1a)

$$A = \frac{1}{3}\left[q^{(1)} \; q^{(2)} \; q^{(3)} \right] \qquad \langle q^{(1)}, q^{(2)} \rangle = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} = -2-2+4$$

$$= 0$$

$$\langle q^{(1)}, q^{(3)} \rangle = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} = -2+4-2 \qquad \langle q^{(2)}, q^{(3)} \rangle = \begin{bmatrix} 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} = 4-2-2$$

$$= 0 \qquad\qquad = 0$$

$$\|q^{(i)}\|_2 = \frac{1}{3}\sqrt{2^2+2^2+(-1)^2} = \frac{\sqrt{9}}{3} = 1 \qquad \therefore \text{Colums of } A \text{ are orthonormal}$$

It is evident that $A = A^T$, $\therefore$ the rows of A are also orthonormal

$\therefore$ A is an orthogonal matix.

Check:

$$A^TA = AA = AA^T = \frac{1}{9}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

1b) Solve V:

$$\det\begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix} = 0 = (1-\lambda)(1-\lambda)^2 \to \lambda = 1 \quad AM(\lambda) = 3$$

$$(\lambda I - A^TA) = \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad V_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad V_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = V^T \qquad \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \sigma_i = \sqrt{\lambda_i}$$

Find U:

$$u^{(i)} = \frac{Av^{(i)}}{\sigma_i} = \frac{Ae^{(i)}}{1} \qquad \therefore u = A$$

$\therefore$ SVD of A is $U = A$, $\Sigma = I$ $V = I$

$$A = \underbrace{\frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}}_{u} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{V} = A \checkmark$$

2a) $\mathcal{R}(A)=\{b\in\mathbb{R}^m : Ax=b,\ x\in\mathbb{R}^n\}$

$\mathcal{R}(AA^T)=\{b\in\mathbb{R}^m : AA^Tx=b,\ x\in\mathbb{R}^m\}$

$b\in\mathcal{R}(A)\to Ax=b \to \text{let } x=A^Ty \to A(A^Ty)=b\to b\in\mathcal{R}(AA^T)$

$b\in\mathcal{R}(AA^T)\to AA^Tx=b\to A(A^Tx)=b\to \text{let } y=A^Tx\to Ay=b\to b\in\mathcal{R}(A)$

Since any $b\in\mathcal{R}(A)$ is also in $\mathcal{R}(AA^T)$ and any $b\in\mathcal{R}(AA^T)$ is also in $\mathcal{R}(A)$,

$\mathcal{R}(A)=\mathcal{R}(AA^T)$. QED

b) Note AA^T and A^TA have the same eigen values.

$A=\tilde{U}\tilde{\Sigma}\tilde{V}^T$, where $\tilde{U}$ is formed by the eigenvectors of AA^T, $\tilde{V}$ is . . . . A^TA, — Have the Same eigenvalues

and $\tilde{\Sigma}$ is the square root of ordered non-zero eigenvalues of AA^T or A^TA

$$A=\begin{bmatrix} | & & | \\ u^{(1)} & \cdots & u^{(r)} \\ | & & | \end{bmatrix}\begin{bmatrix}\sigma_1 & & 0\\ & \ddots & \\ 0 & & \sigma_r\end{bmatrix}\begin{bmatrix} — & v^{(1)T} & — \\ & \vdots & \\ — & v^{(r)T} & — \end{bmatrix}$$

$\therefore r=$ # of non-zero eigen values in AA^T, Since Σ constructed as $\text{Sqrt}(\lambda_i)$

We have Shown $r=$ # non-zero eigen values in AA^T, but now must show that rank $A=r$

Spectral theorem garuntees that the eigenvectors of a symmetric matrix are orthogonal

↳ $\therefore$ the colums of U and V are individually orthogonal.

in lecture, it was shown that $\mathcal{R}(A)=\text{Span}\{u^{(1)},\dots,u^{(r)}\}$

rank$(A)=\dim \text{Col} A=\dim\mathcal{R}(A)=r$ by orthogonality of $u^{(1)},\dots,u^{(r)}$

$\therefore$ rank$(A)=r=$ # of non-zero eigenvalues in AA^T.

