International Journal of Statistics and Applications 2014, 4(2): 124-133
DOI: 10.5923/j.statistics.20140402.07

# Singular Value Decomposition Compared to cross Product Matrix in an ill Conditioned Regression Model

**Ijomah Maxwell Azubuike**[1], **Opabisi Adeyinka Kosemoni**[2,*]

[1]Dept. of Mathematics/Statistics, University of Port Harcourt, Rivers State
[2]Dept. of Mathematics/Statistics, Federal Polytechnic Nekede, Owerri Imo State

**Abstract** Cross product matrix approach is a more widely used least-squares technique in estimating parameters of a multiple linear regression. However, it is imperative to consider other forms of estimating these parameters. The use of singular value decomposition as an alternative to the widely known cross-product matrix approach as formed one of the basis of discussion in this paper. However, the use of cross product matrix have placed a lot of restrictions on the user of this technique which makes it difficult to estimate parameter values when the predictor variables seem to be highly correlated, this is often the case when Collinearity (Multi-collinearity) exists among the predictor variables. The consequence of this is that certain conditions that makes the cross product matrix applicable are not satisfied, hence the singular value decomposition approach enables users to resolve this difficult. With the aid of empirical data sets, both methods have been compared and difficulties encountered in the use of the former have been resolved using the latter.

**Keywords** Singular value decomposition, Cross product and Collinearity

## 1. Introduction

The use of cross product matrix for obtaining least squares estimates in multiple regression analysis has been in use for a long time. It is about the major statistical technique for "fitting equations to data". However, recent works has also provided modifications of the techniques aimed at increasing its reliability as a data-analytic tool most especially when near-collinearity exists between explanatory variables. When perfect collinearity occurs, the model can be reformulated in an appropriate fashion. A different type of problem occurs when collinearity is near perfect Belsley et al (1980). Collinearity can increase estimates of parameter variance; yield models in which no variable is statistically significant even when the correlation coefficient is large; produce parameter estimates of the "incorrect sign" and of implausible magnitude; create situations in which small changes in the data produce wide swings in parameter estimates; and, in truly extreme cases, prevent the numerical solution of a model. These problems can be severe and sometimes crippling.

Rolf (2002), stated that mathematically, exact collinearities can be eliminated by reducing the number of variables, and near-collinearities can be made to disappear by forming new, orthogonalized and rescaled variables. However, omitting a perfect confounder from the model can only increase the risk for misinterpretation of the influences on the response variable. Replacing an explanatory variable by its residual, in a regression on the others from a set of near-collinear variables, makes the new variable orthogonal to the others, but with comparatively very little variability. Owing to this lack of substantial variability, we are not likely to see much of its possible influence. Mathematical rescaling cannot change this reality, of course. To a large extent it is a matter of judgment what should be called a comparatively large or small variation in different variables and variable combinations, and this ambiguity remains when it comes to shrinkage methods for predictor construction, because these are not invariant under individual rescaling and other non-orthogonal transformations of variables. We could say that the 'near' in near-collinearity is not scale invariant.

The purpose of this paper is to compare the singular value decomposition to that of cross product matrix approach in an ill conditioned regression in order to observe if both approaches vary according to the degree of collinearity among the explanatory variables. The rest of this paper is organized as follows: Section two discusses the methodology, Section three discusses the data analysis and results while concluding remarks are presented in Section four.

## 2. Methodology

The cross-product matrix least-squares estimates approach and singular value decomposition will be employed in obtaining the parameters of the model. A comparison of the

* Corresponding author:
yinksko@yahoo.com (Opabisi Adeyinka Kosemoni)
Published online at http://journal.sapub.org/statistics

two approaches will be looked into with a view of observing if the two approaches will enable us obtain the equal parameter estimates.

The general linear model assumes the form

$$\mu_{Y/x_1,x_2,\ldots x_p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p \quad (1)$$

This model can also be written in the form

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon \quad i = 1,2,\ldots,n \quad (2)$$

### 2.1. Cross-Product Matrix Approach for Obtaining Parameter Estimates

From equation (1) we form the equation below

$$Y = X\beta + e \quad (3)$$

where $Y$ is a $n \times 1$ column vector of responses variable, $X$ is $n \times p$ matrix of $p$ predictor variables of $n$ observations, $\beta$ is a $p \times 1$ column vector and $e$ is a random error of $n \times 1$ dimension associated with $Y$ vector. $Y$ and $X$ are known vectors (observations obtained from a data set), while $\beta$ and $e$ are unknown vectors.

As defined earlier the vectors;

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ : \\ : \\ Y_n \end{bmatrix} \quad (4)$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ : \\ : \\ \beta_p \end{bmatrix} \quad (5)$$

$$e = \begin{bmatrix} e_1 \\ e_2 \\ : \\ : \\ e_n \end{bmatrix} \quad (6)$$

$$X = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{31} & \ldots & X_{p1} \\ 1 & X_{12} & X_{22} & X_{32} & \ldots & X_{p2} \\ 1 & X_{13} & X_{23} & X_{33} & \ldots & X_{p3} \\ \vdots & & & & & \\ 1 & X_{1n} & X_{2n} & X_{3n} & \ldots & X_{pn} \end{bmatrix} \quad (7)$$

From equation (2) the unknown vectors $\beta$ and $e$ will be obtained using the *cross-product matrix*;

