CS663: Digital Image Processing - Homework 4

Harsh | Pranav | Swayam

October 22, 2024

Homework 4 - Question 3

(a)

Let $\mathbf{A}$ be an $m \times n$ matrix. The Singular Value Decomposition (SVD) of $\mathbf{A}$ can be expressed as:

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

where:
- $\mathbf{U}$ is an $m \times m$ orthogonal matrix,
- $\mathbf{S}$ is an $m \times n$ diagonal matrix with non-negative entries (singular values),
- $\mathbf{V}$ is an $n \times n$ orthogonal matrix.

Considering the product $\mathbf{A}^T\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = (\mathbf{U}\mathbf{S}\mathbf{V}^T)^T(\mathbf{U}\mathbf{S}\mathbf{V}^T) = \mathbf{V}\mathbf{S}^T\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{V}\mathbf{S}^T\mathbf{S}\mathbf{V}^T$$

Since $\mathbf{V}$ is orthogonal, the eigenvalues of $\mathbf{A}^T\mathbf{A}$ are the diagonal entries of $\mathbf{S}^T\mathbf{S}$, which are the squares of the singular values. Thus, the non-zero singular values of $\mathbf{A}$ are the square roots of the eigenvalues of $\mathbf{A}^T\mathbf{A}$.
Similarly, for $\mathbf{A}\mathbf{A}^T$:

$$\mathbf{A}\mathbf{A}^T = (\mathbf{U}\mathbf{S}\mathbf{V}^T)(\mathbf{U}\mathbf{S}\mathbf{V}^T)^T = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^T\mathbf{U}^T = \mathbf{U}\mathbf{S}\mathbf{S}^T\mathbf{U}^T$$

The eigenvalues of $\mathbf{A}\mathbf{A}^T$ are also the diagonal entries of $\mathbf{S}\mathbf{S}^T$, confirming that the non-zero singular values correspond to the square roots of the eigenvalues of both matrices.

(b)

The Frobenius norm is defined as:

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

Thus, we have:

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

From the SVD, we know:

$$\mathbf{A} = \mathbf{USV}^T$$

Therefore:

$$\|\mathbf{A}\|_F^2 = \|\mathbf{USV}^T\|_F^2 = \|\mathbf{S}\|_F^2$$

Since both $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, they do not change the Frobenius norm. The Frobenius norm of a diagonal matrix is simply the sum of squares of its diagonal elements (the singular values):

$$\|\mathbf{S}\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$$

Thus, we conclude that:

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$$

(c)

When attempting to reconstruct $\mathbf{A}$ from the product $\mathbf{USV}^T$, the result does not match the original matrix $\mathbf{A}$, causing confusion. This mismatch arises due to sign inconsistencies in the eigenvectors obtained during the eigen decomposition.

Explanation

Eigenvectors can have arbitrary signs: if $\mathbf{u}$ is an eigenvector of $\mathbf{AA}^T$ with eigenvalue λ, then $c\mathbf{u}$ is also an eigenvector with the same eigenvalue. This leads to potential sign discrepancies between the left singular vectors in $\mathbf{U}$ and the right singular vectors in $\mathbf{V}$, causing $\mathbf{A} \neq \mathbf{USV}^T$.

Mathematical Formulation

Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, we know that the SVD of $\mathbf{A}$ is given by:

$$\mathbf{A} = \mathbf{USV}^T$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ are obtained via the eigen decomposition:

$$\mathbf{A}^T\mathbf{A} = \mathbf{VSV}^T \quad \text{and} \quad \mathbf{AA}^T = \mathbf{USU}^T$$

where $\mathbf{S}$ is the diagonal matrix of eigenvalues, and the columns of $\mathbf{U}$ and $\mathbf{V}$ are the eigenvectors corresponding to $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$, respectively.
Sign Inconsistency-
The issue arises because both $\mathbf{u}$ and $-\mathbf{u}$ (similarly, $\mathbf{v}$ and $-\mathbf{v}$) are valid eigenvectors corresponding to the same eigenvalue λ. Therefore, the numerical eigen decomposition may return eigenvectors with arbitrary signs.
When constructing the SVD as $\mathbf{U}\mathbf{S}\mathbf{V}^T$, any sign discrepancy between corresponding columns of $\mathbf{U}$ and $\mathbf{V}$ can result in a mismatch, leading to:

