The 3rd International Conference on Structural Health Monitoring of Intelligent Infrastructure
Vancouver, British Columbia, Canada
November 13-16, 2007

REPRESENTATIVE LEAST SQUARES METHOD FOR SENSOR PLACEMENT

D.S. Li
Dalian University of Technology, China, and University of Siegen, Germany

H.N. Li
Dalian University of Technology, China

C.P. Fritzen
University of Siegen, Germany

Abstract

The ordinary least squares method to estimate parameters is indefinitely affected by outliers. Regression diagnostics and robust regression methods are proposed to tackle this problem. There are cases in which the entire data set for regression estimation is completely reliable, however, only a small subset is allowed to be actually fitted. The original whole data set may come from past experiments, or from a simulation model, and will only be partially feasible in future estimations. We propose a representative least squares method to deal with such situations. The objective of this approach is to obtain an optimal estimator with a small subset to approximate the original least squares estimator with the whole data set. The selected subset is, therefore, required to be representative and sufficiently approaching to portrait the scenario defined by the original full data set. The basic ideas of the representative least squares method are illustrated by a simple linear regression example. The algorithm for finding the representative subset is described, and finally the properties of the representative least squares estimator are discussed. An initiative excitation-based sensor placement algorithm in structural health monitoring is presented. The proposed method can be used for sensor placement for achieving the best identification of modal frequencies and mode shapes in structural health monitoring. We will discuss its applications in further works.

INTRODUCTION

Although there are many sensor placement methods, comparisons show that a good method for sensor placement in a particular application is not necessarily good for another project because optimal sensor locations are depending not only on the structure itself, but also on the cases of loadings and their positions [5,16-21]. However, only a few investigators take the influence of loading situations into consideration when a method of placing sensors is developed. The authors noted that sensors should be placed at different positions if the loadings vary with positions and frequency range during modal experiments [22-23]. It is the authors' opinion that no sensor placement methods can achieve the claimed optimum or sub-optimum without consideration of actual loading conditions for a specified structure. Based on insights gathered from the above reasoning, a representative least squares method for sensor placement is proposed. This paper presents the theory part of the method. Following works will demonstrate its applications in structural health monitoring.

Facing perturbations of outliers in ordinary least squares (OLS) method, there are two kinds of remedy solutions. The first is to use residue analysis and regression diagnostic metrics to detect the degree of abnormality [1]. However, masking effects render such points, in most cases, undetected because of their combination influence [2]. Another solution is to adopt robust regression techniques, which are designed specially to identify outliers [3]. Regression diagnostics and robust estimators share the same objective, but proceed in different directions. In regression diagnostics, the outliers are first identified, and then a routine regression is processed for the good data, whereas the robust estimators fit first the majority of the data and regards the rest points with large residues as outliers. The outliers identified by both methods are deemed to be unfaithful, and usually to be reexamined. Subsequent actions may be taken to analyze the mechanics for the sources of such abnormalities. These are the conventionally applied methods for establishing a linear model in engineering areas. Other variants of linear estimators, which are designed to minimize a subset or all of the regression residuals, fall into range of OLS and robust estimation [4].

There are, however, other engineering cases in which all the data points for regression estimation are completely trustworthy. No data points have gross errors, but only a fraction of the entire data set is permitted in regression. The original whole data set may come from past experiments, or from a simulation model, and will only be partially feasible in the future. The number of data obtainable hereafter in regression is limited and mostly preset by accessible experiment facilities. In this contribution, we propose a representative least squares method to deal with these cases. By this approach, a subset of data, which is expected to approximately characterize the entire data set, is to be selected for regression estimation. The number of data points to be reserved is, in general, more than that of the parameters, and much less than the number of all data points. The fitted data points in the selected subset are required to be representative in the proposed representative least squares method, and are sufficiently approaching to portray the scenarios of the original full data set although they are not necessarily successive. The suggested representative least squares method has great potentials in engineering applications, such as to select a limited number of sensor positions among available candidates, and also in experiment design problems[5,6].

The organization of the paper is as follows. In Section 2, we review the OLS theory and a popular robust estimator, and then present the basic ideas behind the representative least squares method, which is illustrated by a simple linear regression example. Section 3 describes the algorithm for finding the representative subset of the data through genetic algorithms. Finally, properties of the representative least squares estimator and its perspectives in engineering applications are discussed.

