1 Geometric interpretation of the SVD

In this exercise, we explore the geometric interpretation of symmetric matrices 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

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

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

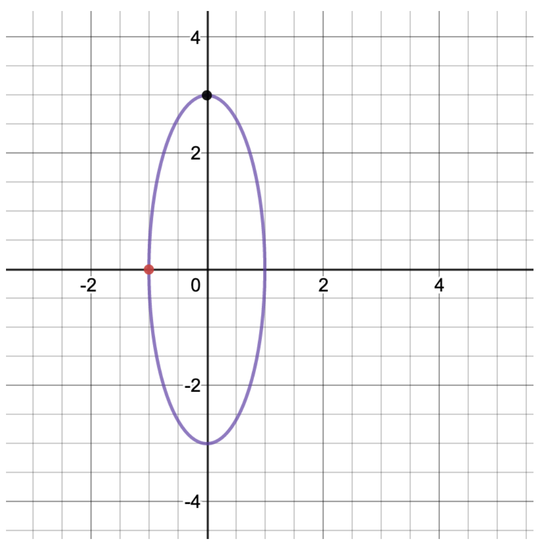
$$S = \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \mid 0 \le \theta < 2\pi \right\}.$$

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

Answer

$$AS = \left\{ \begin{pmatrix} -\sin\theta \\ 3\cos\theta \end{pmatrix} \mid 0 \le \theta < 2\pi \right\}.$$

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



b) Calculate the SVD of A. Write this as a matrix factorization, i.e. $A = U\Sigma V^*$.

Answer

Since A is square, both AA^* and A^*A will be the same size. We arbitrarily choose to start with A^*A.

$$A^*A = \begin{bmatrix} 0 & 3 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

The eigenvalues can be read off as:

$$\lambda_1 = 9, \; \lambda_2 = 1$$

and the corresponding eigenvectors as:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We then calculate the left singular vectors as $\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}$, where $\sigma_i = \sqrt{\lambda_i}$ as usual.

$$\vec{u}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Thus, the SVD of A is given by:

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

c) Consider the columns of the matrices U, V obtained in the previous part, 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 $\sigma_1 \geq \sigma_2$.

Draw in your plot of AS the vectors $\sigma_1\vec{u}_1$ and $\sigma_2\vec{u}_2$, drawn from the origin. What do these vectors correspond to geometrically?

Answer

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

d) **Repeat what you did above for the matrix $A = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$.**

Answer

Here,

$$AS = \left\{ \begin{pmatrix} 2\cos\theta + \sin\theta \\ -2\cos\theta + \sin\theta \end{pmatrix} \mid 0 \leq \theta < 2\pi \right\}.$$

Note that

$$A^*A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$$

which has eigenvalues

$$\lambda_1 = 8, \ \lambda_2 = 2$$

and eigenvectors

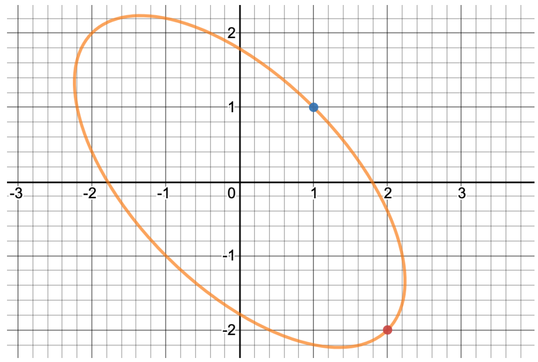
$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Running these through A and dividing by the appropriate singular values $\sigma_1 = 2\sqrt{2}, \ \sigma_2 = \sqrt{2}$:

$$\vec{u}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

The SVD decomposition for A is therefore:

$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$



The ellipse and corresponding points are plotted in the above graph.

Notice that the first vector points along the -45 degree line which corresponds to the major axis for the ellipse while the second is along the 45 degree line which is the minor axis for the ellipse. Once again, it is the U matrix whose constituent vectors give these directions.

Why? We note that we are looking at the image of a circle under A. If we view the SVD as a rotation by V^*, a scaling of the axes by Σ, and then finally a rotation by U, the circle is invariant under the initial rotation V^* by definition. Thus the principle axes of the ellipse are determined by $\sigma_i \vec{u}_i$.

If you want to explore some more on your own, make a wide matrix filled with the output of $A\vec{x}$ for $\vec{x}$ that are drawn randomly from the set S representing the unit circle. (You can create these by using numpy.random.uniforn to draw a bunch of samples uniformly from 0 to 2π, and then generate vectors using cosine and sine.) Then ask numpy to take the SVD of the resulting wide matrix. Look at the U matrix that the SVD returns. You will see that it will be pretty close to what we got above. This reflects the power of the SVD to discover the underlying elliptical structure given a bunch of points. Given the evaluation of a matrix A on a bunch of points drawn uniformly from around the unit circle (and in higher dimensions, from the unit hypersphere) it will reveal the major and minor axes for the relevant ellipsoid along with the shape of the ellipsoid through the singular values.

e) **Consider the case where A is a real $n \times n$ symmetric matrix. What do you observe geometrically in this case?**

Answer

In this case, AA^T and A^TA are equal and therefore have the same eigenvalues and eigenvectors. Hence $U = V$, and geometrically the action of A corresponds to scaling the unit sphere in $\mathbb{R}^n$ along the vectors $\vec{v}_i$ by a factor of σ_i for each i to get a hyperellipse.

$A\vec{x} = U\Sigma V^T\vec{x}$. First, $\vec{x}$ is projected onto the eigenvectors of A^TA, then the different dimensions are scaled by the singular values, and the vector is reconsituted through a linear combination of the same eigenvalues by applying U.

2 SVD and Induced 2-Norm

a) Show that if U is a unitary matrix then for any $\vec{x}$

$$\|U\vec{x}\| = \|\vec{x}\|.$$

Answer

$$\|U\vec{x}\| = \sqrt{(U\vec{x})^*(U\vec{x})} = \sqrt{\vec{x}^*U^*U\vec{x}} = \sqrt{\vec{x}^*\vec{x}} = \|\vec{x}\|$$

b) Find the maximum

$$\max_{\{\vec{x}:\|\vec{x}\|=1\}} \|A\vec{x}\|$$

in terms of the singular values of A.

Answer

If we write A in terms of its SVD $A = U\Sigma V^*$ and introduce a coordinate transformation $\vec{x} = V\vec{y}$ then

$$\begin{aligned}
\max_{\{\vec{x}:\|\vec{x}\|=1\}} \|A\vec{x}\| &= \max_{\{\vec{x}:\|\vec{x}\|=1\}} \|U\Sigma V^*\vec{x}\| \\
&= \max_{\{\vec{y}:\|V\vec{y}\|=1\}} \|U\Sigma V^*V\vec{y}\| \\
&= \max_{\{\vec{y}:\|\vec{y}\|=1\}} \|\Sigma\vec{y}\|
\end{aligned}$$

If the singular values are ordered such that the largest singular value is in the Σ_{11} location than the maximum is achieved at $y = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$ and the value achieved is $\sigma_{\max}(A)$. Thus

$$\max_{\{\vec{x}:\|\vec{x}\|=1\}} \|A\vec{x}\| = \sigma_{\max}(A).$$

c) Find the $\vec{x}$ that maximizes the expression above.

Answer

$$\vec{x} = V\vec{y}_{\max} = \vec{v}_1,$$