Walter(1994).

Principal Axis Summary Statistics of Ill-Conditioning

(Regressors and Response are centered/scaled)

Eigenvalue	Sing.Value	Uncorr.Comp.	Princ.Corr.	t-stat
1	69.05066	8.30967	0.44565	0.95617
2	17.63011	4.19882	-0.11157	-0.12096
3	3.05138	1.74682	0.52973	0.23892
4	0.22392	0.47321	-0.10174	-0.01243
5	0.03828	0.19566	-1.75680	-0.08875
6	0.00565	0.07517	1.98275	0.03848

Adjusted Response Sum-of-Squares: 15.00000
Residual Mean Square for Error: 0.00753
Estimate for Residual Std. Error: 0.08680

Group Two:

Identify Most Likely Path Q-shape

This item uses the closed-form expressions of Obenchain(1981, 1996a) to identify the shrinkage path Q-shape (and the MCAL-extent of shrinkage along that path) which have maximum **classical** (fixed coefficient) normal-theory likelihood of achieving overall minimum MSE risk in estimation of regression coefficients.

For the centered/scaled Longley data, this best path shape turns out to be $Q = -1.5$.

The figure below illustrates the variety of path Q-shapes in the two-dimensional case, $p=2$. Well-known special cases of Q-shape paths (in 2 or more dimensions) are...

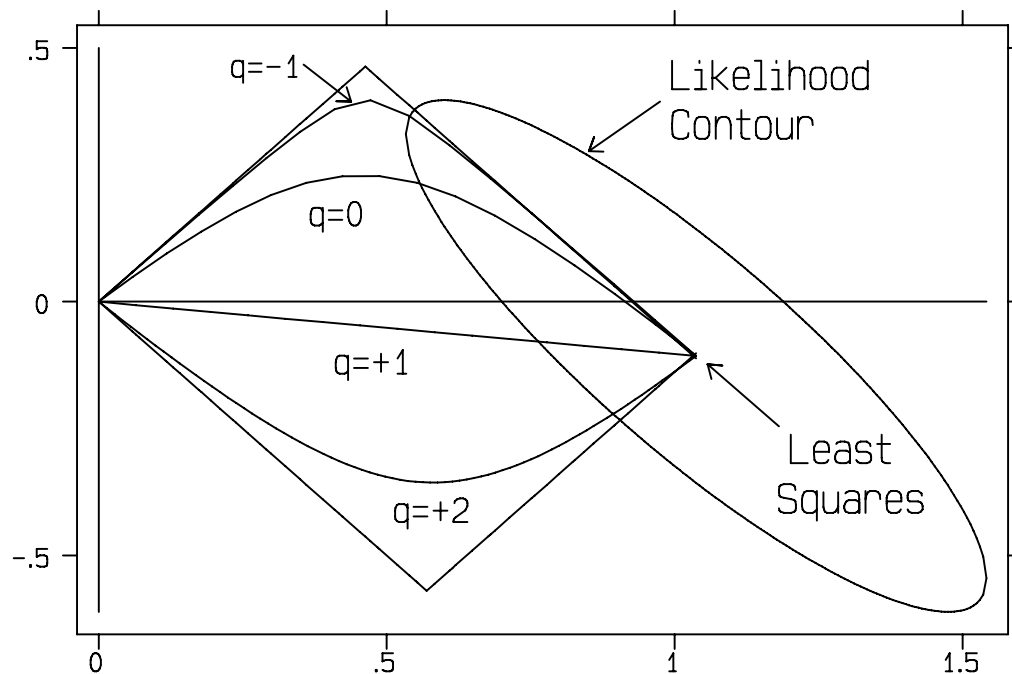
$Q = 0$ for Hoerl-Kennard(1970) “ordinary” ridge regression, and

$Q = +1$ for uniform shrinkage. (See also Figure 1.)

An extremely important limiting case is...

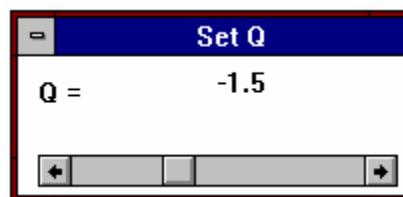
$Q = -\infty$ for principal components regression.

