Discussion 11A

The following notes are useful for this discussion: Note 16

1. **Geometric Interpretation of the SVD**

   In this exercise, we explore the geometric interpretation of matrix transformations and how this connects to the SVD. We consider how a real 2×2 matrix acts on the unit circle, transforming it into an ellipse. It turns out that the principal semiaxes of the resulting ellipse are related to the singular values of the matrix, as well as the vectors in the SVD.

   (a) Consider the real 2×2 matrix

$$A = \begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix}. \tag{1}$$

   Also consider the unit circle in $\mathbb{R}^2$,

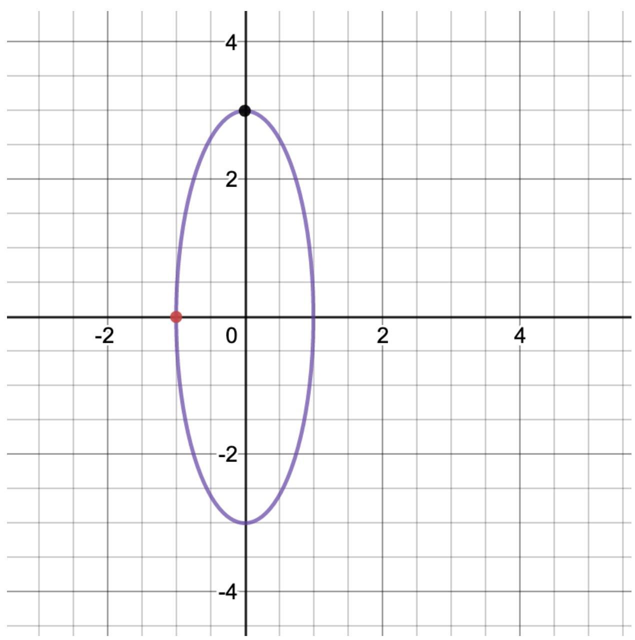
$$S = \left\{ \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \middle| 0 \le \theta < 2\pi \right\}. \tag{2}$$

   Plot the transformed circle, AS, on the $\mathbb{R}^2$ plane.

   Solution:

$$AS = \left\{ \begin{bmatrix} -\sin\theta \\ 3\cos\theta \end{bmatrix} \middle| 0 \le \theta < 2\pi \right\}. \tag{3}$$

   The plot should be the ellipse centered at the origin that passes through the points $(0,3)$, $(0,-3)$, $(-1,0)$, $(1,0)$.



   (b) Now let's consider how this transformation looks in the lens of the SVD. The SVD for matrix A is:

$$A = U\Sigma V^\top = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4}$$

$$A\vec{x} = U\Sigma V^\top \vec{x} = U\left(\Sigma\left(V^\top \vec{x}\right)\right). \tag{5}$$

Let's start by examining the effects of each of these matrices one at a time, right to left, in the same order that they would be applied to a vector $\vec{x}$. **What does the unit circle look like after being transformed by just $V^\top$? Plot $S_1 = V^\top S$ on the $\mathbb{R}^2$ plane. Geometrically speaking, what does $V^\top$ do to any given $\vec{x}$?**

Solution: $V^\top$, being an orthonormal matrix can only rotate or reflect a vector $\vec{x}$. In particular, it applies a rotation or reflection such that the vectors $\vec{v}_i$ in the standard basis are transformed to the elementary vectors $\vec{e}_i$ in the V basis. Note that this matrix cannot do any scaling.

See jupyter notebook for plots.

(c) **What does the unit circle look like after being transformed by $\Sigma V^\top$? Plot $S_2 = \Sigma V^\top S$ on the $\mathbb{R}^2$ plane. Geometrically speaking, what is the Σ matrix doing to any given $V^\top \vec{x}$?**

Solution: The matrix Σ scales vectors that have been transformed into the V basis. In terms of the SVD, it scales the components of $\vec{x}$ in the direction of $\vec{v}_1$ by σ_1, the components in the direction of $\vec{v}_2$ by σ_2, and so on for larger matrices. All of the scaling done by the original matrix A is captured by the Σ matrix.

See jupyter notebook for plots.

(d) **What does the unit circle look like after being transformed by $U\Sigma V^\top$? Plot $S_3 = U\Sigma V^\top S$ on the $\mathbb{R}^2$ plane. Geometrically speaking, what is the U matrix doing to any given $\Sigma V^\top \vec{x}$?**

Solution: U is an orthonormal matrix similar to $V^\top$, and as such can only apply a rotation or reflection to a vector. In the context of the SVD, U rotates or reflects the scaled vectors $\Sigma V^\top \vec{x}$ to their final locations.

See jupyter notebook for plots.

(e) Consider the columns of the matrices U, V from the SVD of A in part (b), and treat them as vectors in $\mathbb{R}^2$. Let $U = (\vec{u}_1\ \vec{u}_2)$, $V = (\vec{v}_1\ \vec{v}_2)$. Let σ_1, σ_2 be the singular values of A, where $\sigma_1 \geq \sigma_2$. **In your plot of AS, draw the vectors $\sigma_1\vec{u}_1$ and $\sigma_2\vec{u}_2$ from the origin. What do these vectors correspond to geometrically?**

Solution: $\sigma_1\vec{u}_1 = (0, -3)$ corresponds to the semi-major axis of the ellipse, while $\sigma_2\vec{u}_2 = (-1, 0)$ corresponds to the semi-minor axis.

See jupyter notebook for plots.

(f) **Repeat parts (b-e) for the following matrices, and note down any interesting things you notice.**

i. A 3D matrix, $X = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

ii. A rotation matrix, $A_1 = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$.

iii. A diagonal matrix, $A_2 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$.

iv. A symmetric matrix, $A_3 = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$.

v. A matrix with non-trivial nullspace, $A_4 = \begin{bmatrix} 4 & 2 \\ -2 & -1 \end{bmatrix}$.

vi. An arbitrary matrix, $A_5 = \begin{bmatrix} 1.6 & 2.4 \\ -0.4 & -1 \end{bmatrix}$.

Solution: See jupyter notebook for plots.

Contributors:

- Neelesh Ramachandran.

- Lynn Chua.
- Shane Barratt.
- Kuan-Yun Lee.
- Anant Sahai.
- Kareem Ahmad.

