

## Some Modifications to Calculate Regression Coefficients in Multiple Linear Regression

S. M. Sajadifar<sup>1\*</sup>, M. Allameh<sup>2</sup>

<sup>1</sup> Department of Industrial Engineering, University of Science and Culture, P. O. Box: 13145-871, Tehran, Iran  
sajadifar@usc.ac.ir

<sup>2</sup> Department of Mathematics, Azad University, P. O. Box: 81595-158, Khorasgan, Isfahan, Iran  
allameh@khuisf.ac.ir

### ABSTRACT

In a multiple linear regression model, there are instances where one has to update the regression parameters. In such models as new data become available, by adding one row to the design matrix, the least-squares estimates for the parameters must be updated to reflect the impact of the new data. We will modify two existing methods of calculating regression coefficients in multiple linear regression to make the computations more efficient. By resorting to an initial solution, we first employ the *Sherman-Morrison* formula to update the inverse of the transpose of the design matrix multiplied by the design matrix. We then modify the calculation of the product of the transpose of design matrix and the design matrix by the Cholesky decomposition method to solve the system. Finally, we compare these two modifications by several appropriate examples.

**Keywords:** Regression, Inverse matrix, Cholesky decomposition, Sherman-Morrison - Woodbury formula

### 1. INTRODUCTION

There are some studies in the literature that employ the Sherman-Morrison formula and the corresponding generalization given by the Sherman-Morrison-Woodbury formula to express the inverse of a matrix after a small rank perturbation in terms of the inverse of the original matrix.

Hager (1989) presents the history of Sherman-Morrison-Woodbury formulas and discusses some of its applications to statistics, networks, structural analysis, asymptotic analysis, optimization, and partial differential equations. He notes that the Sherman-Morrison-Woodbury formulas express the inverse of a matrix after a small rank perturbation in terms of the inverse of the original matrix. Hager in his paper surveys the history of these formulas and examines some applications where these formulas are helpful. Especially he presents an application of the modified formula to estimate some parameters in a linear model as new data are received.

---

\* Corresponding Author

Egidia and Maponia (2006) proposed a new direct method to solve linear systems. Their method is based on the Sherman–Morrison formula and uses a finite iterative scheme. As they say, the Sherman–Morrison formula and the corresponding generalization given by the Sherman–Morrison–Woodbury formula have been used in several applications, such as the solution of special linear systems, the solution of linear systems arising in mathematical programming problems and other interesting applications. There are many aspects of Sherman-Morrison-Woodbury formula which have been addressed in the literature (see: Fitzpatrick and Murphy , 1993, Kok Song Chua, 2003 and Lai and Vemuri, 2003).

Cholesky decomposition, on the other hand, is widely used for solving linear systems. Pourahmadi, et. al (2007) present a method for simultaneous modeling of Cholesky decomposition of several covariance matrices. They highlight the conceptual and computational advantages of the unconstrained parameterization of the Cholesky decomposition and compare the results with those obtained by the classical spectral (eigenvalue) and variance-correlation decompositions. Some other studies related to employing the Cholesky decomposition for solving linear systems are Kaci et. al (2001), Khamis and Kepler (2002), and Huang et al (2006).

Many of the results derived for the simple linear model can be extended to the general linear case. It is not possible to develop the general model conveniently without resorting to matrix notation. We will denote the transpose of an arbitrary matrix  $\mathbf{A}$  by  $\mathbf{A}^t$  and if  $\mathbf{A}$  is a square nonsingular matrix, then we denote its inverse by  $\mathbf{A}^{-1}$ .

Consider the linear regression model

$$\mathbf{Y} = \beta_0 \mathbf{x}_0 + \beta_1 \mathbf{x}_1 + \dots + \beta_p \mathbf{x}_p + \boldsymbol{\varepsilon} \quad (1)$$

and assume that a response  $y_i$  is observed at the values  $x_{i0}, x_{i1}, \dots, x_{ip}$ ,  $i = 1, \dots, n$  with  $n > p + 1$ . That is, assume that

$$E(y_i) = \sum_{j=0}^p \beta_j x_{ij}, \quad \text{Var}(y_i) = \sigma^2, \quad \text{Cov}(y_i, y_k) = 0, \quad i \neq k$$

It is possible to reformulate the model in matrix notation as follows:

$$E(\mathbf{Y}) = \mathbf{X}\mathbf{B} \quad \mathbf{V} = \sigma^2 \mathbf{I}$$

where  $\mathbf{I}$  is the  $n \times n$  identity matrix, and  $\mathbf{Y}$ ,  $\mathbf{B}$ , and  $\mathbf{X}$  are

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{10} & \dots & x_{1p} \\ \vdots & & \vdots \\ x_{n0} & \dots & x_{np} \end{pmatrix}$$