THEORY OF REPRESENTATIVE LEAST SQUARES METHOD

The mathematical concept of least squares dates back to Laplace, Gauss and Legendre[1-3,15] separately. The purpose of a least squares model, which is the cornerstone of classical linear theory, is to fit equations with independent coefficients to dependent responses. Its theory has long been established, and is described here for symbol explanation and to facilitate our development of the representative least squares estimator.

Ordinary Least Squares Method

The classical linear model can be expressed in matrix notation as,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad (1)$$

where $\mathbf{y}$ is an n x 1 vector of observations, $\mathbf{X}$ is an n x p design matrix, $\boldsymbol{\beta}$ is a p x 1 vector of unknown coefficients, and $\mathbf{e}$ is an n x 1 vector of unknown errors. The errors are usually assumed to be independent and identically distributed with zero mean and variance σ^2.

The objective of the classical OLS method is to minimize the summation of the squared residues $\mathbf{r}_i$, which is the difference between the i^{th} response $\mathbf{y}_i$ and its estimator $\hat{\mathbf{y}}_i$,

$$\mathbf{J}_{OLS}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \mathbf{r}_i^{\,2} = \sum_{i=1}^{n} (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2 = \sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{X}_i^{\,T}\boldsymbol{\beta})^2 \qquad (2)$$

The OLS estimator of the unknown coefficients $\boldsymbol{\beta}$ can then be obtained by differentiating the objective function $\mathbf{J}_{OLS}(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and letting it equal to zero as follows,

$$\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}, \qquad (3)$$

where the superscript ^{T} indicates matrix transposition. We assume that the design matrix $\mathbf{X}$ is of full column rank, *i.e.* $rank(\mathbf{X}) = p$, and that $\hat{\boldsymbol{\beta}}_{OLS}$ can be solely determined. For rank deficient design matrix, variable selection and principle component least squares methods may be applied [7].

The OLS estimator $\hat{\boldsymbol{\beta}}_{OLS}$ is unbiased and can be regarded as projection of the response vector $\mathbf{y}$ onto the space spanned by the columns of the design matrix $\mathbf{X}$. It is implicitly assumed in OLS that there is only measurement error in the responses $\mathbf{y}$, and no error in the design matrix $\mathbf{X}$. For cases with errors in both responses $\mathbf{y}$ and the design matrix $\mathbf{X}$, total least squares method can be applied [15].

Least Trimmed Sum of Squares Method

The OLS method is well known to be extremely sensitive to outliers, that one outlier can corrupt the estimation [1-3]. Robust estimation provides an alternative approach to the classical OLS method. The motivation is to produce estimators that are not unduly affected by outliers. Many robust estimators have been proposed, among which the least trimmed sum of squares (LTS) and the least median of squares (LMS) estimators are the most influential [3-4]. To take LTS for example, it attempts to minimize a subset of the residuals and regarded other cases as outliers.
LTS minimizes the objective function given by,

$$\mathbf{J}_{LTS}(\boldsymbol{\beta}) = \sum_{i=1}^{m} (\mathbf{r}_i^{\,2})_{i:n} \qquad (4)$$

where $(\mathbf{r}^2)_{i:n}$ is the i^{th} smallest residual when the residuals are sorted in ascending order, i.e., $(\mathbf{r}^2)_{1:n} \le (\mathbf{r}^2)_{2:n} \le \cdots \le (\mathbf{r}^2)_{n:n}$. Since *m* is the number of data points whose residuals are included in the sum, the LTS estimator basically finds a robust estimation by identifying the *n-m* points having the largest residuals as outliers, and discarding (trimming) them from the data set.

The computation of $\hat{\boldsymbol{\beta}}_{LTS}$ takes much time, and special computer program are usually employed [3]. Other robust estimators are similar to LTS, in that they minimize certain criteria function of residues for a portion of data points in order to avoid strong influence of outliers.

