Question 5

CS663 (Digital Image Processing) Assignment 4

Atishay Jain (210050026)
Cheshta Damor (210050040)
Kanad Shende (210050078)

210050026@iitb.ac.in
210050040@iitb.ac.in
210050078@iitb.ac.in

Contents

PART

I

Question 5

Problem 1

The aim of this exercise is to help you understand the mathematics of SVD more deeply. Do as directed: [5+5+10=20 points]

1. Argue that the non-zero singular values of a matrix $\boldsymbol{A}$ are the square-roots of the eigenvalues of $\boldsymbol{A}\boldsymbol{A}^T$ or $\boldsymbol{A}^T\boldsymbol{A}$. (Make arguments for both)

2. Show that the squared Frobenius norm of a matrix is equal to the sum of squares of its singular values.

3. A students tries to obtain the SVD of a $m \times n$ matrix $\boldsymbol{A}$ using eigendecomposition. For this, the student computes $\boldsymbol{A}^T\boldsymbol{A}$ and assigns the eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$ (computed using the `eig` routine in MATLAB) to be the matrix $\boldsymbol{V}$ consisting of the right singular vectors of $\boldsymbol{A}$. Then the student also computes $\boldsymbol{A}\boldsymbol{A}^T$ and assigns the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$ (computed using the `eig` routine in MATLAB) to be the matrix $\boldsymbol{U}$ consisting of the left singular vectors of $\boldsymbol{A}$. Finally, the student assigns the non-negative square-roots of the eigenvalues (computed using the `eig` routine in MATLAB) of either $\boldsymbol{A}^T\boldsymbol{A}$ or $\boldsymbol{A}\boldsymbol{A}^T$ to be the diagonal matrix $\boldsymbol{S}$ consisting of the singular values of $\boldsymbol{A}$. He/she tries to check his/her code and is surprised to find that $\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ is not equal to $\boldsymbol{A}$. Why could this be happening? What processing (s)he do to the eigenvectors computed in order rectify this error? (Note: please try this on your own in MATLAB.)

SECTION 1

Part-1

Given a m x n matrix A, Singular Value Decomposition (SVD) is a way to factor A as (this decomposition always exists):

$$A = USV^T$$

- $A \in \mathbb{R}^{m \times n}$
- $UU^T = U^TU = I_m, U \in \mathbb{R}^{m \times m}$
- $VV^T = V^TV = I_n, U \in \mathbb{R}^{n \times n}$
- $S \in \mathbb{R}^{m \times n}$

The S matrix contains the singular values of A on its pseudo-diagonal, with zeros elsewhere. Thus,

$$A = USV^T = \begin{bmatrix} u_1 \mid u_2 \mid \cdots \mid u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & & & & \vdots & & \vdots \\ & & \ddots & & & & & 0 \\ \vdots & & & \sigma_r & & & \ddots & \\ 0 & & & & 0 & \vdots & & \vdots \\ 0 & 0 & & \cdots & & \ddots & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_n^T \end{bmatrix}$$

with $u_1, ..., u_m$ being orthonormal columns of U, $\sigma_1, ..., \sigma_r$ being the singular values of A satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $v_1, ..., v_n$ being the orthonormal columns of V^T.
To prove that the non-zero singular values of a matrix A are the square-roots of the eigenvalues of both A^TA and AA^T, we can break the proof into two parts.

SUBSECTION 1.1

Eigenvalues of A^TA

We want to find the eigenvalues of the matrix A^TA. So let's compute A^TA:

$$A^TA = (USV^T)^T(USV^T) = (VS^TU^T)(USV^T) = VS^T(U^TU)SV^T$$

$$A^TA = VS^TSV^T$$

Now let's postmultiply the equation by V

$$A^T AV = VS^T SV^T V = VS^T S$$

Here, $S^T S$ is an n x n diagonal matrix with entries $\sigma_1^2, \sigma_2^2 ..., \sigma_r^2$ on the diagonal (the squares of singular values).
Going further, we can write $A^T AV$ as:

$$A^T AV = \left[\; A^T Av_1 \mid A^T Av_2 \mid \cdots \mid A^T Av_r \mid A^T Av_{r+1} \mid \cdots \mid A^T Av_n \right] \quad (1.1)$$

Also,

$$A^T AV = VS^T S = \left[\; v_1 \mid v_2 \mid \cdots \mid v_r \mid v_{r+1} \mid \cdots \mid v_n \right] \begin{bmatrix} \sigma_1^2 & 0 & \cdots\cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \sigma_r^2 & \\ 0 & & & 0 \\ 0 & 0 & \cdots & & \ddots \end{bmatrix}$$

$$A^T AV = \left[\; \sigma_1^2 v_1 \mid \sigma_2^2 v_2 \mid \cdots \mid \sigma_r^2 v_r \mid 0 \mid \cdots \mid 0 \right] \quad (1.2)$$

Hence, from equation (1.1) and (1.2), we can conclude:

- $A^T Av_1 = \sigma_1^2 v_1, \cdots, A^T Av_r = \sigma_1^2 v_r$, and
- $A^T Av_{r+1} = 0, \cdots, A^T Av_{r+(n-r)} = 0$

That is, the σ_i^2s are the eigenvalues of $A^T A$ with v_i as the corresponding eigenvectors. Hence, the non-zero singular values (σ_i) of A are the square-roots of the eigenvalues of $A^T A$.

