Wenzhi Tseng

Prerequisite information:

First, the following definition:

An $n \times n$ matrix Q is orthogonal if any of the following are true (the following conditions are all equivalent, so if you satisfy one, you satisfy all the rest):

a. The columns of Q form an orthonormal basis of $\mathbb{R}^n$. That is, the columns of Q are all orthogonal to each other and are all of length one.
   That is, $\langle Q\boldsymbol{e_i}, Q\boldsymbol{e_j}\rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$.
b. The rows of Q also form an orthonormal basis of $\mathbb{R}^n$. That is, the rows of Q are all orthogonal to each other and are all of length one.
c. $Q^{-1} = Q^T$. That is, the inverse of Q is just the transpose of Q. This is important because computing the inverse of a matrix is very tedious, but taking the transpose is a piece of cake. Especially for a computer.

The spectral theorem you learned says every symmetric matrix A can be eigendecomposed as $A = Q^T DQ$ where D is a diagonal matrix and Q is an orthogonal matrix (i.e. $Q^{-1} = Q^T$). Furthermore, the columns of Q are the eigenvectors of A, normalized to have length one, so Q is a "change of basis" matrix where the basis vectors are the eigenvectors of A. The eigenvalues of A are precisely the diagonal entries of D, so we can think of D as a "stretch" matrix where the i^{th} column of Q (the i^{th} eigenvector) is being stretched by the i^{th} diagonal entry of D (the i^{th} eigenvalue). However, since the columns of Q are length one, Q tends to be a rotation or a reflection matrix of sorts (Q does not stretch vectors in any direction, only D is doing that).

HOWEVER: you may see some texts about spectral theorem say that symmetric matrix A is decomposed into $A = U^* DU$ where D is a diagonal matrix and U is a unitary matrix. What is a unitary matrix? It is just an orthogonal matrix generalized to the complex numbers. The thing is, the *true* spectral theorem concerns itself with matrices with *complex* entries, and when this happens, instead of $U^{-1} = U^T$, the theorem says $U^{-1} = U^*$ where U^* is the *conjugate transpose* of U. However, all you have to know is that if all entries of A are real numbers, then U^* is the same as U^T. So, when you read about spectral theorem and about SVD, if you see unitary matrices, they are just orthogonal matrices but for complex numbers. Unitary matrices have the same properties as orthogonal matrices (in fact, orthogonal matrices are a subset of the unitary matrices) but for completion's sake, I will provide the definition of a unitary matrix below:

An $n \times n$ matrix U is unitary if any of the following are true (the following conditions are all equivalent, so if you satisfy one, you satisfy all the rest):

a. The columns of U form an orthonormal basis of $\mathbb{C}^n$. That is, the columns of U are all orthogonal to each other and are all of length one.
   That is, $\langle U\boldsymbol{e_i}, U\boldsymbol{e_j}\rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$.
b. The rows of U also form an orthonormal basis of $\mathbb{C}^n$. That is, the rows of U are all orthogonal to each other and are all of length one.
c. $U^{-1} = U^*$. That is, the inverse of U is just the conjugate transpose of U. If all entries of U are real numbers, then $U^* = U^T$.

For this document, I will assume all entries of A are real, and instead of U^*, I will write U^T. However, you should be aware that it is possible for a matrix A to have only real entries but for its eigenvalues to be complex numbers. The matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has only real entries but you can verify that the its eigenvalues are i and $-i$.

The takeaway is, every symmetric matrix can be written as a rotation/reflection (Q) followed by a stretch (D) followed by undoing that rotation/reflection (Q^T). However, this spectral decomposition (also called eigendecomposition) of a symmetric matrix doesn't exist for other matrices in general. For example, what if A is not symmetric? Even worse, what if A is not even a square matrix? Can we still write these matrices as a rotation followed by stretch followed by another rotation? The answer is yes, and the decomposition is called the singular value decomposition (SVD). The spectral decomposition is in fact just a special case of the SVD for symmetric matrices.