Representative Least Squares Method

The theories of OLS and robust estimation methods have been soundly established, and have been successfully applied to solve common engineering problems. However, there are cases in practice, where OLS, robust estimation and their variants don't cover, and therefore cannot be applied. This is the topic and focus of the current work. In

these cases, we are aware in prior that all the data points are faithful, which all contribute valuable information to our model. Let $\mathbf{Z}_i$ denote the i^{th} case [$\mathbf{Z}_i^T$, $\mathbf{y}_i$] corresponding to the i^{th} row $\mathbf{X}_i^T$ of $\mathbf{X}$ and the i^{th} element of $\mathbf{y}$. The cases $\mathbf{Z}$ may come from past experiments or simulations of an analytical model. Only a portion of all cases, say $m(n/2 \le m \le n)$, is feasible to be obtained in further experiments because of the limitation of experiment facilities or unavailability of some observations. The problem arises, then, how to select a quantity of *m* cases for a most informative future experiment. This is equally to say, how to find *m* cases to obtain the best OLS estimator in regression.

Here, 'Best' means the estimation from the *m* cases among the collections of all subsets of size *m* is the nearest to the previous OLS estimator with all the cases. It equals to minimize an objective function of a scaled distance between partial and full estimates,

$$\mathbf{J}_{RLS}(\hat{\boldsymbol{\beta}}_S) = (\hat{\boldsymbol{\beta}}_S - \hat{\boldsymbol{\beta}}_{OLS})^T \mathbf{W} (\hat{\boldsymbol{\beta}}_S - \hat{\boldsymbol{\beta}}_{OLS}) \qquad (5)$$

where $\hat{\boldsymbol{\beta}}_S$ is an OLS estimator with a subset of *m* cases $\hat{\boldsymbol{\beta}}_S = (\mathbf{X}_S{}^T \mathbf{X}_S)^{-1} \mathbf{X}_S{}^T \mathbf{y}_S$, and the subscript $_S$ denotes a set of *m* cases including possible combinations of selecting *m* out of *n* rows. $\mathbf{W}$ is a weight matrix. If $\mathbf{W}$ is an identity matrix, the objective function in Eq.(5) is simply an Euclidean distance, whereas if $\mathbf{W}$ is the inverse of the variance $\hat{\boldsymbol{\beta}}_{OLS}$ ($\text{var}(\hat{\boldsymbol{\beta}}_{OLS}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$), the objective function in Eq.(5) is a squared Mahalanobis distance. If not indicated otherwise, a squared Mahalanobis distance is applied later in this paper. A squared Mahalanobis distance is preferred because it takes into account the correlations of the cases and is scale-invariant, *i.e.* not dependent on the scale of measurements.

The criteria defined in Eq.(5) is quite different from that in Eq.(2) and Eq.(4), and is tentatively termed representative least squares (RLS) method. The definition of the RLS estimator is due to its representative feature of the whole data set as an analog to representatives of a population sample, and it reduces to an OLS estimator when the optimal subset equals to the whole given data set. As the name 'representative least squares' indicates, a subset of cases is supposed to approximately characterize the entire data set. The selected cases in RLS are not necessarily successive, but they are required to be representative and sufficiently approaching to portray the scenarios of the original whole data set. The objective function in Eq.(5) can also be rewritten in another form as follows,

$$\begin{aligned} \mathbf{J}_{RLS}(\hat{\boldsymbol{\beta}}_S) &= (\hat{\boldsymbol{\beta}}_S - \hat{\boldsymbol{\beta}}_{OLS})^T \mathbf{X}^T \mathbf{X} (\hat{\boldsymbol{\beta}}_S - \hat{\boldsymbol{\beta}}_{OLS}) = (\mathbf{X}\hat{\boldsymbol{\beta}}_S - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})^T (\mathbf{X}\hat{\boldsymbol{\beta}}_S - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS}) \\ &= \sum_{i=1}^{n} (\hat{\mathbf{y}}_{S,i} - \hat{\mathbf{y}}_{OLS,i})^2 = \sum_{i=1}^{n} \hat{\mathbf{r}}_i^{\,2} \end{aligned} \qquad (6)$$

where the $\hat{\mathbf{r}}_i^{\,2}$ is the i^{th} residue between the i^{th} response predication using a partial estimator $\hat{\boldsymbol{\beta}}_S$ and that using the OLS estimator $\hat{\boldsymbol{\beta}}_{OLS}$, $\hat{\mathbf{r}}_i = \hat{\mathbf{y}}_{S,i} - \hat{\mathbf{y}}_{OLS,i}$. The predication residue $\hat{\mathbf{r}}_i^{\,2}$ has a similar form as a normal OLS residue in Eq.(2), and measures response forecast difference between the entire data set and a subset with *m* cases. In Eq.(6), a constant σ is not included for simplicity since it does not affect the result. The objective function of RLS in Eq.(6) takes a comparable form as that of OLS in Eq.(2) and LTS in Eq.(4). The proposed RLS has its merit.