SUBSECTION 1.2

Eigenvalues of AA^T

We want to find the eigenvalues of the matrix AA^T. So let's compute AA^T:

$$AA^T = (USV^T)(USV^T)^T = (USV^T)(VS^T U^T) = US(V^T V)S^T U^T$$

$$AA^T = USS^T U^T$$

Now let's postmultiply the equation by U

$$AA^T U = USS^T U^T U = USS^T$$

Here, SS^T is an m x m diagonal matrix with entries $\sigma_1^2, \sigma_2^2 ..., \sigma_r^2$ on the diagonal (the squares of singular values).
Going further, we can write $AA^T U$ as:

$$AA^T U = \left[\; AA^T u_1 \mid AA^T u_2 \mid \cdots \mid AA^T u_r \mid AA^T u_{r+1} \mid \cdots \mid AA^T u_m \right] \quad (1.3)$$

Also,

$$AA^TU = USS^T = \left[\, u_1 \mid u_2 \mid \cdots \mid u_r \mid u_{r+1} \mid \cdots \mid u_m \right] \begin{bmatrix} \sigma_1^2 & 0 & \cdots\cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \sigma_r^2 & \\ 0 & & & 0 \\ 0 & 0 & \cdots & \ddots \end{bmatrix}$$

$$AA^TU = \left[\, \sigma_1^2 u_1 \mid \sigma_2^2 u_2 \mid \cdots \mid \sigma_r^2 u_r \mid 0 \mid \cdots \mid 0 \right] \tag{1.4}$$

Hence, from equation (1.3) and (1.4), we can conclude:

- $AA^Tu_1 = \sigma_1^2 u_1, \cdots, AA^Tu_r = \sigma_1^2 u_r$, and
- $AA^Tu_{r+1} = 0, \cdots, AA^Tu_{r+(m-r)} = 0$

Thus, the first r vectors u_i are the eigenvectors of AA^T with the eigenvalues σ_i^2. Hence, the non-zero singular values (σ_i) of A are the square-roots of the eigenvalues of AA^T.

SECTION 2

Part-2

We need to show that the squared Frobenius norm of a matrix is equal to the sum of squares of its singular values.

PROOF The squared Frobenius norm of a $m \times n$ matrix A is defined as the sum of the squares of its entries:

$$\|A\|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2$$

here a_{ij} is the element of A at i^{th} row anf j^{th} column.
Using matrix multiplication, we can see that the trace of A^TA is the sum of the squares of elements of the matrix A.
Therefore we can write:

$$\|A\|_F^2 = trace(A^TA)$$

Further using SVD of A, we get:

$$\|A\|_F^2 = trace((USV^T)^T(USV^T))$$

$$= trace((VS^TU^T)(USV^T))$$

$$= trace(VS^T(U^TU)SV^T)$$

$$= trace(VS^TSV^T)$$

Now, we use the property that the trace of a product of matrices is invariant under cyclic permutation i.e, trace(ABC) = trace(CAB). Therefore,

$$= trace(VV^TS^TS)$$

$$= trace(S^TS)$$

In part-1, we have seen that S^TS is an diagonal matrix with entries $\sigma_1^2, \sigma_2^2 ..., \sigma_r^2$ on the diagonal. Hence

$$\|A\|_F^2 = \sum_i \sigma_i^2$$

This shows that the squared Frobenius norm of a matrix is equal to the sum of squares of its singular values □

Hence proved.

Section 3

Part-3

I tried to perform the SVD decomposition in MATLAB and came across 2 major issues for the last 2 outputs being different.

(a) The eig command returns the eigenvectors in ascending order of magnitude of eigenvalues, i.e. $v_1 \leq v_2 ... \leq v_k$. But our implementation requires this order to be in descending form, thus we need to reorganize these eigenvectors appropriately, which can be achieved by flipping the matrices U, V, and S column-wise.

(b) Secondly, each eigenvector with unit magnitude can be oriented in 2 directions. This leads to a possibility of 2^k solutions in the eigenvector matrix solution set. For SVD decomposition, let u be the eigenvector for U and v be the eigenvector for V corresponding to a eigenvalue λ. The condition required to be satisfied is $u = \frac{Av}{\sqrt{\lambda}}$, which won't always hold true for all the above solutions in our solution set. Hence in order to correct this, U can be calculated as $\frac{AV}{\sqrt{D}}$ instead of using the eig function in MATLAB. This will help in resolving the sign issue. Another cumbersome method could be iterating through all permutations of signs until we find the right fit.