Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REVIEW ON MATRIX PSEUDO-INVERSE USING SINGULAR VALUE DECOMPOSITION-SVD AND APPLICATION TO REGRESSION

[1]BASHAR SAMI NAYYEF AL-DABBAGH, [2]ZINAH MOHSIN ARKAH, [3]RAJA DAAMI RESAN, [4]LAITH ALZUBAIDI
[1]Autonomous University of Madrid, Spain
[2,4]Iraqi Commission for Computer and Informatics (ICCI), Baghdad, Iraq
[3]University of Information Technology & Communications, Baghdad, Iraq
Email: [1]bashar.nayyef@estudiante.uam.es, [2]zeinaitbi@yahoo.com, [3]jojo20072008@yahoo.com, [4]laithicci@gmail.com

ABSTRACT

Singular Value Decomposition (SVD) is one of the most factorization of the real or complex mathematical matrix problems. In this paper, one of the most significant applications of the Signa gular Value Decomposition (SVD) which is the Matrix decomposition is being selected to be described and explained as a regression model. The experimental results show that the SVD regression using Matrix-Pseudo Inverse results are more realistic and nearly as expected that the simple regression model when the results have been compared between the simple regression model and the SVD regression model based on the Matrix-Pseudo Inverse model based on implement them on the same dataset (data points). In this paper, two main cases are discussed. The first one is the insertable matrix pseudo-inverse, and the non-invertible matrix pseudo-inverse. Both cases are mainly discussed with a relative example given which shows that main approach that is used to compute based on the Singular Value Decomposition.

Keywords: *Singular Value Decomposition, SVD, Matrix Decomposition, Matrix-Pseudo Inverse, Regression.*

1. INTRODUCTION

In a mathematical application such as data analysis, and particular linear algebra, there are different important techniques that are used mainly with matrixes. One of the most mathematical application is the matrix pseudoinverse. Matric pseudoinverse is defined as a generalization of the inverse matrix [1]. One of the most widely type of matrix pseudoinverse is the matrix pseudo-inverse using Singular Value Decomposition (SVD) [2] [3] [4] [5] which has been described by [6] [7] [8]. Earlier in 1903, Ivar Fredholm [9] had introduced the main concept of the matrix pseudoinverse of integral operation. That was when Ivar Fredholm [9] had referring to a matrix, and that was depending on the generalization aspect term of the pseudoinverse without basing on the further specification.

The term generalization of matrix- inverse is sometimes as a synonym for matrix-pseudoinverse. There are many applications that the matrix pseudo-inverse is mainly used. In general, a common technique of pseudoinverse is to compute a 'best fit' which means that the least square solution. In other word, it means that the least square solution that is generated as a solution to the linear equation system that has lacks for finding a unique solution [6]. Another mathematical application that the matrix-inverse is used for is to Euclidean solution. In this case, the matrix pseudo-inverse is used to find the minimum “Euclidean” norm solution to the linear equation system to generate multiple solutions [8].

In linear algebra, the matrix pseudo-inverse facilitates the
statement and proof of results. Matrix pseudo-inverse is defined for all matrixes that’s the entire components are real and/or complex number. It is tetchily can be computed by using the Singular Value Decomposition (SVD) [9].

2. BACKGROUND THEORY (SINGULAR VALUE DECOMPOSITION-SVD)

Singular value decomposition (SVD) is defined as a one of the most popular unsupervised data

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

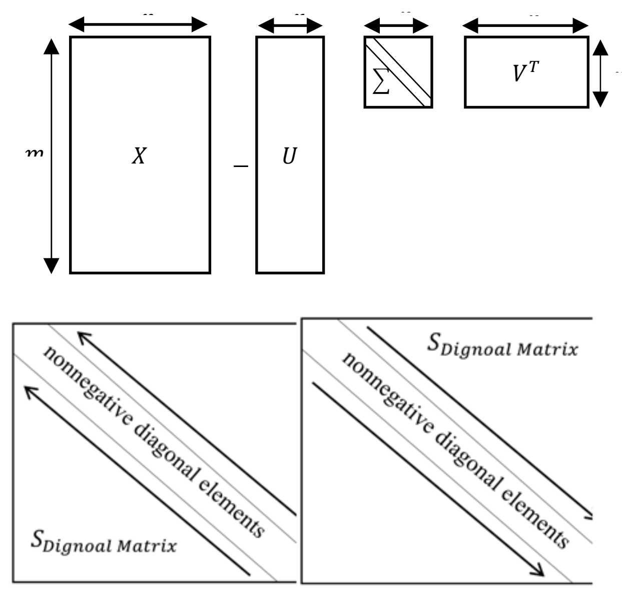
mining algorithms. It such a significant algorithm that is mainly used for higher dimensionality data (feature space) projection. It also, one of the most appropriate mapping tool that is used for mapping the higher dimensionality data space or (vector space) to another dimension. Moreover, Singular Value Decomposition (SVD) illustrate as the most useful method for analyzing and mapping the data (feature vector space) in one dimension (one vector space) into another space such as higher dimensionality space (with different dimension) [10].

Most linear equations simulation systems rely on the Singular Value Decomposition (SVD) for analyzing and mapping data space. In this matter, the SVD allows the linear equations simulation systems to exact representation of any matrix and makes it easy to eliminate and simulate by based on the important parts of that representation. It based on produce an approximate matrix representation with any desired number of dimensions [11]. Of course, the fewer the dimensions that have been chosen, the less accurate will be the approximation. SVD is a technical useful of number application which includes the analysis trick of the two-way variables (tables) evaluation. Although, SVD in an experimental design, empirical fitting of any function, and regression [12].

