EECS 16B Designing Information Devices and Systems II
Fall 2021 Discussion Worksheet **Discussion 10B**

The following notes are useful for this discussion: Note 15, Note 16

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

Define the matrix

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

(a) In this part, we will find the full SVD of A in steps.

Solution: In this subpart to calculate the full SVD, we will follow the algorithm of the SVD Note. We select Method 1 (computing using $A^\top A$) since A is "tall", and $A^\top A$ is smaller than $AA^\top$.

(i) **Compute $A^\top A$ and find its eigenvalues.**

Solution: First, we compute

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

$$= \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}. \tag{3}$$

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

(ii) **Find orthonormal eigenvectors $\vec{v}_i$ of $A^\top A$ (right singular vectors, columns of V).**

Solution: We can find the corresponding (unit) eigenvectors for the above eigenvalues in the usual way, by computing $\text{Null}\left(A^\top A - \lambda_1 I\right)$ and $\text{Null}\left(A^\top A - \lambda_2 I\right)$. This yields that:

$$\vec{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}. \tag{4}$$

(iii) **Find singular values, $\sigma_i = \sqrt{\lambda_i}$.**

Solution: A has one nonzero singular value $\sqrt{18} = 3\sqrt{2}$, and the other singular value is zero.

(iv) **Find the orthonormal vectors $\vec{u}_i$ (and for nonzero σ, you can use $\vec{v}_i$).**

Hint: given $\vec{v}_k$ corresponding to nonzero σ, we can compute $\vec{u}_k = \frac{1}{\sigma_k} A\vec{v}_k$.

Another hint: How can we extend a basis, and why is that needed here? Note what the Jupyter notebook contains.

Solution: We obtain:

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix} \tag{5}$$

To complete a basis of $\mathbb{R}^3$ as required for the full SVD, we can do Gram-Schmidt using the Jupyter notebook to get

$$\vec{u}_2 = \begin{bmatrix} \frac{\sqrt{8}}{3} \\ \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{6}$$

(v) **Use the previous parts to write the full SVD of A.**

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

$$\underbrace{A = 3\sqrt{2}\begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{compact SVD}} = \underbrace{\begin{bmatrix} -\frac{1}{3} & \frac{\sqrt{8}}{3} & 0 \\ \frac{2}{3} & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{2}{3} & -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{full SVD}}. \tag{7}$$

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.

(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?**

Solution: The SVD that Jupyter notebook gives is 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.

(b) **Find the rank of A.**

Solution: A has 1 nonzero singular value. So A has rank 1.

(c) **Find a basis for the range (or column space) of A.**

Solution: We know if $A = U\Sigma V^\top$ is an SVD, then the columns of U with nonzero corresponding singular values are a basis for the column space of A. Any columns corresponding to $\sigma = 0$ cannot add to the span. Therefore, matching terms with the SVD of A,

$$\text{range}(A) = \text{span}\left(\left\{\begin{bmatrix} -1/3 \\ 2/3 \\ -2/3 \end{bmatrix}\right\}\right) \tag{8}$$

(d) **Find a basis for the null space of A.**

Solution: We know if $A = U\Sigma V^\top$ is a full SVD, then the columns of V with corresponding singular values equal to 0 are a basis for the null space of A. Thus

$$\text{Null}(A) = \text{span}\left(\left\{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\right\}\right) \tag{9}$$

(e) 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 $A^\top$?**

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 $U_\star \Sigma_\star V_\star^\top$, and to find how these new $\star$ variables relate to the originals, we write:

$$A^\top = \left(U\Sigma V^\top\right)^\top = V\Sigma^\top U^\top \tag{10}$$

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

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

$$A^\top = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{\sqrt{8}}{3} & \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{11}$$

2. Understanding the SVD

We can compute the SVD for a wide matrix A with dimension $m \times n$ where $n > m$ using $A^\top A$ with the method covered in lecture. However, when doing so, you may realize that $A^\top A$ is much larger than $AA^\top$ for such wide matrices. This makes it more efficient to find the eigenvalues for $AA^\top$. In this question, we will explore how to compute the SVD using $AA^\top$ instead of $A^\top A$.

(a) **What are the dimensions of $AA^\top$ and $A^\top A$?**

Solution: Since A is $m \times n$, $AA^\top$ is $(m \times n) \times (n \times m)$, which is $m \times m$. Similarly $A^\top A$ is $(n \times m) \times (m \times n)$ which is $n \times n$.

(b) **Given that the SVD of A is $A = U\Sigma V^\top$, find a symbolic expression for $AA^\top$ in terms of U, Σ, $V^\top$.** Simplify where possible!