Marquardt(1970) calls this limit “assigned rank” regression. I have found that the $Q=-5$ path is frequently quite close to this limiting case, numerically. In the above figure, the $Q=-1$ path is already quite close to the limiting case in $p=2$ dimensions.



As a general rule-of-thumb, paths with Q-shapes within the $[-1, +2]$ range generally tend to be fairly “smooth” or have “rounded” corners. Paths with Q-shapes greater than $+2$ or less than -1 can display quite “sharp” corners. In fact, the paths with limiting shape $\pm\infty$ are actually linear splines with join points at integer MCAL values!

Set Shrinkage Path Q-shape



This menu item brings up a slider that allows you to select a shrinkage path Q-shape within $[-5, +5]$ in steps of 0.5.

SUGGESTION: Why not explore the Q-shape path most likely to be MSE optimal?

Set Number of STEPS per Shrinkage Unit

This item brings up a slider that allows you to select a value for STEPS in the [1,100] range. The default value is 4. The total number of values along the shrinkage path at which calculations will be performed will be $\text{STEPS} \times \text{Rank}(X) + 1$. Thus $\text{STEPS} = 1$ generates the fewest calculations, while larger values typically yield more highly detailed TRACE plots.

Identify Most Likely Shrinkage M-extent.

No closed-form expressions exist for the **empirical Bayes** [Efron and Morris(1976)] or **random coefficient** [Golub, Heath and Wahba(1979), Shumway(1982)] maximum likelihood approaches to shrinkage. So this menu item simply performs a search, using both the lattice of steps in MCAL and the Q-shape for the shrinkage path that you selected using other menu items, above. A minus-two-log-likelihood-ratio is also listed for the **classical** (fixed coefficient) approach, but the finite steps-per-MCAL-unit restriction in effect here usually makes these classical calculations somewhat less accurate than those from the “Most Likely Q-shape” option, above.

For the centered/scaled Longley data and path Q-shape = -1.5 , the classical extent of shrinkage most likely to minimize overall MSE risk is approximately $\text{MCAL} = +4$.

According to the “ $2/p$ -ths rule” of Obenchain(1978), the MCAL shrinkage extent for “good” (multivariate) ridge estimates, relative to ordinary-least-squares, is roughly a factor of [two divided by p] times the “optimal” (univariate) MCAL value. Since $\text{rank}(X) = p = 6$ for the Longley data, we see that a somewhat more conservative approach would limit shrinkage for the Longley data to about $\text{MCAL} = (2/6) * 4 = 1.33$.

Group Three:

List Shrinkage Trace Details

This item writes numerical matrix listings of shrinkage results to the Listener, with one row for each step in MCAL value. If these details are of interest to you, you should probably be saving them in a permanent “Dribble” file.

Display Shrinkage TRACE Plots

Use a dialog box to select which of the 5 types of traces you want to view. These trace displays are linked ...with curves numbered 1 to p = number of regressors = number of principal axes with positive variance.

Which TRACES should be plotted?