An RLS estimator $\hat{\boldsymbol{\beta}}_{RLS}$ is, then, obtained and the *m* cases are determined when the objective function in Eq.(6) is minimized. The RLS estimator will have the least predication residual among the range. In cases, it is possible to have several optimum solutions of $\hat{\boldsymbol{\beta}}_{RLS}$. The one with maximal determinant $\det(\mathbf{X}^T \mathbf{X})$ will be chosen because of its smaller variance than the others.

Moreover, the number of cases *m* retained in RLS is more than that or at least equals to that of parameters. Otherwise, no unique solution exists. The number of cases *m* is usually much less than the number of all cases, as

can be seen from the example in Section 3. In practice, the selection of *m* cases in RLS must guarantee that $\hat{\boldsymbol{\beta}}_{RLS}$ is an appropriate approximate of $\hat{\boldsymbol{\beta}}_{OLS}$. Trial and error is recommended to find a suitable quantity and to avoid loss of much information.

A SIMPLE REGRESSION EXAMPLE

While the above mentioned effective independence method has been shown to be effective, it has a crucial weakness. It fails to take into account the knowledge of the loading effects, i.e. the characteristics of the loading cases and the duration of the loadings. However, the responses of a structure are, in fact, totally dependent on the loading patterns. Given a specific loading case, there is a corresponding response pattern, in which several or many particular mode shapes dominate while other mode shapes are inactive. From the viewpoints of the authors, different loading situations for an identical structure demand different deployment configurations of the sensors to be implemented. The modal participation factors are thus of great significance when designing a sensor topology.

Table 1. Two sets of data for the regressions fitted in Fig.1 and Fig.2

No. of data point		1	2	3	4	5	6	7	8	9	10	11
Case 1 &2	X	10	8	13	9	11	14	6	4	12	7	5
Case 1	Y	8.04	6.95	7.58	8.81	8.33	9.96	7.24	4.26	10.84	4.82	5.68
Case 2	Y	7.46	6.77	12.74	7.11	7.81	8.84	6.08	5.39	8.15	6.42	5.73
Leverage values		0.10	0.10	0.24	0.09	0.13	0.32	0.17	0.32	0.17	0.13	0.24

To illustrate the basic ideas and fundamentals of the proposed RLS in Section 2, a simple regression example is employed. Table 1 shows the data for drawing Fig.1 and Fig.2. The data is directly borrowed from Example 5.1 in Weisberg[8]. Fig.1 is drawn from the data set of Case 1 in Table 1, and Fig.2 from the data set of Case 2. The *x* coordinates for both cases are the same, and the difference resides in the response *y*.

Figures 1 and 2 show two different scenarios. In Fig.1, all the data are randomly scattered, and can be said to approximate a normal distribution. However in Fig.2, almost all the data except one (data point 13) are located on a robust regression line (the red dotted line with *m* of 13). The regression line of OLS is the magenta centered line in Fig.2. The robust estimator deviates much from that of the OLS because of their difference in the judgment of all cases. The OLS used all cases for regression, whereas the robust estimation use only 13 out of the total 14 points (without data point 13).

