Understanding SVD

Geometric Perspective

$A \in \mathbb{C}^{m\times n}$ assuming $m \geq n$ (tall rectangle)

(right-mult. by V)

Write $\boxed{A = U\Sigma V^*}$ as $AV = U\Sigma$:

n under our assumption

$$A\mathbf{v}_k = \sigma_k \mathbf{u}_k, \quad k = 1, \ldots, r = \min\{m, n\}.$$

unitary

$U^* U = I_{m\times m}$, $V^* V = I_{n\times n}$

$$\begin{pmatrix} k^{th} \text{ column of} \\ AV \end{pmatrix} = \begin{pmatrix} k^{th} \text{ column of} \\ A\left[\vec{v}_1 \,\middle|\, \vec{v}_2 \,\middle|\, \cdots \,\middle|\, \vec{v}_k \,\middle|\, \cdots \,\middle|\, \vec{v}_n\right] \end{pmatrix} = \begin{pmatrix} k^{th} \text{ column of} \\ \left[A\vec{v}_1 \,\middle|\, A\vec{v}_2 \,\middle|\, \cdots \,\middle|\, A\vec{v}_k \,\middle|\, \cdots \,\middle|\, A\vec{v}_n\right] \end{pmatrix}$$

The image of the unit sphere under any $m \times n$ matrix is a hyperellipse.

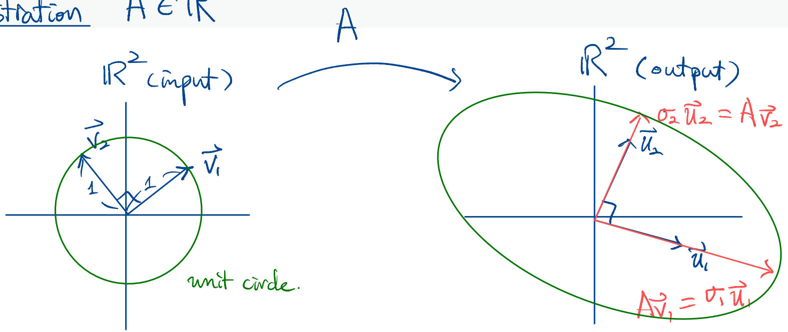
Illustration $A \in \mathbb{R}^{2\times 2}$

$$\begin{cases} A\vec{v}_1 = \sigma_1 \vec{u}_1 \\ A\vec{v}_2 = \sigma_2 \vec{u}_2 \\ \vdots \\ A\vec{v}_n = \sigma_n \vec{u}_n \end{cases}$$

left-singular vectors.

right-singular vectors.

orthonormal



$\vec{v}_1, \vec{v}_2$ ONB

$\vec{u}_1, \vec{u}_2$ ONB

$\sigma_1 \geq \sigma_2 \geq 0$

2-D : [unit circle] [an ellipse]

Generally : [unit sphere] [an hyperellipse]

Algebraic Perspective

— "change of basis" or "coordinate transformation"

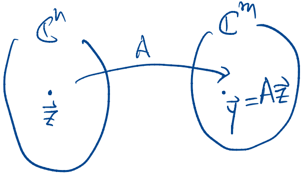
— of interpretation of EVD

Alternately, note that $\mathbf{y} = A\mathbf{z} \in \mathbb{C}^m$ for any $\mathbf{z} \in \mathbb{C}^n$ can be written as

$A : \mathbb{C}^n \longrightarrow \mathbb{C}^m$

(linear transformation)

$$(U^*\mathbf{y}) = \Sigma\,(V^*\mathbf{z})\,.$$



$\vec{y} = A\vec{z}$

$\vec{y} = U\Sigma V^*\vec{z}$ (SVD)