SVD defines a small number of "concepts" that connect the rows and columns of the matrix. We show how eliminating the least important concepts gives us a smaller representation that closely approximates the original matrix. Next, we see how these concepts can be used to query the original matrix m [10].

2.1 SVD Mathematical Definition

Let X be an $m \times n$ matrix and let the rank of X be r. By recalling that the rank of a matrix is the largest number of rows (or equivalently columns) we can choose for which no nonzero linear combination of the rows is the all-zero vector 0 (we say a set of such rows or columns is independent). Then we can find matrices U, Σ, and V as shown in Figure 1 [13].



(a) (b)

Figure. 1. The form of singular-value decomposition, (a) regular based of the SVD algorithms, (b) and (c) evaluation Form-F [13]

Singular value decomposition (SVD) is an eigenvalue/eigenvector mechanics which is similar process of finding the singular value (eigenvector). Although, it is used to find the corresponding singular vectors (Eigenvector) that are mainly yields on matrix decomposition term. This term is more general and flexible matrix decomposition factorization. The term 'singular vector' and 'Eigenvector' will be used in an interchangeably where the Singular Value Decomposition (SVD) of matrix A can be written as it is shown in equation (1) [13].

$$A = USV^T \qquad (1)$$

where U is the orthogonal1 $m \times m$ matrix and the columns of the U are the eigenvectors of AA^T. Moreover, V is the orthogonal of $n \times n$ matrix and the columns of the V are the eigenvector s of the AA^T matrix. However, S is the diagonal eigenvalues (entities) which also called the diagonal sigma's values $\sigma_1, \ldots, \sigma_2$ which are computed based on the square roots of the nonzero eigenvalues of the AA^T and $A^T A$ matrix. Both of them are the singular values of matrix A and they fill the first r places on the main diagonal of S where r is defined as the rank of A [14].

Based on equation (1), the connections with $AA^T$and $A^T A$ can be described and written as the

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

equation (2) shows [14].

$$AA^T = (USV^T)(VS^TU^T) = USS^TU^T \quad (2)$$

similarly, A^TA can be written as equation (3) shows.

$$AA^T = (USV^T)(VS^TU^T) = USS^TU^T \quad (3)$$

By relying on equation (2), U must be the eigenvector matrix AA^T, where the $SS^T$is the eigenvalue matrix that is placed in the middle and it is defined as the $m \times m$ matrix with the eigenvalues $\lambda_1 = \sigma_1^2, \ldots, \lambda_r = \sigma_r^2$. Although, based on the same way and using equation (3) U is defined as the eigenvector matrix for A^TA. The diagonal matrix S^TS has the same property $\lambda_1 = \sigma_1^2, \ldots, \lambda_r = \sigma_r^2$ which is also is defined as the $n \times n$ matrix [14].

Singular-Value Decomposition (SVD) algorithm steps are described in the algorithm (1) below [15].

Where the most important properties of the Singular Value Decomposition (SVD) can be described as the following [15]:

1) U is an $x \times r$ column-orthonormal matrix; that is, each of its columns is a unit vector and the dot product of any two columns is 0.
2) V is an $n \times r$ column-orthonormal matrix, note that we always use V in its transposed form, so it is the rows of V^T that are orthonormal.
3) Σ is a diagonal matrix; that is, all elements not on the main diagonal are 0. The elements of Σ are called the singular values of X.

2.2 SVD Mathematical Example

This section provides such a mathematical example about the Singular Value Decomposition (SVD). In this case, let's assume that we want to find the Find the singular value decomposition of the next matrix which we called matrix A below.

$$A = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}$$

To find the Singular Value Decomposition (SVD) value first we need to compute the eigenvalues by based on using matrix A times the transpose of the matrix A based on the next formula:

$$Eigenvalues\ (A) = A \times A^T$$

After that, it is generally preferred to put them into a (decreasing order) and then find the corresponding unit eigenvectors based on the next formula:

First Find the $A \times A^T$ by using:

$$\boldsymbol{AA^T} = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} \times \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$$

Put them into a decreasing order by using

$$\boldsymbol{AA^T} = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \Longrightarrow det(\boldsymbol{AA^T} - \lambda \boldsymbol{I})$$
$$= det \begin{bmatrix} 8-\lambda & 0 \\ 0 & 8-\lambda \end{bmatrix} = 0$$
$$(8-\lambda)(2-\lambda) = 0 \Longrightarrow \lambda_1 = 8\ ,\ \lambda_2 = 2$$

Find the corresponding unit eigenvectors which are:

$$\boldsymbol{AA^T}\underline{u}_1 = \lambda_1\underline{u}_1 \Longrightarrow \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = 8 \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix}$$
$$\Longrightarrow \begin{matrix} 8 \times u_{11} = 8u_{11} \Longrightarrow u_{11} = 1 \\ 2 \times u_{12} = 2u_{12} \Longrightarrow u_{12} = 0 \end{matrix}$$

Algorithm (1) Singular-Value Decomposition (SVD)