3a) $A^T = \frac{1}{\sqrt{10}}\begin{bmatrix}5 & 3\\0 & 4\end{bmatrix}$ $AA^T = \frac{1}{10}\begin{bmatrix}5 & 0\\3 & 4\end{bmatrix}\begin{bmatrix}5 & 3\\0 & 4\end{bmatrix} = \frac{1}{10}\begin{bmatrix}25 & 15\\15 & 25\end{bmatrix}$

$A^TA = \frac{1}{10}\begin{bmatrix}5 & 3\\0 & 4\end{bmatrix}\begin{bmatrix}5 & 0\\3 & 4\end{bmatrix} = \frac{1}{10}\begin{bmatrix}34 & 12\\12 & 16\end{bmatrix}$

Construct U:

$\det\begin{bmatrix}2.5-\lambda & 1.5\\1.5 & 2.5-\lambda\end{bmatrix} = 0 = (2.5-\lambda)^2 - 1.5^2 = 2.5^2 - 5\lambda + \lambda^2 - 1.5^2$

$= \lambda^2 - 5\lambda + 4 \quad \lambda_1 = 4,\ \lambda_2 = 1$

$\lambda_1 = 4 \begin{bmatrix}-1.5 & 1.5 & | & 0\\1.5 & -1.5 & | & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & | & 0\\1.5 & -1.5 & | & 0\end{bmatrix} \to V_1 = \begin{bmatrix}1\\1\end{bmatrix} \quad u^{(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$

$\lambda_2 = 1 \begin{bmatrix}1.5 & 1.5 & | & 0\\1.5 & 1.5 & | & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & | & 0\\1.5 & 1.5 & | & 0\end{bmatrix} \to V_2 = \begin{bmatrix}1\\-1\end{bmatrix} \quad u^{(2)} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \quad \therefore U = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix} \quad \Sigma = \begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}$

Construct V:

$\det\begin{bmatrix}3.4-\lambda & 1.2\\1.2 & 1.6-\lambda\end{bmatrix} = 0 = (3.4-\lambda)(1.6-\lambda) - 1.2^2 = 3.4\cdot1.6 - 5\lambda + \lambda^2 - 1.2^2$

$= \lambda^2 - 5\lambda + 4 \quad \lambda_1 = 4,\ \lambda_2 = 1$

$\lambda_1 = 4 \begin{bmatrix}-0.6 & 1.2 & | & 0\\1.2 & -2.4 & | & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & | & 0\\1 & -2 & | & 0\end{bmatrix} \to V_1 = \begin{bmatrix}2\\1\end{bmatrix} \quad V^{(1)} = \frac{1}{\sqrt{5}}\begin{bmatrix}2\\1\end{bmatrix}$

$\lambda_2 = 1 \begin{bmatrix}2.4 & 1.2 & | & 0\\1.2 & 0.6 & | & 0\end{bmatrix} \to \begin{bmatrix}0 & 0 & | & 0\\2 & 1 & | & 0\end{bmatrix} \to V_2 = \begin{bmatrix}1\\-2\end{bmatrix} \quad V^{(2)} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\-2\end{bmatrix} \quad \therefore V = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\1 & -2\end{bmatrix} \quad \Sigma = \begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}$

$\therefore A = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\1 & -2\end{bmatrix}^T = \frac{1}{\sqrt{10}}\begin{bmatrix}2 & 1\\2 & -1\end{bmatrix}\begin{bmatrix}2 & 1\\1 & -2\end{bmatrix} = \frac{1}{\sqrt{10}}\begin{bmatrix}5 & 0\\3 & 4\end{bmatrix}$ ✓

b) $A = \frac{1}{\sqrt{10}}\left[2\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}2 & 1\end{bmatrix} + 1\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1 & -2\end{bmatrix}\right]$

$2\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}2 & 1\end{bmatrix} = 2\begin{bmatrix}2 & 1\\2 & 1\end{bmatrix} = \begin{bmatrix}4 & 2\\4 & 2\end{bmatrix}$