Singular Value Decomposition (SVD)

Every $m \times n$ matrix A can be written in the form $U\Sigma V^T$ where U is an $m \times m$ unitary matrix, V is an $n \times n$ unitary matrix, and Σ is an $m \times n$ "rectangular diagonal" matrix. A rectangular diagonal matrix is just like a diagonal matrix but it may or may not be square. We still define the diagonal entries of a rectangular diagonal matrix as the entries a_{ii}. To illustrate this, I have highlighted the diagonal entries in each of the rectangular diagonal matrices below:

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

So rectangular diagonal matrices are still stretch matrices, and they may or may not have zero row(s) or zero column(s) depending on the dimensions of the matrix. Remember that U, V are rotations/reflections of some sort since they are unitary. There is a theorem saying that V being unitary implies V^T is also unitary, so V^T is also a rotation/reflection of sorts. So the SVD shows that any matrix A can be decomposed into a rotation (V^T), followed by a stretch (Σ), followed by another rotation (U).

How do we know the SVD *always* exists for any matrix A? The proof of existence is a little involved and not that important, Gregory Gunderson[1] has all the proofs of the theorems we are about to take for granted here (if you are curious).

So, given that the SVD $A = U\Sigma V^T$ exists, how do we find U, Σ, and V? We start by looking at $A^T A$, you'll see why.

On one hand, $A^T A$ is always symmetric (you can prove this), which implies $A^T A$ has an eigendecomposition by the spectral theorem: to be precise, there exists unitary P and diagonal D such that:

$$A^T A = PDP^T \quad - \text{Eq. (1)}$$

[1] https://gregorygundersen.com/blog/2018/12/20/svd-proof/#4-textbfu_i-is-a-unit-eigenvector-of-aatop

On the other hand, we know the SVD $A = U\Sigma V^T$ exists. Then:

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

We simplify $(U\Sigma V^T)^T$ by using the theorem $(AB)^T = B^T A^T$ twice,

$$\Rightarrow A^T A = (V\Sigma^T U^T)(U\Sigma V^T)$$

$$\Rightarrow A^T A = V\Sigma^T (U^T U)\Sigma V^T$$

But U is unitary so $U^T = U^{-1}$, so $U^T U = U^{-1} U = I$,

$$\Rightarrow A^T A = V\Sigma^T I \Sigma V^T$$

$$\Rightarrow A^T A = V(\Sigma^T \Sigma)V^T \quad - \text{Eq. (2)}$$

Notice that $\Sigma^T\Sigma$ is an $n \times n$ matrix (because Σ is $m \times n$ and thus Σ^T is $n \times m$). As it so happens, $\Sigma^T\Sigma$ is a diagonal matrix. Why? This is one of those things where you just try it out a couple times and you see why it sort of has to be true. For all those example Σ matrices on the previous page, compute $\Sigma^T\Sigma$ and you'll see the result is always a square diagonal matrix.

Now let's set Eq. (1) and Eq. (2) equal to each other, since they are both equal to $A^T A$.

$$PDP^T = V(\Sigma^T\Sigma)V^T$$

D is diagonal, but so is $\Sigma^T\Sigma$. Also, P is unitary, but so is V. Therefore, we can say $D = \Sigma^T\Sigma$ and $P = V$. Furthermore, when you computed all those examples of $\Sigma^T\Sigma$ earlier, you should have noticed that the diagonal entries of Σ are just the square roots of the diagonal entries of D. I give a quick "proof" of this below for the case where $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$ for arbitrary entries σ_1, σ_2 (you should be able to generalize to any size Σ).

$$D = \Sigma^T\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

As you can see, the diagonal entries of Σ are simply the square roots of the diagonal entries of D. But we know that the diagonal entries of D are the eigenvalues of $A^T A$ (this is a consequence of the eigendecomposition of symmetric $A^T A$, review this if you must). That is, if λ_1, λ_2 are the eigenvalues of $A^T A$, we have $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. If we equate our two formulations of D, we get $\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, so it must be that $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_2 = \sqrt{\lambda_2}$. Note, since we said $A^T A$ was positive definite earlier, all its eigenvalues λ_1, λ_2 are nonnegative, so their square roots are guaranteed to be real numbers, so σ_1, σ_2 are real numbers (whew).