The least-square (*LS*) estimates are the values  $\beta_j = \hat{\beta}_j$  that minimize the quantity

$$S = \sum_{i=1}^n \left[ y_i - \sum_{j=0}^p \beta_j x_{ij} \right]^2 = (\mathbf{Y} - \mathbf{XB})^t (\mathbf{Y} - \mathbf{XB})$$

If we set the first partial derivatives of  $S$  with respect to  $\beta_j$ 's,  $j=0,1,\dots,p$ , equal to zero and solve the resulting system of equations, then we obtain the  $LS$  estimates. Specifically, we should solve the following set of equations:

$$\frac{\partial}{\partial \beta_k} S = \sum_{i=1}^n 2 \left[ y_i - \sum_{j=0}^p \beta_j x_{ij} \right] (-x_{ik}) = 0, \quad k = 0,1,\dots,p$$

This system of equations is linear in  $\beta_j$ 's, and it is conveniently expressed in matrix notation as

$$\mathbf{X}^t \mathbf{Y} = \mathbf{X}^t \mathbf{X} \cdot \mathbf{B} \quad (2)$$

Note that if the matrix  $\mathbf{X}^t \mathbf{X}$  is nonsingular, then there exists a unique solution of the form

$$\mathbf{B} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y} \quad (3)$$

Now suppose we receive new data that gives us the relation  $\bar{c} \mathbf{B} = y$ . The new  $\mathbf{X}$  matrix and  $\mathbf{Y}$  vector corresponding to the additional data are  $\mathbf{X}^{\text{new}} = \begin{bmatrix} \mathbf{X} \\ \bar{c} \end{bmatrix}$  and  $\mathbf{Y}^{\text{new}} = \begin{bmatrix} \mathbf{Y} \\ y \end{bmatrix}$  respectively .

If  $\mathbf{W}$  denotes the product  $\mathbf{X}^t \mathbf{X}$  that is inverted in (3), then the new  $\mathbf{W}$  corresponding to the new  $\mathbf{X}$  can be expressed as:

$$\mathbf{W}^{\text{new}} = \mathbf{W}^{\text{old}} + \bar{c}' \bar{c} \quad (4)$$

Therefore, after adding one row to the design matrix, relations (2) and (3) will be changed to:

$$(\mathbf{X}^{\text{new}})^t \mathbf{Y}^{\text{new}} = \mathbf{W}^{\text{new}} \cdot \mathbf{B} \quad (5)$$

and

$$\mathbf{B} = (\mathbf{W}^{\text{new}})^{-1} (\mathbf{X}^{\text{new}})^t \mathbf{Y}^{\text{new}} \quad (6)$$

Unless indicated otherwise, we will assume that  $\mathbf{W} = \mathbf{X}^t \mathbf{X}$  is nonsingular. Of course, a more basic assumption would be that  $\mathbf{X}$  has full rank (for more detail see Bain and Engelhardt ,1992 or Neter et. al, 1996). Thus, one can easily find that  $\mathbf{X}^{\text{new}}$  has also full rank. Hence the matrix  $\mathbf{W}^{\text{new}} = (\mathbf{X}^{\text{new}})^t \mathbf{X}^{\text{new}}$  is nonsingular (Hoffman, 1971).

To solve this new system, which is obtained after adding new data, one can either evaluate  $(\mathbf{W}^{\text{new}})^{-1}$  by the inverse matrix modification formula which will be stated in the next section and

from (6) directly obtain the regression coefficients or without evaluating  $(\mathbf{W}^{new})^{-1}$ , directly from (5) and by decomposing  $\mathbf{W}^{new}$  to the product of the lower and upper triangular matrices solve the system and find the regression coefficients by some simple matrix operations. In this paper we employ the other methods to achieve this purpose (Gloub and Von Loan, 1998 and Rencher, 2000).

To the best of our knowledge, no work has been done on numerical comparison between Cholesky decomposition method and Sherman-Morrison formula for solving linear systems. In this paper, by some appropriate examples, we compare these two methods for updating the least-square estimates for the parameters of multiple linear regression models and show that the Sherman-Morrison formula has considerably more numerical error than the Cholesky decomposition method.

In the next two subsections we will present the Sherman-Morrison formula and the Cholesky decomposition method. In section 2 we will present some useful algorithms followed by two appropriate numerical examples in section 3.

### 1.1. Sherman-Morrison and Woodbury Formulas

Hager (1989) discusses various applications of the Sherman-Morrison and Woodbury formulas. He states the Woodbury and the Sherman-Morrison formula as follows:

If both  $A$  and  $\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U}$  are invertible, then  $\mathbf{A} - \mathbf{U}\mathbf{V}$  is invertible and we have

$$(\mathbf{A} - \mathbf{U}\mathbf{V})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1} \quad (7)$$

Suppose that  $\mathbf{U}$  is  $n \times m$  and  $\mathbf{V}$  is  $m \times n$ .

