T.C.
YILDIZ TECHNICAL UNIVERSITY
FACULTY OF ART AND SCIENCES
DEPARTMENT OF MATHEMATICS

SINGULAR VALUE DECOMPOSITION FOR MATRICES

Prepared By

METEHAN KURTBAŞ

17025004

Prepared at Department of Mathematics

UNDERGRADUATE THESIS

Thesis Advisor: Assoc. Prof. Dr. Eyüp KIZIL

İSTANBUL, 2022

CONTENTS

ACKNOWLEDGEMENTS

I want to thank to my advisor Assoc. Prof. Eyüp KIZIL and my family for their support.

March, 2022

Metehan KURTBAŞ

LIST OF SYMBOLS

A^{T}	Transpose of A
A^{-1}	Inverse matrix of A
$det(A)$	Determinant of A
$M_{\mathrm{m\ x\ n}}$ (R)	The set of all $m \times n$ matrices with real entries
$rank(A)$	Rank of A

LIST OF FIGURES

ABSTRACT

This thesis focuses on singular value decomposition for matrices. We present in the first chapter basic definitions on linear algebra. The second chapter deals with the main subject of our work, namely, singular value decomposition of matrices. We end the thesis with a brief section about some applications of singular value decomposition.

Key Words: *Matrix***,** *Singular Value, Decomposition*

1. PRELIMINARIES

First of all we start with a very basic level of linear algebra knowledge involving matrices.

DEFINITION 1.1 (Matrix) A matrix A is rectangular array or table of objects such as numbers, symbols, expressions, etc arranged in m rows and and n columns, which we denote by

$A = [a_{ij}]_{m \times n}$.

If a matrix A has the same number of rows and columns (i.e., $m = n$), we say it is a square matrix of order n.

Among square matrices, some of them are of particular importance such as diagonal matrices, orthogonal matrices, symmetrix, matrices etc. Hence, we find it convenient to define at least these matrices here.

We let $M_{m \times n}$ (R) denote the set of all $m \times n$ matrices with real entries. In particular, the set

M_n (R) will stand for the set of all square matrices of order n.

Recall that all the entries a_{ij} of a matrix A for which i=j form the main diagonal of the matrix.

DEFINITION 1.2 (Diagonal matrix) A matrix for which all the entries outside the diagonal are zero is called diagonal.

DEFINITION 1.3 (Tranpose of a matrix) The tranpose of a matrix A is obtained by interchanging the rows and columns of A and denoted by A^T.

DEFINITION 1.4 Orthogonal matrix) A square matrix A such that $AA^T = A^TA = I$, where I denotes the identity matrix, is called orthogonal. Equivalently, A is orthogonal if $A^T = A^{-1}$, where A^{-1} means the inverse matrix of A.

DEFINITION 1.5 (Symmetric matrix) A square matrix A whose entries are real numbers is called symmetric (resp. skew-symmetric) if $A^T = A$ (resp. $A^T = - A$).

DEFINITION 1.6 (Rank of a matrix) The maximum number of linearly independent rows (resp. columns) of a matrix A is called rank of A which we denote by rank(A).

It is clear that one might consider some operations also with matrices such as matrix addition, multiplication of a matrix by a scalar and matrix multiplication, etc. The latter one is of most importance for us so that we define it below;

DEFINITION 1.7 (Matrix multiplication) Let A=[a_{ij}]∈ $M_{m \times n}$ (R) and B=[b_{jk}]∈ $M_{n \times p}$ (R) and matrix multiplication fot these two matrices defined as

$$AB = \left[\sum_{j=1}^{n} (a_{ij} b_{jk})\right]_{m \times p}$$

Let S = {1, 2, …, n} be the set of positive integers from 1 to n, arranged in ascending order. A rearrangement $j_1 j_2 \ldots j_n$ of the elements of S is called a *permutation* of S. We denote the set of all permutations of S by S_n.

A permutation $j_1 j_2 \ldots j_n$ is said to have an inversion if a larger integer, j_r precedes a smaller one. j_s. A permutation is called even if the total number of inversions in it is even. or odd if the total number of inversions in it is odd. If n≥2, there are n!/2 even and n!/2 odd permutations in S_n.