Input: Generate Data matrix $\boldsymbol{X}$
Output: New Dimensions $\boldsymbol{C}$
Repeat
  Applying SVD to the matrix X as $X = USV^T$
    $X \rightarrow$ is an $m \times n$ matrix
    $m \rightarrow$ no. of vectors.
    $n \rightarrow$ no. of attributes.
    $U \leftarrow XX^T$matrix of the eigenvectors.
    $S \leftarrow$ is matrix which is diagonal.
    $V \leftarrow$ is matrix the eigenvectors.
  Construct the covariance matrix from this decomposition by
  $XX^T\ XX^T \leftarrow (USV^T)(USV^T)^T\ = (USV^T)(VSU^T)$
  $V\ \rightarrow$ an orthogonal matrix $(V^TV = I), XX^T = US2U^T$
    Compute the square roots of the eigenvalues of XX^T are the singular values
Until Re $x(t)_i$
Represent every transaction over the time interval as a vector $x(t)_i$
Return U^TX

$$\begin{matrix} 8 \times u_{11} = 8u_{11} \Longrightarrow u_{11} = 1 \\ 2 \times u_{12} = 2u_{12} \Longrightarrow u_{12} = 0 \end{matrix} \Longrightarrow \underline{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\boldsymbol{AA^T}\underline{u}_2 = \lambda_2\underline{u}_2 \Longrightarrow \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = 8 \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix}$$
$$\Longrightarrow \begin{matrix} 8 \times u_{21} = 8u_{21} \Longrightarrow u_{21} = 0 \\ 2 \times u_{22} = 2u_{22} \Longrightarrow u_{22} = 1 \end{matrix}$$

$$\begin{matrix} 8 \times u_{21} = 8u_{21} \Rightarrow u_{21} = 0 \\ 2 \times u_{22} = 2u_{22} \Rightarrow u_{22} = 1 \end{matrix} \Rightarrow \underline{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then, compute the matrix U based on:

$$\boldsymbol{U} = [\underline{u}_1 \quad \underline{u}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since, the eigenvalues of the $\boldsymbol{A}^T\boldsymbol{A}$ are the same as the eigenvalues of the$\boldsymbol{A}\boldsymbol{A}^T$, then the eigenvectors of the $\boldsymbol{A}^T\boldsymbol{A}$ is computed based on the following formula:

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} \times \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 0 & 2 \end{bmatrix}$$

Find the corresponding unit eigenvectors v_1 which are:

$$\boldsymbol{A}^T\boldsymbol{A}\underline{v}_1 = \lambda_1 \underline{v}_1 \Rightarrow \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = 8 \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix}$$

$$\Rightarrow \begin{matrix} 5v_{11} + 3v_{12} = 8v_{11} = v_{12} \\ 3v_{11} + 5v_{12} = 8v_{12} = v_{11} \end{matrix}$$

In general, we will define the choice of v_{11}, define of v_{12} as a vice versa. In this case v_{11} and v_{12} can be any numbers, but since the vector v_1 should have length of 1, in this condition the v_{11} and v_{12} are chosen as follows:

$$\|\underline{v}_1\| = 1 \Rightarrow \sqrt{v_{11}^2 + v_{12}^2} = 1 \Rightarrow v_{11} = v_{12} = \frac{1}{\sqrt{2}}$$

$$\Rightarrow \underline{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Then by using the unit eigenvector v_2 is:

$$\boldsymbol{A}^T\boldsymbol{A}v_2 = \lambda_2 \underline{v}_2 \Rightarrow \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = 2 \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix}$$

$$\Rightarrow \begin{matrix} 5v_{21} + 3v_{22} = 2v_{21} = v_{22} = 0 \\ 3v_{21} + 5v_{22} = 2v_{22} = v_{21} = 1 \end{matrix}$$

$$v_{21} = -v_{22} \Rightarrow \underline{v}_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Compute the matrix V which is:

$$\boldsymbol{V} = [\underline{v}_1 \quad \underline{v}_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } \boldsymbol{V}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Then the matrix S is:

$$\boldsymbol{S} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix}$$

Finally, the SVD of the matrix A is:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$$

3. MATRIX (PSEUDO) INVERSE

Matrix pseudo inverse is one the most popular operation of the matrix operations. In additional to that there are many other matrix operations such as matrix inversion, Eigen value decomposition, singular value decomposition which are mainly used in the real -world applications. Some matrix operations are very expensive in time and memory especially when they used to scale the data in a large scale [15].

Matrix pseudo inverse is one of the main techniques of the Matrix decomposing which plays a key role in the modern data. In other world, Matrix pseudo inverse based on the matric decomposing theorem such as SVD (Eigen value decomposition) that is used to make the large-scale matrix computation is possible as much as it could [16].

3.1 Inverted Matrix Pseudo-Inverse Computation

Mathematically, for any square matrix A with size $n \times n$, the matrix inverse is existing if the matrix A is in non-singular rank. In other word, the matrix A with non-singular rank as it shown below in equation (4) [16]:

$$rank(A) = n \tag{4}$$

In this case, let's assume that the inverse matrix A is A^{-1} which be inverted. Then AA^{-1} is technically equates and be equivalent to the following formula that is defined in equation (5) [17]:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$AA^{-1} = A^{-1}A = I_n \quad (5)$$

That means only the squared matrix with full rank can be inverted. In this case, for general rectangle matrix $A^{n\times n}$ with the deficient rank matrix, matrix-inverse is mainly used as a generalization mathematical approach for matrix inverse [15].