We see that (7) provides a formula for the inverse of a matrix after it is modified by  $m$  rank 1 corrections. Notice that the matrix  $\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U}$  is  $m \times m$ . Formula (7) which is frequently called the Woodbury formula is useful in situations where  $m$  is much smaller than  $n$  and the structure of  $\mathbf{A}$  is “nice” so that the effort involved in evaluating the correction  $\mathbf{A}^{-1}\mathbf{U}(\mathbf{I} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$  relative to the effort involved in inverting a general  $n \times n$  matrix is small (Hager, 1989).

In the special case where  $\mathbf{U}$  is a column vector  $\bar{u}$  and  $\mathbf{V}$  is a row vector  $\bar{v}$ , (7) simplifies to

$$(\mathbf{A} - \bar{u}\bar{v})^{-1} = \mathbf{A}^{-1} + \frac{1}{(1 - \bar{v}\mathbf{A}^{-1}\bar{u})} \mathbf{A}^{-1}\bar{u}\bar{v}\mathbf{A}^{-1} \quad (8)$$

Equation (8) is often called the Sherman-Morrison formula. Based on the study of Hager (1989), equation (7) appeared in several papers before Woodbury’s report (1950) while (8) is actually a formula given by Bartlett (1951).

In the context of our modifications, we can replace the vectors  $\bar{u}$  and  $\bar{v}$  in (8) with the vectors “ $-\bar{c}^t$ ” and  $\bar{c}$  respectively and replace the matrix  $\mathbf{A}$  in (8) with the matrix  $\mathbf{W}$  to find from (8) that

$$(\mathbf{W} - (-\bar{c}^t\bar{c}))^{-1} = \mathbf{W}^{-1} + \frac{1}{(1 - \bar{c}\mathbf{W}^{-1}(-\bar{c}^t))} \mathbf{W}^{-1}(-\bar{c}^t)\bar{c}\mathbf{W}^{-1}$$

Or

$$(\mathbf{W}^{new})^{-1} = (\mathbf{W} + \bar{c}^t \bar{c})^{-1} = \mathbf{W}^{-1} - \frac{1}{(1 + \bar{c} \mathbf{W}^{-1} \bar{c}^t)} \mathbf{W}^{-1} \bar{c}^t \bar{c} \mathbf{W}^{-1} \quad (9)$$

where  $\bar{c}$  is the vector of the new data. Since  $\mathbf{W}^{new}$  is nonsingular then it is invertible and the relation (9) is just right.

### 1.2. Cholesky Decomposition Method

This method converts the symmetric matrix  $\mathbf{W}$  or  $\mathbf{W}^{new}$  to a lower triangular matrix  $\mathbf{L}$  such that  $\mathbf{W} = \mathbf{L}\mathbf{L}^t$ . With this conversion the system depicted in (2) (or (5)) can be easily solved. This method is based on the following definitions and theorem.

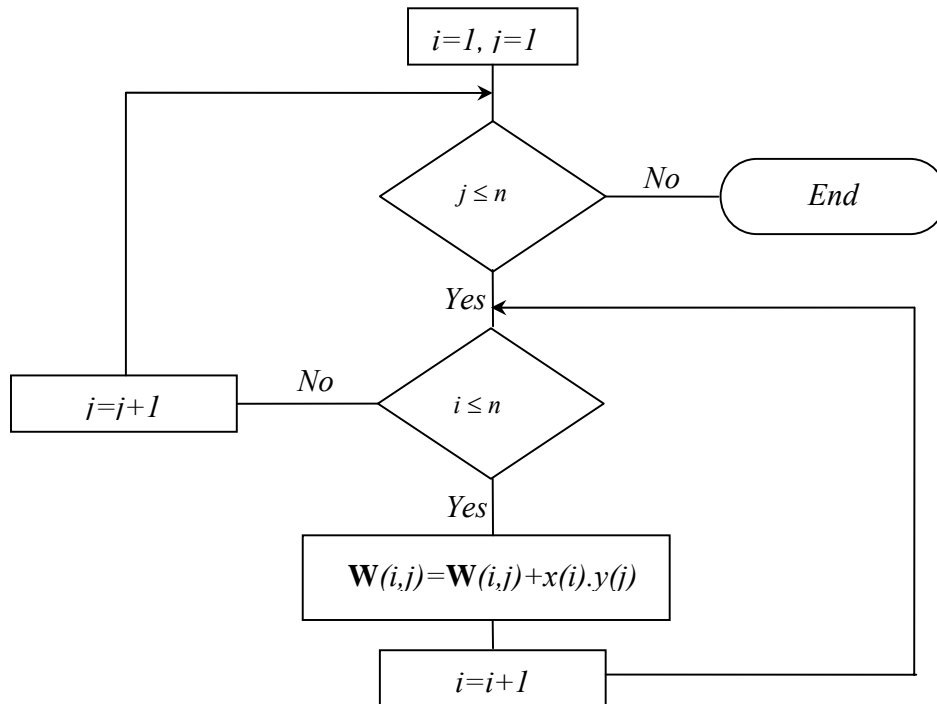


Figure 1: Flowchart of updating the matrix  $\mathbf{W}$

**Definition:** A (complex)  $n \times n$  matrix  $\mathbf{A}$  is said to be positive definite if

- $\mathbf{A} = \mathbf{A}^H$ , i.e  $\mathbf{A}$  is a Hermitian matrix.
- $x^H \mathbf{A} x > 0$  for all  $x \in \mathbf{C}^n$   $x \neq 0$