$$\mathbf{A} \neq \mathbf{U}\mathbf{S}\mathbf{V}^T$$

```
Original matrix A:
    0.9572    0.4218    0.6557    0.6787
    0.4854    0.9157    0.0357    0.7577
    0.8003    0.7922    0.8491    0.7431
    0.1419    0.9595    0.9340    0.3922


Reconstructed matrix A using U * S * V^T (without fixing signs):
    0.2611    1.0889    0.6185    0.5901
    0.5500    0.3692    1.0506    0.3272
    0.6511    1.0116    0.6816    0.7933
    1.0089    0.4681    0.2716    0.8101
```

Figure 1: Incorrect SVD computation

Resolution of the Sign Inconsistency

To resolve this, for each i-th pair of eigenvectors $\mathbf{u}_i$ (from $\mathbf{A}\mathbf{A}^T$) and $\mathbf{v}_i$ (from $\mathbf{A}^T\mathbf{A}$), adjust the signs as follows:

$$\text{If } \operatorname{sign}(\mathbf{u}_i^T \mathbf{A} \mathbf{v}_i) < 0, \text{ set } \mathbf{u}_i \leftarrow -\mathbf{u}_i$$

This ensures the correct reconstruction $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$.

```
Original matrix A:
    0.4387    0.1869    0.7094    0.6551
    0.3816    0.4898    0.7547    0.1626
    0.7655    0.4456    0.2760    0.1190
    0.7952    0.6463    0.6797    0.4984


Reconstructed matrix A using U * S * V^T (after fixing signs):
    0.4387    0.1869    0.7094    0.6551
    0.3816    0.4898    0.7547    0.1626
    0.7655    0.4456    0.2760    0.1190
    0.7952    0.6463    0.6797    0.4984
```

Figure 2: Corrected SVD computation

(d)

(i)

Considering there are appropriate number of elements in the vector,

$$\mathbf{y}^T\mathbf{P}\mathbf{y} = \mathbf{y}^T\mathbf{A}^T\mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{y})^T(\mathbf{A}\mathbf{y}) = \|\mathbf{A}\mathbf{y}\|^2 \geq 0$$
$$\implies \mathbf{y}^T\mathbf{P}\mathbf{y} \geq 0$$

Similarly, for the matrix $\mathbf{Q}$, we have:

$$\mathbf{z}^T\mathbf{Q}\mathbf{z} = \mathbf{z}^T\mathbf{A}\mathbf{A}^T\mathbf{z} = (\mathbf{A}^T\mathbf{z})^T(\mathbf{A}^T\mathbf{z}) = \|\mathbf{A}^T\mathbf{z}\|^2 \geq 0$$
$$\implies \mathbf{z}^T\mathbf{Q}\mathbf{z} \geq 0$$

Now, the eigenvalues of a positive semi-definite matrix are non-negative. And, since $\mathbf{P}$ and $\mathbf{Q}$ are positive semi-definite matrices, as shown above, the eigenvalues of $\mathbf{P}$ and $\mathbf{Q}$ are non-negative.
Proof: Let λ be an eigenvalue of $\mathbf{P}$ and $\mathbf{v}$ be the corresponding eigenvector. Then:

$$\mathbf{P}\mathbf{v} = \lambda\mathbf{v}$$
$$\implies \mathbf{v}^T\mathbf{P}\mathbf{v} = \lambda\mathbf{v}^T\mathbf{v}$$
$$\implies \mathbf{v}^T\mathbf{P}\mathbf{v} = \lambda$$

Since $\mathbf{P}$ is positive semi-definite, $\mathbf{v}^T\mathbf{P}\mathbf{v} \geq 0$, which implies $\lambda \geq 0$. Thus, the eigenvalues of $\mathbf{P}$ are non-negative.
Similarly, for $\mathbf{Q}$, let μ be an eigenvalue of $\mathbf{Q}$ and $\mathbf{w}$ be the corresponding eigenvector. Then:

$$\mathbf{Q}\mathbf{w} = \mu\mathbf{w}$$
$$\implies \mathbf{w}^T\mathbf{Q}\mathbf{w} = \mu\mathbf{w}^T\mathbf{w}$$
$$\implies \mathbf{w}^T\mathbf{Q}\mathbf{w} = \mu$$

Since $\mathbf{Q}$ is positive semi-definite, $\mathbf{w}^T\mathbf{Q}\mathbf{w} \geq 0$, which implies $\mu \geq 0$. Thus, the eigenvalues of $\mathbf{Q}$ are non-negative.

