liureg: A Comprehensive R Package for the Liu Estimation of Linear Regression Model with Collinear Regressors

by Muhammad Imdadullah, Muhammad Aslam, Saima Altaf

Abstract The Liu regression estimator is now a commonly used alternative to avoid the adverse effects of the conventional ordinary least squares estimator in the situations when there exists some considerable degree of multicollinearity among the regressors. There are only few software packages available for estimation of the Liu regression coefficients but with limited methods to estimate the Liu biasing parameter without addressing testing procedures. Our developed package **liureg** can be used to estimate the Liu regression coefficients with a range of different existing biasing parameters and testing of these coefficients with more than 15 Liu related statistics and different graphical displays of these statistics.

Introduction

For data collected either from a designed experiment or from an observational study, the ordinary least square (OLS) method does not provide precise estimates of the effect of any explanatory variable (regressor) when regressors are interdependent (collinear with each other). Consider a multiple linear regression (MLR) model,

$$y = X\beta + \varepsilon,$$

where y is an $n \times 1$ vector of observation on dependent variable, X is known design matrix of order $n \times p$, β is a $p \times 1$ vector of unknown parameters and ε is an $n \times 1$ vector of random errors with mean zero and variance $\sigma^2 I_n$, where I_n is an identity matrix of order n.

The OLS estimator (OLSE) of β is given by

$$\hat{\beta} = (X'X)^{-1}X'y,$$

which depends on the characteristics of the matrix $X'X$. If $X'X$ is ill-conditioned (near dependencies among various regressor of $X'X$) exist or $det(X'X) \approx 0$ then the OLS estimates are sensitive to a number of errors, such as non-significant or imprecise regression coefficients (Kmenta, 1980) with wrong sign and non-uniform eigenvalues spectrum. Moreover, the OLS method yields in high variance of estimates, large standard errors and wide confidence intervals etc.

Researchers may tempt to eliminate regressor(s) causing the problem by consciously removing them from the model or by using some screening method such as stepwise and best subset regression etc. However, these methods can destroy the usefulness of the model by removing relevant regressor(s) from the model. To control variance and instability of the OLS estimates, one may regularize the coefficients, with some regularization methods such as the ridge regression (RR), Lasso regression and Liu regression (LR) methods etc., as alternative to the OLS. Computationally, the RR ($\hat{\beta}_r = (X'X + KI)^{-1}X'y$) suppresses the effects of collinearity and reduces the apparent magnitude of the correlation among regressors in order to obtain more stable estimates of the coefficients than the OLS estimates and it also improves the accuracy of prediction (see Hoerl and Kennard, 1970; Montgomery and Peck, 1982; Myers, 1986; Rawlings et al., 1998; Seber and Lee, 2003; Tripp, 1983, etc.). However, ridge coefficient is a complicated function of k when some popular methods such as given in Golub et al. (1979), Mallows (1973) and McLeod and Xu (2017) etc., are used for (optimal) selection of k. Usually k is quite small in different application, that's why, selection of small k may not be enough to correct the problem of ill-conditioned $X'X$. In such cases, the RR may still be unstable. Similarly, the choice of k belongs to the researcher and also there is no consensus regarding how to select optimal k, therefore, other innovative methods were needed to deal with collinear data. Liu (1993) proposed another biased estimator to mitigate the collinearity effect on regressors. Liu (1993) also discussed some of the properties and methods for suitable selection of biasing parameter used in LR. For further detail see section "Liu regression estimator".

We developed the **liureg** (Imdadullah and Aslam, 2017) package that primarily provide the functionality of Liu related computations. The package, **liureg** provides the most complete suite of tools for the LR available in R, comparable to those listed in Table 1. For package development and R documentation, we followed Hadley (2015); Leisch (2008); Team (2015). The **ridge** package by Cule and De Iorio (2012), **lmridge** by Imdadullah and Aslam (2016a) and `lm.ridge` from the **MASS** by Venables and Ripley (2002) also helped us in coding.

	lrmest (Dissanayake and Wijekoon, 2016)	**ltsbase** (Kan et al., 2013)	**liureg**
Standardization of regressors			
	✓	✓	✓
Estimation and testing of Liu coefficient			
Estimation	✓	✓	✓
Testing	✓		✓
SE of coeff.	✓		✓
Liu related statistics			
R^2	✓		✓
Adj-R^2			✓
Variance			✓
Bias^2			✓
MSE			✓
F-test			✓
σ^2			✓
C_L			✓
Effective df			✓
Hat matrix			✓
Var-Cov matrix			✓
VIF			✓
Residuals		✓	✓
Fitted values		✓	✓
Predict values			✓
Liu model selection			
GCV			✓
AIC&BIC			✓
PRESS			✓
Liu related graphs			
Liu trace			✓
Bias, Var, MSE		✓	✓
AIC, BIC			✓

Table 1: Comparison of Liu related R packages

In the available literature, there are only two R packages capable of estimating and/ or testing of the Liu coefficients. The R packages mentioned in Table 1 are compared with our **liureg** package. The **lrmest** package (Dissanayake and Wijekoon, 2016) computes different estimates such as the OLS, ordinary ridge regression (ORR), Liu estimator (LE), LE type-1,2,3, adjusted Liu estimator (ALTE) and their type-1,2,3 etc. Moreover, **lrmest** provides scalar mean square error (MSE), prediction residual error sum of squares (PRESS) values of some of the estimators available in the package **lrmest**. The testing of ridge coefficient is performed on only scalar `k`, however, for vector of `d`, function `liu()` of **lrmest** package returns only MSE along with value of biasing parameter used. The **ltsbase** package (Kan et al., 2013) computes ridge and Liu estimates based on the least trimmed squares (LTS) method. The MSE value from four regression models can be compared on plot if argument `plot=TRUE` in `ltsbase()` function. There are three main functions, (i) `ltsbase()` computes the minimum MSE values for six methods: OLS, ridge, ridge based on LTS, LTS, Liu and Liu based on LTS method for sequences of biasing parameters ranging from 0 to 1, (ii) The `ltsbaseDefault()` function returns the fitted values and residuals of the model having minimum MSE, and (iii) The `ltsbaseSummary()` function returns the regression coefficients and the biasing parameter for the best MSE among the four regression models.

It is important to note that **ltsbase** package displays these statistics for model having minimum MSE (bias and variance are not displayed in output), while our package **liureg** computes these and all other statistics not only for scalar but also for vector biasing parameter.

This paper outlines the collinearity detection methods available in existing literature and use of **mctest** (Imdadullah and Aslam, 2016b) package through illustrative example. To overcome the issue of collinearity effect on regressors a thorough introduction to the Liu regression (LR), properties of the Liu estimator, different methods for the selecting values of d and testing of the Liu coefficients is presented. Finally, estimation of the Liu coefficients, methods of selecting biasing parameter, testing of the Liu coefficients and different Liu related statistics are implemented in R with proposed **liureg** package.