The Moore-Penrose matrix inverse is one of the most widely used pseudo-inverse, which is mathematically defined by [17]:

$$A^{\dagger} = V_A \sum\nolimits_A^{-1} U_A^T \quad (6)$$

In this case, let A be any matrix with size $m \times n$ with the rank p matrix. Then,

$$AA^{\dagger} = U_A \sum\nolimits_A \underbrace{V_A^T V_A}_{=I_p} \sum\nolimits_A^{-1} U_A^T = \underbrace{U_A}_{m\times p} \underbrace{U_A^T}_{p\times m} \quad (7)$$

This presents an orthogonal projection. That because for any matrix B the nain formula of the Moore-Penrose matrix inverse is defined as [17]:

$$AA^{\dagger} = U_A U_A^T B \quad (8)$$

Which is projected of the matrix B onto the column space of the matrix A.

In this case, let's assume that we have a squared matrix A with size 2×2

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Then the pseudoinverse of the matrix A is:

$$A^{+} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Generally, the pseudoinverse of the matrix A which is a zero matrix its transposed. In this case, the uniqueness of the pseudoinverse matrix can be written as:

$$A^{+} = A^{+}AA^{+}$$

Then, sine multiplication it's by a zero matrix would be always produce a zero matrix as it is shown:

- $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ the pseudoinverse is $A^{+} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}$
- Indeed $AA^{+} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ and thus
- $AA^{+}A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = A$
- Similarly, $A^{+}A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and thus
- $A^{+}AA^{+} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} = A^{+}$
- For $A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$, $A^{+} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}$
- For $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$, $A^{+} = \begin{bmatrix} \frac{1}{5} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}$, then the denominators are $5 = 1^2 + 2^2$.
- For $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $A^{+} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$
- For $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$, where the pseudoinverse of the matrix is $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$
- Then, the A^{+}in indeed to $A^{+}A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

3.2 Non-Inverted Matrix Pseudo-Inverse Computation

In In this part we discuss such a different situation where the matrix cannot be inverted because it is singular then in in this case, we use the SVD to get the pseudo-inverse of tis matrix. In order to compute the singular value decomposition (SVD) and the pseudo-inverse of the non-inverted matrix, first we need to define such a complex matrix (non-inverted) where it is defined as a matrix A that has a complex dimension such as $m \times n$. In this case, we assume that the convenient matrix decomposition should be defined. We assume that the matrix dimension which is described by $m \geq n$.

In this case, let A be defined as any matrix with $m \times n$ that contains complex elements. In this case,

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the matrix A can be decomposed as the following formula that is described in Equation (9):

$$A = PJQ^* \tag{9}$$

Where here P and Q are the unitary matrix and also J is defined as a matrix with $m \times n$ dimensions which is called a (bidiagonal matrix) and it is mathematically described in the following form.

$$J = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & . & . & . & 0 \\ 0 & \alpha_2 & \beta_2 & 0 & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & \beta_{n-1} \\ 0 & . & . & . & . & . & \alpha_n \end{bmatrix}$$

To proof that, let assume that $A = A^{(1)}$ and also let's assume that $A^{(3/2)}, A^{(3)}, A^{(n)}, A^{(n+1/2)}$ which will be defined as the following formula in Equation (10) based on Householder transformation [18] [19] [20].

$$A^{(k+1/2)} = P^{(k)}A^{(k)} \tag{10}$$

Where $k = 1,2,\dots,n$, and also be defined as in the following Equation (11):

$$A^{(k+1)} = A^{(k+1/2)}Q^{(k)} \tag{11}$$

Where $k = 1,2,\dots,n-1$. Also, $P^{(k)}$ and $Q^{(k)}$ are Hermitian and also unitary matrices where are define in the following form that are described in Equations (12) and (13):

$$P^{(k)} = I - 2x^{(k)}x^{(k)}, \quad x^{(k)} \times x^{(k)} = 1 \tag{12}$$

$$Q^{(k)} = I - 2y^{(k)}y^{(k)}, \quad y^{(k)} \times y^{(k)} = 1 \tag{13}$$

Then, the unitary transformation $P^{(k)}$ is determine as it is defined by the Equation (14):

$$\alpha_{i,k}^{(k+1/2)} = 0 \quad were \quad i = k+1,\dots,m \tag{14}$$

Where $Q^{(k)}$ is determined based on the following formula in Equation (15):

$$\alpha_{i,k}^{(k+1/2)} = 0 \quad were \quad i = k+2,\dots,n \tag{15}$$

And then $A^{(k+1)}$ has the mathematical form

$$A^{(k+1)} = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & . & & & & & . & 0 \\ 0 & \alpha_2 & \beta_2 & 0 & & & & & & . \\ . & 0 & . & . & & & & & & \\ . & . & . & . & & & & & & \\ & & & \alpha_k & \beta_k & & & & & \\ & & & & x & x & . & . & . \\ & & & & x & x & . & . & . \\ . & & & & . & . & . & . & . \\ 0 & . & & & . & . & . & . & . \end{bmatrix}$$

So, in this case, to distribute the x elements, the new set has been defined which:

$$x_i^{(k)} = 0, \quad were \quad i = 1,2,\dots,k-1 \tag{16}$$

In this situation, since $P^{(k)}$ is a unitary transformation matrix, then the length is preserved and called consequently as it is defined in Equation (17):

$$|\alpha_k|^2 = \sum_{i=k}^{m} \left|\alpha_{i,k}^{(k)}\right|^2 \tag{17}$$

Although, since $P^{(k)}$ is a Hermitian, then

$$P^{(k)}A^{(k+1/2)} = A^{(k)} \tag{18}$$

That is equivalent to:

$$\left(1 - 2|x_k^{(k)}|^2\right)\alpha_k = \alpha_{k,k}^{(k)}$$
$$-2x_i^{(k)}x_k^{(k)}\alpha_k = \alpha_{k,k}^{(k)} \quad where \quad i = k+1,\dots,m \tag{19}$$

