1. **Properties of SVD components.**

Now that you have seen SVD in lecture, this warm-up problem is intended to help you recall some linear algebra properties in order to understand why SVD has its particular form.

Let X be a real matrix of size $m \times n$ (assuming in the beginning $m \geq n$), and let $X = U\Sigma V^T$ be the SVD of X, where U is an $m \times m$ orthonormal matrix, V is an $n \times n$ orthornormal matrix, and Σ is an $m \times n$ block diagonal matrix whose diagonal entries $\sigma_0, \ldots, \sigma_{n-1}$ are the singular values of X.

(a) We know that the singular values are equal to the square roots of the eigenvalues of $X^T X$, or XX^T, i.e., $\sigma_i = \sqrt{\lambda_i}$. In lecture, it was shown that the eigenvalues of $X^T X$, or XX^T are real.
Using a similar technique, show that (1) the eigenvalues of $X^T X$, denoted by $\lambda_0, \cdots, \lambda_{n-1}$, are non-negative, and that (2) $X^T X$ and XX^T have the same set of non-zeros eigenvalues.

(b) We also know that the columns of U are the eigenvectors of XX^T, the columns of V are the eigenvectors of $X^T X$, and that they are orthonormal matrices. Recall (or prove) that an orthonormal transformation preserves inner product (hence norm as well). Consider a matrix as a linear operator that does "rotations", "reflection", and scaling. Which operations can U and V perform? How about Σ?

(c) What would you do if $n \geq m$? What if anything changes in your argument?

(d) Now, what can you say about SVD?

2. **Diagonalization of symmetric matrices by orthonormal eigenvectors.**

You may be wondering why U and V have their specific forms. In this problem, we will show more generally that every real symmetric matrix can be diagonalized by a matrix of its orthonormal eigenvectors. In other words, a symmetric matrix has a full complement of eigenvectors that are all orthogonal to each other.

Note that in homework 3, you will be asked to derive this in a more formal way using induction. Here we will just provide some key steps for a recursive derivation. (That can be turned into an inductive proof.)

In order for you to better understand the involved steps, you can consider a concrete case

$$S_{[3\times 3]} = \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

and figure out the general case by abstracting variables.

(a) Consider a non-zero vector $\vec{u}_0 \in \mathbb{R}^n$. Can you think of a way to extend it to a set of basis of $\mathbb{R}^n$? In other words, find $\vec{u}_1, \cdots, \vec{u}_{n-1}$, such that $\text{span}(\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-1}) = \mathbb{R}^n$. To begin with, consider $[1, -1, 0]^T$.

(b) Can you get an orthonormal basis from what you just constructed?

(c) Now consider a real eigenvalue λ_0, and the corresponding eigenvector $\vec{g}_0 \in \mathbb{R}^n$ of a symmetric matrix $S \in \mathbb{R}^{n\times n}$. From the previous part, we can extend $\vec{g}_0$ to an orthonormal basis of $\mathbb{R}^n$, denoted by

$$V = [\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_{n-1}]$$

where $\vec{v}_0 = \frac{\vec{g}_0}{\|\vec{g}_0\|}$. Compute $V^T SV$ by writing $V = [\vec{v}_0, R]$, where $R \triangleq [\vec{v}_1, \cdots, \vec{v}_{n-1}]$. If you prefer, you can do this and the next question with the concrete $S_{[3\times 3]}$ first.

(d) Define $Q = R^T SR$. Look at the first column and the first row of $V^T SV$ and show that

$$S = V \begin{bmatrix} \lambda_0 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} V^T$$

What can you say about Q?

(e) You have observed that Q is an $(n-1)\times(n-1)$ symmetric matrix. Now, we will perform the same steps (c) and (d) on Q to get:

$$Q = [\vec{u}_0, Y] \begin{bmatrix} \lambda_1 & \vec{0}^T \\ \vec{0} & P \end{bmatrix} [\vec{u}_0, Y]^T$$

where we have taken $\vec{u}_0 \in \mathbb{R}^{n-1}$, a eigenvector Q, associated with eigenvalue λ_1. Again $\vec{u}_0$ is extended into an orthonormal basis $[\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-2}]$ of $\mathbb{R}^{n-1}$. We denote $Y \triangleq [\vec{u}_1, \cdots, \vec{u}_{n-2}]$.
Plug this into S to show that:

$$S = [\vec{v}_0, R\vec{u}_0, RY] \begin{bmatrix} \lambda_0 & 0 & \vec{0}^T \\ 0 & \lambda_1 & \vec{0}^T \\ \vec{0} & \vec{0} & P \end{bmatrix} [\vec{v}_0, R\vec{u}_0, RY]^T$$

Again, using the concrete case may help you first.

(f) Show that the matrix $[\vec{v}_0, R\vec{u}_0, RY]$ is still orthonormal. Moreover, show that $R\vec{u}_0$ is an eigenvector of S corresponding to eigenvalue λ_1.

(g) Perform the above process recursively - what will you get in the end? Comment on how you obtained orthonormal diagonization of S.

Contributors:

- Yuxun Zhou.
- Edward Wang.