Definition: If $A = U\Sigma V^T$ is an SVD of A, the diagonal entries of Σ are conventionally denoted $\sigma_1, \dots, \sigma_n$ and are called the *singular values of A.* Therefore, **the singular values of $\boldsymbol{A}$ are the square roots of the eigenvalues of $\boldsymbol{A^T A}$.** Note that if A is $m \times n$, then $A^T A$ is $n \times n$, hence $A^T A$ has n eigenvalues counting multiplicity, hence there are always n singular values of A counting multiplicity.

As it turns out, the *nonzero* singular values of A are also equal to the square roots of the eigenvalues of AA^T,[2] but using AA^T instead of $A^T A$ can run into problems since AA^T and $A^T A$ are of different sizes, so I may end up with too many or too few singular values equal to zero. To be safe, just eigendecompose $A^T A$ to find the singular values of A.

So, the summary so far: if we want to find V and Σ, we simply find the eigendecomposition of $A^T A = PDP^T$, upon which $\boldsymbol{V = P}$ and **the diagonal entries of $\boldsymbol{\Sigma}$ (aka the singular values of $\boldsymbol{A}$) are the square roots of the eigenvalues of $\boldsymbol{A^T A}$**.
How do we find U then? Now we consider AA^T instead of $A^T A$. Using a similar argument as above, it can be shown that if $AA^T = W\Lambda \mathrm{W}^{\mathrm{T}}$ is the eigendecomposition of AA^T, then $\boldsymbol{U = W}$. So we need the eigendecomposition of AA^T to find U and we need the eigendecomposition of $A^T A$ to find V (and you could actually use either eigendecomposition to find Σ). When assembling your SVD, don't forget to write V^T and not V.
In short,

$$\begin{cases} A^T A = PDP^T \\ AA^T = W\Lambda \mathrm{W}^{\mathrm{T}} \end{cases} \Rightarrow A = W\Sigma P^T = U\Sigma V^T$$

Now, you will unfortunately have to eigendecompose $A^T A$ to find Σ and V. Although you could eigendecompose AA^T to find U as described above, it is rather demoralizing to have to eigendecompose two matrices just to build an SVD. Luckily, if we know Σ and V, there is a shortcut to finding U that does not involve eigendecomposing AA^T. We will show this shortcut in the worked example later on.

To summarize the relations between the SVD of A and the eigendecompositions of $A^T A$ and AA^T, we have the following:

Definition/Theorem: If $A = U\Sigma V^T$ is the SVD of A, the columns of V (which are just the unit[3] eigenvectors of $A^T A$) are called the *right singular vectors* of A. Similarly, the columns of U (which are just the unit eigenvectors of AA^T) are called the *left singular vectors* of A. As mentioned before, the diagonal entries of Σ (which are the singular values of A) are the square roots of the eigenvalues of $A^T A$.

Lastly, by convention, **the singular values of $\boldsymbol{A}$ are ordered from largest to smallest in $\boldsymbol{\Sigma}$**, with the largest entry being at the top left. For example, if A was 3×4 and its singular values were $2, 3, 5$, we would write its SVD $U\Sigma V^T$ as $U\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}V^T$, and NOT as $U\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}V^T$ or some other order.

By the way, the SVD is not necessarily unique even when the singular values are ordered largest to smallest. That is, a matrix A can have two perfectly legitimate SVDs. Recall that in $A = U\Sigma V^T$, V also satisfies $A^T A = VDV^{-1}$. If two eigenvectors of V (i.e. two columns of V) have the same eigenvalue, those two columns can be interchanged freely to generate two versions of V

---

[2] https://math.stackexchange.com/questions/1087064/non-zero-eigenvalues-of-aat-and-ata
[3] Remember, the columns of V and U must be length one because V, U are unitary, so we need unit eigenvectors.

that both form a perfectly legitimate SVD $U\Sigma V^T$ of A (and a perfectly legitimate eigendecomposition VDV^{-1} of A^TA). Therefore, when people say "let $U\Sigma V^T$ be the SVD of A," they really mean "let $U\Sigma V^T$ be *an* SVD of A." However, it should be noted that if all the eigenvalues of A^TA are distinct, then the SVD is indeed unique.