DEFINITION 1.8 (Determinant of a square matrix) Let A = [a_{ij}] be an $m \times n$ matrix. The determinant function, denoted by det, is defined by

$$\det(A) = \sum(\pm) a_{1j1} a_{2j2} \ldots a_{njn}$$

where the summation is over all permutations $j_1 j_2 \ldots j_n$ of the set S = {1, 2, …, n}.The sign is taken as + or - according to whether the permutation $j_1 j_2 \ldots j_n$ is even or odd.

In particular, for 2×2 and 3×3 matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ we have respectively det (A) = ad – bc and det (B) = aei + bfg + cdh – ceg – ahf – ibd). In both expressions, the term of positive sign correspond to even permutation of the set $S = \{1,2\}$ (resp. $S = \{1,2,3\}$) and those of negative signs to odd permutations.

DEFINITION 1.9 (Characteristic polynomial/equation of a matrix) Let A be an $n \times n$ matrix. Then the determinant of the $\lambda I_n - A$ is called the *characteristic polynomial* of A. The equation

$$P(\lambda) = \det(\lambda I_n - A) = 0$$

is called the *characteristic equation* of A.

DEFINITION 1.10 (Eigenvalue and Eigenvector of a matrix) In a square matrix A if there is a column matrice or vector $X \neq 0$ that $AX = \lambda X$ then the scalar λ is called an *eigenvalue* of A and the nonzero vector X is the *eigenvector* corresponding to eigenvalue λ.

The notion of an eigenvalue is important in calculating singular value of a matrix and (See Definition 1.12 below). Moreover, eigenvalues and their corresponding eigenvectors are also important for singular value decomposition of matrices since one might use them to calculate left and right singular matrices (See Definition 1.11).

DEFINITION 1.11 (Left and right singular matrices) Let A be any matrix. Then a matrix U whose columns contain eigenvectors of AA^{T} is said to be *left singular matrix*. Analogously, a matrix V whose columns are eigenvectors of $A^{\mathrm{T}}A$ is called *right singular matrix*.

Its known that symmetric matrices possess real eigenvalues and the eigenvector corresponding to distinct eigenvalues are always orthogonal (See Definition 1.14)

DEFINITION 1.12 (Singular value of a matrix) Let A be a $m \times n$ matrix. Square roots of eigenvalues (which are non-negative) of the symmetric matrix $A^{\mathrm{T}}A$ associated to A are called *singular values* of A.

DEFINITION 1.13 (Real vector space) Let V be set of vectors, c and d be any real numbers. Suppose that the operations $\otimes$ (scaler multiplication) and $\oplus$ (vector addition) on V are closed. Then we say V is a real vector space if satisfies the following properties;

(1) $\boldsymbol{u} \oplus \boldsymbol{v} = \boldsymbol{v} \oplus \boldsymbol{u}$ for every $\boldsymbol{u}, \boldsymbol{v}$ in V

(2) There exists an element $\mathbf{0}$ in V such that $\mathbf{0} \oplus \boldsymbol{u} = \boldsymbol{u} \oplus \mathbf{0}$ for any $\boldsymbol{u}$ in V (We call $\mathbf{0}$ as zero vector.)

(3) $\boldsymbol{u} \oplus (\boldsymbol{v} \oplus \boldsymbol{w}) = (\boldsymbol{u} \oplus \boldsymbol{v}) \oplus \boldsymbol{w}$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in V

(4) For every $\boldsymbol{u}$ in V there exists an element $-\boldsymbol{u}$ in V such that $\boldsymbol{u} \oplus -\boldsymbol{u} = -\boldsymbol{u} \oplus \boldsymbol{u} = \mathbf{0}$.

