Question 3

Anshika Raman
Roll No: 210050014

Kushal Agarwal
Roll No: 210100087

Kavan Vavadiya
Roll No: 210100166

October 23, 2024

(a) Non-zero singular values of A are the square roots of the eigenvalues of AA^T and A^TA.
For a matrix $A \in \mathbb{R}^{m\times n}$, the singular value decomposition (SVD) of A is:

$$A = U\Sigma V^T$$

where $U \in \mathbb{R}^{m\times m}$ is an orthogonal matrix whose columns are the left singular vectors of A, $V \in \mathbb{R}^{n\times n}$ is an orthogonal matrix whose columns are the right singular vectors of A, and $\Sigma \in \mathbb{R}^{m\times n}$ is a diagonal matrix whose diagonal entries $\sigma_1, \sigma_2, \dots, \sigma_{\min(m,n)}$ are the singular values of A.

To show that the non-zero singular values of A are the square roots of the eigenvalues of both A^TA and AA^T:

- Consider the matrix A^TA. We can write:

$$A^TA = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

Since U is orthogonal, $U^TU = I$, so $A^TA = V\Sigma^T\Sigma V^T$. The matrix $\Sigma^T\Sigma$ is diagonal with entries $\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2$. Hence, the eigenvalues of A^TA are the squares of the singular values $\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2$, and the singular values σ_i are the square roots of these eigenvalues.

- Similarly, for the matrix AA^T:

$$AA^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T$$

Since $V^TV = I$, $AA^T = U\Sigma\Sigma^T U^T$. The eigenvalues of AA^T are $\sigma_1^2, \sigma_2^2, \dots, \sigma_{\min(m,n)}^2$, and the singular values are their square roots.

Thus, the non-zero singular values of A are the square roots of the eigenvalues of both A^TA and AA^T.

(b) Squared Frobenius norm of a matrix is equal to the sum of the squares of its singular values.

The Frobenius norm $\|A\|_F$ of a matrix A is defined as:

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$$

Alternatively, the Frobenius norm can be expressed in terms of the singular values of A. Let $\sigma_1, \sigma_2, \dots, \sigma_{\min(m,n)}$ be the singular values of A. Then:

$$\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_{\min(m,n)}^2}$$

To prove this, recall that $A = U\Sigma V^T$. Then the squared Frobenius norm can be written as:

$$\|A\|_F^2 = \sum_{i,j} |a_{ij}|^2 = \text{trace}(A^TA)$$

Using the SVD, we have:

$$A^T A = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T \Sigma V^T$$

Thus,

$$\|A\|_F^2 = \text{trace}(V\Sigma^T \Sigma V^T) = \text{trace}(\Sigma^T \Sigma) = \sum_{i=1}^{\min(m,n)} \sigma_i^2$$

which proves that the squared Frobenius norm is the sum of the squares of the singular values.

(c) Explanation of why $USV^T \neq A$.

The student's approach is almost correct, but there is a key issue that can lead to $USV^T \neq A$. When using the eigenvectors from $A^T A$ and AA^T, the signs of these eigenvectors are not uniquely determined. Eigenvectors can differ by a scalar factor (typically ± 1) and still be valid.

Thus, while the student correctly computes the eigenvectors of $A^T A$ and AA^T, the resulting matrices U and V may not satisfy the SVD condition because the eigenvectors might need to be adjusted in sign to ensure the correct reconstruction of A. To rectify this, the student should ensure that the signs of the eigenvectors in U and V are aligned such that $USV^T = A$.

(d) Properties of matrices $P = A^T A$ and $Q = AA^T$.

(i) Show that $y^T P y \geq 0$ and $z^T Q z \geq 0$.

Let $P = A^T A$ and $Q = AA^T$. Consider:

$$y^T P y = y^T A^T A y = (Ay)^T (Ay) = \|Ay\|^2 \geq 0$$

Similarly,

$$z^T Q z = z^T A A^T z = (A^T z)^T (A^T z) = \|A^T z\|^2 \geq 0$$

Since both expressions are squared norms, they are non-negative. This shows that the eigenvalues of P and Q are non-negative.

(ii) If u is an eigenvector of P with eigenvalue λ, show that Au is an eigenvector of Q with eigenvalue λ.

Let u be an eigenvector of $P = A^T A$ with eigenvalue λ, i.e.,

$$A^T A u = \lambda u$$

Multiplying both sides by A, we get:

$$AA^T (Au) = \lambda (Au)$$

This shows that Au is an eigenvector of $Q = AA^T$ with eigenvalue λ.

Similarly, if v is an eigenvector of Q with eigenvalue μ, we have:

$$AA^T v = \mu v$$

Multiplying both sides by A^T, we get:

$$A^T A (A^T v) = \mu (A^T v)$$

Thus, $A^T v$ is an eigenvector of $P = A^T A$ with eigenvalue μ.

The number of elements in u is n (since u is an eigenvector of $P = A^T A$) and the number of elements in v is m (since v is an eigenvector of $Q = AA^T$).

(iii) If v_i is an eigenvector of Q, define $u_i = \frac{A^T v_i}{\|A^T v_i\|_2}$. Show that $Au_i = \gamma_i v_i$.

By definition,

$$u_i = \frac{A^T v_i}{\|A^T v_i\|_2}$$

Multiplying both sides by A, we get:

$$Au_i = A\left(\frac{A^T v_i}{\|A^T v_i\|_2}\right) = \frac{AA^T v_i}{\|A^T v_i\|_2}$$

Since v_i is an eigenvector of $Q = AA^T$ with eigenvalue μ_i, we have $AA^T v_i = \mu_i v_i$. Therefore:

$$Au_i = \frac{\mu_i v_i}{\|A^T v_i\|_2}$$

Define $\gamma_i = \frac{\mu_i}{\|A^T v_i\|_2}$. Hence, $Au_i = \gamma_i v_i$, as required.

(iv) Orthonormality of eigenvectors and SVD.

From the previous HW, we know that the eigenvectors of symmetric matrices $P = A^T A$ and $Q = AA^T$ are orthonormal. That is, for distinct eigenvalues, $u_i^T u_j = 0$ and $v_i^T v_j = 0$ for $i \neq j$.

Now, define $U = [u_1, u_2, ..., u_m]$ and $V = [v_1, v_2, ..., v_n]$. Then , let Γ be a diagonal matrix with entries $\gamma_1, \gamma_2, \ldots, \gamma_m$, where each γ_i is defined as in part (iii), the expression for $U\Gamma V^T$ is:

$$U\Gamma V^T = \sum_{i=1}^{m} \gamma_i u_i v_i^T$$

where:

- $U = [u_1 \mid u_2 \mid \cdots \mid u_m]$
- $V = [v_1 \mid v_2 \mid \cdots \mid v_n]$
- $\gamma_i = \frac{\mu_i}{\|A^T v_i\|_2}$
- u_i and v_i are the left and right singular vectors of A
- v_i^T is the transpose of the right singular vector.

Thus, the matrix A is expressed as the sum of the outer products of the singular vectors, each weighted by the corresponding singular value γ_i.