Therefore we rewrite (2) since $E(e)=0$ (*Regression Analysis Assumption*)

$$Y = X\beta \quad (8)$$

Multiplying both sides of equation (8) by $X'$

$$X'Y = X'X\beta \quad (9)$$

From equation (9) we form

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (10)$$

Equation (10) forms matrices for obtaining the estimates of the parameters.

$$(X'X) = \begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} & \ldots & \sum_{i=1}^{n} x_{pi} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}^2 & \sum_{i=1}^{n} x_{1i}x_{2i} & & \sum_{i=1}^{n} x_{1i}x_{pi} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{2i}^2 & & \sum_{i=1}^{n} x_{2i}x_{pi} \\ \vdots & & & & \\ \sum_{i=1}^{n} x_{ni} & \sum_{i=1}^{n} x_{ni}x_{1i} & \sum_{i=1}^{n} x_{ni}x_{2i} & & \sum_{i=1}^{n} x_{pi}^2 \end{bmatrix} \quad (11)$$

$$(X'Y) = \begin{bmatrix} \sum_{i=1}^{n} yi \\ \sum_{i=1}^{n} x_{1i}yi \\ \sum_{i=1}^{n} x_{2i}yi \\ \vdots \\ \sum_{i=1}^{n} x_{pi}yi \end{bmatrix} \quad (12)$$

From equation (10) it can be established that $(X'X)^{-1}$ is the *variance/co-variance matrix* of the parameters, which is a useful in testing hypotheses and constructing of confidence intervals.

### 2.2. Singular Value Decomposition Approach

Given $n \times p$ matrix $X$ as earlier defined, it is possible to express each element $x_{ij}$ in the following way:

$$x_{ij} = \varphi_1 u_{1i} v_{1j} + \varphi_2 u_{2i} v_{2j} + \ldots + \varphi_r u_{ri} v_{rj} \quad (13)$$

Hence by relation

$$x_{ij} = \sum_{k=1}^{r} \varphi_k u_{ki} v_{kj} \quad (14)$$

where $\varphi_1 \geq \varphi_2 \geq \ldots \geq \varphi_p$

Equation (14) is known as the *singular value decomposition* (SVD) of matrix *X*. The number of terms in (15) is *r* which is the rank of matrix *X*; *r* cannot exceed *n* or *p* whichever is smaller.

It is assumed that $n \geq p$, and it follows that $r \leq p$. The *r* vectors *u* are orthogonal to each other, as are the *r* vectors *v*. Furthermore, each of these vectors has unit length so that;

$$\sum_i u_{ki}^2 = \sum_j v_{kj}^2 = 1 \quad \text{for all k} \quad (15)$$

In matrix notation we have

$$\underset{n\times p}{X} = \underset{n\times r}{U}\,\underset{r\times r}{\varphi}\,\underset{r\times p}{V} \quad (16)$$

The matrix $\varphi$ is diagonal, and all $\varphi_k$ are positive. The columns of the matrix *U* are the *u* vectors, and the rows of *V'* are the *v* vectors of (12). The orthogonality of the *u* and *v* and their unit length implies the conditions

$$U'U = I \quad (17)$$

$$V'V = I \quad (18)$$

The $\varphi_k$ can be shown to be the square roots of the non zero eigenvalues of the square matrix $X'X$ as well as the square matrix $XX'$. The columns of U are the eigenvectors of $XX'$ and the rows of V' are the eigenvectors of $X'X$. When $r = p$ it is known as a full-rank case. Each element of X is easily reconstructed by multiplying the corresponding elements of the $U, \varphi, V'$ matrices and summing the terms.

From equation (3) $Y = X\beta + e$

We obtain

$$Y = U\varphi V'\beta + e \quad (19)$$

Can be expressed also as

$$Y = U\left(\varphi V'\beta\right) + e$$

Where $\varphi V'\beta$ is a $r \times 1$ matrix, that is a vector of *r* elements. Let us denote the vector by $\alpha$.

$$\alpha = \varphi V'\beta \quad (20)$$

$$Y = U\alpha + e \quad (21)$$

The matrix *Y* and the matrix *U* are known. The least squares solution for the unknown coefficients $\alpha$ is obtained by the usual matrix equation

$$\hat{\alpha} = (U'U)^{-1}U'Y$$

which, as a result of (17) becomes

$$\hat{\alpha} = U'Y \quad (22)$$

This equation is easily solved since $\hat{\alpha}_j = U'Y$ simply the inner product of the matrix *Y* with *j*th vector $u_j$. It follows from (21) that

$$\hat{\alpha} = U'(U\alpha + e)$$

$$\hat{\alpha} = U'U\alpha + U'e$$

$$\hat{\alpha} = \alpha - U'e$$

$$\hat{\alpha} - \alpha = U'e$$

But $E\ (\hat{\alpha}_j - \alpha_j) = 0$

$$\Rightarrow E(\hat{\alpha}_j) = \alpha_j \quad (23)$$

Thus, $\hat{\alpha}_j$ is unbiased. Furthermore, the variance of $\hat{\alpha}_j$ is

$$Var\,(\hat{\alpha}_j) = \mathrm{E}(\hat{\alpha}_j - \alpha_j)^2$$

$$= E\left(\sum_l \sum_t u_{lj} u_{tj} e_l e_t\right)$$

$$= \left(\sum_l u_{ij}^2\right)\sigma^2 = \sigma^2$$

Hence

$$Var(\hat{\alpha}_j) = \sigma^2 \quad (24)$$

Therefore the relationship between $\beta$ and $\alpha$ is given by equation (20), which also holds for the least squares estimates of $\beta$ and $\alpha$

$$\hat{\alpha} = \varphi V'\hat{\beta} \quad (25)$$

In (25), $\varphi V'$ has the dimension $r \times p$. Thus, given the $r$ values of $\hat{\alpha}$, the matrix relation (25) represents *r* equations in *p* unknown parameter estimates $\hat{\beta}$. If *r=p* (the *full-rank case*). The solution is possible and unique.