(5) $c \otimes (\boldsymbol{u} \oplus \boldsymbol{v}) = c \otimes \boldsymbol{u} \oplus c \otimes \boldsymbol{v}$ for every $\boldsymbol{u}, \boldsymbol{v}$ in V

(6) $(c + d) \otimes \boldsymbol{u} = c \otimes \boldsymbol{u} \oplus d \otimes \boldsymbol{u}$ for every $\boldsymbol{u}$ in V

(7) $c \otimes (d \otimes \boldsymbol{u}) = (cd) \otimes \boldsymbol{u}$ for every $\boldsymbol{u}$ in V

(8) $1 \otimes \boldsymbol{u} = \boldsymbol{u}$ for every $\boldsymbol{u}$ in V

DEFINITION 1.14 (Inner product) Let V be a real vector space. An inner product on V is a function that assigns to each ordered pair of vectors $\boldsymbol{u}, \boldsymbol{v}$ in V a real number $(\boldsymbol{u}.\boldsymbol{v})$ satisfying the following properties:

(1) $(\boldsymbol{u}.\boldsymbol{u}) \geq \boldsymbol{0}$; $(\boldsymbol{u}.\boldsymbol{u}) = \boldsymbol{0}$ if and only if $\boldsymbol{u} = \boldsymbol{0}$

(2) $(\boldsymbol{u}.\boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{u})$ for every $\boldsymbol{u}, \boldsymbol{v}$ in V

(3) $(\boldsymbol{u} \oplus \boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{u}.\boldsymbol{w}) + (\boldsymbol{v}.\boldsymbol{w})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in V

(4) $(c \otimes \boldsymbol{u}, \boldsymbol{v}) = c(\boldsymbol{u}.\boldsymbol{v})$ for every $\boldsymbol{u}.\boldsymbol{v}$ in V and c is any real number.

For X and Y n dimension vectors their inner product defined as;

$$< X, Y > = \sum_{i=1}^{n} x_i y_i$$

We noticed that inner product of two vectors is equal to a real number.

X and Y vectors are called perpendicular or orthogonal if their inner product is equal to 0.

A set $\{X_1, X_2, \ldots, X_n\}$ of vectors forms an orthogonal system if the inner product of every two different vector is equal to 0.

If every vector in an orthogonal system $\{X_1, X_2 \ldots , X_n\}$ is unit then $\{X_1, X_2, \ldots , X_n\}$ forms an orthonormal system.

DEFINITION 1.15 (Inner product space) A real vector space equipped with an inner product is called an inner product space.

DEFINITION 1.16 (Orthonormal basis)) Let V be an inner product space and $S = \{u_1, u_2, \ldots , u_n\}$ be an ordered basis for V. We say S is an *orthonormal basis* if it satisfies:

(1) $<u_i , u_j> = 0$ for every $i \neq j$

(2) $<u_i , u_i> = 1$ for $i = 1 \ldots$ n

DEFINITION 1.17 (Gram-Schmidt process)

Theorem: Let V be an inner product space and $V > W \neq \{\mathbf{0}\}$ and number of linear indepently vector in W be m. Then there exists an orthonormal basis $T = \{t_1, t_2, \ldots ,t_m\}$ for W.

To formulate this theorem let W and T be $W = \{w_1, w_2, \ldots ,w_m\}$ and $T = \{t_1, t_2, \ldots ,t_m\}$ then

$$t_m = w_m - \sum_{j=\perp}^{m-1} \frac{\langle w_m, t_j \rangle}{\langle t_j, t_j \rangle} t_j$$

2. SINGULAR VALUE DECOMPOSITION OF MATRICES

There are numerous decompositions for matrices like spectral decomposition, LU decomposition, QR decomposition, and etc. We refer the reader **[1]** for more detailed exposition on such decomposition theorems. Nonetheless, the main purpose here is another kind of decomposition so that we study now singular value decomposition of matrices.

Remember the theorem that we can decompose an $n \times n$ symmetric matrix A as follows

$$A = PDP^T$$

where D is a diagonal matrix and P is an orthogonal matrix. The diagonal entries of D are the eigenvalues of A, $\lambda_1, \lambda_2, \dots, \lambda_n$, and the columns of P are associated orthonormal eigenvectors $x_1, x_2, \dots, x_n$.

Now, M be $m \times n$ real matrix. Then there exist orthogonal matrices U of size $m \times m$ and V of size $n \times n$ such that

$$M = U\,S\,V^T, \tag{1}$$

where S is an $m \times n$ matrix with nondiagonal entries all zero and

$$s_{11} \geq s_{12} \geq \dots \geq s_{pp} \geq 0,$$

where $p = min\{m, n\}$.