$1\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1 & -2\end{bmatrix} = \begin{bmatrix}1 & -2\\-1 & +2\end{bmatrix} \quad \therefore A = \frac{1}{\sqrt{10}}\left[\begin{bmatrix}4 & 2\\4 & 2\end{bmatrix} + \begin{bmatrix}1 & -2\\-1 & +2\end{bmatrix}\right] = \frac{1}{\sqrt{10}}\begin{bmatrix}5 & 0\\3 & 4\end{bmatrix}$ ✓

$$\begin{array}{c} \\ c)\ A = \end{array} \begin{array}{c} U \\ \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \end{array} \begin{array}{c} \Sigma \\ \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} V^T \\ \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}^T \end{array} \qquad \bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 2x_2 \end{bmatrix}$$

$$Ax = U\Sigma V^T x = U\Sigma \bar{x} = U\begin{bmatrix} 2/\sqrt{10}\,\bar{x}_1 & 0 \\ 0 & \bar{x}_2/\sqrt{10} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\frac{2}{\sqrt{10}}(2x_1+x_2) + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\frac{1}{\sqrt{10}}(x_1-2x_2)$$

$$Ax = \frac{1}{\sqrt{10}}\begin{bmatrix} 4x_1 + 2x_2 + x_1 - 2x_2 \\ 4x_1 + 2x_2 - x_1 + 2x_2 \end{bmatrix} = \frac{1}{\sqrt{10}}\begin{bmatrix} 5x_1 \\ 3x_1 + 4x_2 \end{bmatrix}$$

$$Av^{(1)} = A\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \frac{1}{\sqrt{5}}A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{50}}\begin{bmatrix} 5\cdot 2 \\ 3\cdot 2 + 4\cdot 1 \end{bmatrix} = \frac{1}{\sqrt{50}}\begin{bmatrix} 10 \\ 10 \end{bmatrix} \leftarrow Av^{(1)}$$

$$\|Av^{(1)}\|_2 = \sqrt{\frac{10^2}{50} + \frac{10^2}{50}} = 2 \quad \text{— Amplification factor}$$

$\therefore$ 1) Unit input dir is $v^{(1)} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

2) Amplification is 2. — σ_1

3) Resulting output direction is $Av^{(1)} = \frac{1}{\sqrt{50}}\begin{bmatrix} 10 \\ 10 \end{bmatrix}$

$$d)\ v^{(2)} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad Av^{(2)} = \frac{1}{\sqrt{50}}\begin{bmatrix} 5 \\ 3-8 \end{bmatrix} = \frac{1}{\sqrt{50}}\begin{bmatrix} 5 \\ -5 \end{bmatrix} \leftarrow Av^{(1)}$$

$$\|Av^{(2)}\|_2 = \sqrt{\frac{5^2}{50} + \frac{5^2}{50}} = 1 \quad \text{— A.F.}$$

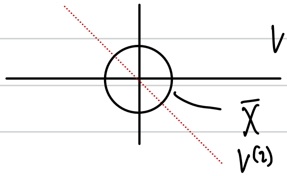
$\therefore$ 1) Unit input dir is $v^{(2)} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix}$

2) Amplification is 1. — σ_2

3) Resulting output direction is $Av^{(2)} = \frac{1}{\sqrt{50}}\begin{bmatrix} 5 \\ -5 \end{bmatrix}$

$$\begin{array}{cccc} & U & \Sigma & V^T \\ e)\ A = \frac{1}{\sqrt{10}} & \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}^T \end{array}$$

$$Ax = U\Sigma \underbrace{V^T x}_{\bar{x}}$$

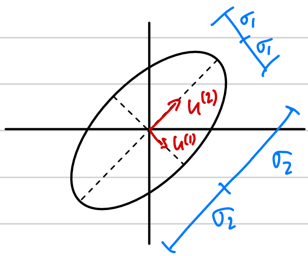


Rotation of the unit circle

$\Sigma \bar{x} \to$ Scaling by σ_1 in $V^{(1)}$, σ_2 in $V^{(2)}$

$U\Sigma\bar{x} \to$ rotation + flip into $u^{(1)}$ and $u^{(2)}$

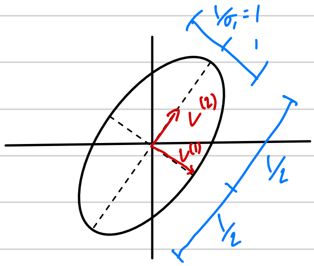
So plotting:



f) $Ax = U\Sigma V^T x$ U, V^T orthogonal by definition of SVD

So $\|Ax\|_2 = \|\Sigma x\|_2 \leq 1$

$\therefore$ axes are $1/\sigma_i$ in $V^{(i)}$



4. $A \in \mathbb{R}^{m \times n}$ $B \in \mathbb{R}^{p \times n}$ $A^TA = B^TB$

Note, Σ_A, Σ_B are both same in entries but differ in shape.