By based on Equations (16), (17), and (19) the possible $x^{(k)}$ is defined. In this case, as a summarization we have:

$$A^{(k+1/2)} = A^{(k)} - x^{(k)}.2\left(x^{(k)} \times A^{(k)}\right) \tag{20}$$

With the following:

$$s_k = \left(\sum_{i=k}^{m} \left|\alpha_{i,k}^{(k)}\right|^2\right)^{1/2} \tag{21}$$

$$\alpha_k = -s_k\left(\frac{\alpha_{i,k}^{(k)}}{\left|\alpha_{i,k}^{(k)}\right|}\right) \tag{22}$$

$$x_k^{(k)} = \left[\frac{1}{2}\left(1 + \frac{\left|\alpha_{i,k}^{(k)}\right|}{s_k}\right)\right]^{1/2} \tag{23}$$

$$c_k = \left(2s_k \frac{\alpha_{i,k}^{(k)}}{\left|\alpha_{i,k}^{(k)}\right|} x_k^{(k)}\right)^{-1} \tag{24}$$

$$x_i^{(k)} = c_k\alpha_{i,k}^{(k)} \quad for\ i > k \tag{25}$$

Then the final formula will be described in the Equation (26):

$$A^{(k+1)} = A^{(k+1/2)} - 2\left(A^{\left(k+\frac{1}{2}\right)} y^{(k)}\right) . y^{(k)}$$

In case of the non-invertible matrix is defined as a matrix that has one side (left or right is invertible). In other word, no-square matrix of full rank is a matrix that has several one-side inverses. For instance:

- A is matrix with size $m \times n$ where $m > n$ in this case, we have a left inverse that is define as the following formula shows:
$$\underbrace{(A^T A)^{-1} A^T}_{A^{-1}_{left}} A = I_n$$
- A is matrix with size $m \times n$ where $m < n$ in this case, we have a right inverse that is define as the following formula shows:
$$I_m = \underbrace{(A^T A)^{-1} A^T A}_{A^{-1}_{left}}$$

In this case, the left inverted side can be used to determine the least norm solution. That means it is originally used as a least square formula for regression matrix, which has any even-one-inverted-side. Although, the Pseudo-inverse approach using SVD can be used for both sides either left or right to find the exists invested side.

For such an example, let's assume and consider that an inverted matrix A is defined as:

$$A{:}\, 2 \times 3 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

So, in this case, $m < n$, and for this reason we have right inverse, which is defined as:

$$A^{-1}_{right} = A^T (AA^T)^{-1}$$

Based on that and by computing the matrix component, we have:

$$AA^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}$$

$$(AA^T)^{-1} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}^{-1} = \frac{1}{54} \begin{bmatrix} 77 & -32 \\ -32 & 14 \end{bmatrix}$$

$$A^T (AA^T)^{-1} = \frac{1}{54} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \times \begin{bmatrix} 77 & -32 \\ -32 & 14 \end{bmatrix}$$

$$= \frac{1}{18} \begin{bmatrix} -17 & 8 \\ -2 & 2 \\ 13 & -4 \end{bmatrix} = A^{-1}_{right}$$

While the left side does not exist because:

$$A^T A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix} \quad (26)$$

This shows that the singular matrix cannot be inverted.

3.3 Algorithm Complexity

In this section, we want to discuss and showing the relational between the algorithm complexity as a function of the matrix size in this case, if the matrix is invertible and has complex values, then it's just the inverse matrix. Finding the pseudo-inverse in this case takes $O(n^{\omega})$ time, where w is the matrix size (multiplication constant) [Theorem 28.2 in Introduction to Algorithms 3rd Edition].

In the other case, if the matrix A has a linearly independent matrix size such as rows or columns and also complex value, then the pseudoinverse matrix can be computed as shown in equation (29).

$$A^{*}(AA^{*})^{-1} \quad (27)$$

Where A^{*} is the conjugate transpose of A. In particular, this implies an $O(n^{\omega})$ time for finding the pseudoinverse of A.

In the mean while. Matrix size (ω) is an infimum. Whenever it is written by $O(n^{\omega})$, which means that for all $\gamma > \omega$ where in case, algorithm running in time of $O_{\gamma}(n^{\gamma})$. For example, if A is the matrix of r and size 2×2. In this case, a normal rank of A is:

$$S = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} T$$

For some cases, the invertible S and T of the appropriate dimensions [21] has normal rank form which is similar to the rank decomposition mentioned in the equation (28) [Wikipedia article].

$$A = XY \quad (28)$$

Where in this case, X has a matrix size r columns, and in the same way Y has also a matrix size r rows. For indeed, the matrix X can be taken as the first r column of the matrix S, also, in the same way, the matrix Y can be taken as the first r rows of the matrix T. In this case, given the decomposition, the formal formula of the pseudoinverse using only the Hermitian adjoint, matrix multiplication and matrix inverse. Therefore, the pseudoinverse can be computed in time $O(n^{\omega})$ [21].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. SVD APPLICATION OF REGRESSION

Linear regression mathematically defined as the way that attacking the certain prediction. More clearly, by considering such a model that define in Equation (29) as it is shown below [21]:

$$Y = \beta_0 + \beta_1 e^{\beta_2 X} + \epsilon$$

Although, linear regression aims to model the main relationship between two variables. It is mainly based on the fitting linear equation that allows us to observe the data. Technically, one variable is considered to be such some explanatory variable while the other variable is considered to be a dependent variable.