Collinearity detection

Diagnosing collinearity is important to many researchers that consists of two related but separate elements (1) detecting the existence of collinear relationship among regressors and (2) assessing the extent to which this relationship has degraded the parameter estimates. There are many diagnostic measures used for detection of collinearity in the existing literature provided by various authors (Belsley et al., 1980; Curto and Pinto, 2011; Farrar and Glauber, 1967; Fox and Weisberg, 2011; Gunst and Mason, 1977; Klein, 1962; Koutsoyiannis, 1977; Kovács et al., 2005; Marquardt, 1970; Theil, 1971). These diagnostics methods assist in determining whether and where some corrective action is necessary (Belsley et al., 1980). Widely used and the most suggested diagnostics are value of pair-wise correlations, variance inflation factor (VIF)/ tolerance (TOL) (Marquardt, 1970), eigenvalues and eigenvectors (Kendall, 1957), CN & CI (Belsley et al., 1980; Chatterjee and Hadi, 2006; Maddala, 1988), Leamer's method (Greene, 2002), Klein's rule (Klein, 1962), the tests proposed by Farrar and Glauber (Farrar and Glauber, 1967), Red indicator (Kovács et al., 2005), corrected VIF (Curto and Pinto, 2011) and Theil's measures (Theil, 1971), also see Imdadullah et al. (2016). All of these diagnostic measures are implemented in a latest developed R package **mctest** (Imdadullah and Aslam, 2016b). We used the Hald dataset (Hald, 1952), for testing collinearity among regressors and then using **liureg** package, computation of the Liu regression coefficients, different Liu related statistics and methods of selection of Liu biasing parameter is performed. For optimal choice of biasing parameter, graphical representation of the Liu coefficients, bias variance trade-off plot and model selection criteria is also performed. The Hald data are about heat evolved during setting of 13 cement mixtures of 4 basic ingredients and used by Hoerl et al. (1975). Each ingredient percentage appears to be rounded down to a full integer. The data set is already bundled in **mctest** and **liureg** package.

Collinearity detection: An example

```
R > data(Hald)
R > x <- Hald[, -1]
R > y <- Hald[,  1]
R > mctest (x, y)

Call:
omcdiag(x = x, y = y, Inter = TRUE, detr = detr, red = red, conf = conf,
        theil = theil, cn = cn)

Overall Multicollinearity Diagnostics

MC Results detection
Determinant |X'X|:         0.0011         1
Farrar Chi-Square:        59.8700         1
Red Indicator:             0.5414         1
Sum of Lambda Inverse:   622.3006         1
Theil's Method:            0.9981         1
Condition Number:        249.5783         1

1 --> COLLINEARITY is detected
0 --> COLLINEARITY is not detected by the test

===================================
Eigvenvalues with INTERCEPT
                 Intercept     X1     X2      X3       X4
Eigenvalues:        4.1197 0.5539 0.2887  0.0376   0.0001
Condition Indeces:  1.0000 2.7272 3.7775 10.4621 249.5783
\end{CodeOutput}
```

The results from all overall collinearity diagnostic measures indicate the existence of collinearity among regressor(s). Since, these results do not tell which regressor(s) are reasons of collinearity, the individual collinearity diagnostic measures can be obtained, such as,

```
> mctest(x = x, y, all = TRUE, type = "i")
```

```
Call:
imcdiag(x = x, y = y, method = method, corr = FALSE, vif = vif,
      tol = tol, conf = conf, cvif = cvif, leamer = leamer, all = all)

All Individual Multicollinearity Diagnostics in 0 or 1

   VIF TOL Wi Fi Leamer CVIF Klein
X1   1   1  1  1      0    0     0
X2   1   1  1  1      1    0     1
X3   1   1  1  1      0    0     0
X4   1   1  1  1      1    0     1

1 --> COLLINEARITY is detected
0 --> COLLINEARITY in not detected by the test

X1 , X2 , X3 , X4 , coefficient(s) are non-significant may be due to multicollinearity

R-square of y on all x: 0.9824

* use method argument to check which regressors may be the reason of collinearity
```

The results from most of the individual collinearity diagnostics suggest that all of the regressors may the reason of collinearity among regressors. The last line of `imcdiag()` function's output suggests that method argument should be used to check which regressors may be the reason of collinearity among different regressors. This finding suggest that one should use regularization method such as LR.

Liu regression estimator

To deal with multicollinear data, Liu (1993) formulated a new class of biased estimators that has combined benefits of ORR by Hoerl and Kennard (1970) and the Stein type estimator Stein (1956), $\hat{\beta}_S = c\hat{\beta}$, where c is parameter $0 < c < 1$ to avoid their disadvantages. The Liu estimator (LE) can be defined as,

$$\begin{aligned} \hat{\beta}_d &= (X'X + I_p)^{-1}(X'y + d\hat{\beta}_{ols}), \\ &= (X'X + I_p)^{-1}(X'X + dI_p)\hat{\beta}_{ols}, \\ &= F_d\hat{\beta}_{ols}, \end{aligned} \tag{1}$$

where d is the Liu parameter also known as the biasing (tuning or shrinkage) parameter and lies between 0 and 1 (i.e. $0 \leq d \leq 1$), I_p is identity matrix of order $p \times p$, and $\hat{\beta}$ is OLSE.

The $\hat{\beta}_d$ is named as the LE by Akdeniz and Kaçįranlar (1995) and Gruber (1998). Recently, in econometrics, engineering and other statistical areas, the LE has produced a number of new techniques and ideas, see for example Akdeniz and Kaçįranlar (2001); Hubert and Wijekoon (2006); Jahufer and Chen (2009, 2011, 2012); Kaçįranlar et al. (1999); Kaçįranlar and Sakalhoğlu (2001); Torigoe and Ujiie (2006).

However, Liu (2011) and Druilhet and Mom (2008) have made statement that the biasing parameter d may lie outside the range given by Liu (1993), that is, it may be less than 0 or greater than 1. The LE is linear transformation of the OLSE, $\hat{\beta}_d = \hat{\beta}_{ols}$.

The suitable selection of d at which MSE is minimum and efficiency of estimators improves as compared to other values of d is the main interest of LE. Liu (1993) provided some important methods for the selection of d and also provided numerical example by iterative minimum MSE method to get the smallest possible value to overcome the problem of collinearity in an effective manner.

Reparameterization

The design matrix $X_{n\times p}$ and response variable $y_{n\times 1}$ should be standardized, scaled or centered first such that information matrix $X'X$ is in the correlation form and vector $X'y$ is in form of correlation among regressors and the response variable. Consider regression model, $y = \beta_0 1 + \widetilde{X}\beta_1 + \varepsilon$, where $\widetilde{X}$ is centered and $1 = c(1, 1, \cdots, 1)'$, while β_0 can be estimated by using $\overline{y}$. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, be the ordered eigenvalues of matrix $\widetilde{X}'\widetilde{X}$ and $q_1, q_2, \cdots, q_p$ be the eigenvectors corresponds to their eigenvalues, such that $Q = (q_1, q_2, \cdots, q_p)$ is orthogonal matrix of $\widetilde{X}'\widetilde{X}$ and $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix}$, therefore, model can be rewritten in canonical form as $y = \beta_0 1 + Z\alpha + \varepsilon$, where $Z = \widetilde{X}Q$ and $\alpha = Q'\beta_1$. Note that, $\Lambda = Z'Z = Q'\widetilde{X}'\widetilde{X}Q$. The estimate of α is $\hat{\alpha} = \Lambda^{-1}Z'y$. Similarly, Eq. 1 can be written in canonical form as,

$$\hat{\alpha}_d = (\Lambda + I_p)^{-1}(Z'y + d\hat{\alpha})$$

Corresponding estimate of $\hat{\beta}_1$ and $\hat{\beta}_d$ can be obtained by following relation of $\hat{\beta}_1 = Q\hat{\alpha}$ and $\hat{\beta}_d = Q\hat{\alpha}_d$, respectively. For simplification of notations, $\widetilde{X}$ and $\hat{\alpha}$ will be represented as X and β, respectively.

