In the next proof, we use the *Cauchy-Schwartz inequality*: for vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, $|\boldsymbol{a}\cdot\boldsymbol{b}| \leq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$. The proof is as follows: Write $\boldsymbol{b} = \boldsymbol{b}^{||\boldsymbol{a}} + \boldsymbol{b}^{\perp a}$. By the Pythagorean Theorem, $\|\boldsymbol{b}\|^2 = \|\boldsymbol{b}^{||\boldsymbol{a}}\|^2 + \|\boldsymbol{b}^{\perp a}\|^2$, so $\|\boldsymbol{b}\|^2 \geq \|\boldsymbol{b}^{||\boldsymbol{a}}\|^2 = \|\frac{\boldsymbol{b}\cdot\boldsymbol{a}}{\boldsymbol{a}\cdot\boldsymbol{a}}\boldsymbol{a}\|^2 = \left(\frac{\boldsymbol{b}\cdot\boldsymbol{a}}{\|\boldsymbol{a}\|^2}\right)^2\|\boldsymbol{a}\|^2 = \frac{(\boldsymbol{b}\cdot\boldsymbol{a})^2}{\|\boldsymbol{a}\|^2}$, so $\|\boldsymbol{b}\|^2\|\boldsymbol{a}\|^2 \geq (\boldsymbol{b}\cdot\boldsymbol{a})^2$, which proves the inequality.

Property S3 of the singular value decomposition states that the matrix U of left singular vectors is column-orthogonal. We now prove that property.

The left singular vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ have norm one by construction. We need to show that they are mutually orthogonal. We prove by induction on k that, for $i = 1, 2, \ldots, k$, the vector $\boldsymbol{u}_i$ is orthogonal to $\boldsymbol{u}_{i+1}, \ldots, \boldsymbol{u}_r$. Setting $k = r$ proves the desired result.

By definition of the singular vectors and values,

$$AV = \left[\begin{array}{c|c|c|c|c|c|c} \\ \sigma_1\boldsymbol{u}_1 & \cdots & \sigma_{k-1}\boldsymbol{u}_{k-1} & \sigma_k\boldsymbol{u}_k & \sigma_{k+1}\boldsymbol{u}_{k+1} & \cdots & \sigma_r\boldsymbol{u}_r \\ \\ \end{array}\right]$$

By the inductive hypothesis, $\boldsymbol{u}_k$ is orthogonal to $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}$. Since $\boldsymbol{u}_k$ has norm one, $\boldsymbol{u}_k \cdot \sigma_k\boldsymbol{u}_k = \sigma_k$. Let

$$\begin{aligned} \beta_{k+1} &= \boldsymbol{u}_k \cdot \boldsymbol{u}_{k+1} \\ \beta_{k+2} &= \boldsymbol{u}_k \cdot \boldsymbol{u}_{k+2} \\ &\vdots \\ \beta_r &= \boldsymbol{u}_k \cdot \boldsymbol{u}_r \end{aligned}$$

Then

$$\boldsymbol{u}_k^T AV = \left[\begin{array}{ccccccc} 0 & \cdots & 0 & \sigma_k & \beta_{k+1} & \cdots & \beta_r \end{array}\right] \tag{11.1}$$

Our goal is to show that $\beta_{k+1}, \ldots, \beta_r$ are all zero, for this would show that $\boldsymbol{u}_k$ is orthogonal to $\boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_r$.

Let $\boldsymbol{w} = \left[\begin{array}{ccccccc} 0 & \cdots & 0 & \sigma_k & \beta_{k+1} & \cdots & \beta_r \end{array}\right]$. Then $\|\boldsymbol{w}\|^2 = \sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2$. Since V is column-orthogonal, $\|V\boldsymbol{w}\|^2 = \|\boldsymbol{w}\|^2$, so

$$\|V\boldsymbol{w}\|^2 = \sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2 \tag{11.2}$$

Furthermore, since the first $k-1$ entries of $\boldsymbol{w}$ are zero, the vector $V\boldsymbol{w}$ is a linear combination of the remaining $r-(k-1)$ columns of V. Since the columns of V are mutually orthogonal, $V\boldsymbol{w}$ is orthogonal to $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$. Let $\boldsymbol{v} = V\boldsymbol{w}/\|V\boldsymbol{w}\|$. Then $\boldsymbol{v}$ has norm one and is orthogonal to $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$. We will show that if $\beta_{k+1}, \ldots, \beta_r$ are not all zero then $\|A\boldsymbol{v}\| > \|A\boldsymbol{v}_k\|$, contradicting the choice of $\boldsymbol{v}_k$.

By Equation 11.1, $(\boldsymbol{u}_k^T AV) \cdot \boldsymbol{w} = \sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2$. By the Cauchy-Schwartz Inequality, $|\boldsymbol{u}_k \cdot (AV\boldsymbol{w})| \leq \|\boldsymbol{u}_k\|\,\|AV\boldsymbol{w}\|$, so, since $\|\boldsymbol{u}_k\| = 1$, we infer $\|AV\boldsymbol{w}\| \geq \sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2$. Combining this inequality with Equation 11.2, we obtain

$$\frac{\|AV\boldsymbol{w}\|}{\|V\boldsymbol{w}\|} \geq \frac{\sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2}{\sqrt{\sigma_k^2 + \beta_{k+1}^2 + \cdots + \beta_r^2}}$$

which is greater than σ_k^2 if $\beta_{k+1}, \ldots, \beta_r$ are not all zero. This completes the induction step, and the proof.