In this case *V'* is a $p \times p$ *orthogonal matrix.*

Hence, the solution is given by

$$\hat{\beta} = V'^{-1}\varphi^{-1}\hat{\alpha} = V\varphi^{-1}\hat{\alpha} \quad (26)$$

It will be noted that $\varphi^{-1}\hat{\alpha}$ is $p \times 1$ vector, obtained by dividing each $\hat{\alpha}_j$ by the corresponding $\varphi_j$.

An important use of (26) aside from its supplying the estimates of $\beta$ as functions of $\hat{\alpha}$, is the ready calculation of the variances of the $\hat{\beta}_j$.

In general notation this equation is written as

$$\hat{\beta}_j = \sum_{k=1}^{p} v_{jk} \frac{\hat{\alpha}_k}{\varphi_k} \quad (27)$$

Also since the $\hat{\alpha}_j$ are mutually orthogonal and have all variances $\sigma^2$ we see that

$$Var(\hat{\beta}_j) = \left( \sum_{k=1}^{p} \frac{v_{jk}^2}{\varphi_k^2} \right) \sigma^2 \quad (28)$$

The numerators in each term are the squares of the elements in the first row of the *V* matrix and the denominators are the squares of $\varphi_k$.

# 3. Data Analysis and Discussion

To simplify this exposition, we will consider the data below; which is on the details of measurements taken on 22 adults on lean body mass with height, body mass index and age given in the table below;

**Table 1**

| S.No. | Height | BMI | Age (Yrs) | Lean Body Mass |
|---|---|---|---|---|
| | $X_1$ | $X_2$ | $X_3$ | Y |
| 1 | 163.5 | 16.4 | 27 | 39.9 |
| 2 | 154.1 | 15.5 | 25 | 32.8 |
| 3 | 159.8 | 17.5 | 27 | 39.5 |
| 4 | 159.8 | 18 | 28 | 42 |
| 5 | 160.2 | 15.6 | 26 | 37.2 |
| 6 | 160.4 | 17 | 25 | 39.5 |
| 7 | 164.3 | 18.9 | 28 | 44 |
| 8 | 149.2 | 21 | 26 | 41.3 |
| 9 | 153.6 | 17.8 | 27 | 37.1 |
| 10 | 150.6 | 16.5 | 28 | 34.3 |
| 11 | 170 | 14.8 | 26 | 38.1 |
| 12 | 166.5 | 18.4 | 27 | 46 |
| 13 | 169.2 | 21 | 26 | 51.2 |
| 14 | 156.8 | 19 | 28 | 42 |
| 15 | 151.5 | 17.2 | 27 | 37.1 |
| 16 | 158.6 | 17.2 | 23 | 40.5 |
| 17 | 167.6 | 15.5 | 30 | 38.8 |
| 18 | 166.8 | 16.6 | 24 | 42.8 |
| 19 | 152.1 | 16.8 | 26 | 35.1 |
| 20 | 150.1 | 18.4 | 25 | 37.6 |
| 21 | 156.4 | 16.6 | 24 | 37.4 |
| 22 | 151.7 | 16.5 | 23 | 34.2 |

Dependent variable =Y= Lean body mass
Predictor or independent variables are;
Height (cm)= $X_1$; Body Mass Index=$X_2$; Age in years = $X_3$

### 3.1. Application of Cross-Product Matrix Approach

From equations (7and 4) we define the matrices

$$X = \begin{bmatrix} 1 & 163.5 & 16.4 & 27 \\ 1 & 154.1 & 15.5 & 25 \\ 1 & 159.8 & 17.5 & 27 \\ 1 & 159.8 & 18.0 & 28 \\ 1 & 160.2 & 15.6 & 26 \\ 1 & 160.4 & 17.0 & 25 \\ 1 & 164.3 & 18.9 & 28 \\ 1 & 149.2 & 21.0 & 26 \\ 1 & 153.6 & 17.8 & 27 \\ 1 & 150.6 & 16.5 & 28 \\ 1 & 170.0 & 14.8 & 26 \\ 1 & 166.5 & 18.4 & 27 \\ 1 & 169.2 & 21.0 & 26 \\ 1 & 156.8 & 19.0 & 28 \\ 1 & 151.5 & 17.2 & 27 \\ 1 & 158.6 & 17.2 & 23 \\ 1 & 167.6 & 15.5 & 30 \\ 1 & 166.8 & 16.6 & 24 \\ 1 & 152.1 & 16.8 & 26 \\ 1 & 150.1 & 18.4 & 25 \\ 1 & 156.4 & 16.6 & 24 \\ 1 & 151.7 & 16.5 & 23 \end{bmatrix} \quad Y = \begin{bmatrix} 39.9 \\ 32.8 \\ 39.5 \\ 42.0 \\ 37.2 \\ 39.5 \\ 44.0 \\ 41.3 \\ 37.1 \\ 34.3 \\ 38.1 \\ 46.0 \\ 51.2 \\ 42.0 \\ 37.1 \\ 40.5 \\ 38.8 \\ 42.8 \\ 35.1 \\ 37.6 \\ 37.4 \\ 34.2 \end{bmatrix}$$