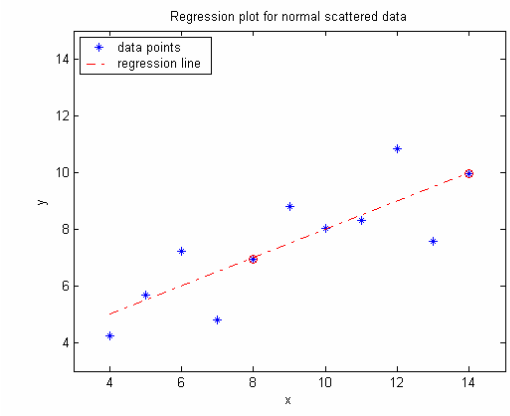

Figure 1. Regression line for normal scattered data

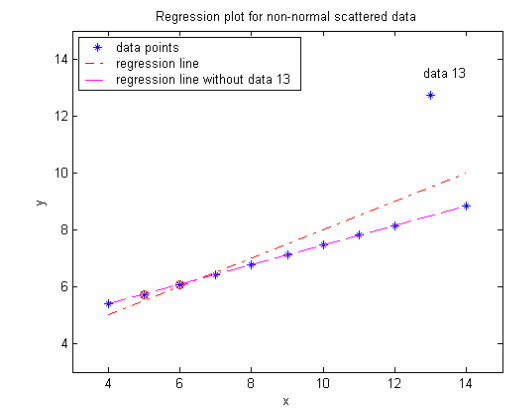

Figure 2. Regression line for non-normal distributed data

We refer to Fig.1 and Fig.2 for explaining the representative feature of RLS method. If *m* equals to 2 in RLS, the selected data points are indicated in both figures by circles. For Case 1, data points No.2 and No.6 are finally

selected, and the regression line fitted by the RLS approximates that by OLS perfectly. For Case 2, data points No.7 and No.11 are selected. This simple regression example demonstrates the basic ideas behind RLS, *i.e.*, to choose the optimal approximaton of the coefficients estimation provided that the number of cases is limited.

The results of RLS can be compared with the diagnostic measures of OLS. The 5th row in Table 1 shows the leverage values of all the data points, which are the diagonal terms of the projection matrix (or hat matrix, $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$). The leverage values indicate the distance of a data point from the bulk of the entire data, and measures their relative importance in fitting the OLS regression line. For instance, data points 4, 5, and 13,14 have large leverage values since they locate near the edge of data set in the sense of *x* coordinate. In addition, data point 13 is separated apparently from all other data points, and has pulled the regression line near itself. The OLS, as well as the robust estimation will first reject data point 13, whereas RLS may keep it. The RLS views from another perspective, and regards that all data points are faithful to reflect the reality, and that a subset has to be chosen to achieve optimal approximation of the coefficient estimation using the whole data set. Therefore, both OLS and robustation do not contribute much for the insight of RLS, and RLS has the right of its own interests.

COMPUTATION OF THE REPRESENTATIVE LEAST SQUARES METHOD

The RLS estimator for a certain data subset can be computed by the stable and efficient Cholesky decomposition since the matrix $\mathbf{X}^T\mathbf{X}$ is symmetric and positive definite [9]. The RLS criterion is not monotonically non-decreasing with *m*, and the RLS fit has to be found by examining all subsets of size *m*. Although many numerical methods such as forward selection, backward elimination, stepwise, and all subset methods in literature [4,9], can be consulted, no one guarantees to achieve a global optimum. In addition, branch and bound algorithms can be feasibly evaluated only if the data set size is small enough. The RLS is, in principle, a problem of combination known as integer programming in optimization. There is no deterministic algorithm, and search methods have to be used. For large problems, an exhaustive search is impractical. We adapt a genetic algorithm to the computation of RLS.

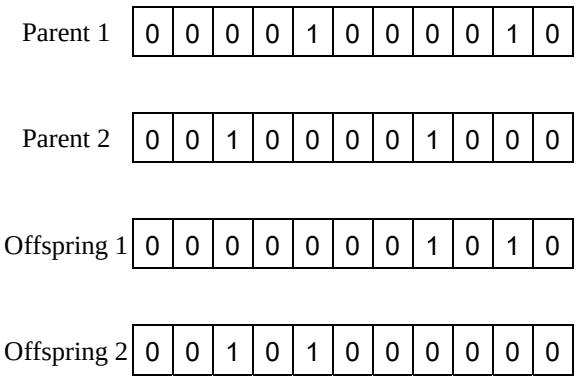

Figure 3. Sketch of crossover template

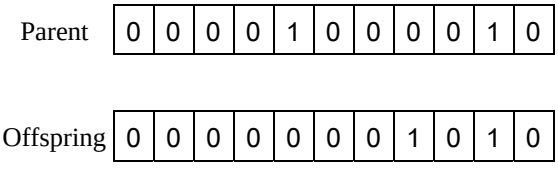

Figure 4. Sketch of mutation scheme

Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover [10-11]. A tailored genetic algorithm is implemented as a computer simulation in which the genome evolves toward better solutions. The genome or chromosomes are a population of abstract representations of candidate solutions to an RLS minimization problem. Each solution is called an individual. Solutions are represented in binary as strings of 0s and 1s in our implementation. The evolution starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of each individual in the population is evaluated. A predefined amount of individuals are stochastically selected according to rose wheel from the current population based on their fitness, and are possibly recombined (crossover) or even mutated to form a new population. The new population is then used in the next iteration of the algorithm.

