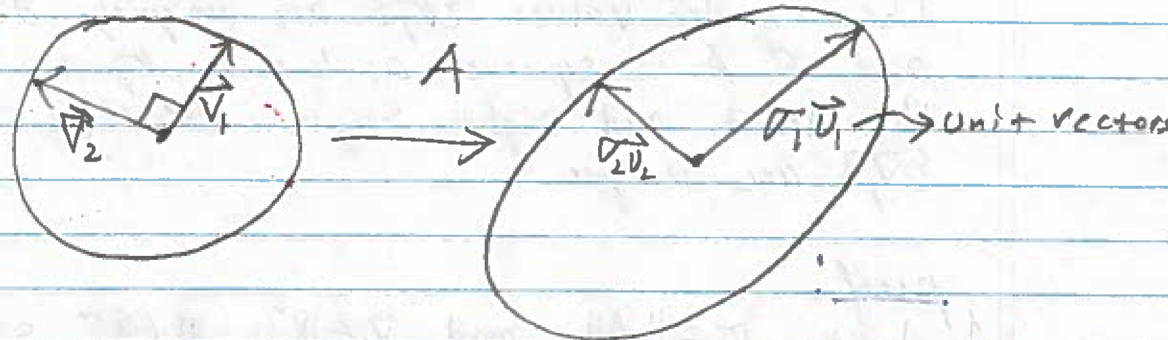
Lecture 4: Singular Value Decomposition



$\sigma_i A^{-1}\vec{u}_i = \vec{v}_i$ (right singular vector)

(left singular vector) ↑ $\vec{u}_i$

σ_i = singular values of A.

$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.

$$\Rightarrow A\vec{v}_j = \sigma_j \vec{u}_j$$

$$\Rightarrow A\begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix}\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix}$$

$$AV = \hat{U}\hat{\Sigma}$$

↑ Unitary!!

$$\Rightarrow A = \hat{U}\hat{\Sigma}V^*$$

This factorization is called the reduced SVD. Full SVD formed by completing $\hat{U}$ into a unitary matrix U and padding Σ with zeroes.

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

rotation ↗   stretch ↑   ↖ rotation

Theorem: Every matrix $A \in \mathbb{R}^{m \times n}$ has an SVD.
The singular values $\{\sigma_j\}$ are uniquely determined and if A is square and the σ_j are distinct, the left and right singular vectors $\{\vec{u}_j\}$ and $\{\vec{v}_j\}$ are unique.

proof:

1. Let $\sigma_1 = \|A\|_2$ and $\vec{v}_1 \in \mathbb{R}^n$, $\vec{u}_1 \in \mathbb{R}^m$ such that $\|\vec{v}_1\|_2 = \|\vec{u}_1\|_2 = 1$ and
$$A\vec{v}_1 = \sigma_1 \vec{u}_1.$$

2. Extend $\vec{v}_1$ to an orthonormal basis $\{\vec{v}_j\}$ of $\mathbb{R}^n$ and $\vec{u}_1$ to an orthonormal basis $\{\vec{u}_j\}$. Let

$$U_1 = \left[u_1 \middle| \cdots \middle| u_m \right], \quad V_1 = \left[v_1 \middle| \cdots \middle| v_n \right]$$

$$\Rightarrow U_1^* A V_1 = \begin{bmatrix} \sigma_1 & \vec{w}^* \\ 0 & B \end{bmatrix} = S$$

$$\left\| \begin{bmatrix} \sigma_1 & \vec{w}^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \vec{w} \end{bmatrix} \right\|_2 \geq \sigma_1^2 + \vec{w}^* \vec{w} = (\sigma_1^2 + \vec{w}^* \vec{w})^{1/2} \cdot \left\| \begin{bmatrix} \sigma_1 \\ \vec{w} \end{bmatrix} \right\|_2$$

$$\Rightarrow \|S\|_2 \geq (\sigma_1^2 + \vec{w}^* \vec{w})^{1/2}$$

However since U_1, V_1 are unitary it follows that it follows that $\|S\|_2 = \|A\| = \sigma_1$. Therefore,
$$\vec{w} = 0.$$

3. Induct on rows of matrix to get
$$B = U_2 \Sigma_2 V_2^*,$$
which proves existence.

4. See book for uniqueness.

Low Rank Approximations

Theorem - A is the sum of rank one matrices

$$A = \sum_{j=1}^{r} \sigma_j \vec{u}_j \vec{v}_j^*$$

~~proof~~:

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

$$= U \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} \vec{v}_1^* & \cdots & 0 \end{bmatrix} + \dots + U \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} 0 & \cdots & \vec{v}_n^* \end{bmatrix}$$

$$= U \begin{bmatrix} \sigma_1 \vec{v}_1^* & \cdots & 0 \end{bmatrix} + \dots + U \begin{bmatrix} 0 & \cdots & \sigma_n \vec{v}_n^* \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^* + \dots + \sigma_n \vec{u}_n \vec{v}_n.$$

Theorem - For any ν with $0 \le \nu \le r$, define

$$A_\nu = \sum_{j=1}^{\nu} \sigma_j u_j v_j^*,$$

then if $\nu < \min\{m,n\}$.

$$\|A - A_\nu\|_2 = \min_{\substack{B \in \mathbb{R}^{m \times n} \\ \operatorname{rank}(B) \le \nu}} \|A - B\|_2 = \sigma_{\nu+1}$$