Solution:

$$AA^\top = U\Sigma \underbrace{V^\top V}_{I} \Sigma^\top U^\top \tag{12}$$

$$= U\Sigma\Sigma^\top U^\top \tag{13}$$

(c) Using the solution to the previous part, **how can we find a U and Σ from $AA^\top$?** *Hint: first, think about matrix dimensions. Next, consider the properties of the SVD, and what each matrix signifies. Another Hint: you may want to compute for yourself, based on the structure of Σ, what $\Sigma^\top\Sigma$ and $\Sigma\Sigma^\top$ are.*

Solution: Knowing that $AA^\top$ is a symmetric matrix, we know that its normalized eigenvectors will be orthonormal.

From the properties of the SVD, we know that U is an orthonormal matrix of dimension $m \times m$ and $\Sigma\Sigma^\top$ is an $m \times m$ diagonal matrix, with the entries on the diagonal being ${\sigma_i}^2$. Each σ_i is a *singular value* of A.

We can calculate U by diagonalizing the symmetric matrix $AA^\top$. By the spectral theorem for real symmetric matrices, we will get an orthonormal basis of eigenvectors. The square root of the corresponding eigenvalues of $AA^\top$ will give us the singular values σ_i.

We can then construct Σ by putting these on the diagonal of an otherwise zero matrix with the same dimensions as A, and the corresponding eigenvectors will form the U matrix.

(d) Now that we have found the singular values σ_i and the corresponding vectors $\vec{u}_i$ in the matrix U, **can you find the corresponding vectors $\vec{v}_i$ in V?** *Hint: Apply the definition of an eigenvector. What do the $\vec{v}_i$ vectors signify with regards to $A^\top A$?*

Solution: We know everything except for V. In particular, we know $\vec{u}_i$ is an eigenvector of $AA^\top$ with eigenvalue σ_i^2. Then

$$\begin{aligned} AA^\top \vec{u}_i &= \sigma_i^2 \vec{u}_i && (14)\\ A^\top AA^\top \vec{u}_i &= A^\top(\sigma_i^2 \vec{u}_i) && (15)\\ A^\top A(A^\top \vec{u}_i) &= \sigma_i^2(A^\top \vec{u}_i). && (16) \end{aligned}$$

So we see that $A^\top \vec{u}_i$ is an eigenvector of $A^\top A$ with eigenvalue σ_i^2. Define $\vec{v}_i = \frac{A^\top \vec{u}_i}{\|A^\top \vec{u}_i\|}$. Then

$$\begin{aligned} \vec{v}_i &= \frac{A^\top \vec{u}_i}{\|A^\top \vec{u}_i\|} && (17)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{\|A^\top \vec{u}_i\|^2}} && (18)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{(A^\top \vec{u}_i)^\top (A^\top \vec{u}_i)}} && (19)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{\vec{u}_i^\top AA^\top \vec{u}_i}} && (20)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{\vec{u}_i^\top \sigma_i^2 \vec{u}_i}} && (21)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{\sigma_i^2 \|\vec{u}_i\|^2}} && (22)\\ &= \frac{A^\top \vec{u}_i}{\sqrt{\sigma_i^2}} && (23)\\ &= \frac{A^\top \vec{u}_i}{\sigma_i}. && (24) \end{aligned}$$

(e) Now we have a way to find the vectors $\vec{v}_i$ in matrix V! **Use the fact that the vectors $\vec{u}_i, \vec{u}_j$ are orthonormal to show that $\vec{v}_i, \vec{j}$ in V (corresponding to nonzero σ_i, σ_j and $i, j \leq n$) are orthonormal by direct computation.**

Solution: To verify that $\vec{v}_i$ in V are orthonormal, we must show that:

i. $\vec{v}_i$ are mutually orthogonal

ii. each $\vec{v}_i$ has norm 1.

Orthogonality:

To show orthogonality, we must show that any two vectors $\vec{v}_i = \frac{A^\top \vec{u}_i}{\sigma_i}$ and $\vec{v}_j = \frac{A^\top \vec{u}_j}{\sigma_j}$, with $i \neq j$, have an inner product of zero. Writing the inner product out:

$$\vec{v}_i^\top \vec{v}_j = \frac{\vec{u}_i^\top A}{\sigma_i} \frac{A^\top \vec{u}_j}{\sigma_j} \tag{25}$$

$$= \frac{\vec{u}_i^\top A A^\top \vec{u}_j}{\sigma_i \sigma_j} \tag{26}$$

$$= \frac{(\sigma_j)^2 \vec{u}_i^\top \vec{u}_j}{\sigma_i \sigma_j} \tag{27}$$

$$= 0 \tag{28}$$

In going from eq. (26) to eq. (27), we could have substituted the matrix product $AA^\top$ with the answer of part c) and simplified. Here, we recognize that the inner matrix $\Sigma\Sigma^\top$ is diagonal with σ_i on the diagonals. This is because we know that $\vec{u_i}$ and $\vec{u_j}$ are orthonormal as they are eigenvectors of a symmetric matrix $AA^\top$.