Lastly, a good theorem to know is that the singular values of a symmetric matrix are equal to the absolute value of its eigenvalues.

Theorem: If A is symmetric with eigenvalues $\lambda_1, \ldots, \lambda_n$ and singular values $\sigma_1, \ldots, \sigma_n$, then $\sigma_i = |\lambda_i|$ for all $i = 1, \ldots, n$. If you are curious, this is because the eigenvalues of A^TA end up being λ_i^2 when A is symmetric, so the singular values of A are just $\sqrt{\lambda_i^2} = |\lambda_i|$. (By convention, we never take the negative root)

Worked Example: Find the SVD of $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$.[4]

Solution: So $A = U\Sigma V^T$ and we have to find U, Σ, and V. If you go back and read the yellow highlighted portion, we see that we can eigendecompose A^TA to get V and Σ.

$$A^TA = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

You learned how to eigendecompose in class, so I won't show all the steps here. We can calculate that the eigenvalues of A^TA are $\lambda = 0, 9, 25$. Recall that the unit eigenvector(s) of A^TA corresponding to the eigenvalue λ are found by row reducing $A - \lambda I$, finding the vector(s) that span $\ker(A + \lambda I)$, and normalizing them to have length one. If you do this, you will find that one choice of unit eigenvectors corresponding to the eigenvalues $\lambda_1 = 25, \lambda_2 = 9$, and $\lambda_3 = 0$ is $\boldsymbol{v_1} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \boldsymbol{v_2} = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix}$, and $\boldsymbol{v_3} = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$ respectively. Recall these vectors, which are the unit eigenvectors of A^TA, are defined to be the *right singular vectors* of A.

Recall the entries of Σ are the square roots of the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of A^TA. So the entries of Σ are $\sigma_1 = 5, \sigma_2 = 3, \sigma_3 = 0$. These are the singular values of A. Recall that Σ has the same dimensions as A. We arrange the singular values greatest to least,

$$\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

But the last singular value 0 doesn't quite "fit." But that's okay, we just put in the nonzero singular values. Why is this okay? Well, the matrix A^TA maps vectors to $\mathbb{R}^3$ (it has three rows)

[4] https://www.d.umn.edu/~mhampton/m4326svd_example.pdf

but A itself only maps vectors to $\mathbb{R}^2$ (since it only has two rows), so A^TA sort of has an "extra" dimension that A lacks, and this is reflected in that extra dimension being associated with this singular value of 0. Just know that we don't have to worry about it when we construct Σ.

Since we arranged the singular values in the order $\sigma_1 = 5, \sigma_2 = 3, \sigma_3 = 0$, we arrange v_1, v_2, v_3 in V in the corresponding order:

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \end{bmatrix}$$

Recall that if we eigendecomposed A^TA as $A^TA = PDP^{-1}$, then the above matrix V equals P.

It remains to find U, which you will recall we could find by eigendecomposing AA^T. But I also mentioned a shortcut that finds U that avoids that. The shortcut is the following theorem:

Theorem:
Let $A = U\Sigma V^T$ be the SVD for the $m \times n$ matrix A. (So U is $m \times m$, V is $n \times n$, and Σ is $m \times n$)
Let $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$ be the i^{th} column of U and V respectively, and let $\sigma_1, \dots, \sigma_n$ be the singular values of Σ (so these are the diagonal entries of Σ). Then $\boldsymbol{u_i} = \frac{1}{\sigma_i} A\boldsymbol{v_i}$ for all i.

We will prove this theorem after the worked example. For now, the important part is that we can now calculate each column of U using the simple formula $\boldsymbol{u_i} = \frac{1}{\sigma_i} A\boldsymbol{v_i}$. Since our matrix A is 2×3, we know U is 2×2 so we only calculate $\boldsymbol{u_1} = \frac{1}{\sigma_1} A\boldsymbol{v_1}$ and $\boldsymbol{u_2} = \frac{1}{\sigma_2} A\boldsymbol{v_2}$ where σ_1, σ_2 correspond to $\boldsymbol{v_1}, \boldsymbol{v_2}$. Even though third singular value of A is $\sigma_3 = 0$, we don't need it to construct U for the reasons described earlier.

$$\boldsymbol{u_1} = \frac{1}{\sigma_1} A\boldsymbol{v_1} = \frac{1}{5}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$\boldsymbol{u_2} = \frac{1}{\sigma_2} A\boldsymbol{v_2} = \frac{1}{3}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}\begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

Recall that $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are the *left singular vectors* of A.

