

Discussion 12B

1. Computing the SVD: A “Tall” Matrix Example

Define the matrix

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}. \quad (1)$$

Here, we expect $U \in \mathbb{R}^{3 \times 3}$, $\Sigma \in \mathbb{R}^{3 \times 2}$, and $V \in \mathbb{R}^{2 \times 2}$ (recall that U and V must be square since they are orthonormal matrices).

In this problem, we will walk through the SVD algorithm, prove some important theorems about the SVD matrices and column/null spaces, and consider an alternate way to approach the SVD.

- (a) Let’s start by trying to write A as an outer product in the form of $\sigma \vec{u} \vec{v}^\top$ where both \vec{u} and \vec{v}^\top have unit norm. (HINT: Are the columns of A linearly independent or dependent? What does that tell us about how we can represent them?)

Solution: Notice that the columns of A are linearly dependent meaning that they can be written as some scalar multiple of the same vector. If we were to say that

$$A = [\vec{a}_1 \quad \vec{a}_2] \quad (2)$$

then we can say that $\vec{a}_1 = \alpha_1 \vec{x}$ and $\vec{a}_2 = \alpha_2 \vec{x}$. For our matrix A , we notice that both our columns lie in the span of the vector:

$$\vec{x} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \quad (3)$$

which means we can represent $\vec{a}_1 = 1\vec{x}$ and $\vec{a}_2 = -1\vec{x}$. Stacking these together we get that A can be represented as:

$$A = \vec{x} \begin{bmatrix} 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \quad (4)$$

Now that we have A as an outer product we need to normalize our vectors which will get us:

$$A = 3 \begin{bmatrix} \frac{1}{3} \\ \frac{-2}{3} \\ \frac{2}{3} \end{bmatrix} \sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = 3\sqrt{2} \begin{bmatrix} \frac{1}{3} \\ \frac{-2}{3} \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad (5)$$

where $\sigma = 3\sqrt{2}$, $\vec{u} = \begin{bmatrix} \frac{1}{3} \\ \frac{-2}{3} \\ \frac{2}{3} \end{bmatrix}$, and $\vec{v}^\top = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$. This is the outer product form of our SVD!

We will now use the general algorithm to solve for the SVD of the same matrix.

- (b) In this part, we will walk through Algorithm 7 in [Note 16](#). This algorithm applies for a general matrix $A \in \mathbb{R}^{m \times n}$.

- i. Find $r := \text{rank}(A)$. Compute $A^\top A$ and diagonalize it using the spectral theorem (i.e. find V and Λ).
- ii. Unpack $V := [V_r \quad V_{n-r}]$ and unpack $\Lambda := \begin{bmatrix} \Lambda_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix}$.
- iii. Find $\Sigma_r := \Lambda_r^{1/2}$ and then find $\Sigma := \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$.
- iv. Find $U_r := AV_r \Sigma_r^{-1}$, where $U_r \in \mathbb{R}^{3 \times 1}$ and then extend the basis defined by columns of U_r to find $U \in \mathbb{R}^{3 \times 3}$.
(HINT: How can we extend a basis, and why is that needed here?)
- v. Use the previous parts to write the full SVD of A .
- vi. If we were to calculate the SVD of our matrix using a calculator, are we *guaranteed* to always get the same SVD? Why or why not?

Solution: Part 1.(b)i

We can immediately see that the first column is a multiple of the second, so $r = \text{rank}(A) = 1$. Next, we compute

$$A^\top A = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \quad (6)$$

$$= \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix} \quad (7)$$

The eigenvalues of $A^\top A$ are the roots of $(\lambda - 9)^2 - 81 = 0$, and therefore, $\lambda_1 = 18$ and $\lambda_2 = 0$. The corresponding orthonormal eigenvectors are

$$\vec{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad (8)$$

Hence,

$$V = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (9)$$

$$\Lambda = \begin{bmatrix} 18 & 0 \\ 0 & 0 \end{bmatrix} \quad (10)$$

Note that the diagonal entries of Λ are sorted in decreasing order (i.e. $\Lambda_{11} \geq \Lambda_{22}$)