(ii)

Let $\mathbf{u}$ be an eigenvector of $\mathbf{P}$ with eigenvalue λ. Then:

$$\mathbf{P}\mathbf{u} = \lambda\mathbf{u}$$
$$\therefore \mathbf{A}\mathbf{P}\mathbf{u} = \mathbf{A}(\lambda\mathbf{u}) = \lambda\mathbf{A}\mathbf{u}$$
$$\implies \lambda\mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{A}^T\mathbf{A}\mathbf{u} = (\mathbf{A}\mathbf{A}^T)\mathbf{A}\mathbf{u} = \mathbf{Q}\mathbf{A}\mathbf{u}$$

Thus, $\mathbf{A}\mathbf{u}$ is an eigenvector of $\mathbf{Q}$ with eigenvalue λ.
Similarly, let $\mathbf{v}$ be an eigenvector of $\mathbf{Q}$ with eigenvalue μ. Then:

$$\mathbf{Q}\mathbf{v} = \mu\mathbf{v}$$
$$\therefore \mathbf{A}^T\mathbf{Q}\mathbf{v} = \mathbf{A}^T(\mu\mathbf{v}) = \mu\mathbf{A}^T\mathbf{v}$$
$$\implies \mu\mathbf{A}^T\mathbf{v} = \mathbf{A}^T\mathbf{A}\mathbf{A}^T\mathbf{v} = (\mathbf{A}^T\mathbf{A})\mathbf{A}^T\mathbf{v} = \mathbf{P}\mathbf{A}^T\mathbf{v}$$

Thus, $\mathbf{A}^T\mathbf{v}$ is an eigenvector of $\mathbf{P}$ with eigenvalue μ.
The number of elements in $\mathbf{u}$ and $\mathbf{v}$ are the same as the number of rows in $\mathbf{P}$ and $\mathbf{Q}$, respectively.

(iii)

$$\mathbf{Q}\mathbf{v}_i = \mu_i\mathbf{v}_i$$
$$\implies \mathbf{A}\mathbf{A}^T\mathbf{v}_i = \mathbf{A}(\mathbf{A}^T\mathbf{v}_i) = \mu_i\mathbf{v}_i$$
$$\implies \mathbf{A}\mathbf{u}_i = \frac{\mu_i\mathbf{v}_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$$

Let $\gamma_i = \frac{\mu_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$. Then:

$$\mathbf{A}\mathbf{u}_i = \gamma_i\mathbf{v}_i$$

where, γ_i is non-negative because μ_i is real, non-negative and $\|\mathbf{A}^T\mathbf{v}_i\|$ is non-negative.
Thus, there exists a real, non-negative γ_i such that $\mathbf{A}\mathbf{u}_i = \gamma_i\mathbf{v}_i$.

(iv)

$$\mathbf{A}\mathbf{u}_i = \gamma_i\mathbf{v}_i \quad \forall 1 \leq i \leq m$$
$$\mathbf{A}\mathbf{u}_i = 0 \quad \forall m+1 \leq i \leq n$$
$$\implies \mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Gamma}$$

where, $\mathbf{V}$ is an $n \times n$ orthonormal matrix with columns as $\mathbf{v}_i$, $\mathbf{U}$ is an $m \times m$ matrix with columns as $\mathbf{u}_i$, and $\mathbf{\Gamma}$ is an $m \times n$ matrix with diagonal entries as γ_i and all other entries as 0.

$$\implies \mathbf{A} = \mathbf{U}\mathbf{\Gamma}\mathbf{V}^T$$

where the values in $\mathbf{\Gamma}$ are non-negative as proved above.