☒ MCAL vs Ridge Coefficients

☒ MCAL vs Scaled MSE Risk

☒ MCAL vs Excess Eigenvalues

☒ MCAL vs Inferior Direction

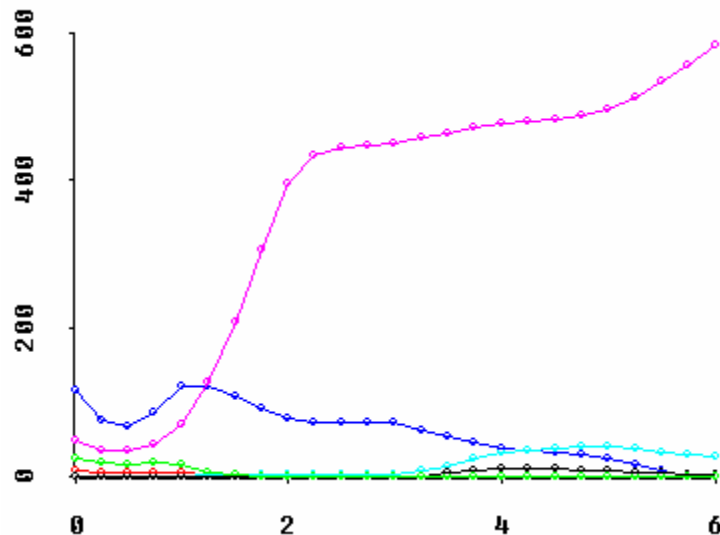
☒ MCAL vs Shrinkage Pattern

OK Cancel

bstar: fitted shrinkage coefficients

The coefficient trace shows how point estimates change as shrinkage along a Q-shape path occurs. The trace curve for any regression coefficient estimate that becomes numerically "stable" will remain fairly **straight** for the rest of its journey to $\beta_i^* = 0$ at $\text{MCAL}=p$. Unstable coefficient estimates usually tend to change more quickly than stable coefficients, possibly switching numerical sign or oscillating as shrinkage occurs. (See Figure 2, above for the Longley data.)

risk: estimated scaled (relative) MSE risk



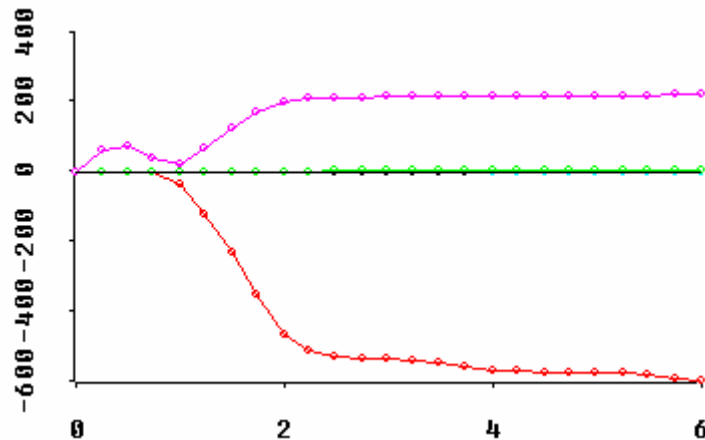
This trace gives normal distribution theory, "modified" maximum likelihood estimates, Obenchain(1978), of "scaled" MSE risk as Q-shape shrinkage occurs.

Relative risk is "scaled" by dividing a risk estimate by an estimate of the error variance. In other words, scaled risk expresses imprecision in fitted coefficients in units that are multiples of the variance of a single observation.

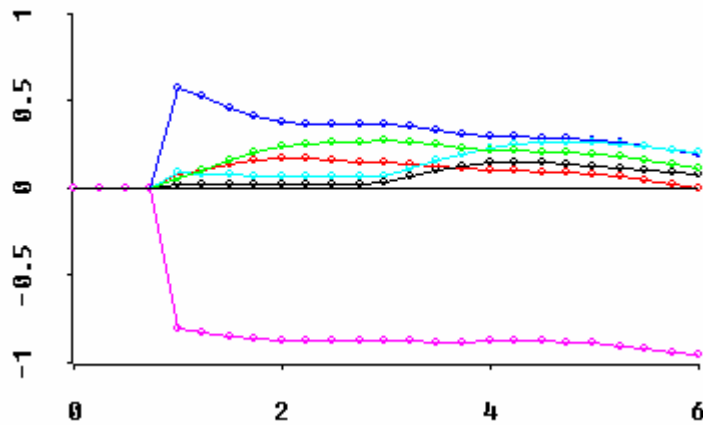
Maximum likelihood scaled risk estimates are first "modified" so as to be **unbiased**. Then they are adjusted upward, if necessary, to have **correct range** relative to a known lower bound on scaled risk, which re-introduces some bias.

exev: estimated Eigenvalues of [VAR(ols) - MSE(ridge)] matrix

This trace plots the EigenValues of the estimated difference in scaled MSE risk matrices, ordinary-least-squares minus ridge. As long as all EigenValues are zero or positive, there is reason to hope that the corresponding ridge estimators yield smaller MSE risk than Least Squares for **all directions in p-space** (i.e. all possible linear combinations.) However, as shrinkage continues, **at most one** negative EigenValue will appear, Obenchain(1978).