Genetic algorithms, categorized as global search heuristics, are expected to find the global optimum solution [12]. We describe in follows briefly the coding, crossover and mutation parts of our algorithm. Other common parts, such as initialization, selection, are similar to a general GA approach.

Coding Part

A binary coding is embedded, and the coding length is n (the number of rows of the design matrix). If the ith bit of a string is 1, then the ith row of the design matrix is selected. The number of bits 1 in a string equals to m (the number of rows of the design matrix to be selected). For instance, a string 000010000010 indicates that the fifth and tenth rows are selected, and that m equals 2, and n equals to 11 respectively.

Crossover Part

A crossover template is tailored for the RLS problem. First, two parent strings are compared and searched for a length of template, in which both strings have the same number of 1s. Although the length of a crossover template variates with different parent combinations, such a template is alwasys guranteed to be found. The two parent strings exchange the selected section of the crossover template, and give birth to two offsprings. Fig.3 illustrate this process. The use of crossover template assures the consitency of the number of 1s in an offspring string inherited from parents.

Table 2. Cases selected for RLS estimators using genetic algorithm

Number of cases	Data cases selected by genetic algorithm	Values of the objective function
2	2,6	0.02695
3	1,2,6	0.00279
4	1,2,5,6	0.02872
5	2,5,6,7,8	0.02661
6	3,4,5, 6, 8, 9	0.02005
7	3,4,6,7,8,9,10	0.01360
8	3,4,5,7,8,9,10,11	0.00106
9	2,3,4,5,7,8,9,10,11	0.00010
10	2,3,4,5,6,7,8,9,10,11	0.00019

An epoch approach is embedded in the algorithm[13-14]. The idea is based on the observation of brief bursts of change between epochs in natural evolution, and innovation is then introduced. In this way, epoch behaviour aids to enhence genome diversity in a generation and as well to avoid premature. Moreover, the seach spead is improved by using elitism, by which the best member in a generation is ensured to propagate into the next generation. In this manner, the possibility of the algorithm to locating the true global optimum is increased. These genetic algorithm techniques have been coded into a matlab program, which can be obtained by contacting the first author.

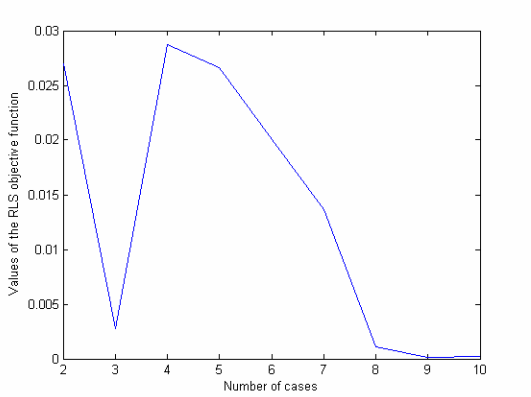

Figure 5. Values of the RLS objective function for increasing number of cases

Table 2 shows the computation results using the genetic algorithm described above. The number of cases regressed for RLS estimators range from 2 to 10 for Case 1 of Table 1. The data cases selected in a smaller data subset are not necessarily included in a relative larger dataset. For instance, The Case 2 are chosen for *m* of 2,3,4,and 5 respectively, but not selected for m of 6. It is noted that the search result using the genetic algorithm for *m* of 2 agrees exactly with that of a traditional combinational search, which is simple for this small-scale problem. This validates the effectiveness of the techniques used in the code genetic algorithm program.

The values of the objective function according to Eq.(6) are listed in the 3rd column of Table 2, and are also drawn in Fig.5. The curve has a decrease trend with the increase of the number of the cases *m*. This is evident since the RLS estimator with more cases is expected to approximate the OLS estimator with all the cases with more accuracy [4]. However, the curve is not assured to be monotonically decreasing as stated in the first paragraph of this section. For instance, the RLS estimator with 3 cases is better than that with neighboring cases. The fluctuation of objective function may indicate certain potential applications in choosing an optimum number of cases for RLS, which deserves further research.