**Theorem:** (Stoer and Bulirsch, 1993) For each  $n \times n$  positive definite matrix  $\mathbf{A}$  there is a unique  $n \times n$  lower triangular matrix  $\mathbf{L}$  ( $l_{ik} = 0$  for  $k > i$ ) with  $l_{ii} > 0$ ,  $i = 0, 1, \dots, n$ , satisfying  $\mathbf{A} = \mathbf{L}\mathbf{L}^H$ . If  $\mathbf{A}$  is real, so is  $\mathbf{L}$ .

## 2. ALGORITHMS

In this section we will examine two algorithms for updating the design matrix and decomposing the matrix  $\mathbf{W}$  to lower and upper triangular matrices (for more details see Golub and Von Loan (1989)).

**Algorithm 1:** Updating the general matrix  $\mathbf{W}$  for computing  $\mathbf{W}^{new}$

Figure 1 shows the algorithm flowchart for computation of the outer product update through which we can compute  $\mathbf{W}^{new}$  (see (4)).

**Algorithm 2.** Cholesky decomposition:

Figure 2 shows the flowchart of the algorithm for decomposing a Hermitian positive definite matrix  $\mathbf{A}_{n \times n}$ .

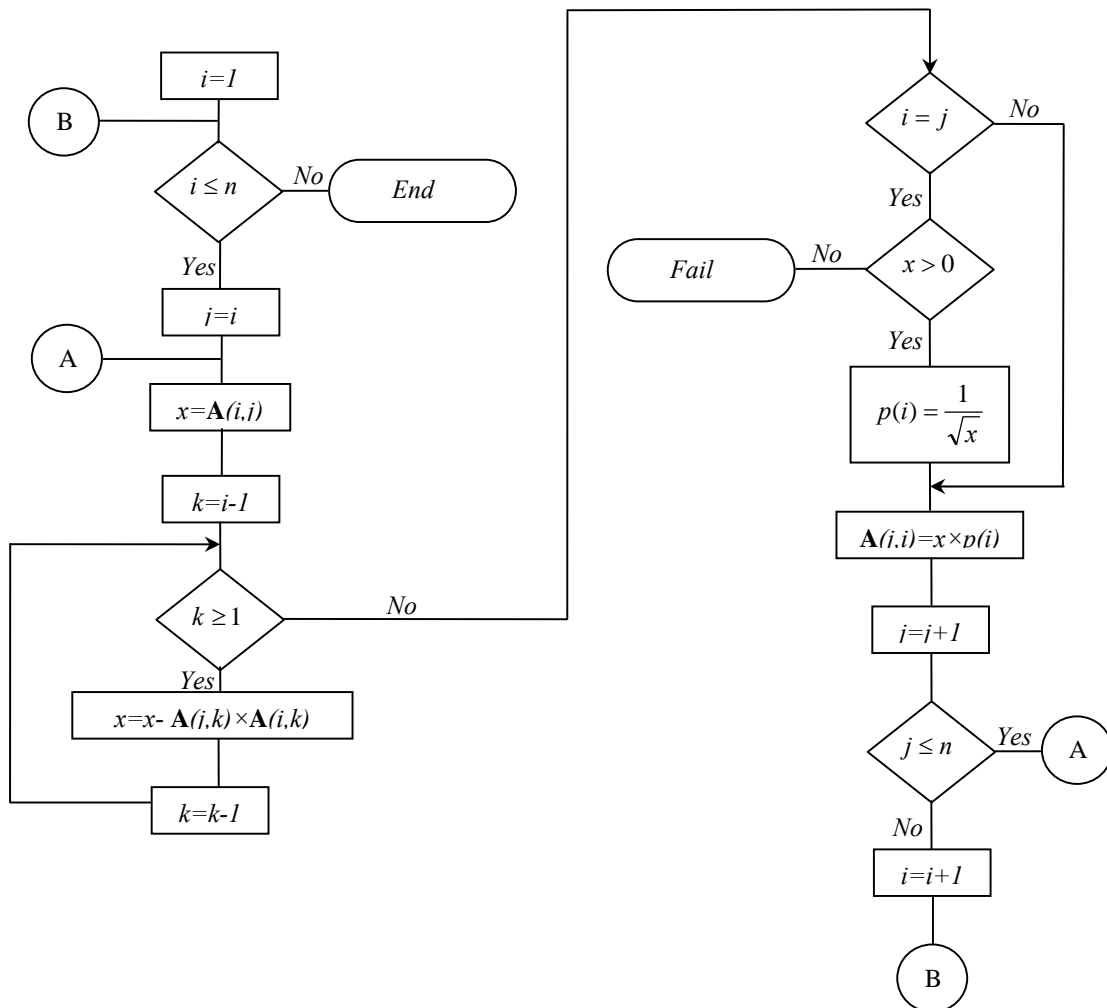


Figure 2: Flowchart of decomposing a Hermitian positive definite matrix  $\mathbf{A}_{n \times n}$

