Discussion 10B

The following note is useful for this discussion: Note 16

1. Computing the SVD: A "Tall" Matrix Example

Define the matrix

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}. \tag{1}$$

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) 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 $V := \begin{bmatrix} V_r & V_{n-r} \end{bmatrix}$ 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. **Use the Jupyter notebook to run the code cell that calls `numpy.linalg.svd` on A. What is the result? Does it match our result above?**

(b) We now want to create the SVD of $A^\top$. Rather than repeating all of the steps in the algorithm, feel free to use the Jupyter notebook for this subpart (which defines a `numpy.linalg.svd` command). **What are the relationships between the matrices composing A and the matrices composing $A^\top$?**

(c) **Show, for a general matrix $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = r$ and $A = U\Sigma V^\top$, that $\operatorname{Null}(A) = \operatorname{Col}(V_{n-r})$. Then, find a basis for the null space of A in eq. (1).** *(HINT: How do we show two sets are equal? Try and use that approach here. Consider the outer product summation form for the SVD. Also, consider using the rank-nullity theorem that $\dim \operatorname{Col}(A) + \dim \operatorname{Null}(A) = n$.)*

(d) **Show, for a general matrix** $A \in \mathbb{R}^{m \times n}$ **with** $\text{rank}(A) = r$ **and** $A = U\Sigma V^\top$**, that** $\text{Col}(A) = \text{Col}(U_r)$**. Then, find a basis for the range (or column space) of** A**.**

(e) **(PRACTICE) Show, for a general matrix** $A \in \mathbb{R}^{m \times n}$ **with** $\text{rank}(A) = r$ **and** $A = U\Sigma V^\top$**, that** $\text{Null}(A^\top) = \text{Col}(U_{m-r})$ **and** $\text{Col}(A^\top) = \text{Col}(V_r)$**. Then show:**

i. $\dim \text{Col}(A) + \dim \text{Null}(A^\top) = n$**,**

ii. and $\text{Col}(A)$ **and** $\text{Null}(A^\top)$ **are orthogonal.**

(f) Suppose A was a wide matrix. Instead of finding $A^\top A$, we may want to find the SVD by computing $AA^\top$. The original Algorithm 7 from Note 16, in its entirety, is shown below:

Algorithm 1 Constructing the SVD

function FULLSVD($A \in \mathbb{R}^{m \times n}$)
    $r := \textsc{Rank}(A)$
    $(V, \Lambda) := \textsc{Diagonalize}(A^\top A)$ ▷ Sorted so that $\Lambda_{11} \geq \cdots \geq \Lambda_{nn}$
    Unpack $V := \begin{bmatrix} V_r & V_{n-r} \end{bmatrix}$
    Unpack $\Lambda := \begin{bmatrix} \Lambda_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix}$
    $\Sigma_r := \Lambda_r^{1/2}$
    Pack $\Sigma := \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$
    $U_r := A V_r \Sigma_r^{-1}$
    $U := \textsc{ExtendBasis}(U_r, \mathbb{R}^m)$
    return (U, Σ, V)
end function

Write a modified version of Algorithm 7 where you solve for the SVD of A using $AA^\top$ instead of $A^\top A$. *(HINT: Consider replacing every instance of "A" in $A^\top A$ with "$A^\top$". What happens? How can we use the result from the 1.b part?)*

Contributors:

- Anish Muthali.
- Neelesh Ramachandran.
- Druv Pai.
- John Maidens.