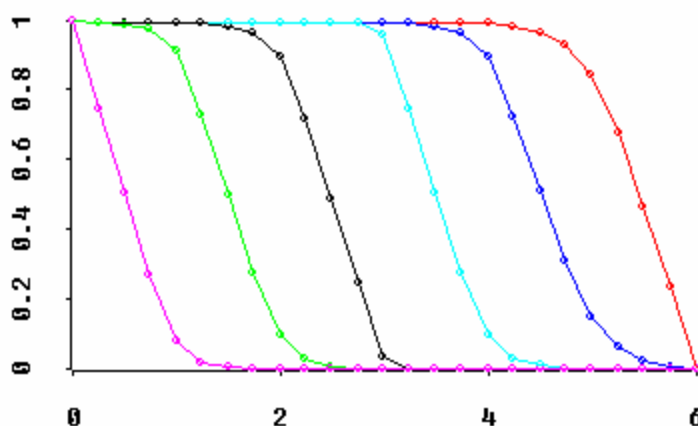
infd: estimated inferior direction cosines



This trace plots the Direction Cosines (normalized EigenVector) corresponding to the negative EigenValue (if any) of the difference in Mean-Squared-Error matrices, OLS - ridge. This direction gives that single linear combination of ridge regression coefficients that not only fails to benefit from Q-shape shrinkage but probably actually suffers increased risk due to shrinkage.

delta: shrinkage delta factors

This trace plots the Delta Shrinkage-Factor Pattern as Q-shape shrinkage occurs. All deltas are equal when $Q=1$; the trailing deltas are small when $Q < 1$; and the leading deltas are small when $Q > 1$.

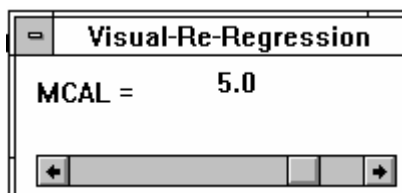


NOTE: Curve numbers on “exev” and “delta” traces refer to principal axes, not to regressors.

Group Four:

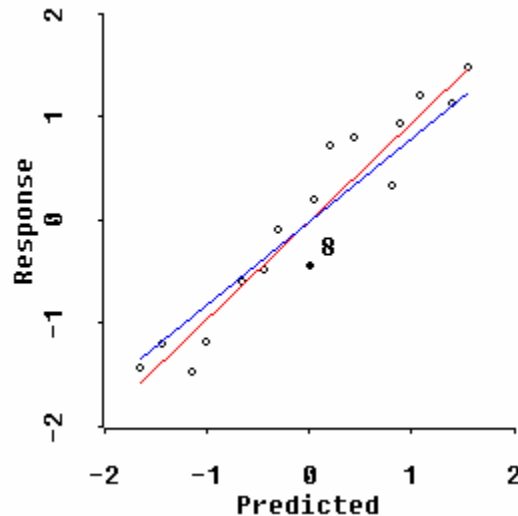
Visual Re-Regression and Influence

This item displays a slider that allows you to move through all computed MCAL steps. Two plots connected to this slider change dynamically as this MCAL = extent-of-shrinkage changes.



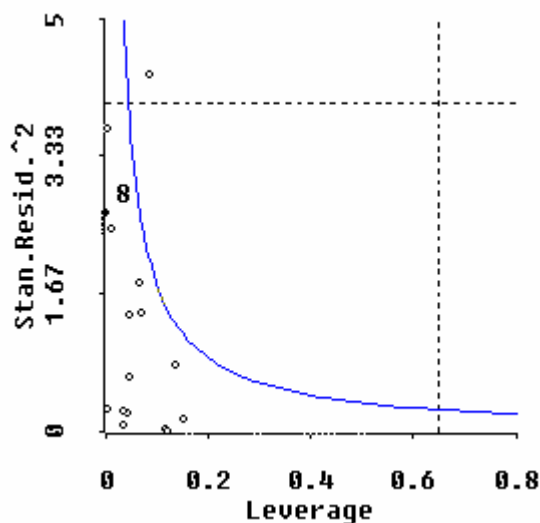
The first plot shows observed response y-values vertically vs. their standardized shrinkage fit-values (i.e. composite regressor x-coordinates) horizontally. The **BLUE**

line on the plot represents this shrinkage fit, while the **RED** “Visual-Re-Regression” line displays the regression of the response y-values onto the composite x-coordinates.