From equation (9) we obtain the cross-product of matrices $(X'X)$ and $X'Y$,

Using MATLAB

$$(X'X) = \begin{bmatrix} 22 & 3492.8 & 382.2 & 576 \\ 3492.8 & 55465.4 & 60655.3 & 91449.2 \\ 382.2 & 60655.3 & 6694.82 & 10012.9 \\ 576 & 91449.2 & 10012.9 & 15146.0 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 868.4 \\ 138242.9 \\ 15183.8 \\ 22763.4 \end{bmatrix}$$

We then obtain

$$(X'X)^{-1} = \begin{bmatrix} 38.9190 & -0.1662 & -0.3633 & -0.2358 \\ -0.1662 & 0.0011 & 0.0006 & -0.0010 \\ -0.3633 & 0.0006 & 0.0187 & -0.0023 \\ -0.2358 & -0.0010 & -0.0023 & 0.0163 \end{bmatrix}$$

$(X'X)^{-1}$ is the variance/co-variance matrix.

$$\hat{\beta} = (X'X)^{-1}X'Y = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} -64.1931 \\ 4.5730 \\ 1.9891 \\ -0.1337 \end{bmatrix}$$

### 3.2. Singular Value Decomposition Approach

From equation (16) $\underset{n\times p}{X} = \underset{n\times r}{U}\underset{r\times r}{\varphi}\underset{r\times p}{V}$

Matrix *X* as defined earlier;

Using MATLAB we obtain the Matrices $\underset{n\times r}{U}$, $\underset{r\times r}{\varphi}$; $\underset{r\times p}{V}$

For U (see appendix)

| | | | |
|---|---|---|---|
| $\varphi$ = 7.5973e+002 | 0 | 0 | 0 |
| 0 | 9.5247e+000 | 0 | 0 |
| 0 | 0 | 7.1451e+00 | 0 |
| 0 | 0 | 0 | 1.6028e-001 |

| | | | | |
|---|---|---|---|---|
| | 6.1689e-003 | 9.8474e-003 | 2.6668e-003 | 9.9993e-001 |
| V = | 9.8100e-001 | -1.9065e-001 | 3.5692e-002 | -4.2698e-003 |
| | 1.0714e-001 | 6.8621e-001 | 7.1941e-001 | -9.3375e-003 |
| | 1.6162e-001 | 7.0191e-001 | -6.9366e-001 | -6.0596e-003 |

Therefore, if $\varphi_k$ is the square roots of the non-zero *eigen-values* of the *square matrix* $X'X$ as well as the square matrix $XX'$.

The *eigen-values* of matrix $XX'$ using MATLAB is obtained below

$X'X = [577190; \quad 90.720; \quad 51.052; \quad 0.025691]$

Redefining the Matrix $\varphi$ by leaving the non-zero vectors; we have

$$\varphi = \begin{bmatrix} 759.73 & 0 & 0 & 0 \\ 0 & 9.5247 & 0 & 0 \\ 0 & 0 & 7.1451 & 0 \\ 0 & 0 & 0 & 0.16028 \end{bmatrix}$$

$\varphi_1 = \sqrt{577190} = 759.73$

$\varphi_2 = \sqrt{90.720} = 9.5247$

$\varphi_3 = \sqrt{51.052} = 7.1451$

$\varphi_4 = \sqrt{0.025691} = 0.16028$

Similarly, the columns *U* are the *eigen-vectors* of $XX'$ and the rows of *V'* are the *eigen-vectors* of $X'X$.

To obtain the parameter estimates, from equation (26)

$$\hat{\beta} = V'^{-1}\varphi^{-1}\hat{\alpha} = V\varphi^{-1}\hat{\alpha}$$

$$\hat{\alpha} = U'Y$$

Where $U$ can be re-defined as column vectors *U* obtained from **SVD** ($\vec{u}_1, \vec{u}_2, \vec{u}_3; \vec{u}_4$) since *r*=4, using MATLAB we have

| $\vec{u}_1$ | $\vec{u}_2$ | $\vec{u}_3$ | $\vec{u}_4$ |
|---|---|---|---|
| 2.1918e-001 | -1.0033e-001 | -1.5284e-001 | -9.3152e-002 |
| 2.0649e-001 | -1.2440e-001 | -9.6255e-002 | 2.8530e-001 |
| 2.1456e-001 | 5.2984e-002 | -6.0573e-002 | -5.8669e-002 |
| 2.1485e-001 | 1.6270e-001 | -1.0731e-001 | -1.2560e-001 |
| 2.1460e-001 | -1.6560e-001 | -1.5280e-001 | 7.9167e-002 |
| 2.1484e-001 | -1.4244e-001 | 8.6245e-002 | 3.0086e-002 |
| 2.2078e-001 | 1.3747e-001 | 5.7847e-003 | -2.9791e-001 |
| 2.0116e-001 | 4.4362e-001 | 3.3596e-001 | 5.7616e-002 |
| 2.0660e-001 | 1.9870e-001 | -6.1339e-002 | 8.9017e-002 |
| 2.0275e-001 | 2.3878e-001 | -3.0430e-001 | 2.0686e-001 |
| 2.2714e-001 | -4.1940e-001 | -1.8439e-001 | -1.3529e-001 |
| 2.2334e-001 | -1.6284e-002 | 6.3514e-002 | -2.8958e-001 |
| 2.2698e-001 | 4.3297e-002 | 4.3587e-001 | -4.7517e-001 |
| 2.1111e-001 | 2.9479e-001 | -2.1612e-002 | -1.0394e-001 |
| 2.0380e-001 | 1.9750e-001 | -1.3224e-001 | 1.7991e-001 |
| 2.1212e-001 | -2.3939e-001 | 2.9155e-001 | 1.4200e-001 |
| 2.2499e-001 | -2.6151e-002 | -5.1423e-001 | -2.6336e-001 |
| 2.2284e-001 | -3.7305e-001 | 1.7502e-001 | -7.9297e-002 |
| 2.0431e-001 | 8.2982e-002 | -7.2436e-002 | 2.2504e-001 |
| 2.0174e-001 | 1.6459e-001 | 1.7575e-001 | 2.2291e-001 |
| 2.0941e-001 | -1.6488e-001 | 1.2307e-001 | 1.9775e-001 |
| 2.0311e-001 | -1.5171e-001 | 1.8661e-001 | 3.6659e-001 |