For instance, to represent such a model that wants to relate the weights of individuals to their heights. In this case, a linear regression is mainly used to do build the main model. In this case, before attempting data fitting using linear model to observe the data. A model should determine the main relationship between the variables in the area of interest. For example, the highest data in the SAT scoring example do the cause higher college grades. In this case, there is some significant association (relationship-predicted) between those two variables [22].

The most significant tool that is mainly used for regression is the SCATTERPLOT. Scatterplot is often used to identify the main protentional relationship (association) between two variables. For instance, Figure (2) shows an example of plot demonstrations that appearance of the relationship between the Size and Age of the paired data [23].

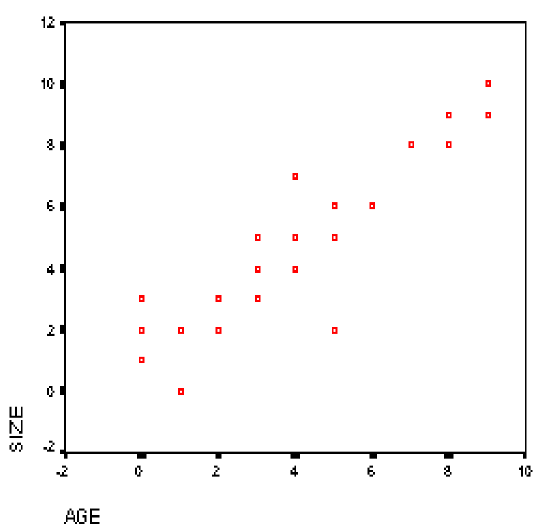


Figure 2. An Example of Scatterplot Regression Demonstrations [23]

For an example of liner regression relationship prediction that is shown in Figure (3) [23].

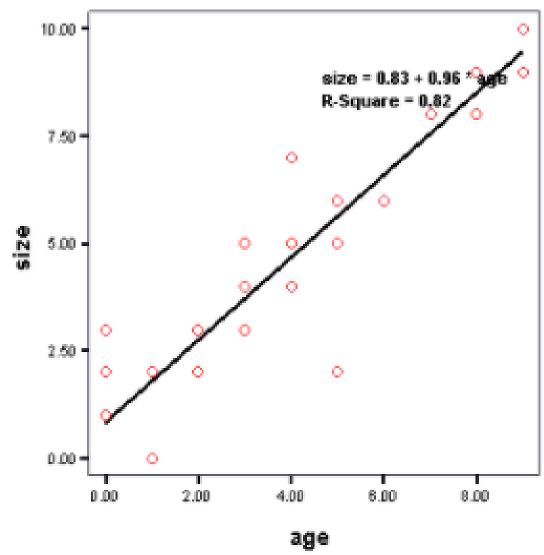


Figure 3. An Example of The Median Scatter Plotting (Line Prediction-Regression) Between Price and Size [23].

In this figure, the median trace (predicted line) clarifies the positive association between both size and price. However, the predicted line is nothing but is plotting of the horizontal x-value (presents size) which is divided into equally spaced segment. In this case, the median of the corresponding y-values (presents price) is plotting in the midpoint of each line segment as in shown in Figure (3).

5. MATRIX BASED SVD-PSEUDO INVERSE LINEAR REGRESSION

The linear regression line fitting or (relation1ship-association prediction) has an equation that formally define as Equation (30) shows below [24].

$$Y = a + bX \tag{30}$$

Were X is the explanatory variable, while Y is the dependent variable. b is the slop of the predicted line, and a is the intercept which means that the value of y when $x = 0$.

Let assume that we have such a matrix A where the dimension of it is m, and it is mathematically represented as

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & t_m \end{bmatrix}$$

While the slop b is also mathematically represented as:

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix}$$

Then the predicted line is mathematically represented based on the formula that is shown above in Equation (28):

$$\min_{x_1 x_2} \sum_{i=1}^{m} (x_1 + t_i \times x_1 - y_i)^2$$

Based on that the matrix A will be computed as:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1.4501 \\ 1.7311 \\ 3.1068 \\ 3.9860 \\ 5.3913 \end{bmatrix}$$

Then SVD Pseudo-Inverse matrix decomposition will be computed as:

$$A^T A = \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}$$

Then, the regression coefficient is also be computed as:

$$c = A^T b = \begin{bmatrix} 16.0620 \\ 58.6367 \end{bmatrix}$$

And based on the main regression formula:

$$5x_1 + 15\,x_2 = 16.0620$$
$$15x_1 + 55\,x_2 = 58.6367$$

Then, the regression value (relationship predicted) are:

$$x_1 = 0.0772$$
$$x_2 = 1.0451$$

Finally, by using the scatter plotting of the predicted value the main egression values will be appear as the blue line between the green circles which are the real values as they shown in Figure (4) [24].

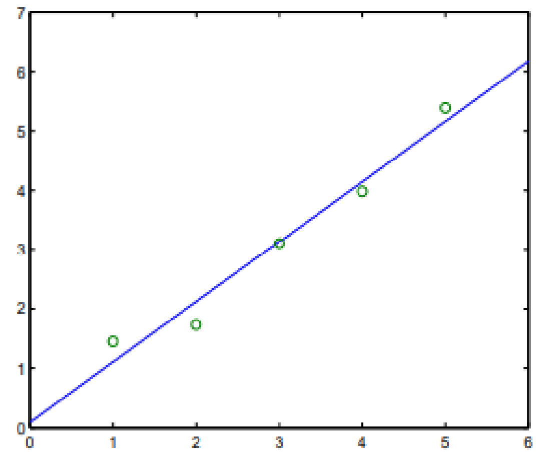


Figure 4. Matrix Pseudo-Inverse Linear Regression Example [20].