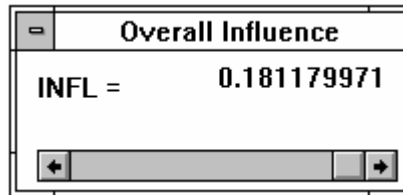


NOTE: The slope of the **BLUE** line decreases as shrinkage becomes extreme, giving the **RED** line the appearance of being a much better fit to the response y-values. Thus this display provides a clear, visual warning when one's shrinkage/ridge estimator becomes too seriously biased.

It is rather simple to spot outlying responses and high leverage regressor combinations on this response vs. composite-predictor plot. Outlying responses have large residuals, and the points with highest leverage are the ones at either (left or right) extreme along the composite x-axis! On the other hand, considerable information can be lost in displaying a multiple regression ($p > 1$) fit using only coordinates along any single ($p = 1$) composite axis. It turns out that $p = \text{one dimensional leverages}$ can be quite distorted!



The second, linked “influence” plot shows squared, standardized residuals vertically vs. regressor-combination leverages (in $p > 1$ dimensions) horizontally. A dashed horizontal line at 4 serves as a “warning line” for outlying residuals, while a dashed vertical line [Chatterjee and Hadi(1988), page 32] serves as a “warning line” for leverages.



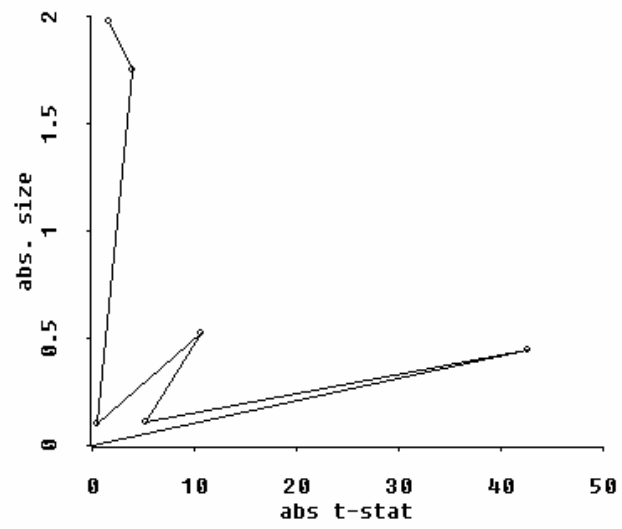
A second slider controls levels of overall Cook-influence ...defined as the product of a squared, standardized residual times its leverage. The corresponding “contours” of constant overall Cook-influence are displayed as HYPERBOLAS in **BLUE**. As ridge/shrinkage occurs, points tend to move to the left and/or upwards on this plot; this shows that ridge/shrinkage tends to reduce the leverage of every regressor combination (row of X) and to (ultimately) increase the size of residuals.

Plot Component Size vs. Significance (or SIZ-SIG plot)

This is a plot of absolute values of t-statistics for OLS coefficients versus the absolute size of these OLS coefficients. When all regressors are uncorrelated (no ill-conditioning is present) and rescaled (i.e. all sample variances are equal), then all of the points in this plot would lie on a single straight line passing through the origin. I.E. numerical size and statistical significance are synonymous in this highly desirable case!

How close are your points to a single, straight line???

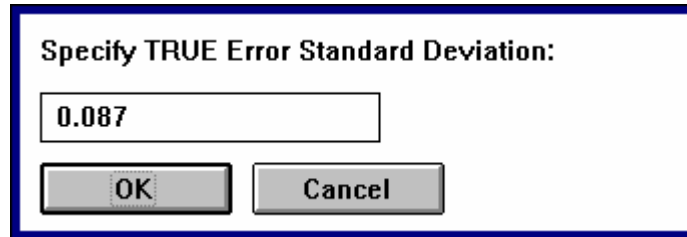
Figure 3: SIZSIG Plot for the Longley Data.