The fitted values of the LE can be found using Eq. 1,

$$\begin{aligned} \hat{y}_d &= X\hat{\beta}_d, \\ &= X(X'X + I_p)^{-1}(X'y + d)\hat{\beta}, \\ &= H_d\, y, \end{aligned}$$

where, H_d is LE the matrix (Liu, 1993; Walker and Birch, 1988, see). It is worthy to note that H_d is not idempotent because it is not projection matrix, therefore it is called quasi-projection matrix.

As $\hat{\beta}_d$ is computed on centered variables, so they need back to the original scale, that is,

$$\hat{\beta} = \left(\frac{\hat{\beta}_{dj}}{S_{xj}}\right),$$

where S_{xj} is scaling method of regressors.

The intercept term for the LE ($\hat{\beta}_{0d}$) can be estimated using the following relation,

$$\begin{aligned} \hat{\beta}_{0d} &= \overline{y} - (\hat{\beta}_{1d}, \cdots, \hat{\beta}_{pd})\overline{x}_j' \;, \\ &= \overline{y} - \sum_{j=1}^{p} \overline{x}_j \hat{\beta}_{jd}. \end{aligned} \tag{2}$$

Properties of the Liu estimator

Like the linear RR, the Liu regression is also the most popular method among biased methods, because of its relation to the OLS and its statistical properties have been studied by Akdeniz and Kaçiranlar (1995, 2001), Arslan and Billor (2000), Kaçiranlar and Sakalhoğlu (2001), Kaçiranlar et al. (1999) and Sakalhoğlu et al. (2001) among many others. Due to comprehensive properties of the LE, researcher have been attracted towards this area of research.

For $d = 1, \hat{\beta}_d = \beta_{ols}$. Therefore, LE is the shrinkage estimator, though biased but has lower MSE than OLS that is, $MSE(\hat{\beta}_d) < MSE(\hat{\beta}_{ols})$ (see Sakalhoğlu et al., 2001, etc.).

Let X_j denote the jth column of $X(j = 1, 2, \cdots, p)$, where $X_j = (x_{1j}, x_{2j}, \cdots, x_{nj})'$. As already discussed, the regressors are centered, thus, the intercept will be zero and can thereby be removed from the model. However, it can be estimated from relation given in Eq. 2. Table 2, lists the Liu properties that are implemented in our newly developed **liureg** package.

Theoretically and practically, the LR is used to propose some new methods for the choice of the biasing parameter d to investigate the properties of LR, since biasing parameter plays a key role while the optimal choice of d is the main issue in this context. In the literature, many methods for selection of appropriate biasing parameter d have been studied by Akdeniz and Özkale (2005), Arslan and Billor (2000), Akdeniz et al. (2006), Özkale and Kaçiranlar (2007) and Liu (1993).

Methods of selecting values of d

The existing methods to select biasing parameter in the LR may not fully address the problem of ill-conditioning when there exists sever multicollinearity, while the appropriate selection of biasing parameter d also remains a problem of interest. The parameter d should be selected when there are improvements in the estimates (have stable estimates) or prediction is improved.

The optimal value of d is one which gives minimum MSE. There is one optimal d for any problem by the analogy with the estimate of k in RR, a wide range of d ($-\infty < d < 1$) give smaller MSE as compared to that of the OLS. For collinear data, a small change in d varies the LR coefficients rapidly. Therefore, a disciplined way of selecting the shrinkage parameter is required that minimizes the

Sr.#	Property	Formula
1)	Linear transformation	The LE is a linear transformation of the OLSE ($\hat{\beta}_d = F_d\hat{\beta}$)
2)	Wide range d	Wide range of d have smaller MSE than the OLS.
3)	Optimal d	An optimal d always exists that gives minimum MSE.
4)	Mean	$E(\hat{\beta}_d) = F_d\beta$, where $F_d = (X'X + I_p)^{-1}(X'X + dI_p)$
5)	Bias	$Bias = Q'(F_d - I_p)\beta$
6)	Var-Cov matrix	$Cov(\hat{\beta}_d) = \sigma^2 F_d(X'X)^{-1}F_d'$
7)	MSE	$MSE(\hat{\beta}_d) = \sigma^2 F_d(X'X)^{-1}F_d + (F_d - I_p)\beta\beta'(F_d - I_p)'$ $= \sigma^2 \sum_{j=1}^{p} \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} + (d-1)^2 \sum_{j=1}^{p} \frac{\beta^2}{(\lambda + 1)^2}$
8)	Effective DF (EDF)	$EDF = trace[XF_d(X'X)^{-1}X']$
9)	Larger regression coeff.	$\hat{\beta}_d'\hat{\beta}_d \geq \hat{\beta}_{ols}'\hat{\beta}_{ols}$
10)	Inflated RSS	$\sum(y - X\hat{\beta}_d)^2$

Table 2: Properties of Liu estimator.

MSE. The biasing parameter d depends on the true regression coefficients (β) and the variance of the residuals σ^2, unfortunately which are unknown but can be estimated from the sample data.

We classified estimation methods as (i) Subjective and (ii) Objective

Subjective methods

In these methods, the selection of d is subjective or of judgmental nature and provides graphical evidence of the effect of collinearity on the regression coefficient estimates and also accounts variation by the LE as compared to the OLSE. In these methods, the reasonable choice of d is done using the Liu trace and plotting of bias, variance and MSE. Like ridge trace, the Liu trace is also a graphical representation of regression coefficients $\hat{\beta}_d$ as a function of d in interval $[-\infty, \infty]$. Similarly, the plotting of bias, variance and MSE from the LE may also be helpful in selecting appropriate value of d. At the cost of bias, optimal d can be selected at which MSE is minimum. All these graphs can be used for selection of optimal (but judgmental) value of d from horizontal axis to assess the effect of collinearity on each of the coefficients. These graphical representations do not provide a unique solution, rather they render a vaguely defined class of acceptable solutions. However, still these traces are still useful graphical representation to check some optimal d.

Objective methods

Objective methods, to some extent are similar to judgmental methods for selection of biasing parameter d, but they require some calculations to obtain these biasing parameters. Table 3 lists widely used biasing parameter d already available in the existing literature. Table 3 also lists other statistics that can be used for the selection of biasing parameter d.

