Theorem 1 *Let $A \in \mathbb{C}^{n\times m}$ be a non zero matrix of rank r. Then A can be expressed as a product*

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

where $U \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{m\times m}$ are unitary matrices and $\Sigma \in \mathbb{R}^{n\times m}$ is a nonsquare diagonal matrix such that

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{bmatrix}, \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Proof: The proof is by induction on the rank of A. If $\mathrm{rank}(A) = 1$, then every column of A is a multiple of some $x \in \mathbb{C}^n \setminus \{0\}$ and there exist $y_1, y_2, \ldots, y_m \in \mathbb{C}$ such that the i^{th} column of A is $(\bar{y}_i)x$. Thus $A = xy^*$ where $y = [y_1 \cdots y_m]^T \in \mathbb{C}^m$. Setting $u = x/\|x\|_2, v = y/\|y\|_2$ and $\sigma = \|x\|_2\|y\|_2$,

$$A = \sigma uv^* = [u\, u_2 \cdots u_n] \begin{bmatrix} \sigma & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} [v\, v_2 \cdots v_m]^* \tag{1}$$

where $\{u_2, \ldots, u_n\} \subset \mathbb{C}^n$, and $\{v_2, \ldots, v_m\} \subset \mathbb{C}^m$ are orthonormal sets such that $\{u, u_2, \ldots, u_n\}$ and $\{v, v_2, \ldots, v_m\}$ are orthonormal bases of $\mathbb{C}^n$ and $\mathbb{C}^m$ respectively. The second equality in (1) shows that the theorem holds for rank 1 matrices.

Now let A be of rank $r > 1$ and assume that all matrices of rank $r-1$ have a decomposition as specified by the theorem. Since $\|A\|_2 = \max\limits_{\|x\|_2=1} \|Ax\|_2$, there exists $v \in \mathbb{C}^m$ with $\|v\|_2 = 1$, such that $Av = w$ and $\|w\|_2 = \|A\|_2$. Thus if $\sigma_1 := \|A\|_2$ and $u := w/\sigma_1$, then $\|u\|_2 = 1$ and $Av = \sigma_1 u$. The sets $\{v\}$ and $\{u\}$ may be extended to form orthonormal bases say, $\{v, v_1, \ldots, v_{m-1}\}$ and $\{u, u_1, \ldots, u_{n-1}\}$ of $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively. Thus if $\hat{V} := [v_1\ v_2\ \ldots\ v_{m-1}]$ and $\hat{U} := [u_1\ u_2\ \ldots\ u_{n-1}]$, we have unitary matrices $V_1 := [v\ \hat{V}]$ and $U_1 = [u\ \hat{U}]$ such that

$$\begin{aligned} U_1^* A V_1 &= \begin{bmatrix} u^* \\ \hat{U}^* \end{bmatrix} A[v\ \hat{V}] \\ &= \begin{bmatrix} u^*Av & u^*A\hat{V} \\ \\ \hat{U}^*Av & \hat{U}^*A\hat{V} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 & w^* \\ \\ \sigma_1\hat{U}^*u & \hat{U}^*A\hat{V} \end{bmatrix} \quad \text{where } w = \hat{V}^*A^*u \\ &= \begin{bmatrix} \sigma_1 & w^* \\ \\ 0 & B \end{bmatrix} \quad \text{(by orthonormality of}\{u, u_1, \ldots, u_{n-1}\}) \end{aligned}$$

where 0 is a column vector of dimension $n-1$, w^* is a row vector of dimension $m-1$ and $B = \hat{U}^*A\hat{V}$ has dimension $n-1 \times m-1$. The rest of the proof consists of establishing that

$w = 0$ so that the induction hypothesis may be invoked on B. Setting $S := \begin{bmatrix} \sigma_1 & w^* \\ 0 & B \end{bmatrix}$, observe that,

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

This implies that

$$\|S\|_2 \geq \sqrt{\sigma_1^2 + w^*w}. \tag{2}$$

But since U_1 and V_1 are unitary matrices,

$$\|S\|_2 = \|U_1^* A V_1\|_2 = \|A\|_2 = \sigma_1. \tag{3}$$

From (2) and (3) we have $w = 0$ so that $S = \begin{bmatrix} \sigma_1 & 0 \\ 0 & B \end{bmatrix}$. Now S has the same rank as A which is r and the first column $\begin{bmatrix} \sigma_1 \\ 0 \end{bmatrix}$ of S is evidently orthogonal to the remaining columns. Therefore it is linearly independent of the remaining columns of S. Therefore exactly $r - 1$ of the remaining columns of S which form the matrix $\begin{bmatrix} 0 \\ B \end{bmatrix}$ are linearly independent. This implies that $\text{rank}(B) = r - 1$. In view of the induction hypothesis, there exist unitary matrices $U_2 \in \mathbb{C}^{n-1 \times n-1}$ and $V \in \mathbb{C}^{m-1 \times m-1}$ and a diagonal matrix

$$\Sigma_2 = \begin{bmatrix} \sigma_2 & & & & & \\ & \sigma_3 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{bmatrix} \in \mathbb{R}^{n-1 \times m-1}$$

where $\sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_{r-1} > 0$ such that $B = U_2 \Sigma_2 V_2^*$ is an SVD of B. Using this fact, we have

$$A = U_1 \begin{bmatrix} \sigma_1 & 0 \\ 0 & U_2^* \Sigma_2 V_2 \end{bmatrix} V_1^* = U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix}^* V_1^*.$$

Thus if $U := U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}$, $V := V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$, then U and V are unitary matrices and Σ is a nonsquare diagonal matrix such that

$$\|\Sigma\|_2 = \|A\|_2 \Rightarrow \max\{\sigma_1, \|\Sigma_2\|_2\} = \sigma_1 \Rightarrow \|\Sigma_2\|_2 \leq \sigma_1.$$

This implies that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and establishes that A satisfies the statement of the theorem. Hence the proof follows by induction. □