Homework 1 Solution

Question 1 By definition of matrix product and matrix addition,

$$\begin{aligned}
\boldsymbol{AB} &= \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_k \end{bmatrix} \begin{bmatrix} \vec{b}_1^\top \\ \vdots \\ \vec{b}_k^\top \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{k1} \\ a_{12} & a_{22} & \dots & a_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kp} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{l=1}^k a_{l1}b_{l1} & \sum_{l=1}^k a_{l1}b_{l2} & \dots & \sum_{l=1}^k a_{l1}b_{lp} \\ \sum_{l=1}^k a_{l2}b_{l1} & \sum_{l=1}^k a_{l2}b_{l2} & \dots & \sum_{l=1}^k a_{l2}b_{lp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{l=1}^k a_{ln}b_{l1} & \sum_{l=1}^k a_{ln}b_{l2} & \cdots & \sum_{l=1}^k a_{ln}b_{lp} \end{bmatrix} \\
&= \sum_{l=1}^k \begin{bmatrix} a_{l1}b_{l1} & a_{l1}b_{l2} & \dots & a_{l1}b_{lp} \\ a_{l2}b_{l1} & a_{l2}b_{l2} & \dots & a_{l2}b_{lp} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ln}b_{l1} & a_{ln}b_{l2} & \cdots & a_{ln}b_{lp} \end{bmatrix} \\
&= \sum_{l=1}^k \begin{bmatrix} a_{l1} \\ a_{l2} \\ \vdots \\ a_{ln} \end{bmatrix} \begin{bmatrix} b_{l1} & b_{l2} & \dots & b_{lp} \end{bmatrix} \\
&= \sum_{l=1}^k \vec{a}_l \vec{b}_l^\top .
\end{aligned}$$

Question 2 By definition of matrix product and scalar-vector multiplication,

$$\begin{aligned}
\boldsymbol{CA} &= \begin{bmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \\
&= \begin{bmatrix} c_1a_{11} & c_1a_{12} & \dots & c_1a_{1p} \\ c_2a_{21} & c_2a_{22} & \dots & c_2a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_na_{n1} & c_na_{n2} & \cdots & c_na_{np} \end{bmatrix}
\end{aligned}$$

Then

$$
\boldsymbol{CAD} = \begin{bmatrix} c_1 a_{11} & c_1 a_{12} & \dots & c_1 a_{1p} \\ c_2 a_{21} & c_2 a_{22} & \dots & c_2 a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_n a_{n1} & c_n a_{n2} & \cdots & c_n a_{np} \end{bmatrix} \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_p \end{bmatrix}
$$

$$
= \begin{bmatrix} c_1 d_1 a_{11} & c1_{d2} a_{21} & \dots & c_1 d_p a_{1p} \\ c_2 d_1 a_{12} & c_2 d_2 a_{22} & \dots & c_2 d_p a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_n d_1 a_{n1} & c_n d_2 a_{n2} & \cdots & c_n d_p a_{np} \end{bmatrix}
$$

Question 3

Proof.

$$
\begin{aligned}
\text{LHS} &= \vec{q}_1 \vec{q}_1^\top + \vec{q}_2 \vec{q}_2^\top + ... + \vec{q}_k \vec{q}_k^\top \\
&= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_k \end{bmatrix} \begin{bmatrix} \vec{q}_1^\top \\ \vec{q}_2^\top \\ \vdots \\ \vec{q}_q^\top \end{bmatrix} \\
&= \boldsymbol{Q}\boldsymbol{Q}^\top
\end{aligned}
$$

Here $\boldsymbol{Q} = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_k \end{bmatrix}$. As $\vec{q}_i,\ i = 1, ..., k$ are unit and pairwise perpendicular, $\boldsymbol{Q}$ is a orthogonal matrix. By the property of orthogonal matrices, we have

$$
\boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I}
$$

which finishes the proof. □

Question 4

(a)

$$
\boldsymbol{A}^\top \boldsymbol{A} = \begin{bmatrix} 54 & -10 \\ -10 & 54 \end{bmatrix}
$$

By definition, the eigenvalues can be solved by $|\boldsymbol{A}^\top \boldsymbol{A} - \lambda \boldsymbol{I}| = 0$, i.e.

$$
(54 - \lambda)^2 - (-10)^2 = 0.
$$

Two solutions are $\lambda_1 = 44, \lambda_2 = 64$.