$$\hat{\alpha} = U'Y = \begin{bmatrix} 185.50 \\ 5.2556 \\ 9.7803 \\ -10.292 \end{bmatrix}$$

Similarly, $V$ can be re-defined as row vectors *V'* obtained from **SVD** ($\vec{v}_1, \vec{v}_2, \vec{v}_3; \vec{v}_4$) since $r$=4, using MATLAB we have

| | | | | |
|---|---|---|---|---|
| $\vec{v}_1$ | 6.1689e-003 | 9.8474e-003 | 2.6668e-003 | 9.9993e-001 |
| $\vec{v}_2$ | 9.8100e-001 | -1.9065e-001 | 3.5692e-002 | -4.2698e-003 |
| $\vec{v}_3$ | 1.0714e-001 | 6.8621e-001 | 7.1941e-001 | -9.3375e-003 |
| $\vec{v}_4$ | 1.6162e-001 | 7.0191e-001 | -6.9366e-001 | -6.0596e-00 |

Recall

$$\varphi = \begin{bmatrix} 759.73 & 0 & 0 & 0 \\ 0 & 9.5247 & 0 & 0 \\ 0 & 0 & 7.1451 & 0 \\ 0 & 0 & 0 & 0.16028 \end{bmatrix}$$

$$\varphi^{-1} = \begin{bmatrix} 0.0013163 & 0 & 0 & 0 \\ 0 & 0.10499 & 0 & 0 \\ 0 & 0 & 0.13996 & 0 \\ 0 & 0 & 0 & 6.2390 \end{bmatrix}$$

$$\hat{\beta} = V\varphi^{-1}U'Y = \begin{bmatrix} 0.0061689 & 0.0098474 & 0.002668 & 0.99999 \\ 0.98100 & -0.19065 & 0.035692 & -0.042698 \\ 0.10714 & 0.68621 & 0.71941 & -0.0093375 \\ 0.16162 & 0.70191 & -0.69366 & -6.0596 \end{bmatrix} \begin{bmatrix} 0.0013163 & 0 & 0 & 0 \\ 0 & 0.10499 & 0 & 0 \\ 0 & 0 & 0.13996 & 0 \\ 0 & 0 & 0 & 6.2390 \end{bmatrix} \begin{bmatrix} 185.50 \\ 5.2556 \\ 9.7803 \\ -10.292 \end{bmatrix}$$

$$\hat{\beta} = V\varphi^{-1}U'Y = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} -64.1931 \\ 4.5730 \\ 1.9891 \\ -0.1337 \end{bmatrix}$$

Similarly, obtaining the variances we have;

$$Var(\hat{\beta}_j) = \left( \sum_{j=1}^{r} \sum_{k=1}^{p} \frac{v_{jk}^2}{\varphi_k^2} \right) \sigma^2$$

Using MATLAB

$$Var(\hat{\beta}_0) = \left[ \frac{v_{11}^2}{\varphi_1^2} + \frac{v_{12}^2}{\varphi_2^2} + \frac{v_{13}^2}{\varphi_3^2} + \frac{v_{14}^2}{\varphi_4^2} \right] \sigma^2 = 38.919\sigma^2$$

$$Var(\hat{\beta}_1) = \left[ \frac{v_{21}^2}{\varphi_1^2} + \frac{v_{22}^2}{\varphi_2^2} + \frac{v_{23}^2}{\varphi_3^2} + \frac{v_{24}^2}{\varphi_4^2} \right] \sigma^2 = 0.00112\sigma^2$$

$$Var(\hat{\beta}_2) = \left[ \frac{v_{31}^2}{\varphi_1^2} + \frac{v_{32}^2}{\varphi_2^2} + \frac{v_{33}^2}{\varphi_3^2} + \frac{v_{34}^2}{\varphi_4^2} \right] \sigma^2 = 0.0018722\sigma^2$$

$$Var(\hat{\beta}_3) = \left[ \frac{v_{41}^2}{\varphi_1^2} + \frac{v_{42}^2}{\varphi_2^2} + \frac{v_{43}^2}{\varphi_3^2} + \frac{v_{44}^2}{\varphi_4^2} \right] \sigma^2 = 0.162855\sigma^2$$