### 3. NUMERICAL EXAMPLES

In this section we present two appropriate examples. We name the initial matrix  $\mathbf{X}_1$  and once a row is added to it we arrive at matrix  $\mathbf{X}_2$ . All calculations are performed with four digits of accuracy for each of the cases under study.

**Example 1.** Let  $\mathbf{X}_1 = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$  and  $\mathbf{X}_2 = \begin{bmatrix} 2 & -1 \\ 3 & 1 \\ 4 & 0 \end{bmatrix}$ , then we have  $\mathbf{X}_1^t \mathbf{X}_1 = \mathbf{W} = \begin{bmatrix} 13 & 1 \\ 1 & 2 \end{bmatrix}$  and

$$(\mathbf{X}_1^t \mathbf{X}_1)^{-1} = \mathbf{W}^{-1} = \begin{bmatrix} 0.08 & -0.04 \\ -0.04 & 0.52 \end{bmatrix}.$$

Now let exact  $\mathbf{B}$  be  $\mathbf{B} = \begin{bmatrix} 2.005 \\ -4.0001 \end{bmatrix}$ , then  $\mathbf{X}_1^t \mathbf{Y}$  which we call  $\mathbf{R}$  will be equal to  $\begin{bmatrix} 22.0649 \\ -5.9952 \end{bmatrix}$ . So we will have

$$\mathbf{W}^{-1} \cdot \mathbf{R} = \begin{bmatrix} 2.005 \\ -4.0001 \end{bmatrix}.$$

Now with Cholesky decomposition for solving the system depicted in (2) by algorithm 2 one gets

$$\mathbf{X}_1^t \mathbf{X}_1 = \mathbf{W} = \begin{bmatrix} 3.6056 & 0 \\ 0.2774 & 1.3868 \end{bmatrix} \begin{bmatrix} 3.6056 & 0.2774 \\ 0 & 1.3868 \end{bmatrix} = \mathbf{L} \mathbf{L}^t$$

So,

$$\mathbf{L} \mathbf{y} = \mathbf{R}, \mathbf{y} = \begin{bmatrix} 6.1196 \\ -5.5471 \end{bmatrix}, \mathbf{L}^t \mathbf{B}^{\text{Chol.}} = \mathbf{y}, \text{ then } \mathbf{B}^{\text{Chol.}} = \begin{bmatrix} 2.005 \\ -3.9999 \end{bmatrix}.$$

Comparing the real  $\mathbf{B}$  and  $\mathbf{B}$  from Cholesky decomposition method denoted by  $\mathbf{B}^{\text{Chol.}}$ , we can find a small error. Now for  $\mathbf{X}_2$  by using algorithm 1, we have  $\mathbf{X}_2^t \mathbf{X}_2 = \mathbf{W}^{\text{new}} = \begin{bmatrix} 29 & 1 \\ 1 & 2 \end{bmatrix}$  and if the

exact  $\mathbf{B} = \begin{bmatrix} 1.2436 \\ -0.0101 \end{bmatrix}$ , then  $\mathbf{R}$  will be  $\begin{bmatrix} 36.0543 \\ 1.2234 \end{bmatrix}$ .

So from (9),

$$(\mathbf{X}_2^t \mathbf{X}_2)^{-1} = (\mathbf{W}^{\text{new}})^{-1} = \begin{bmatrix} 0.08 & -0.04 \\ -0.04 & 0.52 \end{bmatrix} - \frac{1}{1 + \frac{32}{25}} \begin{bmatrix} 0.32 \\ -0.16 \end{bmatrix} \begin{bmatrix} 0.32 & -0.16 \end{bmatrix} = \begin{bmatrix} 0.0351 & -0.0175 \\ -0.0175 & 0.5088 \end{bmatrix}$$