The diagonal entries of S are singular values of M and U is the left singular matrix of M while V means the right singular matrix of M.

The singular value decomposition of M in (1) can be expressed as the following linear combination:

$$M = \mathrm{col}_1(U)\, s_{11}\mathrm{col}_1(V)^T + \mathrm{col}_2(U)\, s_{22}\mathrm{col}_2(V)^T + \dots + \mathrm{col}_p(U)\, s_{pp}\mathrm{col}_p(V)^T \tag{2}$$

To determine the matrices U, S and V in the singular value decomposition given in (1), we start as follows: An $n \times n$ symmetric matrix related to M is M^TM. By theorem given up there exists an orthogonal $n \times n$ matrix V such that

$$V(M^TM)V^T = D,$$

where D is a diagonal matrix whose diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of M^TM. If we denote column j of V as v_j, then $(M^TM)\, v_j = \lambda_j v_j$. Let's multiply both sides of this equation on the left by v_j^T; then we can rearrange the expression as;

$$v_j^T(M^TM)v_j = \lambda_j v_j^T v_j \quad \text{or} \quad (Mv_j)^T(Mv_j) = \lambda_j v_j^T v_j \quad \text{or} \quad |Mv_j|^2 = \lambda_j |v_j|^2$$

It is clear that length of a vector is nonnegative so the last expression implies that $\lambda_j \geq 0$.

Finally, we determine the $m \times m$ orthogonal matrix U. Given the matrix equation in (1), let us see what the columns of U should look like.

- Since U is to be orthogonal, its columns must be an orthonormal set; hence they arc linearly independent $m \times 1$ vectors .

- The matrix S has the form (block diagonal)

$$S=\begin{pmatrix} \begin{pmatrix} s_{11} & 0 & \dots & 0 \\ 0 & s_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & s_{pp} \end{pmatrix} & \vdots & Op,n-p \\ \cdots & \vdots & \cdots \\ Om-p,p & \vdots & Om-p,n-p \end{pmatrix}$$

where $O_{r,s}$ denotes an $r \times s$ matrix of zeros.

- From (1), $MV = US$, so

$$A\,(v_1\; v_2 \dots v_n) = (u_1\; u_2 \dots u_m)\begin{pmatrix} \begin{pmatrix} s_{11} & 0 & \dots & 0 \\ 0 & s_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & s_{pp} \end{pmatrix} & \vdots & Op,n-p \\ \cdots & \vdots & \cdots \\ Om-p,p & \vdots & Om-p,n-p \end{pmatrix}$$

This implies that we need to require that $Av_j = s_{jj}\, u_j$ for $j = 1,2 \ldots p$

However, U must have m orthonormal columns, and $m \geq p$.

We use Gram-Schmidt process here to obtain the remaining m-p columns of U. (This is necessary only if $m \geq p$.) Since these m-p columns are not unique, matrix U is not unique. (Neither is V if any of the eigenvalues of $M^T M$ are repeated.)

Now, we give an example.

EXAMPLE 2.1 Consider the matrix $M = \begin{pmatrix} 6 & 4 & 4 \\ 4 & 6 & -4 \end{pmatrix}$. It follows that the singular values of M might be obtained through the eigenvalues of MM^T. Hence, we form first the associated symmetric matrix below :

$$MM^T = \begin{pmatrix} 68 & 32 \\ 32 & 68 \end{pmatrix}$$

for which the characteristic polynomial is

$$P(\lambda) = det(MM^T - \lambda I) = \lambda^2 - 136\lambda + 3600 = (\lambda - 100).(\lambda - 36).$$

It follows that the eigenvalues are $\lambda = 100$ and $\lambda = 36$ which means the singular values are 10 and 6, respectively.

Since S contains the eigenvalues of MM^T in its diagonal we obtain the matrix S as follows:

$$S = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 6 & 0 \end{pmatrix}$$

Now, the corresponding eigenvectors of MM^T is needed. For the eigenvalue $\lambda = 100$ we look for a non zero X such that $(M - 100I)X = 0$.