For obtaining co-variances we have;

$$Cov(\hat{\beta}_i \hat{\beta}_j) = \left[ \left( \frac{v_{i1}}{\varphi_i} \right) \left( \frac{v_{j1}}{\varphi_j} \right) + \left( \frac{v_{i2}}{\varphi_i} \right) \left( \frac{v_{j2}}{\varphi_j} \right) + \ldots + \left( \frac{v_{ip}}{\varphi_i} \right) \left( \frac{v_{jp}}{\varphi_j} \right) \right] \sigma^2 \quad for \quad i \neq j$$

In general we have;

$$Cov(\hat{\beta}_i \hat{\beta}_j) = \left( \sum_{k=1}^{p} \frac{v_{ik} v_{jk}}{\varphi_{ii} \varphi_{jj}} \right) \sigma^2$$

$$Cov(\hat{\beta}_1 \hat{\beta}_2) = \left[ \left( \frac{v_{11}}{\varphi_1} \right) \left( \frac{v_{21}}{\varphi_2} \right) + \left( \frac{v_{12}}{\varphi_1} \right) \left( \frac{v_{22}}{\varphi_2} \right) + \left( \frac{v_{13}}{\varphi_1} \right) \left( \frac{v_{23}}{\varphi_2} \right) + \left( \frac{v_{14}}{\varphi_1} \right) \left( \frac{v_{24}}{\varphi_2} \right) \right] \sigma^2 = 0.0011\sigma^2$$

### 3.3. Cross Product Matrix Approach Versus Singular Value Decomposition

We have been able to obtain regression model parameter estimates using both procedures and results are the same in both cases. However, one question most users of cross product matrix might ask is why go through the trouble of decomposing your predictor variables to several matrices when similar estimates are obtainable? The user of cross product matrix must bear in mind that estimates are only obtainable if and only if $X'X$ matrix is nonsingular (i.e $|X'X| \neq 0$). This is not often the case when predictor variables are linearly related, a situation often referred to as Perfect Collinearity or ill condition regression model.

It then follows that, when $X'X$ is singular, we can use

SVD to approximate its inverse by the following matrix;

$$\left(X'X\right)^{-1} = (U\varphi_0 V^{'})^{-1} \approx \left(V\varphi_0^{-1}U'\right) \quad (29)$$

where

$$\varphi_0^{-1} = \begin{cases} \dfrac{1}{\varphi_k} & \varphi_k > 0 \qquad k = 1,2,...,r \\ 0 & otherwise \end{cases}$$

Equation (29) is a modified form of equation (16).

## 4. Conclusions

From the analysis carried out using the cross product matrix and Singular value decomposition approach, parameter estimates are easy and uniquely determined using the cross product approach when there exist no near-collinearity between explanatory variables. On the other hand, singular value decomposition enables us to obtain similar estimates when no near-collinearity between explanatory variables though requires much more computational techniques than cross product matrix approach, but in addition, it enables us to handle the case of near-collinearity between the explanatory variables or ill conditioned regression models by applying the use of SVD to the matrix $\left(X'X\right)$ thus making it possible for users of the cross product matrix approach to obtain parameter estimates after obtaining the inverse of SVD of the matrix $\left(X'X\right)$.

## APPENDIX

U =

Columns 1 through 6

| | | | | | |
|---|---|---|---|---|---|
| 2.1918e-001 | -1.0033e-001 | -1.5284e-001 | -9.3152e-002 | -2.5942e-001 | -2.5375e-001 |
| 2.0649e-001 | -1.2440e-001 | -9.6255e-002 | 2.8530e-001 | -8.9953e-002 | -3.9684e-002 |
| 2.1456e-001 | 5.2984e-002 | -6.0573e-002 | -5.8669e-002 | 1.5979e-001 | 6.4343e-002 |
| 2.1485e-001 | 1.6270e-001 | -1.0731e-001 | -1.2560e-001 | -1.2751e-001 | 9.3815e-002 |
| 2.1460e-001 | -1.6560e-001 | -1.5280e-001 | 7.9167e-002 | 9.0586e-001 | -5.0455e-002 |
| 2.1484e-001 | -1.4244e-001 | 8.6245e-002 | 3.0086e-002 | -5.5846e-002 | 9.2971e-001 |
| 2.2078e-001 | 1.3747e-001 | 5.7847e-003 | -2.9791e-001 | 1.0136e-002 | -2.3695e-002 |
| 2.0116e-001 | 4.4362e-001 | 3.3596e-001 | 5.7616e-002 | 6.6050e-002 | -3.4494e-002 |
| 2.0660e-001 | 1.9870e-001 | -6.1339e-002 | 8.9017e-002 | -2.5645e-002 | -2.1685e-002 |
| 2.0275e-001 | 2.3878e-001 | -3.0430e-001 | 2.0686e-001 | -6.0958e-002 | 3.1926e-003 |
| 2.2714e-001 | -4.1940e-001 | -1.8439e-001 | -1.3529e-001 | -1.1792e-001 | -6.7756e-002 |
| 2.2334e-001 | -1.6284e-002 | 6.3514e-002 | -2.8958e-001 | -7.9227e-003 | -4.5728e-002 |
| 2.2698e-001 | 4.3297e-002 | 4.3587e-001 | -4.7517e-001 | 6.5942e-002 | -7.1055e-002 |
| 2.1111e-001 | 2.9479e-001 | -2.1612e-002 | -1.0394e-001 | 1.3143e-002 | -1.0061e-002 |
| 2.0380e-001 | 1.9750e-001 | -1.3224e-001 | 1.7991e-001 | -4.3426e-002 | -1.7385e-002 |
| 2.1212e-001 | -2.3939e-001 | 2.9155e-001 | 1.4200e-001 | -5.6346e-002 | -1.0394e-001 |
| 2.2499e-001 | -2.6151e-002 | -5.1423e-001 | -2.6336e-001 | -8.4263e-002 | 9.6729e-003 |
| 2.2284e-001 | -3.7305e-001 | 1.7502e-001 | -7.9297e-002 | -7.0962e-002 | -1.0008e-001 |
| 2.0431e-001 | 8.2982e-002 | -7.2436e-002 | 2.2504e-001 | -5.8511e-002 | -3.6570e-002 |
| 2.0174e-001 | 1.6459e-001 | 1.7575e-001 | 2.2291e-001 | -1.4275e-002 | -5.2537e-002 |
| 2.0941e-001 | -1.6488e-001 | 1.2307e-001 | 1.9775e-001 | -7.0906e-002 | -8.1066e-002 |
| 2.0311e-001 | -1.5171e-001 | 1.8661e-001 | 3.6659e-001 | -7.7066e-002 | -9.0792e-002 |