Which leads to

$$\mathbf{B}^{S-M} = (\mathbf{W}^{\text{new}})^{-1} \mathbf{R} = \begin{bmatrix} 0.0351 & -0.0175 \\ -0.0175 & 0.5088 \end{bmatrix} \begin{bmatrix} 36.0543 \\ 1.2234 \end{bmatrix} = \begin{bmatrix} 1.2441 \\ -0.0085 \end{bmatrix}$$

where  $\mathbf{B}^{S-M}$  denotes the solution of the linear system depicted by (6) whose corresponding  $(\mathbf{W}^{\text{new}})^{-1}$  comes from the Sherman-Morrison formula. and by the Cholesky decomposition method we have

$$\mathbf{X}_2^t \mathbf{X}_2 = \mathbf{W}^{\text{new}} = \begin{bmatrix} 5.3852 & 0 \\ 0.1857 & 1.4020 \end{bmatrix} \begin{bmatrix} 5.3852 & 0.1857 \\ 0 & 1.4020 \end{bmatrix} = \mathbf{L}\mathbf{L}^t.$$

$$\text{So, } \mathbf{L}\mathbf{y} = \begin{bmatrix} 36.0543 \\ 1.2234 \end{bmatrix}, \text{ then } \mathbf{y} = \begin{bmatrix} 6.6951 \\ -0.0142 \end{bmatrix} \text{ and since } \mathbf{L}^t \mathbf{B}^{\text{Chol.}} = \mathbf{y}, \text{ thus } \mathbf{B}^{\text{Chol.}} = \begin{bmatrix} 1.2436 \\ -0.0101 \end{bmatrix}.$$

If  $\mathbf{e}^{S-M} = |\mathbf{B} - \mathbf{B}^{S-M}|$  and  $\mathbf{e}^{\text{Chol.}} = |\mathbf{B} - \mathbf{B}^{\text{Chol.}}|$  denote the errors corresponding to the Sherman-Morrison formula and Cholesky decomposition method respectively, we will find in example 1 after receiving new data, that  $\mathbf{0} = \mathbf{e}^{\text{Chol.}} < \mathbf{e}^{S-M}$ .

**Example 2.** Let

$$\mathbf{X}_1 = \begin{bmatrix} 2.4752 & -0.1896 & 1 & 1.0073 & 4.7896 \\ -6.7352 & -8.96 & 4.2356 & -0.0001 & 62.0079 \\ 47.7868 & -35.5691 & -23.7998 & 41.0973 & 15.3714 \\ -14.6933 & -32.1950 & -19.0205 & 23.0820 & 0.4232 \\ 45.1882 & 13.2548 & 28.6621 & -34.2043 & -23.2021 \end{bmatrix}.$$

So if  $\mathbf{B}$  is

$$\mathbf{B} = [1.07 \quad -0.6742 \quad 2 \quad -0.7352 \quad 9]$$

then

$$\mathbf{R} = [-353.3117 \quad -13504.8443 \quad -5941.7900 \quad 12540.3001 \quad 41483.7387]^t.$$

Now we have  $\mathbf{B} = \mathbf{W}^{-1} \cdot \mathbf{R}$ , thus

$$\mathbf{B} = \begin{bmatrix} 0.0128 & 0.0019 & -0.1062 & -0.0776 & 0.0090 \\ 0.0019 & 0.0028 & -0.0135 & -0.0079 & 0.0014 \\ -0.1062 & -0.0135 & 0.9020 & 0.6614 & -0.0760 \\ -0.0776 & -0.0079 & 0.6614 & 0.4867 & -0.0556 \\ 0.0090 & 0.0014 & -0.0760 & -0.0556 & 0.0067 \end{bmatrix} \cdot \mathbf{R} = \begin{bmatrix} 1.0629 \\ 0.7382 \\ 1.7329 \\ 1.0735 \\ 10.1898 \end{bmatrix}$$



and by the Cholesky decomposition method we have

$$\mathbf{X}_1^t \mathbf{X}_1 = \mathbf{W} = \begin{bmatrix} 67.7712 & & & & \\ -8.3788 & 49.8747 & & & \\ 6.0689 & 37.1235 & 18.8064 & & \\ 1.2044 & -53.1008 & -22.9986 & 6.5295 & \\ -10.7112 & -30.3592 & 22.3631 & 54.3754 & 12.2276 \end{bmatrix} \mathbf{L}^t.$$

Since  $\mathbf{L}\mathbf{y} = \mathbf{R}$ , thus,  $\mathbf{y} = [-5.2133 \quad -271.6513 \quad 221.9721 \quad 494.1765 \quad 110.0586]^t$  and

$$\mathbf{L}^t \mathbf{B} = \mathbf{y}, \text{ then } \mathbf{B} = \begin{bmatrix} 1.0711 \\ -0.6741 \\ 1.9904 \\ 0.7281 \\ 9.0008 \end{bmatrix}.$$

Now add the row,  $[25.9999 \quad -15.0563 \quad 9.3248 \quad -29.7953 \quad -36.2248]$  to  $\mathbf{X}_1$  and call the modified matrix  $\mathbf{X}_2$ . If exact  $\mathbf{B}$  is

$$\mathbf{B} = [5.071 \quad -3.9942 \quad -0.0073 \quad 1.1123 \quad 4.1]^t,$$

then

$$\mathbf{R} = [22936.9659 \quad -22060.0619 \quad -10852.5632 \quad 20421.8202 \quad 22279.2954]^t$$