6. EXPERIMENTAL RESULTS

It is being clear that the SVD can be used in regression to analysis the data and determine (predict) the relationship between the data variables. The experimental results have been implemented using python and it has been based on the open source python code that has been implemented by Austin [24]. The data point that has been used in these experimental results are shown in Table (1).

Table 1:The Original data points

Original x	Original y
0	2.2
1	2.2
2	1
3	3
4	3
5	4
6	3
7	6
8	6
9	7
10	11
11	12
12	14
13	10
14	11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In this case, we will do both ways of determining the linear regression to determine and identical the results.

6.1 Linear Regression Experiential Results

The simple linear regression prediction results for the input data is shown in Table (2) and Figure (5).

Table 2:Simple Linear Regression Results

Original x	Original y	Linear fit x	Linear fit y
0	2.2	0	0.138108004
1	2.2	1	1.039831482
2	1	2	1.941554966
3	3	3	2.843278437
4	3	4	3.745001915
5	4	5	4.646725393
6	3	6	5.544844887
7	6	7	6.450172348
8	6	8	7.351589582
9	7	9	8.253619303
10	11	10	9.155342781
11	12	11	10.05706626
12	14	12	10.95878974
13	10	13	11.86051321
14	11	14	12.76223669

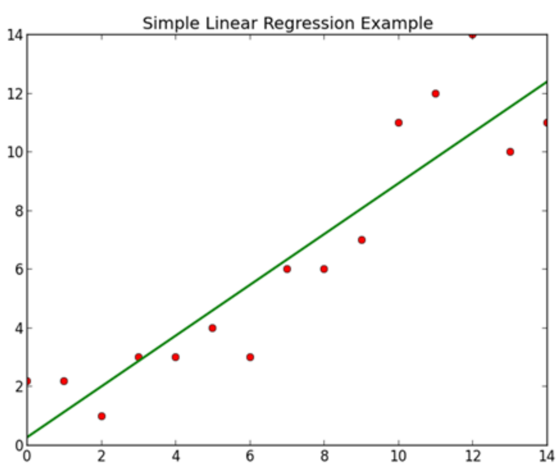


Figure 5. Simple Linear Regression Experimental Results [21].

It is clearly seen that the difference between the original points and the predicted line using simple linear regression as it shown in Figure (6).

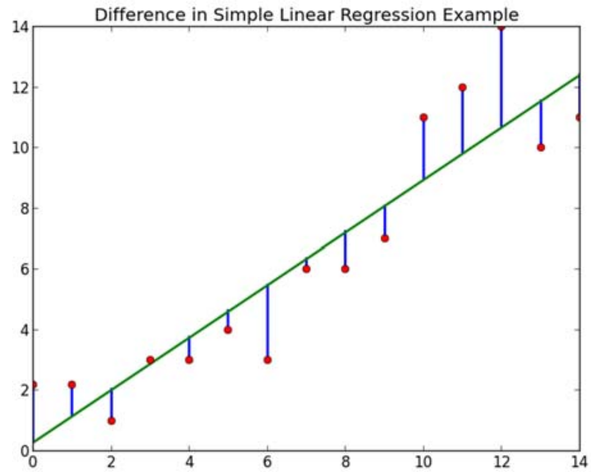


Figure 6. Different between the Original Points and the Predicted Lines using Simple Regression [24].

6.2 Linear Regression Experiential Results

To compare the simple regression results and the SVD regression using matrix-pseudo inverse, we take the same input data and used the SVD regression. The results shown in Table (3) and Figure (7) as well.

Table 3:SVD-Pseudo-Inverse Regression Results

Original x	Original y	Linear fit x	Linear fit y
0	2.2	0.13810800	0.115676843
1	2.2	1.03938314	1.024566326
2	1	1.94155496	1.933455809
3	3	2.84327843	2.842345292
4	3	3.74500191	3.751234776
5	4	4.64672539	4.660124259
6	3	5.54844887	5.569013742
7	6	6.45017234	6.477903225
8	6	7.35189582	7.386792708
9	7	8.25361930	8.295688219
10	11	9.15534278	9.204571674
11	12	10.0570662	10.11346116
12	14	10.9587897	11.02235064
13	10	11.8605132	11.93`24012
14	11	12.7622366	12.84012961

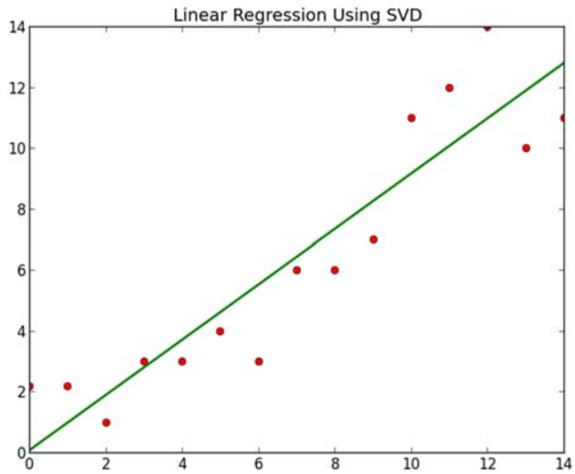


Figure 7. SVD- Regression Experimental Results based on the Matrix-Pseudo Inverse Approach [24].