Group Five:

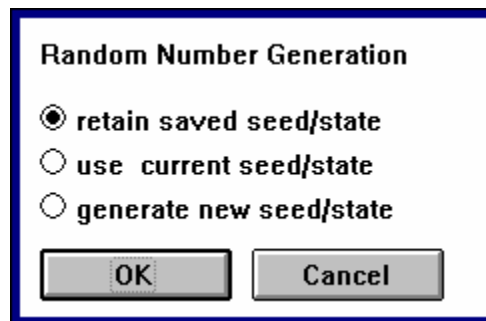
Specify True Values of Parameters

Use dialog boxes to specify TRUE values for the uncorrelated components and the error standard deviation. **Default values** come either from the OLS estimates from the current model or else from a previous invocation of this menu item for the current model object.



A dialog box titled "Specify TRUE Error Standard Deviation:". It contains a text input field with the value "0.087". Below the input field are two buttons: "OK" and "Cancel".

This item also controls XLisp-Stat random number generator **seed/state** options:



A dialog box titled "Random Number Generation". It contains three radio button options: "retain saved seed/state" (which is selected), "use current seed/state", and "generate new seed/state". Below the options are two buttons: "OK" and "Cancel".

These options apply each time a RXridge simulation generates a normal-theory response y-vector. Option "0" means that the same y-vector will be generated every time!

A fourth possibility is to use a XLisp command to manually reset the saved seed/state:

(send my-model :RXridge-seed #\$(1 #(w x y z)))

where w, x, y and z are numbers with as many as ten digits. You will probably want to use copy/paste for this rather than simply "make up" values because each XLisp implementation has its own rules on allowed #\$(1 #(w x y z)) state-object combinations.

List True Shrinkage Details

More details ...again, for possible saving in a permanent “dribble” file. If you are not interested in details, select the next menu item below instead of this one. However, calculations will have to be REDONE if you select this item AFTER the “display” item below.

Display Expected Traces, True MSE Risks

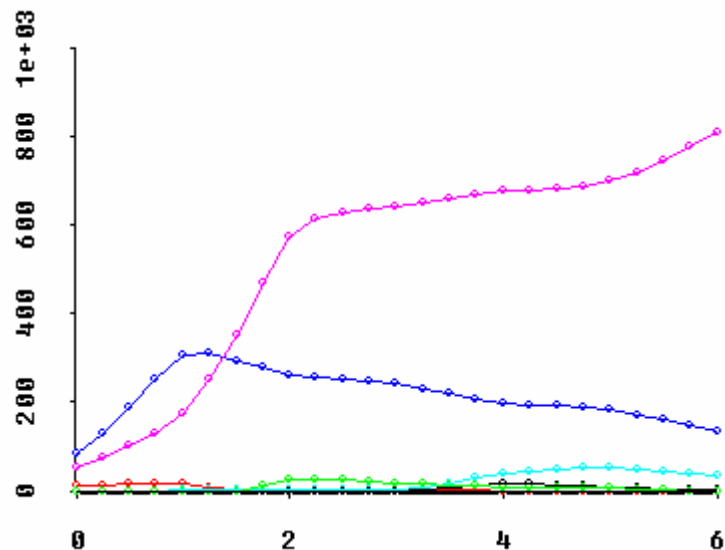
View Expected Coefficient Traces and Exact, TRUE MSE Risks associated with shrinkage along a path of specified Q-shape (when the true standard deviation and uncorrelated components are as specified above.)

List Shrinkage Simulation Details

Simulation and SE Loss details ...again, for possible saving in a permanent “dribble” file. Results will depend upon the true standard deviation, true uncorrelated components, and the initial XLisp random seed/state ...all controlled by the “Specify True Values” menu item. Calculations include Fitted Coefficients and Exact Squared Error LOSSES (not Risks = Expected Losses) associated with shrinkage along a path of specified Q-shape. With seed/state options 1=>continue and 2=>new, you should get DIFFERENT response y-values fitted coefficients, etc. each time you invoke this item.