For the eigenvalue 100;

$(M - 100I)X = 0$ then

$X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ unit form of its is; $X = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$

For the eigenvalue 36;

$(M\text{-}36I)\ X = 0$ then

$X = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ unit form of its is; $X = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{pmatrix}$

Note: If we couldn't find these eigenvectors as unit form we must use Gram-Schmidt process.

Now we need to find the eigenvalues of $M^{T}M$.

$M^{T}M = \begin{pmatrix} 52 & 48 & 8 \\ 48 & 52 & -8 \\ 8 & -8 & 32 \end{pmatrix}$

$P(\lambda) = det = -\lambda^3 + 136\lambda - 3600\,\lambda = -\lambda\,.(\lambda - 100).(\lambda - 36)$

$\lambda = 0,\ \lambda = 36,\ \lambda = 100$

Now, we can find eigenvectors of $M^{T}M$.

For the eigenvalue 0;

$(M^{T}M - 0.I).\ X = 0$ then

$X = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$ unit form of its is; $X = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}$

For the eigenvalue 36;

$(M^{T}M - 36.I).\ X = 0$ then

$$X = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \text{ unit form of its is; } X = \begin{pmatrix} \frac{\sqrt{2}}{6} \\ \frac{-\sqrt{2}}{6} \\ \frac{2\sqrt{2}}{3} \end{pmatrix}$$

For the eigenvalue 100;

(M^T M – 100.I). $X = 0$ then

$$X = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ unit form of its is; } X = \begin{pmatrix} \frac{-2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

Finally we find;

$$U = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \qquad S = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 6 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & -\frac{2}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

3. SOME APPLICATIONS ON COMPUTER ENVIROMENT

Digitized images are big matrix of numbers. Due to singular value decomposition (SVD) we can drop cost of storage and we can get better display resolution with using rank of matrix notion.

In this section 3 examples of these applications are given.