Testing of the Liu coefficients

Testing of the Liu coefficients is performed by following Aslam (2014) and Halawa and El-Bassiouni (2000). For testing $H_0 : \beta_{dj} = 0$ against $\beta_{dj} \neq 0$, the non-exact t-statistics defined by Halawa and El-Bassiouni (2000) is,

$$T_d = \frac{\hat{\beta}_{dj}}{SE(\hat{\beta}_d j)},$$

where $\hat{\beta}_{dj}$ is the jth Liu coefficient estimate and $SE(\hat{\beta}_{dj})$ is an estimate of standard error, which is the square root of the jth diagonal element of the covariance matrix of LE, see property # 6 in Table 2.

The statistics T_{dj} is assumed to follow Student's t distribution with $(n-p)$ df (Halawa and El-Bassiouni, 2000). Hastie and Tibshirani (1990) and Cule and De Iorio (2012) suggested to use df from $(n - trace(H_d))$. For large sample size, the asymptotic distribution of this statistic is normal (Halawa and El-Bassiouni, 2000).

For testing overall significance of vector of LE ($\hat{\beta}_d$) with $E(\hat{\beta}_d) = F_d\,\beta$ and $Cov(\hat{\beta}_d)$, the F-statistic

Sr.#	Formula	Reference
1)	$d_{opt} = \frac{\sum_{j=1}^{p}\left[\frac{\alpha_j^2-\sigma^2}{(\lambda_j+1)^2}\right]}{\sum_{j=1}^{p}\left[\frac{\sigma^2+\lambda_j\alpha_j^2}{\lambda_j(\lambda_j+1)^2}\right]}$	Liu (1993)
2)	$\hat{d} = 1-\hat{\sigma}^2\left[\frac{\sum_{j=1}^{p}\frac{1}{\lambda_j(\lambda_j+1)}}{\sum_{j=1}^{p}\frac{\hat{\alpha}_j^2}{(\lambda_j+1)^2}}\right]$	Liu (1993)
3)	$\hat{d}_{imp} = \frac{\sum_{i=1}^{n}\frac{\widetilde{e}}{1-g_{ii}}\left(\frac{\widetilde{e}_i}{1-h_{1-ii}}-\frac{\hat{e}_i}{1-h_{ii}}\right)}{\sum_{i=1}^{n}\left(\frac{\widetilde{e}}{1-g_{ii}}-\frac{\hat{e}_i}{1-h_{ii}}\right)^2}$, where, $\hat{e} = y_i - x_i'(X'X - x_ix_i')^{-1}(X'y - x_iy_i)$, $\widetilde{e} = y_i - x_i'(X'X + I_p - x_ix_i')^{-1}(X'y - x_iy_i)$, $G = X(X'X+I_p)^{-1}X'$, and $H \cong X(X'X)^{-1}X'$	Liu (2011)
4)	$PRESS_d = \sum_{i=1}^{n}(\hat{e}_{d(i)})^2$, where $\hat{e}_{d_{(i)}} = \frac{\hat{e}_{e_i}}{1-h_{1-ii}} - \frac{\hat{e}_i}{(1-h_{1-ii})(1-h_{ii})}(h_{1-ii}-\widetilde{h}_{d-ii})$, $\hat{e}_{d_i} = y_i - \hat{y}_{d_i}$, $\widetilde{H}_{d-ii}$diagonal elements from Liu hat matrix, $h_{ii} = x_i'(X'X)^{-1}x_i$, and $h_{1-ii} = x_i'(X'X+I)^{-1}x_i$	Özkale and Kaçiranlar (2007)
5)	$C_L = \frac{SSR_d}{\hat{\sigma}^2} + 2\,trace(\widetilde{H}_d) - (n-2)$, where, $\widetilde{H}_d$ is hat matrix of LE	Mallows (1973)
6)	$GCV = \frac{SSR_d}{(n-[1+trace(\widetilde{H}_d)])^2}$	Liu (1993)
7)	$AIC = n\,log(RSS) + 2df$, $BIC = n\,log(RSS) + df\,log(n)$, where $df = trace(H_d)$	

Table 3: Different available methods to estimate d.

is,

$$F = \frac{1}{p}(\hat{\beta}_d - F_d\beta)'(Cov(\hat{\beta}_d))^{-1}(\hat{\beta}_d - F_d\beta)$$

The standard error of $\hat{\beta}_d$ is computed by considering variance of the estimator, given in Eq. 2 and then taking square root of this variance, that is,

$$S.E(\hat{\beta}_{0d}) = \sqrt{Var(\overline{y}) + \overline{X}_j^2\,diag[Cov(\hat{\beta}_d)]} \tag{3}$$

The R package, liureg

Our R package **liureg** contains functions related to fitting of the LR model and provides a simple way of obtaining the estimates of LR coefficients, testing of the Liu coefficients and computation of different Liu related statistics, helpful for selection of optimal biasing parameter d. The package computes different Liu related measures available for the selection of biasing parameter d and value of different biasing parameter proposed by some researchers, available in the literature.

The `liureg` objects contain a set of standard methods such as `print()`, `summary()`, `plot()` and `predict()` etc. Therefore, inferences can be made easily using `summary` method for assessing the estimates of regression coefficients, their standard errors, t-values and their respective p-values. The default function `liu` which calls `liuest()` to perform required computations and estimation for given values of non-stochastic biasing parameter d. The syntax of default function is,

```
liu(formula,data,scaling=("centered","sc","scaled"),d,...)
```

The `liu()` function has following four arguments shown in Table 4:

Argument	Description
`formula`	symbolic representation for LR model of the form, response $\sim$ predictors.
`data`	contains the variables that have to be used in LR model.
`d`	biasing parameter, may be a scalar or vector. If d value is not provided, $d = 1$ will be used as default value, i.e., the OLS results will be produced.
`scaling`	The methods for scaling of predictors. The `centered` option, centers the predictors, suggested by Liu (1993) and it is default scaling option, the `sc` option scales the predictors in correlation form as described in Belsley (1991); Draper and Smith (1998) and `scaled` option standardizes the predictors having zero mean and unit variance.

Table 4: Description of `liu()` function arguments.

The `liu()` function returns an object of class `"liu"`. The function `summary()`, `dest()` and `lstats()` etc., are used to compute and print a summary of the LR results, list of biasing parameter by Liu (1993, 2011) and Liu related statistics such as estimated squared bias, R^2 and variance etc., after bias is introduced in regression model. An object of class `"liu"` is a list that contains the following components in Table 5:

Object	Description
coef	A named vector of fitted Liu coefficients.
lfit	Matrix of Liu fitted values for each biasing parameter d.
mf	Actual data used.
xm	A vector of means of design matrix X.
y	The centered response variable.
xscale	The scales used to standardize the predictors.
xs	The scaled matrix of predictors.
scaling	The method of scaling used to standardized the predictors.
d	The LR biasing parameter(s).
Inter	Whether intercept is included in the model or not.
call	The matached call.
terms	The `terms` object used.

Table 5: Objects from `"liu"` class.

Table 6 lists the functions and methods available in **liureg** package.

