Computational Techniques Assessed CW3

Solution 1

Let $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{2\times2}$ and $\mathbf{V} \in \mathbb{R}^{3\times3}$ and both are orthogonal matrices.

We can obtain $\mathbf{V}$ by finding the eigenvectors of $\mathbf{A}^T\mathbf{A}$, which together will form an orthogonal matrix of $\mathbf{A}^T\mathbf{A}$.

We first find the matrix $\mathbf{A}\mathbf{A}^T$ since its dimension is smaller than $\mathbf{A}^T\mathbf{A}$

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

The eigenvalues of $\mathbf{A}\mathbf{A}^T$ can be found by solving $\det(\mathbf{A}\mathbf{A}^T - \lambda\mathbf{I}) = 0$

$$\begin{aligned} \det\left(\begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix}\right) &= 0 \\ \begin{vmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{vmatrix} &= 0 \\ (17-\lambda)^2 - 8^2 &= 0 \\ (17-\lambda)^2 &= 8^2 \\ \lambda &= 9, 25 \end{aligned}$$

We then find the eigenvectors of matrix $\mathbf{A}^T\mathbf{A}$ corresponding to these eigenvalues

$$\begin{aligned} (\mathbf{A}\mathbf{A}^T - 9\mathbf{I})\mathbf{x} &= 0 \\ \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}\mathbf{x} &= 0 \\ \mathbf{x} &= \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \\ E_9 &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} (\mathbf{A}\mathbf{A}^T - 25\mathbf{I})\mathbf{x} &= 0 \\ \begin{bmatrix} -8 & 8 \\ 8 & -8 \end{bmatrix}\mathbf{x} &= 0 \\ \mathbf{x} &= \operatorname{span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \\ E_{25} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

We have now obtained $\mathbf{U}$

$$\mathbf{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Now we use the fact that $\mathbf{US} = \mathbf{AV}$ to compute $\mathbf{V}$

$$\mathbf{v}_1 = \frac{1}{\sigma_1}\mathbf{A}^T\mathbf{u}_1$$

$$\mathbf{v}_1 = \frac{1}{5\sqrt{2}}\begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\mathbf{v}_2 = \frac{1}{\sigma_2}\mathbf{A}^T\mathbf{u}_2$$

$$\mathbf{v}_2 = \frac{1}{3\sqrt{2}}\begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{4}{3\sqrt{2}} \end{bmatrix}$$

We can then construct $\mathbf{v}_3$ by using the following property, since $\mathbf{V}$ is an orthogonal matrix

$$\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2$$

By using the cross product:

$$\mathbf{v}_{3,normalized} = \mathbf{v}_{1,normalized} \times \mathbf{v}_{2,normalized} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{4}{3\sqrt{2}} \end{vmatrix} = -\frac{2}{3}\mathbf{i} + \frac{2}{3}\mathbf{j} - \frac{1}{3}\mathbf{k} \Rightarrow \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}$$

Hence we have also obtained $\mathbf{V}$

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_{1,normalized} & \mathbf{v}_{2,normalized} & \mathbf{v}_{3,normalized} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{bmatrix}$$

and also $\mathbf{S}$

$$\mathbf{S} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

Therefore the singular value decomposition of $\mathbf{A}$ is as follows

$$
\begin{aligned}
\mathbf{A} = \mathbf{USV}^T &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}5 & 0 & 0\\ 0 & 3 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}}\\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{bmatrix}\\
&= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}\frac{5}{\sqrt{2}} & \frac{5}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \frac{4}{\sqrt{2}}\end{bmatrix}\\
&= \begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}\frac{5}{2} & \frac{5}{2} & 0\\ \frac{1}{2} & -\frac{1}{2} & 2\end{bmatrix}\\
&= \begin{bmatrix}3 & 2 & 2\\ 2 & 3 & -2\end{bmatrix} = \mathbf{A}
\end{aligned}
$$

Solution 2

i) From the definition of matrix multiplication we know that

$$
\begin{aligned}
\mathbf{BC} = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{c=1}^{k} b_{ic}c_{cj} &= \begin{bmatrix} b_{11}c_{11} + b_{12}c_{21} + \ldots + b_{1k}c_{k1} & b_{11}c_{12} + b_{12}c_{22} + \ldots + b_{1k}c_{k2} & \ldots\\ b_{21}c_{11} + b_{22}c_{21} + \ldots + b_{2k}c_{k1} & b_{21}c_{12} + b_{22}c_{22} + \ldots + b_{2k}c_{k2} & \ldots\\ \ldots & \ldots & \ldots \end{bmatrix}\\
&= \begin{bmatrix} b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} & \ldots\\ b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} & \ldots\\ b_{31}c_{11} & b_{31}c_{12} & b_{31}c_{13} & \ldots\\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} + \begin{bmatrix} b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} & \ldots\\ b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{23} & \ldots\\ b_{32}c_{21} & b_{32}c_{22} & b_{32}c_{23} & \ldots\\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} + \ldots\\
&= b_1c^1 + b_2c^2 + b_3c^3 + \ldots + b_kc^k\\
&= \sum_{i=1}^{k} b_ic^i
\end{aligned}
$$