From (9) we have

$$(\mathbf{W}^{\text{new}})^{-1} = \mathbf{W}^{-1} - \frac{1}{1+140.9109} \begin{bmatrix} 1.6900 & 0.0858 & -14.4307 & -10.6841 & 1.1935 \\ 0.0858 & 0.0044 & -0.7326 & -0.5424 & 0.0606 \\ -14.4307 & -0.7326 & 123.2211 & 91.2295 & -10.1914 \\ -10.6841 & -0.5424 & 91.2295 & 67.5437 & -7.5454 \\ 1.1935 & 0.0606 & -10.1914 & -7.5454 & 0.8429 \end{bmatrix}.$$

So,

$$(\mathbf{W}^{\text{new}})^{-1} = \begin{bmatrix} 0.0009 & 0.0013 & -0.0045 & -0.0023 & 0.0006 \\ 0.0013 & 0.0028 & -0.0083 & -0.0041 & 0.0010 \\ -0.0045 & -0.0083 & 0.0337 & 0.0185 & -0.0042 \\ -0.0023 & -0.0041 & 0.0185 & 0.0107 & -0.0024 \\ 0.0006 & 0.0010 & -0.0042 & -0.0024 & 0.0008 \end{bmatrix}.$$

Now  $\mathbf{B}$  is

$$\mathbf{B} = (\mathbf{W}^{\text{new}})^{-1} \mathbf{R} = \begin{bmatrix} 7.1991 \\ -3.3240 \\ -1.6186 \\ 1.9620 \\ 6.0940 \end{bmatrix}.$$

But by the Cholesky decomposition method one gets

$$\mathbf{X}_2^t \mathbf{X}_2 = \mathbf{W}^{\text{new}} = \begin{bmatrix} 72.5874 & & & & \\ -13.2159 & 51.0854 & & & \\ 9.0062 & 34.8299 & 23.6936 & & \\ -9.5479 & -45.7285 & -42.0208 & 18.2338 & \\ -22.9758 & -23.1503 & -4.0525 & 58.7350 & 37.7479 \end{bmatrix} \cdot \mathbf{L}^t.$$

$$\mathbf{L}\mathbf{y} = \mathbf{R}, \text{ then } \mathbf{y} = \begin{bmatrix} 315.9910 \\ -350.0796 \\ -63.5279 \\ 261.0955 \\ 154.7663 \end{bmatrix} \text{ and since } \mathbf{L}^t \mathbf{B}^{\text{Chol.}} = \mathbf{y}, \text{ we will have } \mathbf{B}^{\text{Chol.}} = \begin{bmatrix} 5.0710 \\ -3.9943 \\ -0.0071 \\ 1.1123 \\ 4.1000 \end{bmatrix}.$$

Table 1. Comparison between regression coefficient results from (9) and Cholesky with the exact solution

	Exact $\mathbf{B}$	$\mathbf{B}$ from (9)	$\mathbf{B}$ from Cholesky	$\mathbf{e}^{\text{S-M}}$	$\mathbf{e}^{\text{Chol.}}$
Example1 with $\mathbf{X}_1$	$\begin{pmatrix} 2.0050 \\ -4.0001 \end{pmatrix}$	$\begin{pmatrix} 2.0050 \\ -4.0001 \end{pmatrix}$	$\begin{pmatrix} 2.0050 \\ -3.9999 \end{pmatrix}$	$\begin{pmatrix} 0.0000 \\ 0.0000 \end{pmatrix}$	$\begin{pmatrix} 0.0000 \\ 0.0002 \end{pmatrix}$
Example1 with $\mathbf{X}_2$	$\begin{pmatrix} 1.2436 \\ -0.0101 \end{pmatrix}$	$\begin{pmatrix} 1.2441 \\ -0.0085 \end{pmatrix}$	$\begin{pmatrix} 1.2436 \\ -0.0101 \end{pmatrix}$	$\begin{pmatrix} 0.0005 \\ 0.0016 \end{pmatrix}$	$\begin{pmatrix} 0.0000 \\ 0.0000 \end{pmatrix}$
Example2 with $\mathbf{X}_1$	$\begin{pmatrix} 1.0700 \\ -0.6742 \\ 2.0000 \\ 0.7352 \\ 9.0000 \end{pmatrix}$	$\begin{pmatrix} 1.0629 \\ 0.7382 \\ 1.7329 \\ 1.0735 \\ 10.1898 \end{pmatrix}$	$\begin{pmatrix} 1.0711 \\ -0.6741 \\ 1.9904 \\ 0.7281 \\ 9.0008 \end{pmatrix}$	$\begin{pmatrix} 0.0071 \\ 1.4124 \\ 0.2671 \\ 0.3383 \\ 1.1898 \end{pmatrix}$	$\begin{pmatrix} 0.0011 \\ 0.0001 \\ 0.0096 \\ 0.0071 \\ 0.0008 \end{pmatrix}$
Example2 with $\mathbf{X}_2$	$\begin{pmatrix} 5.0710 \\ -3.9942 \\ -0.0073 \\ 1.1123 \\ 4.1000 \end{pmatrix}$	$\begin{pmatrix} 7.1991 \\ -3.3240 \\ -1.6186 \\ 1.9620 \\ 6.0940 \end{pmatrix}$	$\begin{pmatrix} 5.0710 \\ -3.9943 \\ -0.0071 \\ 1.1123 \\ 4.1000 \end{pmatrix}$	$\begin{pmatrix} 2.1281 \\ 0.6702 \\ 1.6113 \\ 0.8497 \\ 1.9940 \end{pmatrix}$	$\begin{pmatrix} 0.0000 \\ 0.0001 \\ 0.0002 \\ 0.0000 \\ 0.0000 \end{pmatrix}$

Table 1 shows the results of these two examples.

