EECS 16B
Worksheet 8

1. Upper Triangularization

Recall that before we solved the system of differential equation $\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) + \vec{b}u(t)$ using the change of basis V which contained the eigenvectors of A. The result of the transformation showed that $V^{-1}AV = D$ became the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A

However, the underlying assumption for this transformation is that our matrix A was diagonalizable. In such a case that our matrix A is not diagonalizable, we can still solve the system of differential equation using upper triangularization by solving it "bottom-up" using backwards substitution.

(a) Give an example of a 2×2 matrix that is not diagonalizable. Explain why we can't use our original trick $D = V^{-1}AV$.

(b) For this given upper triangular system of differential equations

$$\frac{d}{dt}\vec{x}(t) = \begin{bmatrix} \lambda_1 & a & \dots & A_{1,n-1} & A_{1,n} \\ 0 & \lambda_2 & \dots & A_{2,n-1} & A_{2,n} \\ \vdots & 0 & \ddots & \lambda_{n-1} & A_{n-1,n} \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix} \vec{x}(t)$$

Explain how we can use upper triangularization to solve the system of differential equation.
Hint: Which differential equation can we solve immediately? Can we use that solution to backward substitute in anyway?

(c) We will now perform upper triangularization on a non-diagonalizable 2×2 matrix.

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$$

Show that $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A and find its corresponding eigenvalue.

(d) Using this eigenvector and some appropriate basis vectors of R^2, find an orthonormal basis U for the column space of A using the Gram-Schmidt algorithm. What do you notice about running the Gram-Schmidt algorithm with more vectors that needs to span a certain subspace? How does Gram-Schmidt handle this?

(e) The next step is to compute $Q = U^T AU$. Compute Q. What do you notice about the matrix Q? What is interesting about the diagonals of Q?

(f) Notice that we have computed $U^T AU = Q = \begin{bmatrix} \lambda_1 & 2 \\ 0 & \lambda_2 \end{bmatrix}$. Show that you can write $A = U \begin{bmatrix} \lambda_1 & 2 \\ 0 & \lambda_2 \end{bmatrix} U^T$.

Hint: What is special about the matrix U*?*

(g) Now let's extend this to a 3×3 matrix. Play around with the Jupyter Notebook to see the process of upper triangularizing a 3×3 matrix.

(h) Notice that we got lucky that we only needed to run Gram-Schmidt once to find a basis which upper triangularizes our 2×2 matrix A. For higher dimension matrices like the 3×3 matrix in the Jupyter Notebook, however, this is not the case and multiple iterations of part (d)-(f) is required to construct such basis. For a general matrix $n \times n$, how many times do we need to perform Gram-Schmidt in order to construct a basis that upper triangularizes A?

Hint: In the 3×3 *matrix example in the Jupyter Notebook, how many times was Gram-Schmidt algorithm ran? Why did it run that many times?*

(i) In your own words, write down the general algorithm for upper triangularizing an $n \times n$ matrix A.

2. Spectral Intuition

An amazing result in Linear Algebra is the Spectral Theorem which says that any symmetric matrix is orthogonally diagonalizable. This means that a symmetric matrix will always have n linearly independent eigenvectors that are all mutually orthogonal. We will show that some of these properties are true for the symmetric matrix $A^T A$ to help motivate the SVD.

(a) Show that $A^T A$ is a symmetric matrix.

(b) Show that every eigenvalue λ_i of $A^T A$ is greater than or equal to zero.
Hint: Consider $\|A\vec{v}\|_2^2$. *where* $\vec{v}$ *is an eigenvector of* $A^T A$ *with eigenvalue* λ.

(c) Show that if λ_i and λ_j are distinct eigenvalues of $A^T A$, then the respective eigenvectors $\vec{v}_i$ and $\vec{v}_j$ are orthogonal.
Hint: Write out the eigenvector relationships: $A^T A\vec{v}_i = \lambda_i \vec{v}_i$ *and* $A^T A\vec{v}_j = \lambda_i \vec{v}_j$ *and then try taking the transpose of the second equation.*

(d) Show that if A^TA has a repeated eigenvalue, λ, meaning the eigenspace of λ has dimension greater than or equal to two, we can pick an orthonormal basis for the eigenspace.