It is easily to notice that the SVD regression based on Matrix-Pseudo-Inverse approach prediction results are nearly identical that the simple regression results as expected.

7. CONCLUSION

In this paper, we present such a mathematical description of one of the most significant application of the Singular Value Decomposing Approach (SVD). Matrix-Pseudo Inverse of SVD regression has been selected as one application of the SVD. The mathematical examples show that the SVD regression approach based on the Matrix-Pseudo Inverse is more realistic than the simple regression approach when it has been tested and compared with the simple regression. The SVD regression approach mainly based on the Matrix-Decomposition which makes the regression of such a real complex problem more realistic and expected than the simple approach which is the simple regression approach. However, main significant case which the pseudo-inverse of the non-invertible matrix based on the Singular Value Decomposition (SVD) is discussed in this paper to show how the pseudo-inverse of the non-squared matrix can be computed based on using the component of the SVD.

REFERENCES

[1] Moore, E. H, "Bulletin of the American Mathematical Society", 26 (9): 394–95. doi:10.1090/S0002-9904-1920-03322-7.

[2] Lecturer: Edo Liberty, Lecture 7: Singular Value Decomposition- Algorithms in Data Mining course.

[3] https://www.genealogy.math.ndsu.nodak.edu/id.php?id=20648.

[4] J. C. A. Barata1 and M. S. Hussein, The Moore-Penrose Pseudoinverse. A Tutorial Review of the Theory", Mathematical Physics, DOI: 10.1007/s13538-011-0052-z, Oct 2011.

[5] R. PENROSE," A GENERALIZED INVERSE FOR MATRICES", Downloaded from https://www.cambridge.org/core. IP address: 66.86.130.194, on 14 Mar 2018 at 02:57:42, subject to the Cambridge Core terms of use.

[6] Golub, Gene H.; Loan, "Matrix computations", (3rd ed.). Baltimore: Johns Hopkins. pp. 257–258. ISBN 0-8018-5414-8.

[7] Rakočević, Vladimir, "On continuity of the Moore–Penrose and Drazin inverses", Matematički Vesnik. 49: 163–72.

[8] Fallucchi Francesca, Fabio Massimo Zanzotto," SVD Feature Selection for Probabilistic Taxonomy Learning", GEMS '09 Proceedings of the Workshop on Geometrical Models of Natural Language Semantics, Pages 66-73, 2009.

[9] Golub, Gene H.; Kahan, William (1965). "Calculating the singular values and pseudo-inverse of a matrix". Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis. 2 (2): 205–224. doi:10.1137/0702016. JSTOR 2949777.

[10] Charu C. Aggarwal, "Data Mining:", Charu C. Aggarwal, IBM T. J. Watson Research Center, Yorktown Heights, New York, March 8, 2015.

[11] Introduction to Machine Learning Second Edition.

[12] Liu huiqing," effective use of data mining technologies on biological and clinical data", National University of Singapore, 2004.

[13] Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien," Semi-Supervised Learning", The MIT Press Cambridge, Massachusetts, London, England, 2006 Massachusetts Institute of Technology.

[14] Alex Smola and S.V.N. Vishwanathan," Introduction to Machine Learning", published by the press syndicate of the university of Cambridge The Pitt Building, Trumpington Street, Cambridge, United Kingdom, cambridge university press, The Edinburgh Building, Cambridge CB2 2RU, UK, 40 West 20th Street, New York, NY 10011{4211, USA, 477 Williamstown Road, Port Melbourne, VIC 3207, Australia Ruiz de Alarc_on 13, 28014 Madrid, Spain Dock House, The Waterfront,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Cape Town 8001, South Africa, http://www.cambridge.org, Cambridge University Press 2008.
[15] Shai Shalev-Shwartz and Shai Ben-David," Understanding Machine Learning: From Theory to Algorithms", Published 2014 by Cambridge University Press, http://www.cs.huji.ac.il/~shais/Understanding MachineLearning.
[16] Adi Ben-Israel and Thomas N.E. Greville. Generalized Inverses: Theory and Applications. Second Edition. Springer, 2003.
[17] Christos Boutsidis and David P Woodruff. Communication-optimal distributed principal component analysis in the column-partition model. arXiv preprint arXiv:1504.06729, 2015.
[18] A. S. Housetenes, "Unitary Triangularization of a Non-Symmetric Matrix", J. Assoc. Compute. March., 5, 1985, pp. 51-90.
[19] D. D. Morrison, "Remark on The Unitary Triangularization of a Nonsymmetric Matrix", J. Assoc. Comput. March., 7, 1960, pp. 185-186.
[20] G. Golub and W. Kahan, "Calculating the Singular Value and Pseudo-Inverse of a Matrix", J. Siam, Vol.2, No.2, USA, 1985.
[21] Sanjeev Arora and Boaz Barak," Computational Complexity: A Modern Approach", Princeton University complexitybook@gmail.com.
[22] Sharyn O'Halloran, "Lecture 2 Linear Regression: A Model for the Mean", available at, http://www.columbia.edu/~so33/SusDev/Lecture2.pdf.
[23] http://www.ltcconline.net/greenl/courses/201/regression/scatter.htm.
[24] https://machinelearningmastery.com/solve-linear-regression-using linear-algebra.
[25] http://www.kean.edu/~fosborne/bstat/09rc.html.