Display Simulated Traces, True SE Losses

View shrinkage trace plots for the simulated normal-theory response vector from the MOST RECENT invocation of the above menu item for the current model object. If you skipped the previous item because you weren't interested in details, the necessary calculations will be triggered only ONCE (with most printing turned off.) A second invocation of this menu item will NOT trigger generation of new simulation results (even when the 1=>continue or 2=>new seed/state options are in effect.) To view dynamic simulation results, make certain the 1=>continue or 2=>new seed/state options are in effect and select this simulation “Display” item only once after each re-invocation of the simulation “List” item, above.



This item also creates a RXridge main-menu named “SIM” to analyze simulated response y-values as if the true standard deviation and true uncorrelated components were unknown!!! The corresponding model object is always named simply “sim-reg”, so multiple invocations must either OVERWRITE the previous “sim-reg” or be DISCARDED.

XLISP-STAT		
File	Edit	Longley
SIM		
Window		
Choice of Regressor-Response Scaling		
Compute Principal Axis Summary		
918.7190949	0.	Identify Most Likely Path Q-shape
926.9533542	0.	Set Shrinkage Path Q-shape
928.6386273	0.	Set Number of STEPS per Shrinkage Unit
929.0885590	0.	Identify Most Likely Shrinkage M-extent
931.7820920	5.	List Shrinkage Trace Details
938.4928603	15.	Display Shrinkage TRACE Plots
948.6118096	41.	Visual Re-Regression and Influence
958.9860799	142.	Plot Component Size vs. Significance
963.5822430	483.	Specify True Values of Parameters
965.1167012	1220.	List True Shrinkage Details
966.1515876	2876.	Display Expected Traces, True MSE Risks
967.3867193	7253.	List Shrinkage Simulation Details
969.8208175	1.	Display Simulated Traces, True SE Losses
974.7211408	4.	Remove this MENU
982.6621917	1.	
993.8039371	inf	
Least Squares Esti		
Constant		

Group Six:

Remove this MENU

...The End!!!

Acknowledgments

I wish to thank Bernhard Walter for porting my **softRX freeware** routines for S-plus to XLisp-Stat. Bernhard's initial translation required many basic changes in computational strategy/tactics and resulted in creation of OBJECT ORIENTED XLisp-Stat code. Technically speaking, this conversion required creating many "new methods" for the XLisp-Stat's built-in "regression-model-proto." I found that XLisp-Stat coding presented me a very steep learning curve indeed, and I certainly would not have come this far this fast without the "running head-start" Bernhard provided. Thanks again, Bernhard!!!

Bernhard's XLisp-Stat code included an example of "dynamic graphics" for my Visual Re-Regression concept. (My VRR proposals remain unpublished, Obenchain (1996b), but I sent a copy of the first 200+ pages for my book-in-progress, **Shrinkage Regression: ridge, BLUP, Bayes and Stein**, to Bernhard in March 1994.) The most drastic changes that I (Bob Obenchain) have made to Bernhard's XLisp-Stat code consist of modifications to his implementation of "linked" VRR and Outlier/Leverage plots. These plots change dynamically as the user chooses parameter values from a pair of SLIDERS: (i) the value of the MCAL shrinkage-extent parameter and (2) the overall-influence level of Cook-like, hyperbolic contours on the Leverage vs Squared, Standardized Residual plot.

Bibliography

Chatterjee, S. and Hadi, A. S. (1988). **Sensitivity Analysis in Regression**. New York: John Wiley.

Cook, R. D. and Weisberg, S. (1994). **Introduction to Regression Graphics**. New York: John Wiley.

Cook, R. D. (1977). "Detection of influential observations in linear regression." **Technometrics** 19, 15-18.

Efron B. and Morris, C. N. (1976). "Discussion" (of Dempster, Schatzoff and Wermuth.) **Journal American Statistical Association** 72, 91-93.