CONCLUSIONS

A representative least squares method is proposed, and shows its potential for sensor placement in structural health monitoring. The objective function of RLS is similar to that of ordinary least squares method and that of least trimmed squares method. The RLS method consists of finding a data subset with a definite number of observations. The subset has the property that the distance between the OLS estimator with the entire data set and that with this subset is minimal. The basic ideas are exposed by a simple linear regression example, and properties of the OLS estimator are discussed.

A combinoselective genetic algorithm is specially designed to compute the RLS estimator. Techniques such as crossover template, epoch mutation and elite selection are embedded to achieve stable convergence and efficiency. The tailored Matlab program is to share in the technical community.

ACKNOWLEDGEMENTS

The authors acknowledge for the joint support of China Natural Science Foundation (No. 50408031) and German Academic Exchange Service (DAAD) for the present work.

REFERENCES

[1] Cook, R. D. and Weisberg, S. Residuals and Influence in Regression. Chapman and Hall, London, 1982.

[2] Peno D., Yohai V.J., A fast procedure for robust estimation and diagnostics in large regression problems, J. Am. Statist. Assoc. 1999(94):434-445.

[3] Rousseeuw P. J. and Leroy A. M., Robust Regression and Outlier Detection, Wiley, New York,1986.

[4] Olive, D.J. Applied Robust Statistics, unpublished document available from (http://www.math.siu.edu/olive) ,2006.

[5] Li D.S., Li H.N., and Fritzen, C.P. On the Physical Significance of the Norm Based Sensor Placement Method, in Proceedings of the Third European Workshop on Structural Health Monitoring, Alfredo Güemes Eds., Granada, Spain, July 5-7, 2006, pp. 1542-1550.

[6] Peter Goos, The Optimal Design of Blocked and Split-Plot Experiments, Springer, New York, 2002.

[7] Jürgen G., Linear regression . Berlin,Springer, 2003.

[8] Weisburg, S. Applied Linear Regression, Wiley, New York, 1980.

[9] Åke Björck, Numerical Methods for Least Squares, SIAM, 1996.

[10] Coley, D.A., An Introduction to Genetic Algorithms for Scientists and Engineers, World Scientific Publishing Company, 1997.

[11] Guo, H.Y., Zhang, L.,Zhang, L. L. and Zhou, J.X., Optimal placement of sensors for structural health monitoring using improved genetic algorithms, Smart materials and structures, 2004(13):528-534.

[12] Mitchell, M., An Introduction to Genetic Algorithms, MIT Press, Cambridge, 1996.

[13] Goldberg, D.E. The Design of Innovation: Lessons from and for Competent Genetic Algorithms Boston, MA: Kluwer Academic Publishers, 2002.

[14] van Nimwegen, E., Crutchfield, J.P., and Mitchell,M., Statistical dynamics of the Royal Road genetic algorithm, Theoretical Computer Science, 1999,229(1-2): 41-102.

[15]Nievergelt Y., A tutorial history of least squares with applications to astronomy and geodesy, Journal of Computational and Applied Mathematics, 2000,121(1):37-72.

[16]F.E.Udwadia, Methodology for optimum sensor locations for parameter identification in dynamic systems. Journal of Engineering Mechanics 120(2) (1994) 368-390.

[17]D.C.Kammer, Sensor placement for on-orbit modal identification and correlation of large space structures. J Guidance, Control Dynam 14 (1991) 251–259.

[18] M.Papadopoulos and E.Garcia, Sensor placement methodologies for dynamic testing. AIAA Journal 36(2) (1998) 256-263.

[19] C.R.Pickrel, A practical approach to modal pretest design. Mechanical Systems and Signal processing 13(2) (1999) 271-295.

[20] C.Papadimitriou, Optimal sensor placement methodology for parametric identification of structural systems, Journal of Sound and Vibration 278(4) (2004) 923-947. doi:10.1016/ j.jsv. 2003. 10. 063.

[21]E.Heredia-Zavoni, and L.Esteva, Optimal instrumentation of uncertain structural systems subject to earthquake ground motions, Earthquake Engineering and Structural Dynamics 27(4) (1998) 343-362.

[22] D.S. Li, H.N. Li, and C.P.Fritzen, The connection between Effective Independence and Modal Kinetic Energy methods for sensor placement, submitted to Journal of Sound and Vibration.

[23] Hong-Nan Li, Dong-Sheng Li and Gang-Bing Song, Recent applications of fiber optic sensors to health monitoring in civil engineering, Engineering Structures, 2004;26(11):1647-1657.