More on the SVD.

$[A]_{m\times n}$

Defn: Frobenius norm: matrix norm defined by –

$$\|A\|_F = \left(\sum_j^n \|a_j\|_2^2\right)^{1/2}. \qquad a_j = j^{th} \text{ row of } A.$$

$$= \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{1/2}.$$

$$= \sqrt{tr(A^*A)} \qquad \text{(exercise)}$$

$$= \sqrt{tr(AA^*)}$$

Exercise: For any $A \in \mathbb{C}^{m\times n}$ & unitary $Q \in \mathbb{C}^{m\times m}$,

$$\|QA\|_2 = \|A\|_2 \quad \& \quad \|QA\|_F = \|A\|_F.$$

2-norm & Frobenius norm are invariant under multi. by unitary matrices.

Let $\underset{(m\times n)}{A} = \underset{m\times r}{U}\,\underset{r\times r}{\Sigma}\,\underset{r\times n}{V^*}$ be the SVD of A.

Result 1: The rank of A = # non-zero singular values of A. $= r$

Proof: $\text{rank}(A) = \text{rank}(U\Sigma V^*)$

$= \text{rank}(\Sigma)$

$= r.$

If B is full rank, then $\text{rank}(AB) = \text{rank}(BA) = \text{rank } A.$

If $A = U \Sigma V^*$ & $\operatorname{rank} A = r$, then -

Result 2: $\operatorname{range}(A) = \langle u_1, \dots, u_r \rangle$, $\operatorname{null}(A) = \langle v_{r+1}, \dots, v_n \rangle$

Exercise.

Result 3: $\|A\|_2 = \sigma_1$ & $\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$

Pf: $\|A\|_2 = \sigma_1$ already proven.

$$\|A\|_F = \|U \Sigma V^*\|_F = \|\Sigma\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2} \quad \square.$$

Result 4: If $A = A^*$, then the singular values of A are the absolute values of the eigenvalues of A.

(Exercise).

Result 5: for $\underset{A = U\Sigma V^*}{A \in \mathbb{C}^{m \times m}}$, $|\det A| = \prod_{i=1}^{m} \sigma_i$.

Theorem (Existence & uniqueness of SVD)

Every matrix $A \in \mathbb{C}^{m \times n}$ has a singular value decomposition
$$A = U \Sigma V^*.$$
The singular values $\{\sigma_j\}$ are uniquely determined.
If A is square & the σ_j are distinct, then $\{u_j\}$ & $\{v_j\}$ are uniquely determined upto complex scalar factors of absolute value 1.

$U \quad U'$

$u_j \quad z u_j'$

$|z| = 1.$

Proof: Consider the function $v \mapsto \|Av\|_2$ on the unit sphere in $\mathbb{C}^n$.

This is a continuous function on a compact set, hence it attains its bounds.

Upper bound on $\{\|Av\|_2 \,/\, v \in \underset{\text{i.e. } \|v\|_2 = 1}{\text{unit sphere}}\} = \|A\|_2$

$\therefore \exists$ a vector $v_1 \in \mathbb{C}^n$ with $\|v_1\|_2 = 1$ s.t. $\|A v_1\|_2 = \|A\|_2 = \sigma_1$ ✓

Thus we found a vector v_1 for which $\|Av_1\|_2 = \sigma_1$.

Let $u_1 = \dfrac{Av_1}{\|Av_1\|} = \dfrac{Av_1}{\|Av_1\|} = \dfrac{Av_1}{\sigma_1}$ $(\because \|u_1\| = 1)$

$A \in \mathbb{C}^{m\times n}$

$\therefore \boxed{Av_1 = \sigma_1 u_1}$ ☆ $v_1 \in \mathbb{C}^n$, $u_1 \in \mathbb{C}^m$.

Let $\{v_1, \dots, v_n\} \in \mathbb{C}^n$ & $\{u_1, \dots, u_m\} \in \mathbb{C}^m$ be an extension of v_1 & u_1 to orthonormal bases of $\mathbb{C}^n$ & $\mathbb{C}^m$ resp.

Let $\mathcal{U}_1 = [u_1 | \cdots | u_m]$ & $V_1 = [v_1 | \cdots | v_n]$

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

Let $S = \mathcal{U}_1^* A V_1 = \begin{bmatrix} u_1^* \\ \vdots \\ u_m^* \end{bmatrix} A \, [v_1 | \cdots | v_n]$

$$= \left[\begin{array}{c|c} \sigma_1 & w^* \\ \hline 0 & B \end{array}\right] \quad \begin{bmatrix} \sigma_1 \\ w \end{bmatrix}$$

Claim: $w = 0$.

$$\left\| S \cdot \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} \sigma_1^2 + w^* w \\ Bw \end{bmatrix} \right\|_2 \geq \sqrt{(\sigma_1 + w^* w)^2} = \left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2$$

$$\therefore \|S\|_2 \geq \left\| S \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2 \geq \left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2$$

$\|S\|_2 \geq \|Sv\|_2$

But $\|S\|_2 = \|A\|_2 = \sigma_1$

$\therefore w = 0$.