Thus for all $i \neq j$,

$$\vec{v}_i^\top \vec{v}_j = 0 \tag{29}$$

Norm of 1: If we follow the steps above with $i = j$, then we see that:

$$\vec{v}_i^\top \vec{v}_j = \vec{v}_i^\top \vec{v}_i \tag{30}$$

$$= \frac{(\sigma_i)^2 \vec{u}_i^\top \vec{u}_i}{\sigma_i \sigma_i} \tag{31}$$

$$= \frac{(\sigma_i)^2}{(\sigma_i)^2} \vec{u}_i^\top \vec{u}_i \tag{32}$$

$$= 1 \tag{33}$$

(f) **[Practice]** Given that $A = U\Sigma V^\top$, **verify that the vectors after the first n vectors in V are in the nullspace of A.**

Solution: First, we have to consider the case where the $\vec{v}_i$ vectors from the previous part do not span the entire space; here, we would use Gram-Schmidt to extend.

If we append the standard basis for n-dimensional space, and orthonormalize, this will give us the desired result. The augmented collection of $n + m$ vectors certainly spans the whole space, and so after orthonormalization, we will have a collection of orthonormal vectors that spans the whole space. Along the way, some vectors will be found to be linearly dependent on those that came before — this is fine, we'll discard these. At the end, we will have n orthonormal vectors, the first set of which are the original $\vec{v}_i$.

Now, let $V = \begin{bmatrix} V_s & R \end{bmatrix}$ where V_s are the $\{\vec{v}_i\}$ we found using the $\{\vec{u}_i\}$ and R is composed of the remaining vectors found using Gram-Schmidt. Let S be an $m \times m$ diagonal square matrix with σ_i on the diagonal (σ_i is allowed to be zero) such that $\Sigma = \begin{bmatrix} S & 0 \end{bmatrix}$ where 0 denotes filling in the remaining matrix dimensions with zeros.

$$A = U\Sigma V^\top = U \begin{bmatrix} S & 0 \end{bmatrix} \begin{bmatrix} V_s^\top \\ R^\top \end{bmatrix} \tag{34}$$

And so:

$$\begin{aligned} AR &= U \begin{bmatrix} S & 0 \end{bmatrix} \begin{bmatrix} V_s^\top \\ R^\top \end{bmatrix} R & (35) \\ &= U \begin{bmatrix} S & 0 \end{bmatrix} \begin{bmatrix} 0 \\ R^\top R \end{bmatrix} & (36) \\ &= U \begin{bmatrix} 0 \end{bmatrix} & (37) \\ &= 0 & (38) \end{aligned}$$

Thus, everything in the subspace spanned by R maps to $\vec{0}$, and this shows that the subspace is in $\text{Null}(A)$.

(g) **[Practice]** Using the previous parts of this question and what you learned from lecture, **write out a procedure on how to find the SVD for *any* matrix.**

Solution: We calculate the SVD of matrix A as follows.

i. Pick $A^\top A$ or $AA^\top$ — whichever one is smaller.

ii. i. If using $A^\top A$, find the eigenvalues λ_i of $A^\top A$ and order them, so that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$.

   If using $AA^\top$, find its eigenvalues $\lambda_1, \ldots, \lambda_m$ and order them the same way.

   ii. If using $A^\top A$, find orthonormal eigenvectors $\vec{v}_i$ such that

$$A^\top A\vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, \ldots, r \tag{39}$$

   If using $AA^\top$, find orthonormal eigenvectors $\vec{u}_i$ such that

$$AA^\top \vec{u}_i = \lambda_i \vec{u}_i, \quad i = 1, \ldots, r \tag{40}$$

   iii. Set $\sigma_i = \sqrt{\lambda_i}$.

   If using $A^\top A$, obtain $\vec{u}_i$ from $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i, \quad i = 1, \ldots, r$.

   If using $AA^\top$, obtain $\vec{v}_i$ from $\vec{v}_i = \frac{1}{\sigma_i} A^\top \vec{u}_i, \quad i = 1, \ldots, r$.

iii. If you want to completely construct the U or V matrix, complete the basis (or columns of the appropriate matrix) using Gram-Schmidt to get a full orthonormal matrix.

The full matrix form of SVD is taken to better understand the matrix A in terms of the 3 nice matrices U, Σ, V. Often in practice, we do not completely construct the U and V matrices. After all, in many applications, we don't need all the vectors.

Contributors:

- Neelesh Ramachandran.
- Druv Pai.
- John Maidens.
- Nikhil Shinde.
- Siddharth Iyer.
- Jane Liang.