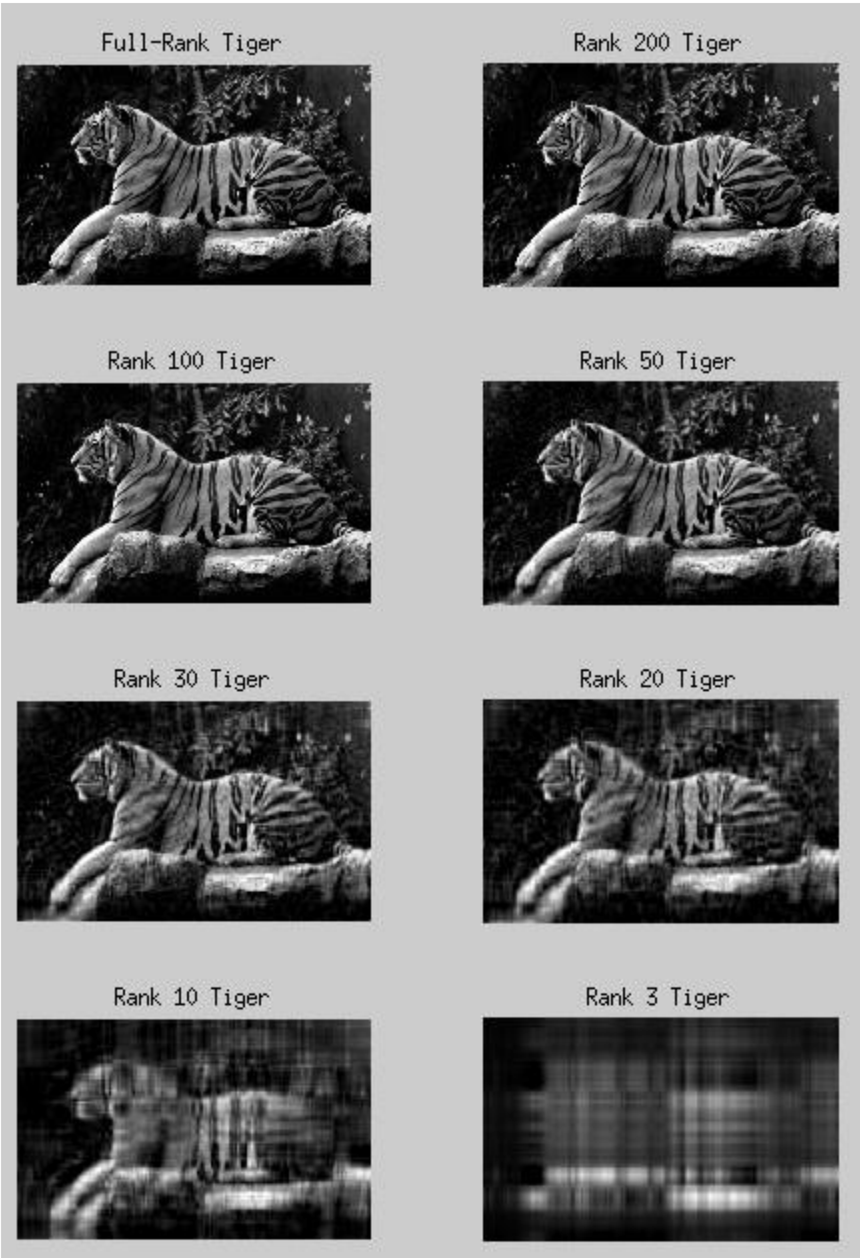

Figure 3.1 Application of SVD with low rank approximation on tiger image. **[2]**

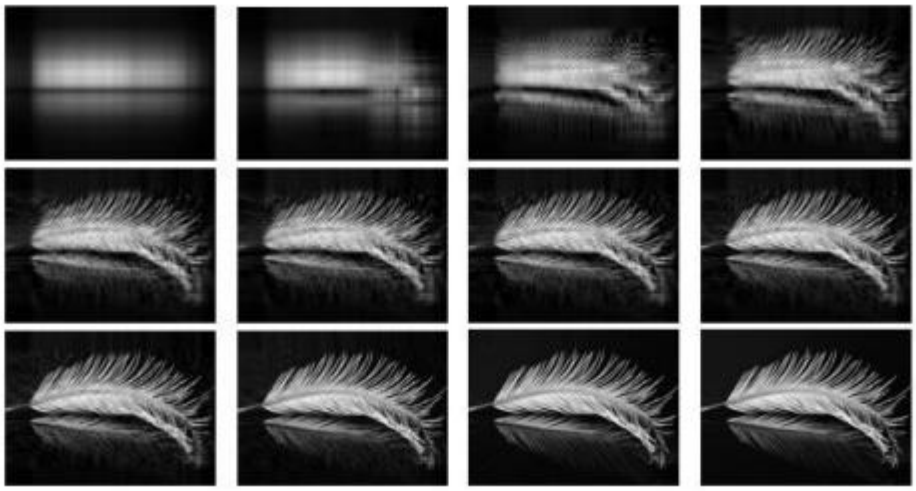

Figure 3.2 Application of SVD with low rank approximation on plume image. **[3]**

