Last name ______________________

First name ______________________

LARSON—MATH 310—CLASSROOM WORKSHEET 27
Why SVD Works

- Suppose A is $m \times n$ with rank r. So, AA^T and A^TA have rank r.
- Suppose that $A = U\Sigma V^T$, with U and V orthogonal (thus square) and Σ all-0 except possible entires on the diagonal written in non-increasing order.
- Then $AV = U\Sigma$.
- Let the (unit) columns of V be $\vec{v_1} \ldots \vec{v_r} \ldots \vec{v_n}$, and let the (unit) columns of U be $\vec{u_1} \ldots \vec{u_r} \ldots \vec{u_m}$. Let the diagonal elements of Σ (that is, $\Sigma_{i,i}$) be σ_i (with $\sigma_1 \geq \ldots \geq \sigma_r \geq \ldots \geq \sigma_m$).
- Then $AV = U\Sigma$ is $A[\vec{v_1} \ldots \vec{v_r} \ldots \vec{v_n}] = [\vec{u_1} \ldots \vec{u_r} \ldots \vec{v_m}]diag(\sigma_1, \ldots, \sigma_r, \ldots, \sigma_n)$.
- So, $[A\vec{v_1} \ldots A\vec{v_r} \ldots A\vec{v_n}] = [\sigma_1\vec{u_1} \ldots \sigma_r\vec{u_r} \ldots \sigma_m]$.
- Since the columns of U are orthogonal they are linearly independent. Since AV has rank r, only r of the columns of $U\Sigma$ are linearly independent, which means $m - r$ of the σ_i's must be 0. Since these were listed in non-increasing order, it must be that $\sigma_1, \ldots, \sigma_r$ are positive and $\sigma_{r+1}, \ldots, \sigma_m$ equal 0.
- It then follows that $A\vec{v_1} = \sigma_1\vec{u_1}, \ldots,\ A\vec{v_r} = \sigma_r\vec{u_r}$ and $A\vec{v_{r+1}} = 0,\ \ldots,\ A\vec{v_n} = 0$.
- Since $\vec{v_{r+1}}, \ldots, \vec{v_n}$ are orthogonal, they are linearly independent and then a basis for the nullspace of A.
- Where do the the σ's come from? Their squares must be eigenvalues of A^TA as $A^TA = (U\Sigma V^T)^T(U\Sigma V^T)$.
- So, $A^TA = V diag(\sigma_1^2, \ldots, \sigma_r^2)V^T$.
- And then, $(A^TA)V = V diag(\sigma_1^2, \ldots, \sigma_r^2)V^T$.
- So, $(A^TA)[\vec{v_1} \ldots \vec{v_r} \ldots \vec{v_n}] = [\vec{v_1} \ldots \vec{v_r} \ldots \vec{v_n}]diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- Finally, we get $(A^TA)\vec{v_1} = \sigma_1^2\vec{v_1}, \ldots,\ (A^TA)\vec{v_r} = \sigma_r^2\vec{v_r}$, and $(A^TA)\vec{v_{r+1}} = 0,\ \ldots,\ (A^TA)\vec{v_n} = 0$.
- Where do the $\vec{v_i}$'s come from (for $i = 1, \ldots, r$)? They are the (unit) eigenvectors corresponding to σ^2 for A^TA.
- Let $u_i = \frac{1}{\sigma_i}A\vec{v_i}$ (for $i = 1, \ldots, r$). By construction $A\vec{v_i} = \sigma_i\vec{u_i}$, as needed.
- Are the $\vec{u_i}$'s unit? $||\vec{u_i}|| = ||\frac{1}{\sigma_i}A\vec{v_i}|| = \frac{1}{\sigma_i}||A\vec{v_i}|| = \frac{1}{\sigma_i}\sqrt{(A\vec{v_i})^T(A\vec{v_i})} = \frac{1}{\sigma_i}\sqrt{\vec{v_i}^T A^TA\vec{v_i}} = \frac{1}{\sigma_i}\sqrt{\sigma_i^2\vec{v_i}^T\vec{v_i}} = \frac{1}{\sigma_i}\sigma_i = 1$.
- Are the $\vec{u_i}$'s orthogonal? They are eigenvectors of AA^T corresponding to different eigenvalues (or, if the subspace corresponding to an eigenvalue has dimension greater than 1, then made orthogonal using Gram-Schmidt.

Let $A = \begin{bmatrix} 3 & 0 \\ 4 & 0 \end{bmatrix}$.

We will find the **singular value decomposition** (SVD) of an $m \times n$ matrix A. That is we want to write $A = U\Sigma V^T$ where U and V are orthogonal matrices and Σ is a "diagonal" matrix (all-zeros except singular values on the diagonal).

1. What is m, n and the rank r of the given matrix A?

2. If A is 2×2, what sizes do U, Σ and V have to be?

3. Find AA^T, its eigenvalues σ_i^2 and unit eigenvectors $\vec{v_i}$ to get V. If there are any zero eigenvalues of AA^T, there will be $n - r$ of them, and we will find a basis for the nullspace, then use Gram-Schmidt to get an orthonormal basis for the null space and call those vectors $\vec{v_{r+1}}, \ldots, \vec{v_n}$.

4. Check that V is orthogonal.

5. What are the positive singular values $\sigma_1 \geq \ldots \geq \sigma_r$? Find Σ.

6. Find $A^T A$ and check that its rank is r. We proved $A^T A$ is symmetric with nonnegative eigenvalues (this means it is also *positive semi-definite*). Why will it have the same eigenvalues as AA^T?

7. Find the unit eigenvectors $\vec{u_i}$ to get U.

8. Check that U is orthogonal.

9. Check that $A\vec{v_i} = \frac{1}{\sigma_i} u_i$ for $i = 1 \ldots r$.

10. Check that $A = U\Sigma V^T$.