Rendel Deguia

8.2.5 Suppose $A = UDV^T$ where $A \in \mathbb{R}^{m \times n}$, $r = \text{rank}(A)$, and:

- $V = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n] \in \mathbb{R}^{n \times n} \ni \{\vec{v}_i\}_{i=1}^{n}$ are eigenvectors of A^TA

- $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0) \in \mathbb{R}^{m \times n}$ such that $\{\sigma_i\}_{i=1}^{r}$ are the singular values of A

- $U = [\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m] \in \mathbb{R}^{m \times m}$ where $\vec{u}_j = \begin{cases} \dfrac{A\vec{v}_j}{\|A\vec{v}_j\|}, & 1 \le j \le r \\ 0, & r < j \le m \end{cases}$

Then, assuming $r \le \min(m, n)$, and defining

$U_r = [\vec{u}_1, \dots, \vec{u}_r]$, $V_r = [\vec{v}_1, \dots, \vec{v}_r]$ and $D_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$

where $U_r \in \mathbb{R}^{m \times r}$, $V_r \in \mathbb{R}^{n \times r}$, it follows $A = U_r D_r V_r^T$.

$\boxed{\text{Proof}}$

$$A = UDV^T \Rightarrow A = [\vec{u}_1, \dots, \vec{u}_m] \underbrace{\left.\begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & \cdots & & 0 \\ 0 & \cdots & & 0 \end{bmatrix}\right\}m}_{n} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}$$

$$\Rightarrow A = \overbrace{[\sigma_1\vec{u}_1, \dots, \sigma_r\vec{u}_r, 0, \dots, 0]} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} \Rightarrow A = \sigma_1\vec{u}_1\vec{v}_1^T + \dots + \sigma_r\vec{u}_r\vec{v}_r^T + 0 + \dots + 0$$

$$\Rightarrow A = \sum_{k=1}^{r} \sigma_k \vec{u}_k \vec{v}_k^T \quad (1)$$

Now, let $B = U_r D_r V_r^T \Rightarrow B = [\vec{u}_1, \dots, \vec{u}_r] \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix}$

$$\Rightarrow B = [\sigma_1\vec{u}_1, \dots, \sigma_r\vec{u}_r] \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix} = \sigma_1\vec{u}_1\vec{v}_1^T + \dots + \sigma_r\vec{u}_r\vec{v}_r^T \Rightarrow B = \sum_{k=1}^{r} \sigma_k \vec{u}_k \vec{v}_k^T \quad (2)$$

By equations (1) and (2), we see that $A = B$; thus $A = U_r D_r V_r^T$ $\boxed{\text{QED}}$

8.2.16

In order to evaluate $||B||_2$, we must find the eigenvalues of B^TB; the largest of these eigenvalues, λ_{max}, corresponds to the l_2 norm of B where $||B||_2 = \sqrt{\lambda_{max}}$

Likewise, to evaluate $||UB||_2$, we must find the eigenvalues of $(UB)^T(UB)$. However, note that:

$(UB)^T(UB) = B^T(U^TU)B$ and since $U^TU = I$,

$B^T(U^TU)B = B^TIB = B^TB$. Thus, finding the eigenvalues of $(UB)^T(UB)$ is the same as finding the eigenvalues of B^TB, so we must also have $||UB||_2 = \sqrt{\lambda_{max}}$ and hence,

$$||UB||_2 = ||B||_2$$