Names: ______________________________

NetIDs: ______________________________

Discussion section: D ________ **Group number:** ________

MATH 257 - WORKSHEET 14

Given a real $m \times n$ matrix A, recall we have a linear transformation $L\colon \mathbb{R}^n \to \mathbb{R}^m$ defined by $L(\mathbf{x}) = A\mathbf{x}$. When A is also square and symmetric, the Spectral Theorem assures us an orthonormal eigenbasis $\mathcal{B}$ for A. This implies that $L_{\mathcal{B},\mathcal{B}}$ is a diagonal matrix with diagonal entries the eigenvalues of A. Note that the basis $\mathcal{B}$ is doubly nice: (1) it is orthonormal, and (2) it is an eigenbasis for A. For general rectangular matrices, can we find nice ordered bases $\mathcal{F}, \mathcal{G}$ for $\mathbb{R}^n, \mathbb{R}^m$ respectively so that $L_{\mathcal{G},\mathcal{F}}$ is diagonal, i.e. having non-zero entries only along the main-diagonal?
In this worksheet we will investigate how to do this by finding certain orthonormal bases (**ONB**) for each of the four fundamenatal subspaces so that they together "play well" with the matrix A to yield a coordinate matrix that is diagonal. This will lead us to the powerful **Singular Value Decomposition** (**SVD**).
Note: Occasionally we will use brackets to denote matrices but this has no additional mathematical meaning. It is for ease of typesetting.

(1) First, an example: Consider the matrix $A = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & -1 & -1 \end{pmatrix}$ which has the following four fundamental subspaces:

$$Col(A^T) = span\{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}\} \qquad Nul(A) = span\{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}\}$$

$$Col(A) = span\{\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}\} \qquad Nul(A^T) = span\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\}$$

where the given tuples are basis vectors for their respective subspaces. Let

$$\mathcal{F} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right)$$

$$\mathcal{G} = (\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) = \left(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right)$$

(a) Assuming $\mathcal{G}, \mathcal{F}$ are bases for $\mathbb{R}^m, \mathbb{R}^n$ respectively, find $L_{\mathcal{G},\mathcal{F}}$.

(b) (**Optional**, slight messy) What if we apply the Gram-Schmidt process to get orthonormal bases $\hat{\mathcal{F}}, \hat{\mathcal{G}}$? What is $L_{\hat{\mathcal{G}},\hat{\mathcal{F}}}$.

(2) We didn't quite achieve our goal, but the preceding example(s) have been informative. Let's consider what the previous example(s) indicate.

(a) First, let's clear up the assumption we used earlier: Suppose $rk(A) = r$ with $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_r)$ an ordered ONB for $Col(A^T)$ and $\mathcal{B} = (\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$ an ordered ONB for $Nul(A)$. Explain why $\mathcal{F} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is an ordered ONB for $\mathbb{R}^n$. Similarly, if $\mathcal{C} = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ is an ordered ONB for $Col(A)$ and $\mathcal{D} = (\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m)$ an ordered ONB for $Nul(A^T)$, then $\mathcal{G} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ is an ordered ONB for $\mathbb{R}^m$.

(b) Next, notice that we could choose *any* ONB for each of $Nul(A), Nul(A^T)$ without changing the fact that $L_{\mathcal{G},\mathcal{F}} = \left[\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array}\right]_{m\times n}$ for some matrix $B_{r\times r}$. Why is this true?

(c) If we want B diagonal, we need ONB $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ for $Col(A^T)$ and ONB $(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ for $Col(A)$ compatible in the sense that for each $i = 1, \ldots, r$ the tuple $A\mathbf{v}_i$ is a multiple of $\mathbf{u}_i$. Show that this implies the collection $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is *orthogonal*.

(3) The preceding result implies for $i \neq j$ we have $0 = (A\mathbf{v}_i) \cdot (A\mathbf{v}_j) = \mathbf{v}_i^T(A^TA)\mathbf{v}_j$ so that we want to choose ONB $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ for $Col(A^T)$ with the property that $0 = \mathbf{v}_i \cdot (A^TA)\mathbf{v}_j$ whenever $i \neq j$.

(a) Show that this condition implies that the $\mathbf{v}_i$ are eigenvectors of A^TA.

(b) This directs our attention to the matrix A^TA. Recall from the lecture modules that $\text{Nul}(A^TA) = \text{Nul}(A)$. Explain why an ONB for the 0-eigenspace of A^TA is an ONB for $Nul(A)$ and why $rk(A^TA) = r$

(c) Since A^TA is symmetric, the Spectral Theorem yields an orthonormal eigenbasis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$. Suppose $\mathcal{B} = (\mathbf{v}_{r+1}, \dots, \mathbf{v}_n)$ is an ONB for $Nul(A^TA)$. Why is $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ an ONB for $Col(A^T)$?

(4) Let $\mathcal{F} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ orthonormal eigenbasis for $A^T A$, which we now recognize as merging ONB for each of $Col(A^T)$ and $\mathrm{Nul}(A)$.

(a) By our earlier work, we know that $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is an orthogonal collection in $Col(A)$. Show that the collection is actually an orthogonal basis for $Col(A)$.

(b) Find the lengths $\|A\mathbf{v}_i\|$ and so obtain an ONB $\mathcal{C} = (\mathbf{w}_1, \ldots, \mathbf{w}_r)$ for $Col(A)$.

(c) Pick any ONB $\mathcal{D} = (\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n)$ for $Nul(A^T)$. Let $\mathcal{G} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Find $L_{\mathcal{G},\mathcal{F}}$.

(d) Let $\mathcal{E}_n, \mathcal{E}_m$ denote the standard bases in $\mathbb{R}^n, \mathbb{R}^m$ respectively. Since $L_{\mathcal{E}_m,\mathcal{E}_n} = A$, use the change of coordinates formula $L_{\mathcal{E}_m,\mathcal{E}_n} = I_{\mathcal{E}_m,\mathcal{G}} L_{\mathcal{G},\mathcal{F}} I_{\mathcal{F},\mathcal{E}_n}$ to write A as a product of three matrices.

(5) **Congratulations!** You've just rediscovered Singular Value Decomposition (SVD). The positive values $\sigma_i = \sqrt{\lambda_i}$ are called **singular values** and we usually order the basis for $Col(A^T)$ so that the singular values will be arranged in non-increasing order $\sigma_1 \geq \ldots \geq \sigma_r$ within the diagonal block D. We often write $\Sigma = L_{\mathcal{G},\mathcal{F}}$ so that SVD is written $A = U\Sigma V^T$. Is SVD unique, i.e. given A are there unique U, V such that $A = U\Sigma V^T$ with the diagonal entries of Σ in non-increasing order?

Epilogue: We see that Singular Value Decomposition of real matrices is not magic but rather the nigh inexorable conclusion of a specific goal and specific approach:

- Goal: Find bases $\mathcal{F}, \mathcal{G}$ of $\mathbb{R}^n, \mathbb{R}^m$ respectively such that $L_{\mathcal{G},\mathcal{F}}$ is diagonal.
- Approach: choose appropriate ONB for each of the four fundamental subspaces to obtain the goal.

These observations support SVD's claim as a beautiful and significant result of linear algebra.