The Liu Package Implementation in R

The use of **liureg** is explained through examples by using Hald data.

```
> library(liureg)
> mod <- liu(y ~ X1 + X2 + X3 + X4, data = as.data.frame(Hald),
+  scaling = "centered", d = seq(0, 1, 0.01) )
```

The output of linear LR from `liu()` function is assigned to an object `mod`. The first argument of function is `formula`, which is used to specify the required linear LR model for the data provided as second argument. By simply typing the object `mod` at R prompt will yields objects of class `"liu"` with de-scaled coefficients. The output (de-scaled coefficients) from above command is only for few selected biasing parameter values.

```
Call:
liu.default(formula = y ~ ., data = as.data.frame(Hald), d = c(0,
0.01, 0.49, 0.5, 0.9, 1))

       Intercept       X1       X2        X3        X4
d=0     75.01755 1.41348 0.38190 -0.03582 -0.27032
```

Functions	Description
Liu coefficient estimation and testing	
`liuest()`	The main model fitting function for implementation of LR models in R.
`coef()`	Display de-scaled Liu coefficients.
`liu()`	Generic function and default method that calls `liuest()` and returns an object of S3 class `"liu"` with different set of methods to standard generics. It has a print method for display of Liu de-scaled coefficients.
`summary()`	Standard LR output (coefficient estimates, scaled coefficient estimates, standard errors, t-value and p-values); returns an object of class `"summary.liu"` containing the relative summary statistics and have a print method.
Residuals, fitted values and prediction	
`predict()`	Produces predicted value(s) by evaluating `liuest()` in the frame `newdata`.
`fitted()`	Displays Liu fitted values for observed data.
`residuals()`	Displays Liu residuals values.
`press()`	Generic function that computes prediction residuals error sum of squares (PRESS) for Liu coefficients.
Methods to estimate d	
`dest()`	Displays various d (biasing parameter) values from different authors available in literature and have a print method.
Liu statistics	
`vcov()`	Displays associated Var-Cov matrix with matching Liu parameter d values.
`hatl()`	Generic function that displays hat matrix from LR.
`infoliu()`	Generic function that compute information criteria AIC and BIC.
`lstats()`	Generic function that displays different statistics of LR such as MSE, squared bias, R^2 etc., and have print method.
Liu plots	
`plot()`	Liu coefficient trace plot against biasing parameter d.
`plot.biasliu()`	Bias, variance, and MSE plot as function of d.
`plot.infoliu()`	Plot of AIC and BIC against d.

Table 6: Functions and methods in **liureg** package.

```
d=0.01  74.89142 1.41486 0.38318 -0.03445 -0.26905
d=0.49  68.83758 1.48092 0.44475  0.03167 -0.20845
d=0.5   68.71146 1.48229 0.44603  0.03304 -0.20719
d=0.9   63.66659 1.53734 0.49734  0.08814 -0.15669
d=1     62.40537 1.55110 0.51017  0.10191 -0.14406
```

To obtain Liu scaled coefficients `mod$coef` can be used.

```
> mod$coef
          d=0      d=0.01      d=0.49       d=0.5       d=0.9        d=1
X1  1.41348287  1.41485907  1.48091656  1.48229276  1.53734067  1.5511026
X2  0.38189878  0.38318147  0.44475049  0.44603318  0.49734070  0.5101676
X3 -0.03582438 -0.03444704  0.03166517  0.03304251  0.08813603  0.1019094
X4 -0.27031652 -0.26905396 -0.20845133 -0.20718877 -0.15668658 -0.1440610
```

The object of class `"liu"` returns components such as `lfit`,`d` and `coef` etc. For fitted Liu model, generic method `summary` is used to investigate the Liu coefficients. The parameter estimates of Liu model are summarized using a matrix of 5 column namely *estimates, estimates(Sc), StdErr (Sc), t-values (Sc)* and *P(>|t|)*. Following results shown are only for `d=-1.47218` which produces minimum MSE as compared to the others given in argument.

```
> summary(mod)

Call:
liu.default(formula = y ~ ., data = as.data.frame(Hald), d = -1.47218)

Coefficients for Liu parameter d= -1.47218
         Estimate Estimate (Sc) StdErr (Sc) t-val (Sc) Pr(>|t|)
```

```
Intercept  93.5849       93.5849     15.6226      5.990 2.09e-09 ***
X1          1.2109        1.2109      0.2711      4.466 7.97e-06 ***
X2          0.1931        0.1931      0.2595      0.744   0.4568
X3         -0.2386       -0.2386      0.2671     -0.893   0.3717
X4         -0.4562       -0.4562      0.2507     -1.820   0.0688 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Liu Summary
           R2     adj-R2 F      AIC   BIC   MSE
d=-1.47218 0.9819 0.8372 127.8 23.95 59.18 0.7047
```

The summary() function also displays Liu related R^2, adjusted-R^2, F-test, AIC, BIC and minimum MSE at certain d given in liu().

The dest() function which works with Liu fitted model, computes different biasing parameters developed by researchers, see Table 3. The list of different d values (5 in numbers) may help in deciding the amount of bias needs to be introduced in LR. The biasing parameters by Liu (1993, 2011) includes d_{CL}, d_{mm}, d_{opt}, d_{ILE} and GCV for appropriate selection of d.

```
> dest(mod)

Liu biasing parameter d
           d values
dmm        -5.91524
dcl        -5.66240
dopt       -1.47218
dILE       -0.83461
min GCV at  1.00000
```

The argument of dest() is object of class "liu".

The lstats() can be used to compute different statistics of given Liu biasing parameter as argument of function liu. The Liu statistics are MSE, squared bias, F-statistics, Liu variance, degrees of freedom (df) by Hastie and Tibshirani (1990), and R^2 etc. Following are results using lstats() for some $d = -1.47218, -0.06, 0, 0.1, 0.5, 1$.

```
> lstats(mod)

Liu Regression Statistics:

              EDF Sigma2     CL    VAR Bias^2    MSE        F     R2 adj-R2
d=-1.47218 9.4135 5.2173 5.0880 0.2750 0.4297 0.7047 127.8388 0.9819 0.8372
d=-0.06    9.0760 5.2989 5.5077 1.0195 0.0790 1.0985 125.8693 0.9823 0.8406
d=0        9.0677 5.3010 5.5315 1.0625 0.0703 1.1328 125.8194 0.9823 0.8407
d=0.1      9.0548 5.3043 5.5722 1.1362 0.0569 1.1931 125.7427 0.9823 0.8408
d=0.5      9.0169 5.3139 5.7488 1.4561 0.0176 1.4737 125.5157 0.9824 0.8412
d=1        9.0000 5.3182 6.0000 1.9119 0.0000 1.9119 125.4141 0.9824 0.8414

minimum MSE occurred at d= -1.47218
```

The lstats() also displays the value of d which produces minimum MSE among all provided values of d as argument in liu() function.

