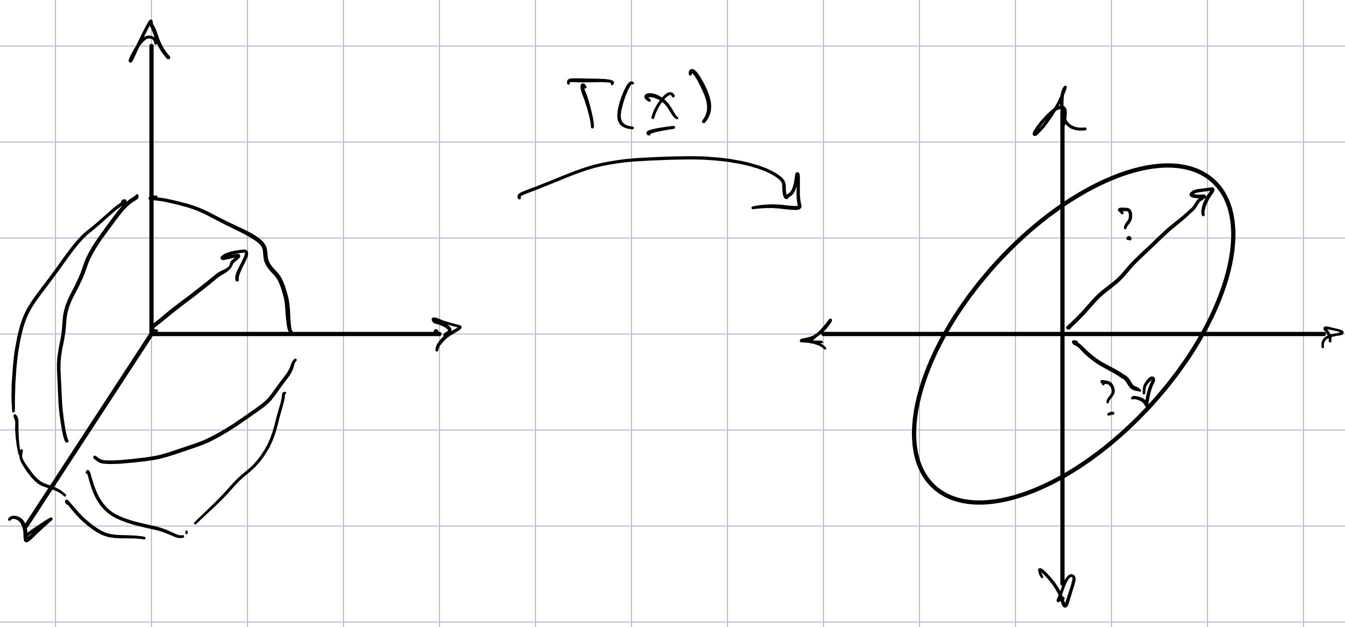
Singular Value Decomposition (SVD)

Motivation

Consider a linear transform $T(\underline{x}) : \mathbb{R}^3 \to \mathbb{R}^2$
where $T(\underline{x}) = A\underline{x}$ and $A \in \mathbb{R}^{2\times 3}$.

It may be of interest to determine for $\|\underline{x}\| = 1$ (unit sphere) the max./min. length and direction of $A\underline{x}$.



unit sphere

mapped region

We note, to determine the largest/smallest length, we want to max/min $\|A\underline{x}\|$.

We may also max/min $\|A\underline{x}\|^2$. We take:

$$\|A\underline{x}\|^2 = (A\underline{x})^T A\underline{x} = \underline{x}^T(A^TA)\underline{x}$$

Where we note that $(A^TA)^T = A^T(A^T)^T = A^TA$ is symmetric, and so the problem is reduced to a constrained optimization, finding the max/min of $Q(\underline{x}) = \underline{x}^T(A^TA)\underline{x}$ subject to the constraint $\|\underline{x}\| = 1$.

From the theory provided in the constrained optimization notes, the max/min length and direction are respectively given by the max/min eigenvalue and eigenvector of A^TA.

SVD Procedure

We wish to decompose $A \in \mathbb{R}^{m \times n}$ which has rank r $\ni$

$$A = \mathcal{U}\Sigma V^T$$

where $\Sigma \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$.
matrices U and V are orthogonal.
The Σ matrix is structured $\ni$

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \leftarrow m-r \text{ rows}$$
$$\uparrow n-r \text{ columns}$$

where $D \in \mathbb{R}^{r\times r}$ is a diagonal matrix composed of the <u>singular values</u> of A

$$D = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix}$$

and the singular values are ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.
Additionally, the columns of U are referred to as the <u>left singular vectors</u> of A and the columns of V are referred to as the <u>right singular vectors</u> of A.