- Gibbons, D. G. (1981). "A simulation study of some ridge estimators." **Journal of the American Statistical Association** 76, 131-139.
- Goldstein M. and Smith, A. F. M (1974). "Ridge-type estimators for regression analysis." **Journal of the Royal Statistical Society B** 36, 284-291.
- Golub, G. H., Heath, M., and Wahba, G. (1979). "Generalized cross- validation as a method for choosing a good ridge parameter." **Technometrics** 21, 215-223.
- Hoerl, A. E. and Kennard, R. W. (1970a). "Ridge regression: biased estimation for nonorthogonal problems." **Technometrics** 12, 55-67.
- Hoerl, A. E. and Kennard, R. W. (1970b). "Ridge regression: applications to nonorthogonal problems." **Technometrics** 12, 69-82.
- Longley, J. W. (1967). "An appraisal of least squares programs for the electronic computer from the point of view of the user." **Journal of the American Statistical Association** 62, 819-841.
- Mallows, C. L. (1973). "Some comments on Cp." **Technometrics** 15, 661-677.
- Marquardt, D. W. (1970). "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation." **Technometrics** 12, 591-612.
- Marquardt, D. W. (1980). "Comment: You should standardize the predictor variables in your regression models." **Journal of the American Statistical Association** 75, 87-91.
- Nash, J. C. (1992). "Statistical shareware: illustrations from regression techniques." **The American Statistician** 46, 312-318. (A review article that includes coverage of softRX freeware.)
- Obenchain, R. L. and Vinod, H. (1974). "Estimates of partial derivatives from ridge regression on ill-conditioned data." **NBER-NSF Seminar on Bayesian Inference in Econometrics**, Ann-Arbor, Michigan.
- Obenchain, R. L. (1975). "Ridge analysis following a preliminary test of the shrunken hypothesis." **Technometrics** 17, 431-441. (Discussion: McDonald, G. C., 443-445; classical likelihood monitoring along any shrinkage/ridge path.)
- Obenchain, R. L. (1977). "Classical F-tests and confidence regions for ridge regression." **Technometrics** 19, 429-439.
- Obenchain, R. L. (1978). "Good and optimal ridge estimators." **Annals of Statistics** 6, 1111-1121. (The "ridge-function" theorem; maximum likelihood estimation of the "inferior direction" and of scaled MSE risk along any direction in p-

- dimensional regression coefficient space; MSE risk optimality of 0 or of $X' y$ along directions orthogonal to or parallel to β , respectively.)
- Obenchain, R. L. (1981). "Maximum likelihood ridge regression and the shrinkage pattern alternatives." **I.M.S. Bulletin** 10, 37; Abstract 81t-23. (67 page review article.)
- Obenchain, R. L. (1984). "Maximum likelihood ridge displays." **Communications in Statistics A** 13, 227-240. (Proceedings of the Fordham Ridge Symposium, ed. H. D. Vinod; illustrations of usage of RXridge SAS/IML freeware.)
- Obenchain, R. L. (1995). "Maximum likelihood ridge regression." **Stata Technical Bulletin** 28, 22-36. (Introduction to basic shrinkage/ridge concepts; illustrations of usage of RXridge .ADO freeware.)
- Obenchain, R. L. (1996a). "Maximum likelihood shrinkage in regression." submitted to... **The American Statistician**. (Closed form expressions for classical, fixed coefficient, maximum likelihood estimation within the 2-parameter shrinkage family.)
- Obenchain, R. L. (1996b). "Influential observations in ridge regression." submitted to... **Technometrics**. (Exposition of basic "Visual Re-Regression" concepts.)
- Obenchain, R. L. (1996c). **Shrinkage Regression: ridge, BLUP, Bayes and Stein**. Preliminary draft (200+ pages.)
- Obenchain, R. L. (2005). **RXshrink**: R package for maximum likelihood shrinkage in generalized (2-parameter) ridge and least angle regression (LAR.) <http://www.r-project.org>
- Shumway, R. H. (1982). "Maximum likelihood estimation of the ridge parameter in linear regression." Technical Report, Department of Statistics, University of California at Davis.
- Theil, H. (1963). "On the use of incomplete prior information in regression analysis." **Journal of the American Statistical Association** 58, 401-414.
- Tierney, Luke. (1990). **LISP-STAT: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics**. New York: John Wiley & Sons.
- Walter, Bernhard (1994). "XLisp-Stat code for shrinkage/ridge regression." Technische Universität München. [internet: walter@pollux.edv.agrar.tu-muenchen.de]