where b_i denotes the ith column of $\mathbf{B}$ and c^i denotes the ith row of $\mathbf{C}$.

ii) From the nature of singular value decomposition we know that $\mathbf{S} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_r, \sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_{min(m,n)})$ where $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_{min(m,n)=0}$.

Hence we can obtain $\mathbf{US}$ using the statement from i):

$$
\mathbf{US} = \sum_{i=1}^{m} u_is^i
$$

where u_i denotes the ith column of $\mathbf{U}$ and s^i denotes the ith row of $\mathbf{S}$. Since only the diagonal element of $\mathbf{S}$ has value σ_i and all other entries are 0, we know that

$$
\mathbf{US} = \sum_{i=1}^{m} u_is^i = \sigma_1u_1 + \sigma_2u_2 + \ldots + \sigma_mu_m = \sum_{i=1}^{m}\sigma_iu_i
$$

which means that the ith column of $\mathbf{US}$ is σ_iu_i.

Again from i), we can then obtain $\mathbf{USV}^T$

$$
\mathbf{A} = \mathbf{USV}^T = \mathbf{US}\cdot\mathbf{V}^T = \sum_{i=1}^{n}\sigma_iu_iv^i
$$

where v^i denotes the ith row of $\mathbf{V}^T$. After transpose, we know that the ith row of $\mathbf{V}^T$ is v_i^T. Therefore

$$\mathbf{A} = \mathbf{USV}^T = \mathbf{US} \cdot \mathbf{V}^T = \sum_{i=1}^{n} \sigma_i u_i v^i = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

Solution 3

i) We substitude w with Vx and let $S = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$

$$\begin{aligned} w_{(1)} &= \arg\max\{w^T A^T A w : w \in \mathbb{R}^n, \|w\| = 1\} \\ &= \arg\max\{(Vx)^T A^T A V x : Vx \in \mathbb{R}^n, \|Vx\| = 1\} \\ &= \arg\max\{(x^T V^T A^T A V x : Vx \in \mathbb{R}^n, \|Vx\| = 1\} \\ &= \arg\max\{(x^T S^T S x : Vx \in \mathbb{R}^n, \|Vx\| = 1\} \end{aligned}$$

We then expand $x^T S^T Sx$:

$$\begin{aligned} x^T S^T S x &= \begin{bmatrix} x_1 & x_2 & x_3 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 & 0 & \ldots \\ 0 & \sigma_2^2 & 0 & \ldots \\ 0 & 0 & \sigma_3^2 & \ldots \\ \ldots & \ldots & \ldots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1\sigma_1^2 & x_2\sigma_2^2 & x_3\sigma_3^2 & \ldots & x_n\sigma_n^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_n \end{bmatrix} \\ &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + \ldots + x_n^2\sigma_n^2 \end{aligned}$$

Since $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq \sigma_n \geq 0$ and $\|Vx\| = 1$, we know that the largest singular value is σ_1, therefore we can find the maximum value by setting $x_1 = 1$ and $x_2 = x_3 = \ldots = x_n = 0$.

Hence, $x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}$

We then put x back into $w = Vx$ and get

$$w = Vx = \begin{bmatrix} v_1, v_2, \ldots, v_n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix} = v_1$$

where v_i denotes the ith column of V.

Hence, we have proved that $w_{(1)} = v_1$

ii) From i) we know that we can find a matrix V such that V^T gives the singular value decomposition of A because of the fact that $V^T A^T AV = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_n^2)$. Then, we have also shown that using V we can obtain $w_{(1)} = v_1$ where $x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \end{bmatrix}$ and $w = Vx$.

From Problem 2 we know that if we can find a singular value decomposition $A = USV^T$, we have

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T + \ldots + \sigma_r u_r v_r^T$$

where r is the rank of A and $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$.

Since we represent A in the direction of first principal component, $v_2 = v_3 = \ldots = \mathbf{0}$ as indicated in Problem 3 i). Hence, A becomes only the first term in the SVD formula

$$A \approx \sigma_1 u_1 v_1^T$$