Columns 7 through 12

| | | | | | |
|---|---|---|---|---|---|
| -2.0563e-001 | -1.1662e-002 | -1.1176e-001 | -7.9904e-002 | -3.9978e-001 | -2.6341e-001 |
| 1.8904e-001 | -1.8500e-001 | -1.7674e-001 | -3.0346e-001 | 1.5916e-001 | 2.1895e-001 |
| -2.8609e-001 | -5.0254e-001 | -1.8018e-001 | -1.1408e-001 | 3.1170e-001 | -1.6081e-001 |
| -5.7587e-002 | 2.4761e-001 | -9.3851e-002 | -3.2185e-001 | -1.4021e-001 | 2.1933e-002 |
| 7.7369e-003 | 1.2483e-001 | 4.9506e-003 | -2.5765e-002 | -1.5463e-001 | -1.9202e-002 |
| -1.3568e-002 | 2.2735e-002 | 6.8298e-003 | 3.4293e-002 | -1.0387e-001 | -4.7332e-002 |
| 8.7541e-001 | -9.8719e-002 | -3.2982e-002 | 1.1017e-002 | -1.2077e-002 | -1.1135e-001 |
| -8.7396e-002 | 6.3080e-001 | -1.1601e-001 | -4.4559e-002 | 1.9673e-001 | -4.9951e-002 |
| -3.8100e-002 | -9.5241e-002 | 9.2808e-001 | -8.8530e-002 | 2.0426e-002 | -1.7839e-002 |
| -8.2065e-003 | -2.2288e-002 | -9.1754e-002 | 8.2374e-001 | -3.7696e-004 | 2.7158e-002 |

| | | | | | |
|---|---|---|---|---|---|
| -1.0741e-002 | 2.6093e-001 | 7.2331e-002 | 6.3004e-002 | 7.1954e-001 | -6.9517e-002 |
| -1.0497e-001 | -4.1392e-002 | 1.3387e-004 | 5.8004e-002 | -6.8058e-002 | 8.8390e-001 |
| -1.6713e-001 | -2.1520e-001 | 5.0654e-003 | 1.6592e-001 | 1.6126e-002 | -1.8427e-001 |
| -9.7268e-002 | -1.6405e-001 | -7.9391e-002 | -6.6117e-002 | 6.0315e-002 | -6.2916e-002 |
| -1.4286e-002 | -6.5930e-002 | -7.8589e-002 | -1.2140e-001 | 1.1386e-002 | 7.8889e-003 |
| 1.9511e-002 | -1.3103e-003 | 2.3447e-002 | 8.4567e-002 | -1.0384e-001 | -3.8444e-002 |
| -8.1291e-002 | 1.7944e-001 | -8.3011e-003 | -8.0935e-002 | -1.7500e-001 | -6.8310e-002 |
| -1.3262e-002 | 1.0482e-001 | 6.4364e-002 | 1.2728e-001 | -1.9311e-001 | -8.2227e-002 |
| 9.1428e-003 | -2.8503e-002 | -5.6166e-002 | -8.9461e-002 | -2.5480e-002 | 1.1946e-002 |
| -8.5395e-003 | -1.6270e-001 | -6.9618e-002 | -4.8964e-002 | 5.4475e-002 | -5.2841e-003 |
| 2.8199e-002 | 2.4743e-002 | 1.4063e-003 | 1.7827e-002 | -9.9202e-002 | -1.2382e-002 |
| 6.5035e-002 | -1.3797e-003 | -1.1275e-002 | -6.1909e-004 | -7.4207e-002 | 2.1457e-002 |