$$\underset{m \times m}{U_A}\,\underset{m \times n}{\Sigma_A}\,\underset{n \times n}{V^T} = A \qquad B = \underset{p \times p}{U_B}\,\underset{p \times n}{\Sigma_B}\,\underset{n \times n}{V^T}$$

$\max \|Ax\|_2$

s.t. $\|Bx\|_2 = 1$

By definition of SVD U, V orthogonal:

$Ax = U_A \Sigma_A V^T x = U_A \Sigma_A y$ $\quad Bx = U_B \Sigma_B V^T x = U_B \Sigma_B y$

$\|Ax\|_2^2 = \|U_A \Sigma_A y\|_2^2 = \|\Sigma_A y\|_2^2 = \sum_{i=1}^{n} \sigma_{Ai}^2 y_i^2$

$\|Bx\|_2^2 = \|U_B \Sigma_B y\|_2^2 = \|\Sigma_B y\|_2^2 = \sum_{i=1}^{n} \sigma_{Bi}^2 y_i^2$

Now consider equivalent problem:

$\max \sum_{i=1}^{n} \sigma_{Ai}^2 y_i^2$

s.t. $\sum_{i=1}^{n} \sigma_{Bi}^2 y_i^2 = 1 \rightarrow$ to satisfy, we pick $y = \begin{bmatrix} 1/\sigma_{B_1} \\ \vdots \\ 1/\sigma_{B_n} \end{bmatrix} \rightarrow \sum_{i=1}^{n} \sigma_{Bi}^2 \left(\frac{1}{\sigma_{Bi}}\right)^2 = 1$

equivalently:

$\max \sum_{i=1}^{n} \frac{\sigma_{Ai}^2}{\sigma_{Bi}^2}$ Put all budget in jth $\rightarrow \max_j \frac{\sigma_{Aj}}{\sigma_{Bj}}$

Optimal y is norm in jth coord $\rightarrow v_j$ by definition.

$\therefore$ optimal $\boxed{x^* = \frac{v_j}{\sigma_{Bj}}}$

The max norm is $\boxed{\|Ax^*\|_2 = \frac{\sigma_{Aj}}{\sigma_{Bj}}}$

Application Problems

1a)

The interpertation of u_l for $l = 1, \cdots, r$ could be an embedding of a semantic theme or idea. This follows from the shape of U, $U \in \mathbb{R}^{n \times n}$ — the word embedding dimension. V_l can be interperted as an encodings of u_i's (themes) for a document l. This follows from $V \in \mathbb{R}^{m \times m}$, the document dimension. From this understanding, it is clear that Σ provides some signifisance or weight for the concepts in each document.

When r is small and u_l, v_l are sparse, it means there are few concepts and thus few document-concept relationships. And that there are only a few specific terms that repsent u_l, and only a few specific documents are associated with the lth concept.

b) The rank-k approximation $\tilde{M}_k$ is $\tilde{M}_k = U_k \Sigma_k V_k^T$

We project q and d_j onto $R(\tilde{M}_k)$ using U_k^T:

$$q' = U_k^T q = \begin{bmatrix} \langle u^{(1)}, q \rangle \\ \vdots \\ \langle u^{(k)}, q \rangle \end{bmatrix} \quad \text{and} \quad d_j' = U_k^T d_j = \begin{bmatrix} \langle u^{(1)}, d_j \rangle \\ \vdots \\ \langle u^{(k)}, d_j \rangle \end{bmatrix}$$

Now we compute cosine similarity in $R(M_k)$:

$$\cos\theta = \frac{q' \cdot d_j'}{\|q\|_2 \|d_j\|_2} = \frac{U_k^T q \, U_k^T d_i}{\|U_k^T q\|_2 \|U_k^T d_i\|}$$