$\downarrow$ left-mult. by U^*

$U^*\vec{y} = \underbrace{U^*U}_{= I}\Sigma V^*\vec{z}$ ✓

$U^*\vec{y} = \Sigma V^*\vec{z}$

$\underbrace{U^{-1}\vec{y}}_{\text{coord. of } \vec{y} \text{ w.r.t. } U\text{-basis (ONB)}} = \Sigma \underbrace{V^{-1}\vec{z}}_{\text{coord. of } \vec{z} \text{ w.r.t. } V\text{-basis (ONB)}}$ ($\because U, V$ are unitary)

> Any matrix $A \in \mathbb{C}^{m\times n}$ can be viewed as a diagonal transformation from $\mathbb{C}^n$ (source space) to $\mathbb{C}^m$ (target space) with respect to suitably chosen orthonormal bases for both spaces.

left & right singular vectors of A

SVD vs. EVD

EVD exists for square matrices (diagonalizable)

Recall that a diagonalizable $A = VDV^{-1} \in \mathbb{C}^{n\times n}$ satisfies

$$\mathbf{y} = A\mathbf{z} \quad \longrightarrow \quad \left(V^{-1}\mathbf{y}\right) = D\left(V^{-1}\mathbf{z}\right).$$

(independent rescaling of coordinates)

This allowed us to view any diagonalizable square matrix $A \in \mathbb{C}^{n\times n}$ as a diagonal transformation from $\mathbb{C}^n$ to itself[1] with respect to the basis formed by a set of eigenvector of A.

"eigenbasis"

Differences.

output input

- **Basis:** SVD uses two ONBs (left and right singular vectors); EVD uses one, usually non-orthogonal basis (eigenvectors).
- **Universality:** all matrices have an SVD; not all matrices have an EVD.
- **Utility:** SVD is useful in problems involving the behavior of A or A^+; EVD is relevant to problems involving A^k.

pseudoinverse.

[1]The source and the target spaces of the transformation coincide.

Properties of SVD

SVD and the 2-Norm

#5 of HW 11.

$\|A\|_2 = \max_{\|\vec{x}\|_2=1} \|A\vec{x}\|_2$

$= \max \sqrt{\lambda(A^*A)}$

set of eigenvalues.

Theorem 1

Let $A \in \mathbb{C}^{m\times n}$ have an SVD $A = U\Sigma V^*$. Then

1. $\|A\|_2 = \sigma_1$ *and* $\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}$ σ_n ($r = \min\{m,n\}$)
2. *The rank of A is the number of nonzero singular values.*
3. *Let $r = \min\{m,n\}$. Then*

pseudoinverse.

Cond. number based on 2-norm.

$$\kappa_2(A) = \|A\|_2\|A^+\|_2 = \frac{\sigma_1}{\sigma_r}.$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_n \\ \hline & 0 & \end{bmatrix} \in \mathbb{R}^{m\times n}, \quad m \geq n$$

$$\underbrace{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n}_{\text{singular values of } A.} \geq 0$$

Connection to EVD $\longrightarrow$ this gives us a way to find singular values/vectors.

Let $A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ and $B = A^*A$. Observe that

- $B \in \mathbb{C}^{n\times n}$ is a *hermitian matrix*[1], *i.e.*, $B^* = B$.

$$B^* = (A^*A)^* = A^*(A^*)^* = A^*A = B$$

- B has an EVD:

$$B = A^*A$$
$$= (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^* U^* U \Sigma V^*$$

$= \Sigma^T$ since Σ is real

$\hookrightarrow = I$

$$= V\Sigma^T\Sigma V^{-1} = VDV^{-1}$$

Note

$$\Sigma^T\Sigma = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

$\rightarrow$ • The squares of singular values of A are eigenvalues of B.

$\rightarrow$ • An EVD of $B = A^*A$ reveals the singular values and a set of right singular vectors of A.

- By these, we know how to find singular values and right singular vectors of A.
- How about left singular vectors? Similar argument w/ AA^*.

[1]This is the $\mathbb{C}$-extension of real symmetric matrices.

Connection to EVD (cont')

Theorem 2

The nonzero singular values of $A \in \mathbb{C}^{m\times n}$ are the square roots of the nonzero eigenvalues of A^*A or AA^*.

Proof in the previous page.