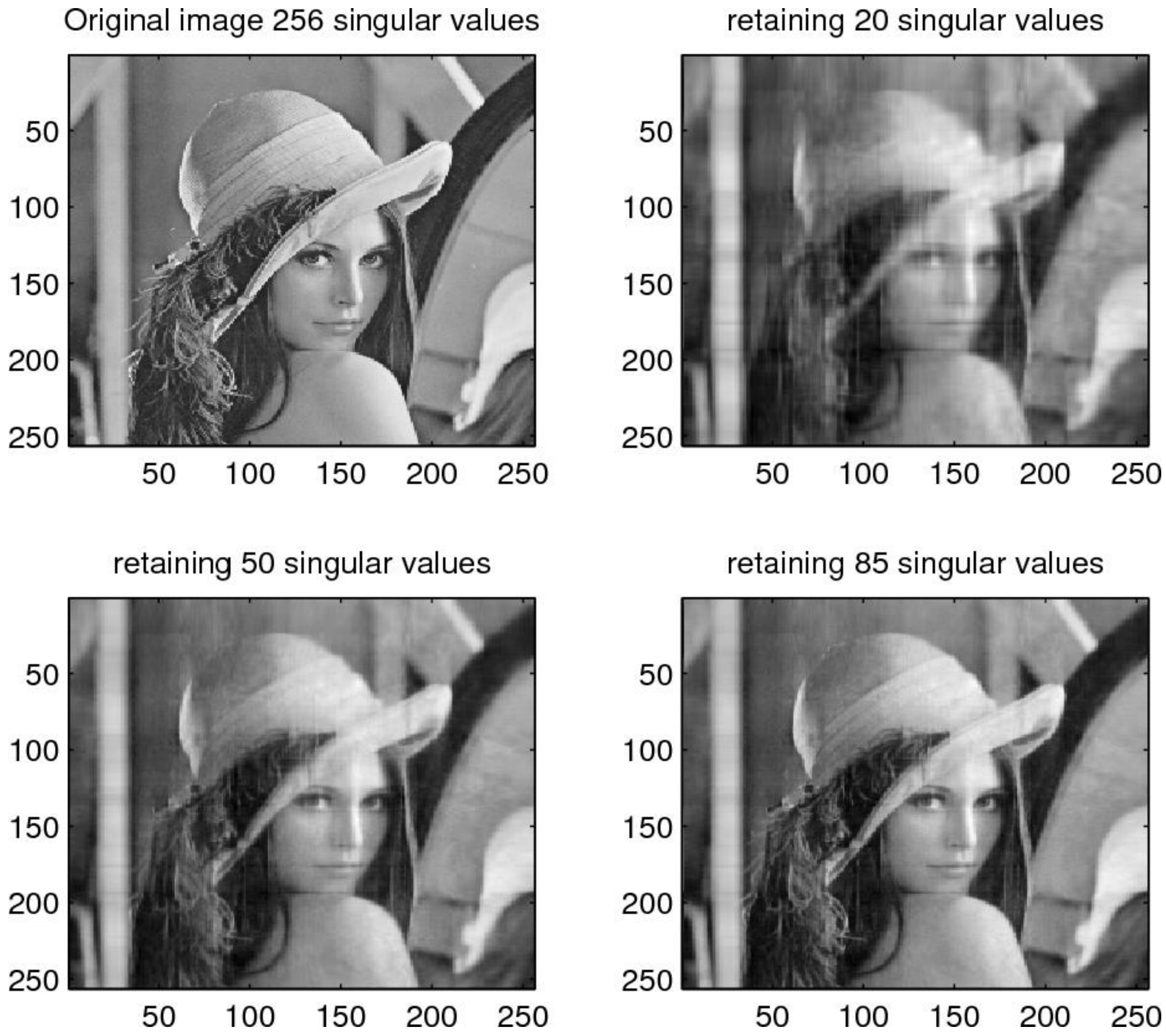

Figure 3.3 Application of SVD with low rank approximation on human face image. **[4]**

RESULTS

Singular value decomposition (SVD) is more than abstract mathematics. It is a way to increasing welfare level and making human life’s easier via applinyg it on industry like social media sites, immediare data providers on finance and etc.

REFERENCES

[1] Kolman B., Hill D. R., Elemantery Linear Algebra with Applications, 9th Edition, Pearson Education Inc, New Jersey, 2007

[2] https://andrew.gibiansky.com/blog/mathematics/cool-linear-algebra-singular-value-decomposition/

[3] https://www.lagrange.edu/academics/undergraduate/undergraduate-research/citations/18-Citations2020.Compton.pdf

[4] https://www.semanticscholar.org/paper/Singular-Value-Decomposition-in-Image-Noise-and-Workalemahu/8af232fb669dfb8c069bceee0ac67a8406d21162

[5] https://www.youtube.com/watch?v=mBcLRGuAFUk

CV

Personal Information

Name Surname: Metehan KURTBAŞ

Date of birth: 19.02.1997

E-mail: kurtbasmetehan@gmail.com

Education

YTU - Mathematics 2017-…

METU - School of Foreign Languages 2015-2016

Kars Fen High School 2011-2015

Work Experience

I'm involved in Istanbul Stock Exchange, FX, Crypto Currencies, NASDAQ via crypto currency platforms and trade there.

I work as a data analyser with the help of my knowledge of mathematics on social media as I have majored in Maths. I am helping to increase real visits. Also, sometimes I give content suggestions to people or companies.

I worked as the creator of recording systems in some companies. I mean I didn't code. I created the theoric basic of recording system thanks to my expertise in maths.