To find the SVD of matrix A, we

take the following steps:

1. Find an orthogonal diagonalization of A^TA.
2. Set up V and Σ
3. Construct $\mathcal{U}$

To illustrate this approach, we consider the following examples:

<u>Example</u>: Compute the SVD of

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

First we compute A^TA:

$$A^TA = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

The eigenvalues are $\lambda_1 = 360$, $\lambda_2 = 90$, $\lambda_3 = 0$, and corresponding eigenvectors are:

$$\underline{v}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \underline{v}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}, \quad \underline{v}_3 = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

Therefore the V matrix is:

$$V = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$$

And the Σ matrix is composed of the singular values of A, by taking the square of the eigenvalues of A^TA:

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = 0.$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} D & 0 \end{bmatrix} = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

We now construct the U matrix, by taking,

$$\underline{u}_1 = \frac{1}{\sigma_1} A\underline{v}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\underline{u}_2 = \frac{1}{\sigma_2} A \underline{v}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

Therefore,

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}$$

All together, we have,

$$\underbrace{\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}}_{A} = \underbrace{\frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}}_{V^T}$$

$$A = U \Sigma V^T$$

<u>Example</u>: Find the SVD of $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$

Note the rank of A is 1. Computing $A^T A$.

$$A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

the eigenvalues of $A^T A$ are then 18 and 0.
with corresponding eigenvectors:

$$\underline{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

which compose the columns of V.

$$V = [\underline{v}_1 \ \ \underline{v}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

The singular values of A are the square of the eigenvalues of $A^T A$.
For this A matrix, the singular values are $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = 0$.
Therefore,

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

To construct $\mathcal{U}$, compute $A\underline{v}_1$ and $A\underline{v}_2$.

$$A\underline{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ -4 \\ \ldots \end{bmatrix}, \quad A\underline{v}_2 = \begin{bmatrix} 0 \\ 0 \\ \ldots \end{bmatrix}$$

$\sqrt{2} \left[4 \right] \qquad \left[0 \right]$

As a check, you may verify that $\|A\underline{v}_1\| = \sigma_1$ and that $\|A\underline{v}_2\| = 0$.

Now the first column of U can be computed as:

$$\underline{u}_1 = \frac{1}{\sigma_1} A\underline{v}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

Since $\sigma_2 = 0$, we cannot find $\underline{u}_2$ and $\underline{u}_3$ in the same manner. We instead find them by extending $\underline{u}_1$ and forming an orthonormal basis for $\mathbb{R}^3$, that is to say, $\underline{u}_1^T\underline{u}_2 = 0$, $\underline{u}_1^T\underline{u}_3 = 0$, and $\underline{u}_2^T\underline{u}_3 = 0$.

First we find two vectors, $\underline{w}_2$ and $\underline{w}_3$, that are orthogonal to $\underline{u}_1$. The elements of these vectors must satisfy,

$x_1 - 2x_2 + 2x_3 = 0$. We have

$\left[2 \right. \qquad \left[-2 \right.$

$$\underline{w}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \underline{w}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Now ensuring that these vectors are orthogonal to one another and $\underline{u}_1$. Note, $\underline{u}_1 \cdot \underline{w}_2 = 0$ as well as $\underline{u}_1 \cdot \underline{w}_3 = 0$, by their construction. We need only to adjust $\underline{w}_3$ such that $\underline{w}_2 \cdot \underline{w}_3 = 0$.

$$\hat{\underline{w}}_3 = \underline{w}_3 - \frac{\underline{w}_2 \cdot \underline{w}_3}{\underline{w}_2 \cdot \underline{w}_2} \underline{w}_2 = \begin{bmatrix} -2/5 \\ 4/5 \\ 1 \end{bmatrix}$$

Normalizing $\underline{w}_2$ and $\hat{\underline{w}}_3$, we have,

$$\underline{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} \quad \text{and} \quad \underline{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

Finally, we have,

$$\begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3/\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Pseudo-Inverse of A

Let $r = \text{rank } A$, then partition U and V into submatrices whose first blocks contain r columns:

$$U = [U_r \quad U_{m-r}]$$
$$V = [V_r \quad V_{n-r}]$$

where, $U_r = [\underline{u}_1 \cdots \underline{u}_r]$ and $V_r = [\underline{v}_1 \cdots \underline{v}_r]$ so then $U_r \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{n \times r}$.

All together, the SVD for A looks like:

$$A = [U_r \quad U_{m-r}] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T$$

which is called the <u>reduced SVD</u> of A.

Note: Since the diagonal entries of D are non-zero, D is invertible. The following matrix is then called the pseudo-inverse of A

$$A^{+} = V_r D^{-1} U_r^{T}$$

Least Squares Solution

Given the equation $A\underline{x} = \underline{b}$, where for a Least-Squares problem, it is common for A to not be square.

We can use the pseudo-inverse of A to give the Least-Squares solution.

$$\hat{\underline{x}} = A^{+}\underline{b} = V_r D^{-1} U_r^{T} \underline{b}$$

So then, from the SVD,

$$\begin{aligned} A\hat{\underline{x}} &= (U_r D V_r^{T})(V_r D^{-1} U_r^{T} \underline{b}) \\ &= U_r D D^{-1} U_r^{T} \underline{b} \\ &= U_r U_r^{T} \underline{b} \end{aligned}$$

where $\hat{\underline{x}}$ is an approximation to $A\underline{x} = \underline{b}$

that minimizes the error $\|A\underline{\hat{x}} - \underline{b}\|$.