7 Singular Value Decomposition, Trace and Eigenvalues

7.1 Singular Value Decomposition

From last time, we had the following:

Theorem 7.1
Let A be $m \times n$. Then we can write

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

Here, U is an orthogonal $m \times m$ matrix (rows and columns are unit vectors, and all rows and columns are orthogonal, note that this matrix is not unique, we used gram-schmidt to find orthogonal matrices).
V is an orthogonal $n \times n$ matrix (not unique)
Σ is an $m \times n$ matrix whose upper left submatrix has <u>positive</u> entries that are non-increasing. For $s_1 \geq s_2 \geq \cdots \geq s_k \geq 0$, we have

$$\Sigma = \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & 0 \\ & & s_k & \\ \hline & 0 & & 0 \end{array}\right]$$

Note 7.2

1. The columns of V are the normalized eigenvectors of A^TA
2. The **singular values** $s_1, \cdots, s_k$ are the <u>square roots</u> of the <u>shared</u> eigenvalues of A^TA and AA^T
3. $V^T = V^{-1}$, $U^T = U^{-1}$ (orthogonal)
4. AA^T is symmetric $\implies$ spectral theorem holds $\implies$ all eigenvalues are real and form an orthonormal basis of eigenvectors (very important for this class)
5. All the eigenvalues of A^TA are nonnegative.
6. $\{A\vec{v}_1, \cdots, A\vec{v}_n\}$ forms an orthogonal set, where $\vec{v}_1, \cdots, \vec{v}_n$ is an orthonormal basis of eigenvectors.
7. AA^T and A^TA share the same non-zero eigenvalues.

Proof. of 5. Let $\vec{v}$ be an eigenvector with eigenvalue $\lambda \neq 0$ for A^TA.

Remember that $||\vec{v}||^2 = \vec{v} \cdot \vec{v} = \vec{v}^T\vec{v}$, and because $A^TA\vec{v} = \lambda\vec{v}$:

$$\begin{aligned} ||A\vec{v}||^2 &= (A\vec{v})^T(A\vec{v}) \\ &= \vec{v}^TA^TA\vec{v} \\ &= \vec{v}^T(\lambda\vec{v}) \\ &= \lambda(\vec{v}^T\vec{v}) \\ &= \lambda||\vec{v}||^2 \geq 0 \end{aligned}$$

Which implies that $\lambda \geq 0$. □

Example 7.3

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

We find

$$V = \begin{bmatrix} 1/\sqrt{5} & 0 & -\sqrt{2}/5 \\ 0 & 1 & 0 \\ 2/\sqrt{5} & 0 & 1\sqrt{5} \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Remember that we must take the transpose of V.

What are the non-zero eigenvalues of $A^T A$? By Σ, we get that the squares of $\sqrt{5}$ and 1, which are 5 and 1, are the non-zero eigenvalues. The other eigenvalue of $A^T A$ should be 0.

Moreover, one can compute

$$A^T A = \begin{bmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{bmatrix} = 5 \begin{bmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{bmatrix}$$

And we find that this is a column vector for V.

I believe that the columns of V are eigenvectors, which correspond to the eigenvalues found in Σ. The first column is an eigenvector with eigenvalue 5, second column is an eigenvector with eigenvalue 1, and third column is an eigenvector with eigenvalue 0.

What is matrix U?
Since $V^T = V^{-1} \implies A = U\Sigma V^T \implies AV = U\Sigma$, we have that

$$A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{v}_m \\ \downarrow & & \downarrow \end{bmatrix} \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & 0 \\ & & s_k & \\ \hline & 0 & & 0 \end{array}\right]$$

And thus

$$\begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} s_1\vec{u}_1 & s_2\vec{u}_2 & \cdots & s_k\vec{v}_m & \cdots & ??? \\ \downarrow & \downarrow & & \downarrow & & \end{bmatrix}$$

Thus, we solve $A\vec{v}_i = s_i\vec{u}_i$, $1 \le i \le k$ to find the vectors $\vec{u}_i$ (may need Gram-Schmidt to get this).

Theorem 7.4
The trace of A, denoted $\text{tr}(A)$, is the sum of the diagonal entries of A. We have the following:

1. $\text{tr}(A) =$ sum of the eigenvalues of A (including multiplicity)
2. $\det(A) =$ product of the eigenvalues (including multiplicity)

Example 7.5

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

A has eigenvalues 1, 1, and 10.
$\text{tr}(A) = 12 =$ sum of eigenvalues.
$\det(A) = 10 =$ product of eigenvalues.

Proof. Let $P_A(x) = \det(A - xI)$ be the characteristic polynomial of A. Let $\lambda_1, \cdots, \lambda_n$ be its eigenvalues.

Remember that we find eigenvalues by setting the characteristic polynomial equal to 0, then solving for x.

So, the eigenvalues are the roots of $P_A(x)$.

We know that $P_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x) = (-x)^n + (\lambda_1 + \lambda_2 + \cdots + \lambda_n)(-x)^{n-1} + \cdots$

Alternatively, we can use the cofactor expansion.

$$\det(A - xI) = \det\left(\begin{bmatrix} a_{11} - x & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & a_{23} & \cdots & \\ a_{31} & a_{32} & a_{33} - x & \cdots & \\ \vdots & & & & \\ & & \cdots & & a_{nn} - x \end{bmatrix}\right)$$

By the expansion, what is the coefficient of $(-x)^{n-1}$?
If we decide to start our cofactor expansion at a_{13}, we get rid of row 1 and column 3, our cofactor expansion looks something like $a_{13}(\cdots)$, which means that we lose two x terms, meaning there is no way a_{13} could contribute to the $(-x)^{n-1}$ term.

So, the only way we can get the coefficient of $(-x)^{n-1}$ is by starting at an element like $a_{11} - x$, then going to $a_{22} - x$, etc.

So, it is only ever the diagonal entries that will ever contribute to the $(-x)^{n-1}$ term. So, this implies that

$$(a_{11} - x)(a_{22} - x)\cdots(a_{nn} - x) = (-x)^n + (a_{11} + a_{22} + \cdots)(-x)^{n-1}$$

Where $(a_{11} + a_{22} + \cdots)$ is the trace. $\square$