The residuals, fitted values from the LR and predicted values of response variable y can be computed using fuctions residuals(), fitted() and predict(), respectively. To obtain Var-Cov and Hat matrix, the function vcov() and hatl() can be used. Note that df are computed by following Hastie and Tibshirani (1990). The results for Var-Cov and diagonal elements of the hat matrix from vcov() and hatl() functions are given below for $d = -1.47218$.

```
> vcov(liu(y ~ ., as.data.frame(Hald), d = -1.47218))
$`d=-1.47218`
           X1         X2         X3         X4
X1 0.07351333 0.04805778 0.06567391 0.04874902
X2 0.04805778 0.06732869 0.05192626 0.06412284
X3 0.06567391 0.05192626 0.07134433 0.05149914
X4 0.04874902 0.06412284 0.05149914 0.06284562
```

```
> diag(hatl(liu(y ~ ., as.data.frame(Hald), d = -1.47218)))
         1          2          3          4          5          6          7
0.43522319 0.22023015 0.21341231 0.18535953 0.27191765 0.04296839 0.28798591
         8          9         10         11         12         13
0.30622895 0.15028900 0.59103231 0.30392765 0.14087610 0.18778716
```

Following are use of some functions to compute different Liu related statistics. For detail description of these function/command, see **liureg** package documentation.

```
> hatl(mod)
> halt(mod)[[1]]
> diag(hatl(mod)[[1]])
> vcov(mod)
> residual(mod)
> fitted(mod)
> predict(mod)
> lstats(mod)$lEDF
> lstats(mod)$var
```

For given values of X such as for first five rows of X matrix, the predicted values for some $d = -1.47218, -0.06, 0, 0.1, 0.5, 1$ will be,

```
> predict(mod, newdata = as.data.frame(Hald[1 : 5, -1]))

  d=-1.47218   d=-0.06       d=0     d=0.1     d=0.5       d=1
1   78.27798  78.40208  78.40736  78.41615  78.45130  78.49524
2   73.09404  72.91968  72.91227  72.89992  72.85053  72.78880
3  106.68373 106.27656 106.25926 106.23043 106.11510 105.97094
4   89.54007  89.41842  89.41325  89.40463  89.37017  89.32710
5   95.61470  95.63443  95.63527  95.63667  95.64226  95.64924
```

The model selection criteria's of AIC and BIC can be computed using `infoliu()` function for each value of d used in argument of `liu()`. For some $d = -1.47218, -0.06, 0.5, 1$, the AIC and BIC values are,

```
> infoliu(liu(y ~ ., as.data.frame(Hald), d = c(-1.47218, -0.06, 0.5, 1)))
                AIC      BIC
d=-1.47218 23.95378 59.18349
d=-0.06    24.43818 59.88178
d=0.5      24.69007 60.21849
d=1        24.94429 60.54843
```

The effect of multicollinearity on the coefficient estimates can be identified by using different graphical displays such as Liu trace (see Figure 1), plotting of bias, variance and MSE against d (see Figure 2) and information criteria against df (Figure 3). These graphical displays are (judgmental) methods for selection of optimal biasing parameter d.

```
> mod <- liu(y~., as.data.frame(Hald), d = seq(-5, 5, .001) )
> plot(mod)
> plot.biasliu(mod)
> plot.infoliu(mod)
```

Summary

liureg package provides the most complete suite of tools for LR available in R, comparable to those available as listed in Table 1. We have implemented functions in such a manner that the Liu coefficients, testing of these coefficients, different Liu related statistics and biasing parameter from different existing methods from authors see Table 3. We have greatly increased the Liu related statistics and different graphical methods for the selection of biasing parameter d for **liureg** package in R.

Up to now, a complete suite of tools for LR was not available for an open source or paid version of statistical software packages, resulting in reduced awareness and used of developed Liu related statistics. The package **liureg** provides a complete open source suite of tools for the computation of Liu coefficients estimation, testing and different statistics. We believe the availability of these tools will lead to increase utilization and better Liu related practices.

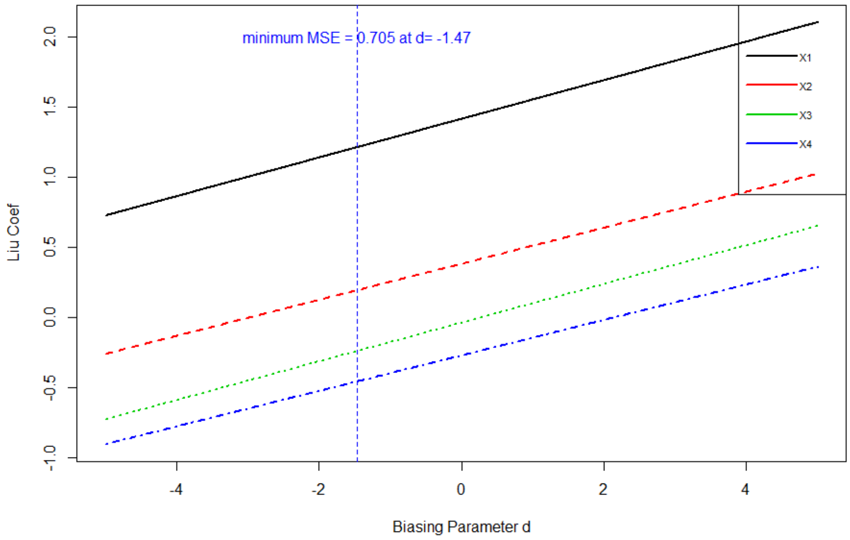

Figure 1: Liu trace: Liu coefficient against biasing parameter d.

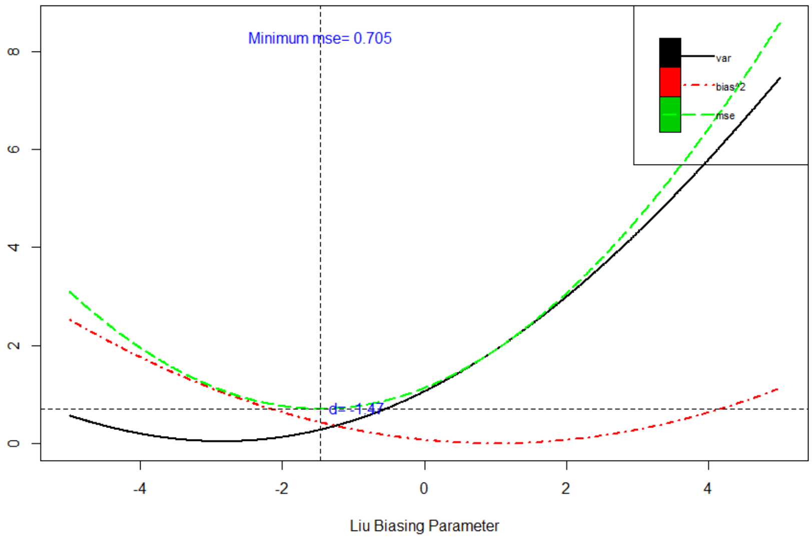

Figure 2: Bias, variance trade-off.