So $S = \mathcal{U}_1^* A V_1 = \left[\begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & B \end{array}\right]$

By induction hypothesis, B has an SVD $B = \mathcal{U}_2 \Sigma_2 V_2^*$

Consider $\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathcal{U}_2 \end{array}\right] \left[\begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & \Sigma_2 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V_2^* \end{array}\right] = \left[\begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & B \end{array}\right]$

$= S = \mathcal{U}_1^* A V_1$

$$\therefore A = \mathcal{U}_1 S V_1^* = \underbrace{\mathcal{U}_1 \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathcal{U}_2\end{array}\right]}_{\mathcal{U}} \underbrace{\left[\begin{array}{c|c}\sigma_1 & 0 \\ \hline 0 & \Sigma_2\end{array}\right]}_{\Sigma} \underbrace{\left[\begin{array}{c|c} 1 & 0\\ \hline 0 & V_2^*\end{array}\right] V_1^*}_{V^*}$$

This gives the SVD of A.

(Uniqueness - exercise)

Writing A as a sum of rank-1 matrices using SVD:

If $A = \mathcal{U} \Sigma V^*$

then $$A = \underbrace{u_1 \sigma_1 v_1^*} + \underbrace{u_2 \sigma_2 v_2^*} + \cdots + \underbrace{u_r \sigma_r v_r^*}$$

rank-1 matrices.

$$= \sum_{i=1}^{r} u_i \sigma_i v_i^*$$

For $1 \le k \le r$, let $A_k = \sum_{i=1}^{k} u_i \sigma_i v_i^*$. $(\text{rank } A_k = k)$

<u>Theorem</u>: (Eckhart - Young theorem) A_k is the "best" rank k approximation to A. i.e. if B is any other rank k matrix in $\mathbb{C}^{m \times n}$, then

$$\| A - A_k \|_2 \le \| A - B \|_2$$

More precisely, let rank $A = r$, then for any $0 \le k \le r$,

define $A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^*$;

(if $p = \min\{m,n\}$, define $\sigma_{p+1} = 0$), then

$$\| A - A_k \|_2 = \inf_{\substack{B \in \mathbb{C}^{m \times n} \\ \text{rank} B \le k}} \| A - B \|_2 = \sigma_{k+1}.$$

$$\boxed{\underset{m \times n}{A} = \mathcal{U} \Sigma V^* = \underset{m \times n}{\left[\begin{array}{c|c|c} u_1 & \cdots & u_n \end{array}\right]} \underset{n \times n}{\begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}} \left[\begin{array}{c} v_1^* \\ \hline \vdots \\ \hline v_n^* \end{array}\right]} = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots + \sigma_r u_r v_r^* + \underbrace{\cdots}_{\text{zero matrices}}$$

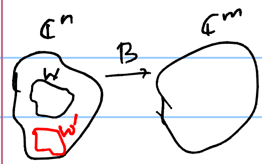
all are rank 1 matrices.

$0 \le k \le r, \qquad A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^*$

$$A - A_k = \sum_{j=1}^{r} \sigma_j u_j v_j^* - \sum_{j=1}^{k} \sigma_j u_j v_j^* = \sum_{j=k+1}^{r} \sigma_j u_j v_j^*$$

So $\|A - A_k\| = \sigma_{k+1}$.

Proof: Suppose there is some $B \in \mathbb{C}^{m \times n}$ with rank $B \le k$ such that $\|A - B\|_2 < \|A - A_k\|_2 = \sigma_{k+1}$



rank $B \le k \Rightarrow$ nullity $B \ge n-k$

$\therefore$ There is an $(n-k)$-dim'l subspace W of $\mathbb{C}^n$ s.t. $Bw = 0 \ \forall w \in W$.

So for any $w \in W$, $\|Aw\|_2 = \|(A-B)w\|_2 \le \|A-B\|_2 \cdot \|w\|_2 < \sigma_{k+1} \cdot \|w\|_2$.

But the subspace W' of $\mathbb{C}^n$ spanned by the first $(k+1)$ columns of V (i.e. first $k+1$ right singular vectors) is a $(k+1)$-dim'l subspace whose elements w' satisfy

$$\|Aw'\|_2 \ge \sigma_{k+1} \|w'\|_2$$

i.e. $\frac{\|Aw'\|_2}{\|w'\|_2} \ge \sigma_{k+1} = \|A - A_k\|_2$ by defn. of 2 norm

So $W \cap W'$ must be $\{0\}$, but $\dim W + \dim W' = (n-k) + (k+1)$ which exceeds n, a contradiction. $\square$