So $U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$. Then the full SVD $A = U\Sigma V^T$ looks like:

Wenzhi Tseng

$$\underbrace{\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & 2/3 & -1/3 \end{bmatrix}}_{V^T}$$

Don't forget to take the transpose of V!

Proof of Theorem: We start from $A = U\Sigma V^T$. We multiply $U^{-1} = U^T$ both sides on the left,

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

We now take the transpose of both sides and use the identity $(AB)^T = B^T A^T$,

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

Multiply by A both sides on the left,

$$\Rightarrow AA^T U = AV\Sigma^T$$

But by plugging in $U\Sigma V^T$ for A, we get:

$$\Rightarrow (U\Sigma V^T)(U\Sigma V^T)^T U = AV\Sigma^T$$

$$\Rightarrow U\Sigma V^T V\Sigma^T U^T U = AV\Sigma^T$$

But since U, V are unitary, we have $V^T = V^{-1}$ and $U^T = U^{-1}$,

$$\Rightarrow U\Sigma\Sigma^T = AV\Sigma^T$$

We would like to "cancel" out the Σ^T on both sides, but since Σ^T is not necessarily invertible (it might not even be square), we need to do more work. I will now assume A is 2×3 for clarity's sake (so Σ is also 2×3) but the proof for any size matrix is analogous.

Let $\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$. Then, our matrix equation above becomes:

$$U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = AV \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow U \underbrace{\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}}_{\Sigma\Sigma^T} = AV \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$$

Assuming σ_1, σ_2 are both nonzero, the matrix $\Sigma\Sigma^T = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ is invertible and its inverse is $(\Sigma\Sigma^T)^{-1} = \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix}$. We multiply this both sides on the right,

$$\Rightarrow U = AV \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix}$$

$$\Rightarrow U = AV \begin{bmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \\ 0 & 0 \end{bmatrix}$$

Wenzhi Tseng

Let $U = \begin{bmatrix} | & | \\ \boldsymbol{u_1} & \boldsymbol{u_2} \\ | & | \end{bmatrix}_{2\times 2}$ and $V = \begin{bmatrix} | & | & | \\ \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} \\ | & | & | \end{bmatrix}_{3\times 3}$ (so $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$ are the i^{th} columns of U and V respectively). You can verify using matrix multiplication that $AV = \begin{bmatrix} | & | & | \\ A\boldsymbol{v_1} & A\boldsymbol{v_2} & A\boldsymbol{v_3} \\ | & | & | \end{bmatrix}_{2\times 3}$ (so $A\boldsymbol{v_i}$ is the i^{th} column of AV).

$$\Rightarrow \begin{bmatrix} | & | \\ \boldsymbol{u_1} & \boldsymbol{u_2} \\ | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ A\boldsymbol{v_1} & A\boldsymbol{v_2} & A\boldsymbol{v_3} \\ | & | & | \end{bmatrix} \begin{bmatrix} 1/\sigma_1 & 0 \\ 0 & 1/\sigma_2 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} | & | \\ \boldsymbol{u_1} & \boldsymbol{u_2} \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ \frac{1}{\sigma_1} A\boldsymbol{v_1} & \frac{1}{\sigma_2} A\boldsymbol{v_2} \\ | & | \end{bmatrix}$$

Hence we have $\boldsymbol{u_i} = \frac{1}{\sigma_i} A\boldsymbol{v_i}$ as desired.

The following sources are excellent (albeit more involved) texts regarding the SVD and its applications. Use them if they help you. Not all of it is relevant.
https://graphics.stanford.edu/courses/cs205a-13-fall/assets/notes/chapter6.pdf
http://web.cs.iastate.edu/~cs577/handouts/svd.pdf
https://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/book-chapter-4.pdf