As can be seen, in both examples, the errors associated with the Cholesky decomposition are considerably less than the ones associated with the Sherman-Morrison formula.

#### 4. CONCLUSIONS

In this paper, using the *Sherman-Morrison* formula we first updated the  $(\mathbf{X}^t\mathbf{X})^{-1}$  matrix after receiving new data by resorting to an initial inverse solution and then modified the calculation of  $(\mathbf{X}^t\mathbf{X})$  by utilizing the Cholesky decomposition method to solve the system. In this study we compared these two modifications. Resorting to some numerical examples it can be seen that using the Cholesky decomposition method to update the regression coefficients leads to considerably less error than using the *Sherman-Morrison* formula. Especially when the size of predicting variables in the multiple linear regression is too large, this error can lead to an irremediable size in forecasting the response variable. One way to overcome this fault can be resorting to much higher accuracy in performing the calculations. Of course it should be noted that even though the error associated with the method based on the *Sherman-Morrison* formula is greater than the one incurred by the Cholesky decomposition method, the former method has the advantage of dealing with linear relations only. This fact can significantly affect the time complexity corresponding to these two methods.

#### REFERENCES

- [1] Bain L.J., Engelhardt M. (1992), Introduction to Probability and Mathematical Statistics; 2<sup>nd</sup> edition, PWS-KENT.
- [2] Bartlett M.S. (1951), An Inverse Matrix Adjustment Arising in Discriminant analysis; *Ann. Math. Statist* 22; 107-111.
- [3] Egidi N., Maponi P. (2006), A Sherman-Morrison Approach to the Solution of Linear Systems; *Journal of Computational and Applied Mathematics* 189; 703-718.
- [4] Fitzpatrick P., Murphy C.C. (1993), Solution of Linear Systems of Equations in the Presence of Two Transient Hardware Faults; *IEE Proceedings E: Computers and Digital Techniques* 140(5); 247-254.
- [5] Golub G.H., Von Loan C.F. (1989), Matrix computations, 2<sup>nd</sup> edition; The Johns Hopkins University press.
- [6] Hager W.W. (1989), updating the Inverse of a Matrix; *SIREV* 31(2); 221-239.
- [7] Hoffman K. (1971), Linear Algebra, 2nd edition, Prentice Hall, Inc.,NJ.
- [8] Huang J. Z., Liu N., Pourahmadi M. and Liu L. (2006), Covariance Matrix Selection and Estimation Via Penalized Normal Likelihood, *Biometrika*; 93(1); 85-98.
- [9] Kaci M., Sadoun T., Moussaceb K. and Akvoune N. (2001), Modelling of degradation of unstabilized and HALS-stabilized LDPE films under thermo-oxidation and natural weathering conditions; *Journal of Applied Polymer Science*; 82(13); 3284-3292.
- [10] Khamis H. and Kepler M. (2002), Multivariate Cubic Spline Smoothing in Multiple Prediction; *Computer Methods and Programs in biomedicine*; 67(2); 131-136

- [11] Kok Song Chua (2003), Efficient Computations for Large Least Square Support Vector Machine Classifiers; *Pattern Recognition Letters*; 24(1-3); 75-80.
- [12] Lai S. H. and Vemuri B. C. (2003), Sherman-Morrison-Woodbury-Based Algorithms for the Surface Smoothing Problem; *Linear Algebra and its Application*; 265(1-3); 203-229.
- [13] Neter J., Kutner M. H., Nachtsheim C. J. and Wasserman W. (1996), Applied linear regression models, 3<sup>rd</sup> edition, McGraw-Hill; New York
- [14] Pourahmadi M., Daniels M. J. and Park T. (2007), Simultaneous Modelling of the Cholesky Decomposition of Several Covariance Matrices; *Journal of Multivariate Analysis*; 98(3); 568-587.
- [15] Rencher, A., C. (2000), Linear Models in Statistics, John Wiley & Sons, New York.
- [16] Stoer J. and Bulirsch R. (1993), Introduction to Numerical analysis, 2<sup>nd</sup> edition, Springer, New York.
- [17] Woodbury M. (1950), Inverting Modified Matrices; *Memorandum Rept. 42, Statistical Research Group*, Princeton University, Princeton, NJ.