(e) It can be shown through induction that the matrix A^TA is has n linearly independent eigenvectors. Conclude by showing that we can pick n mutually orthonormal eigenvectors for the matrix A^TA. *HINT: Recall what gram-schmidt does.*

3. (Optional) Basic SVD Practice

This is a review of the individidual steps of finding the Singular Value Decomposition of an $m \times n$ matrix A.

The final answer will be of the form $A = U\Sigma V^T$ where U is a $m \times m$ orthonormal matrix, V is a $n \times n$ orthonormal matrix, and Σ is a $m \times n$ matrix that is a diagonal matrix with 0s padded on the right or below depending on the dimensions m and n.

We give the following procedure to compute the SVD of a $m \times n$ matrix A.

(i) **Step 1: Compute the symmetric matrix A^TA or AA^T.**

A^TA will be of dimension $n \times n$, and AA^T will be of dimension $m \times m$.

For a tall, skinny matrix, where $m > n$, the SVD will be easier to calculate using A^TA while for a short, fat matrix, where $m < n$, the SVD will be easier to calculate using AA^T.

(ii) **Step 2: Find the eigenvalues and eigenvectors of A^TA or AA^T.**

If $m > n$, find the eigenvalues $(\lambda_1, \ldots, \lambda_n)$ and eigenvectors $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$ of A^TA.
If $m < n$, we find the eigenvalues $(\lambda_1, \ldots, \lambda_m)$ and eigenvectors $(\vec{u}_1,\ \vec{u}_2, \ldots, \vec{v}_m)$ of AA^T.
By the spectral theorem for real symmetric matrices, these eigenvectors are orthonormal.

(iii) **Step 3: Compute the singular values $\sigma_i = \sqrt{\lambda_i}$ where λ_i are the sorted in descending order eigenvalues of A^TA or AA^T.**

We know these are all non-negative because $(A\vec{v}_i)^T(A\vec{v}_i) = \|A\vec{v}_i\|^2$ and $(A\vec{v}_i)^T(A\vec{v}_i) = \vec{v}_i^T(A^TA)\vec{v}_i = \lambda_i\vec{v}_i^T\vec{v}_i = \lambda_i$. The corresponding normalized eigenvectors $\vec{v}_i$ form the V matrix.

(iv) **Step 4: Find the corresponding vectors of the U or V matrix**

If $m > n$, we use the nonzero values of σ_i, and $\vec{v}_i$, to find corresponding vectors of the U matrix, $\vec{u}_i$ by computing $\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}$.
If $m < n$, we use the nonzero values of σ_i and $\vec{u}_i$, to find corresponding vectors of the V matrix, $\vec{v}_i$ by computing $\vec{v}_i = \frac{A^T\vec{u}_i}{\sigma_i}$
These are normalized since $\sigma_i = \|A\vec{v}_i\| = \|A^T\vec{u}_i\|$ by the argument above, and orthogonal since $(A\vec{v}_i)^T(A\vec{v}_j) = \vec{v}_i^T(A^TA)\vec{v}_j = \lambda_j\vec{v}_i^T\vec{v}_j = 0$ if $i \neq j$, since V is an orthonormal matrix.

(v) **Step 5 (for finding the full SVD): Use Gram-Schmidt to complete the U or V matrix**

If $m > n$ we can complete the U matrix by finding $\vec{u}_{n+1}, \ldots, \vec{u}_m$ through Gram-Schmidt.
If $m < n$ we will complete the V matrix by finding $\vec{v}_{m+1}, \ldots, \vec{v}_n$ through Gram-Schmidt.

Alternatively we can solve for $\vec{v}_i$ by computing the null-space of A or $\vec{u}_i$ by computing the null-space of A^T and then performing Gram-Schmidt on the basis for the respective null-space.

Recall the following SVD forms:

Compact SVD	$A = U_r\Sigma_r{V_r}^T$
Outer Product SVD	$\sum_{i=1}^{r}\sigma_i\vec{u}_i\vec{v}_i^T$
Full SVD	$A = U\Sigma V^T$

(a) Given the matrix:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Compute the following:

i. compact SVD

ii. outer product SVD

iii. full SVD

Contributors:

- Nareauphol Liu.
- Taejin Hwang.