Part 1.(b)ii

Since $r = 1$, we know that $\Lambda_r \in \mathbb{R}^{1 \times 1}$. To unpack it from our computation of Λ , we can pattern match to see that

$$\Lambda_r = [18] \quad (11)$$

Next, we must find V_r and V_{n-r} . Here, V_r will contain the eigenvectors that correspond to the nonzero eigenvalues, and V_{n-r} will contain all other eigenvectors. Thus,

$$V_r = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad (12)$$

$$V_{n-r} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad (13)$$

Part 1.(b)iii

When raising a diagonal matrix to the power $\frac{1}{2}$, we can take the square root of all the diagonal elements. Hence,

$$\Sigma_r = \Lambda_r^{1/2} = \begin{bmatrix} 3\sqrt{2} \end{bmatrix} \quad (14)$$

Now, recall that $\Sigma \in \mathbb{R}^{3 \times 2}$. Hence, we can pattern match to the form given in part 1.(b)iii to obtain

$$\Sigma = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (15)$$

Part 1.(b)iv

Plugging in the given formula to find U_r , we obtain:

$$U_r = AV_r \Sigma_r^{-1} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix} \quad (16)$$

To find U , we must extend the basis formed by the column space of U_r using Gram-Schmidt. This allows us to find two more orthonormal vectors that we can stack in the columns of U to give us an orthonormal matrix. We can do Gram-Schmidt using Python to get

$$\vec{u}_2 = \begin{bmatrix} \frac{\sqrt{8}}{3} \\ \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \end{bmatrix} \quad \vec{u}_3 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad (17)$$

Hence,

$$U = [U_r \quad \vec{u}_2 \quad \vec{u}_3] = \begin{bmatrix} -\frac{1}{3} & \frac{\sqrt{8}}{3} & 0 \\ \frac{2}{3} & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{2}{3} & -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (18)$$

Part 1.(b)v

Finally, we compose this information, and write that A can be decomposed as:

$$A = 3\sqrt{2} \underbrace{\begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{compact SVD}} = \underbrace{\begin{bmatrix} -\frac{1}{3} & \frac{\sqrt{8}}{3} & 0 \\ \frac{2}{3} & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{2}{3} & -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{full SVD}}. \quad (19)$$

The full SVD representation of A is given below. Note that the full SVD and compact SVD represent the same matrix; the compact form merely omits the columns/rows of U or V which will hit the zero entries of Σ .

Part 1.(b)vi

The SVD that a calculator outputs might be different because of the non-uniqueness of Gram-Schmidt. We can extend a given set of vectors to an orthonormal basis in an infinite number of ways, so the SVD is not unique. Furthermore, it is important to note that the extended columns of U only ever multiply with the zero-entries of Σ . So, they cannot impact the final result of A . However, it is still critical that all the columns of U are in fact mutually orthogonal and normalized.

- (c) We now want to create the SVD of A^\top . **What are the relationships between the matrices composing A and the matrices composing A^\top ?**

Solution: We know that A has an SVD representation of $U\Sigma V^\top$ as we solved for above. One natural approach to solving for the SVD of A^\top is to take the transpose of the SVD terms, and “reassign variables”. That is, we can say that A^\top has SVD $\tilde{U}\tilde{\Sigma}\tilde{V}^\top$ and pattern match to U, Σ, V from before:

$$A^\top = (U\Sigma V^\top)^\top = V\Sigma^\top U^\top \quad (20)$$

Now, pattern-matching, we can say that $\tilde{U} = V, \tilde{\Sigma} = \Sigma^\top, \tilde{V}^\top = U^\top \implies \tilde{V} = U$. Note how the roles have exchanged, and Σ is transposed.

We can now write the full SVD of A^\top (feel free to confirm that the multiplication yields the right result):

$$A^\top = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{\sqrt{8}}{3} & \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (21)$$

Contributors:

- Anish Muthali.
- Neelesh Ramachandran.
- Druv Pai.
- John Maidens.
- Nikhil Jain.
- Chancharik Mitra.