Columns 13 through 18

| | | | | | |
|---|---|---|---|---|---|
| -2.6541e-001 | -1.0847e-001 | -9.7139e-002 | -2.6659e-001 | -2.7594e-001 | -3.6280e-001 |
| 3.9033e-001 | -2.2610e-002 | -2.6175e-001 | -1.0902e-001 | 1.7034e-001 | 1.1514e-001 |
| -3.7089e-001 | -3.4639e-001 | -1.2845e-001 | 1.4301e-001 | -5.2961e-003 | 2.0635e-001 |
| 3.5256e-001 | -8.8927e-002 | -1.5296e-001 | 3.1651e-001 | -5.2891e-001 | 2.0339e-001 |
| 5.6167e-002 | 3.6840e-002 | -7.9737e-003 | -4.2781e-002 | -1.1247e-001 | -8.8458e-002 |
| -6.6335e-002 | 2.0381e-002 | 1.2894e-002 | -1.0363e-001 | -4.4438e-003 | -1.2221e-001 |
| -1.9257e-001 | -9.2371e-002 | -5.2971e-003 | -3.7693e-003 | -5.6463e-002 | -3.3130e-002 |
| -2.0716e-001 | -1.5457e-001 | -9.4563e-002 | -8.6753e-002 | 1.8213e-001 | 3.0569e-002 |
| -7.7035e-003 | -7.1324e-002 | -7.8725e-002 | -1.3066e-002 | -2.7443e-002 | 1.6562e-002 |
| 1.4003e-001 | -6.6707e-002 | -1.2301e-001 | 5.2035e-002 | -1.2070e-001 | 7.5013e-002 |
| 4.2824e-004 | 9.1309e-002 | 7.3509e-002 | -5.2204e-002 | -1.8543e-001 | -1.7687e-001 |
| -2.0073e-001 | -4.7460e-002 | 3.0321e-002 | -4.5500e-002 | -4.5382e-002 | -9.2560e-002 |
| 5.6091e-001 | -9.0653e-002 | 7.3494e-002 | -1.2577e-001 | 1.1037e-001 | -1.4270e-001 |
| -1.1284e-001 | 8.8077e-001 | -6.8211e-002 | 9.6493e-003 | -2.4827e-002 | 3.2286e-002 |
| 6.0082e-002 | -6.0674e-002 | 9.0298e-001 | -1.6992e-003 | -4.8255e-002 | 3.2119e-002 |
| -9.7622e-002 | 5.5666e-002 | 3.3331e-002 | 8.1046e-001 | 9.6659e-002 | -1.8209e-001 |
| 5.2957e-002 | -3.0266e-002 | -1.6591e-002 | 1.1798e-001 | 6.8587e-001 | -3.1227e-004 |
| -1.3044e-001 | 7.5821e-002 | 8.3913e-002 | -1.5477e-001 | -3.7085e-003 | 7.8241e-001 |
| 6.3236e-002 | -2.3769e-002 | -7.4739e-002 | -4.1520e-002 | -2.6709e-002 | -1.3011e-002 |
| -4.8230e-002 | -5.3047e-002 | -7.2141e-002 | -1.0050e-001 | 9.8657e-002 | -2.6826e-002 |
| -9.8357e-003 | 4.2584e-002 | -3.4775e-003 | -1.3448e-001 | 3.3608e-002 | -1.2749e-001 |
| 3.3066e-002 | 5.3870e-002 | -2.5407e-002 | -1.6760e-001 | 8.8344e-002 | -1.2538e-001 |

Columns 19 through 22

| | | | |
|---|---|---|---|
| -1.3261e-001 | -9.5016e-002 | -2.3085e-001 | -1.9307e-001 |
| -2.7377e-001 | -2.7858e-001 | -1.8403e-001 | -3.3621e-001 |
| -2.4624e-002 | -1.5640e-001 | 1.3789e-001 | 1.7406e-001 |
| -7.3824e-002 | 1.5199e-001 | 1.4882e-001 | 2.2419e-001 |
| -2.6064e-002 | 3.1884e-002 | -5.2754e-002 | -4.2582e-002 |
| -1.3196e-002 | -1.6748e-002 | -7.5895e-002 | -7.6738e-002 |
| 1.2486e-002 | -1.6966e-002 | 1.5674e-002 | 4.9412e-002 |
| -7.6492e-002 | -2.0768e-001 | -5.3691e-002 | -9.7789e-002 |
| -6.5864e-002 | -7.2328e-002 | -2.8957e-002 | -4.0485e-002 |
| -9.8265e-002 | -4.7230e-002 | -9.3922e-003 | -2.0550e-002 |
| 3.9936e-002 | 1.2747e-001 | -4.1415e-002 | 7.1315e-003 |
| 2.9698e-002 | 4.9351e-003 | -9.7022e-003 | 2.5172e-002 |
| 6.6780e-002 | -6.0562e-002 | -2.6274e-002 | 3.8834e-003 |
| -4.0716e-002 | -7.5073e-002 | 5.6543e-003 | 1.1298e-002 |
| -8.4060e-002 | -7.0147e-002 | -3.3408e-002 | -5.1240e-002 |
| -1.6881e-002 | -5.3808e-002 | -1.3200e-001 | -1.5658e-001 |
| 1.0103e-002 | 1.2514e-001 | 6.5458e-002 | 1.3645e-001 |

| | | | |
|---|---|---|---|
| 3.4628e-002 | 3.3110e-002 | -9.5252e-002 | -7.9614e-002 |
| 9.2222e-001 | -6.2954e-002 | -6.0715e-002 | -8.3601e-002 |
| -7.9626e-002 | 8.6636e-001 | -8.6349e-002 | -1.2891e-001 |
| -3.9873e-002 | -4.3391e-002 | 8.9049e-001 | -1.3248e-001 |
| -6.9988e-002 | -8.3998e-002 | -1.4330e-001 | 8.0825e-001 |

# REFERENCES

[1] Belsley D.A., Kuh E.; and Welsch,R.E. (1980), Identifying Influential Data and Sources of Collinearity, New York: John Wiley.

[2] Rolf S., (2002), Collinearity, Encyclopedia of Environmetrics, John Wiley & Sons, Ltd, Chichester, 1, 365 – 366.