Bibliography

F. Akdeniz and S. Kaçiranlar. On the Almost Unbiased Generalized Liu Estimators and Unbiased Estimation of the Bias and MSE. *Communications in Statistics-Theory and Methods*, 24(7):1789–1797, 1995. URL `http://doi.org/10.1080/03610929508831585`. [p4, 5]

F. Akdeniz and S. Kaçiranlar. More on the New Biased Estimator in Linear Regression. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 63(3):321–325, 2001. [p4, 5]

F. Akdeniz and R. M. Özkale. The Distribution of Stochastic Shrinkage Biasing Parameter of the Liu Type Estimator. *Applied Mathematics and Computation*, 163(1):29–38, 2005. URL `https://doi.org/10.1016/j.amc.2004.03.026`. [p5]

F. Akdeniz, G. P. H. Styan, and H. J. Werner. The General Expression for the Moments of the Stochastic Shrinkage Parameters of the Liu Type Estimator. *Communications in Statistics-Theory and Methods*, 35 (3):423–437, 2006. URL `http://doi.org/10.1080/03610920500476572`. [p5]

O. Arslan and N. Billor. Robust Liu Estimator for Regression Based on an M-Estimator. *Journal of Applied Statistics*, 27(1):39–47, 2000. URL `http://doi.org/10.1080/02664760021817`. [p5]

M. Aslam. Using Heteroscedasticity-Consistent Standard Errors for the Linear Regression Model with Correlated Regressors. *Communications in Statistics-Simulation and Computation*, 43(10):2353–2373, 2014. URL `http://doi.org/10.1080/03610918.2012.750354`. [p6]

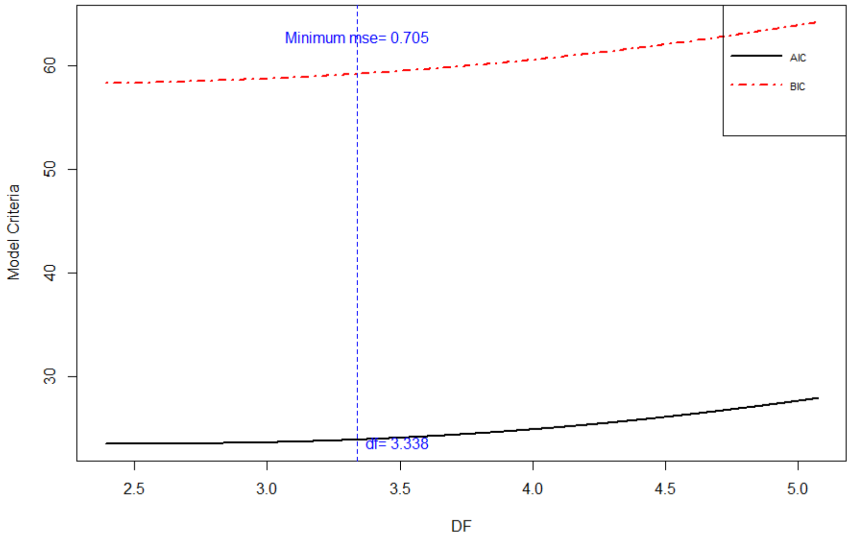

Figure 3: Information Criteria against df.

D. A. Belsley. A Guide to using the Collinearity Diagnostics. *Computer Science in Economics and Management*, 4(1):33–50, 1991. URL https://doi.org/10.1007/BF00426854. [p8]

D. A. Belsley, E. Kuh, and R. E. Welsch. *Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York, 1980. chap. 3. [p3]

S. Chatterjee and A. S. Hadi. *Regression Analysis by Example*. John Wiley & Sons, 4th edition, 2006. [p3]

E. Cule and M. De Iorio. A Semi-Automatic Method to Guide the Choice of Ridge Regression. *Annals of Applied Statistics*, arxiv:1205.0686v1, 2012. URL https://arxiv.org/abs/1205.0686v1. [p1, 6]

J. D. Curto and J. C. Pinto. The Corrected VIF (CVIF). *Journal of Applied Statistics*, 38(7):1499–1507, 2011. URL http://doi.org/10.1080/02664763.2010.505956. [p3]

A. Dissanayake and P. Wijekoon. *lrmest: Different Types of Estimators to Deal with Multicollinearity*, 2016. URL https://CRAN.R-project.org/package=lrmest. R package version 3.0. [p2]

N. R. Draper and H. Smith. *Applied Regression Analysis*. John Wiley & Sons, New York, 2nd edition, 1998. [p8]

P. Druilhet and A. Mom. Shrinkage Structure in Biased Regression. *Journal of Multivariate Analysis*, 99 (2):232–244, 2008. URL https://doi.org/10.1016/j.jmva.2006.06.011. [p4]

D. E. Farrar and R. R. Glauber. Multicollinearity in Regression Analysis: The Problem Revisted. *The Review of Economics and Statistics*, 49(1):92–107, 1967. URL http://doi.org/10.2307/1937887. [p3]

J. Fox and S. Weisberg. *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, second edition, 2011. URL http://socserv.socsci.mcmaster.ca/jfox/Books/Companion. [p3]

G. Golub, G. Wahba, and C. Heath. Generalized Cross Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, 21(2):215–223, 1979. URL http://doi.org/10.2307/1268518. [p1]

W. H. Greene. *Econometric Analysis*. Prentic Hall, New Jersey, 5th edition, 2002. [p3]

M. Gruber. *Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimator*. Marcel Dekker, Inc., New York, 1998. [p4]

R. F. Gunst and R. L. Mason. Advantages of Examining Multicollinearities in Regression Analysis. *Biometrics*, 33(1):249–260, 1977. URL http://doi.org/10.2307/2529320. [p3]

W. Hadley. *R Packages: Organize, Test, Document, and Share Your Code*. O'Reilly Media, 2015. [p1]

A. M. Halawa and M. Y. El-Bassiouni. Tests of Regression Coefficients Under Ridge Regression Models. *Journal of Statistical-Computation and Simulation*, 65(1–4):341–356, 2000. URL http://doi.org/10.1080/00949650008812006. [p6]

A. Hald. *Statistical Theory with Engineering Applications*. John Wiley & Sons, New York, 1952. [p3]

T. Hastie and R. Tibshirani. *Generalized Additive Models*. Chapman & Hall, 1990. [p6, 10]

A. Hoerl and R. W. Kennard. Ridge Regression: Biased Estimation of Nonorthogonal Problems. *Technometrics*, 12(1):55–67, 1970. URL http://doi.org/10.1080/00401706.1970.10488634. [p1, 4]

A. E. Hoerl, R. W. Kennard, and K. F. Baldwin. Ridge Regression: Some Simulations. *Communications in Statistics*, 4(2):105–123, 1975. URL http://doi.org/10.1080/03610927508827232. [p3]

M. Hubert and P. Wijekoon. Improvement of the Liu Estimator in Linear Regression Model. *Journal of Sta- tistical Papers*, 47(3):471–479, 2006. URL https://doi.org/10.1007/s00362-006-0300-4. [p4]

M. Imdadullah and D. M. Aslam. *lmridge: Linear Ridge Regression with Ridge Penalty and Ridge Statistics*, 2016a. URL https://CRAN.R-project.org/package=lmridge. R package version 1.0. [p1]

M. Imdadullah and D. M. Aslam. *liureg: Liu Regression with Liu Biasing Parameters and Statistics*, 2017. URL https://CRAN.R-project.org/package=liureg. R package version 1.1.1. [p1]

M. Imdadullah and M. Aslam. *mctest: Multicollinearity Diagnostic Measures*, 2016b. URL https://CRAN.R-project.org/package=mctest. R package version 1.1. [p2, 3]

M. Imdadullah, M. Aslam, and S. Altaf. mctest: An R Package for Detection of Collinearity Among Regressors. *The R Journal*, online published paper, 2016. URL https://journal.r-project.org/archive/accepted/imdadullah-aslam-altaf.pdf. [p3]

A. Jahufer and J. Chen. Assessing Global Influential Observations in Modified Ridge Regression. *Statistics and Probability Letters*, 79(4):513–518, 2009. URL https://doi.org/10.1016/j.spl.2008.09.019. [p4]

A. Jahufer and J. Chen. Measuring Local Influential Observations in Modified Ridge Regression. *Journal of Data Science*, 9(3):359–372, 2011. [p4]

A. Jahufer and J. Chen. Identifying Local Influential Observations in Liu Estimator. *Journal of Metrika*, 75(3):425–438, 2012. URL https://doi.org/10.1007/s00184-010-0334-4. [p4]

S. Kaçıranlar and S. Sakalhoğlu. Combining the LIU Estimator and the Principal Component Regression Estimator. *Communications in Statistics-Theory and Methods*, 30(12):2699–2706, 2001. URL http://doi.org/10.1081/STA-100108454. [p4, 5]

S. Kaçıranlar, S. Sakalhoğlu, F. Akdeniz, G. Styan, and H. Werner. A new Biased Estimator in Linear Regression and a Detailed Analysis of the Widely-Analysed Dataset on Portland Cement. *Sankhyā: The Indian Journal of Statistics, Series B*, 61(B3):443–459, 1999. [p4, 5]

B. Kan, O. Alpu, and B. Yazici. Robust Ridge and Liu Estimator for Regression Based on the LTS Estimator. *Journal of Applied Statistics*, 40(3):644–665, 2013. URL http://doi.org/10.1080/02664763.2012.750285. [p2]

M. G. Kendall. *A Course in Multivariate Analysis*. Griffin, London, 1957. pp. 70–75. [p3]

L. R. Klein. *An Introduction to Econometrics*. Prentic-Hall, Englewood, Cliffs, N. J., 1962. pp. 101. [p3]

J. Kmenta. *Elements of Econometrics*. Macmillan Publishing Company, New York, 2nd edition, 1980. pp. 431. [p1]

A. Koutsoyiannis. *Theory of Econometrics*. Macmillan Education Limited, 1977. [p3]

P. Kovács, T. Petres, and Tóth. A new Measure of Multicollinearity in Linear Regression Models. *International Statistical Review / Revue Internationale de Statistique*, 73(3):405–412, 2005. URL http://doi.org/10.1111/j.1751-5823.2005.tb00156.x. [p3]

F. Leisch. *Creating R Packages: A Tutorial*. Compstat 2008-Proceedings in Computational Statistics, Physica Verlage, Heidelberg, Germay, 2008, 2008. URL ftp://cran.r-project.org/pub/R/doc/contrib/Leisch-CreatingPackages.pdf. [p1]

K. Liu. A new Class of Biased Estimate in Linear Regression. *Communications in Statistics-Theory and Methods*, 22(2):393–402, 1993. URL http://doi.org/10.1080/03610929308831027. [p1, 4, 5, 7, 8, 10]

X.-Q. Liu. Improved Liu Estimator in a Linear Regression Model. *Journal of Statistical Planning and Inference*, 141(1):189–196, 2011. URL https://doi.org/10.1016/j.jspi.2010.05.030. [p4, 7, 8, 10]

G. S. Maddala. *Introduction to Econometrics*. Macmillan, New York, 1988. [p3]

C. L. Mallows. Some Comments on Cp. *Technometrics*, 15(4):661–675, 1973. URL http://doi.org/10.2307/1267380. [p1, 7]

D. W. Marquardt. Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics*, 12(3):591–612, 1970. URL http://doi.org/10.2307/1267205. [p3]

A. McLeod and C. Xu. *bestglm: Best Bubset GLM*, 2017. URL https://CRAN.R-project.org/package=bestglm. R package version 0.36. [p1]

D. C. Montgomery and E. A. Peck. *Introduction to Linear Regression Analysis*. John Wiley & Sons, New York, 1982. [p1]

R. H. Myers. *Classical and Modern Regression with Application*. PWS-KENT Publishing Company, 2 edition, 1986. [p1]

R. M. Özkale and S. Kaçiranlar. A Prediction-Oriented Criterion for Choosing the Biasing Parameter in Liu Estimation. *Communications in Statistics-Theory and Methods*, 36(10):1889–1903, 2007. URL http://doi.org/10.1080/03610920601126522. [p5, 7]

J. O. Rawlings, S. G. Pantula, and D. A. Dickey. *Applied Regression Analysis: A Research Tool*. Springer-Verlag, New York, 2nd edition, 1998. [p1]

S. Sakalhoğlu, S. Kaçiranlar, and F. Akdeniz. Mean Squared Error Comparisons of Some Biased Regression Estimators. *Communications in Statistics-Theory and Methods*, 30(2):347–361, 2001. URL http://doi.org/10.1081/STA-100002036. [p5]

G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. John Wiley & Sons, New Jersey, 2 edition, 2003. [p1]

C. Stein. Inadmissibility of Usual Estimator for the Mean of a Multivariate Normal Distribution. In *Proc. Third Berkeley Symp. Mathemat. Statist, Probab.*, pages 197–206, Berkeley, 1956. University of California Press. [p4]

R. C. Team. *Writing R Extensions*. R Foundation for Statistical Computing, 2015. Version R 3.2.3. [p1]

H. Theil. *Principles of Econometrics*. John Wiley & Sons, New York, 1971. [p3]

N. Torigoe and K. Ujiie. On the Restricted Liu Estimator in the Gauss-Markov Model. *Communications in Statistics–Theory and Methods*, 35(9):1713–1722, 2006. URL http://doi.org/10.1080/03610920600683754. [p4]

R. E. Tripp. *Non-stochastic Ridge Regression and Effective Rank of the Regressors Matrix*. Ph.d. thesis, Department of Statistic, Virginia Polytechnic Institute and State University., 1983. [p1]

W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer-Verlag, New York, fourth edition, 2002. URL http://www.stats.ox.ac.uk/pub/MASS4. ISBN 0-387-95457-0. [p1]

E. Walker and J. B. Birch. Influence Measures in Ridge Regression. *Technometrics*, 30(2):221–227, 1988. URL http://doi.org/10.2307/1270168. [p5]

Muhammad Imdadullah
Ph.D scholar (Statistics),
Department of Statistics
Bahauddin Zakariya University, Multan, Pakistan
ORCiD ID: 0000-0002-1315-491X
mimdadasad@gmail.com

Muhammad Aslam
Associate Professor,
Department of Statistics
Bahauddin Zakariya University, Multan, Pakistan
aslamasadi@bzu.edu.pk

Saima Altaf
Assistant Professor,
Department of Statistics
Bahaudding Zakariya University, Multan, Pakistan
drsaimaaltaf27@gmail.com