The eigenvector corresponding to λ_1, λ_2 can be solved by $(\boldsymbol{A}^\top \boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{v}_1 = 0$, $(\boldsymbol{A}^\top \boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_2 = 0$ respectively. Let $\boldsymbol{v}_1 = [v_{11} \quad v_{12}]^\top$, we first solve $\boldsymbol{v}_1$ from

$$
\begin{aligned}
10v_{11} - 10v_{12} &= 0 \\
-10v_{11} + 10v_{12} &= 0.
\end{aligned}
$$

Although there are two equations, they are linearly dependent since $(\boldsymbol{A}^\top \boldsymbol{A} - \lambda_1 \boldsymbol{I})$ is of rank 1, which means there are infinite solutions. Additionally, the eigenvectors need to be unit in spectral decomposition. First we can let $v_{11} = 1$ then $v_{12} = 1$ from the equations. Then we normalize $\boldsymbol{v}_1$ by dividing each element of it with the norm of $\boldsymbol{v}_1$, $\sqrt{2}$, where we get

$$
\boldsymbol{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.
$$

To get the second eigenvector, we may use $(\boldsymbol{A}^\top \boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_2 = 0$ and solve it the same way as solving $\boldsymbol{v}_1$. An alternative approach in two dimensional eigenvector (2×2 matrix) case is to use the property that eigenvectors associated with distinct eigenvalues are mutually orthogonal (this is true for any dimension). This suggests $[v_{12} \quad -v_{11}]^\top$ is what we are looking for. It is a unit vector since $v_{11}^2 + v_{12}^2 = 1$ and $[v_{12} \quad -v_{11}]\,[v_{11} \quad v_{12}]^\top = v_{11}v_{12} - v_{11}v_{12} = 0$ tells us it is orthogonal to $\boldsymbol{v}_1$. Also in two dimensional case the orthogonal unit vector is unique with respect to multiplying -1 (either $[-v_{12} \quad v_{11}]^\top$ or $[v_{12} \quad -v_{11}]^\top$ is eigenvector corresponding to λ_2). So we get

$$\boldsymbol{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Then we have a spectral decomposition of $\boldsymbol{A}^\top \boldsymbol{A}$,

$$\boldsymbol{A}^\top \boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 44 & 0 \\ 0 & 64 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

(b) From the spectral decomposition of $\boldsymbol{A}^\top \boldsymbol{A}$, we have

$$(\boldsymbol{A}^\top \boldsymbol{A})^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{44} & 0 \\ 0 & \frac{1}{64} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.019176136 & 0.003551136 \\ 0.003551136 & 0.019176136 \end{bmatrix}$$

$$(\boldsymbol{A}^\top \boldsymbol{A})^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 44^{-\frac{1}{2}} & 0 \\ 0 & 64^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.13787784 & 0.01287784 \\ 0.01287784 & 0.13787784 \end{bmatrix}$$

Question 5

(a) *Proof.* Since $\boldsymbol{S}$, $\boldsymbol{D}$ and $\boldsymbol{C}$ are all invertible matrices, $(\boldsymbol{CSD}^T)^{-1} = (\boldsymbol{D}^T)^{-1}\boldsymbol{S}^{-1}\boldsymbol{C}^{-1}$. Then

$$(\boldsymbol{D}\vec{x})^\top (\boldsymbol{CSD}^\top)^{-1}(\boldsymbol{C}\vec{y}) = \vec{x}^\top \boldsymbol{D}^\top (\boldsymbol{D}^\top)^{-1}\boldsymbol{S}^{-1}\boldsymbol{C}^{-1}\boldsymbol{C}\vec{y} = \vec{x}^\top \boldsymbol{S}^{-1}\vec{y}$$

□

(b)

$$\boldsymbol{C}\vec{x} = [3 \quad 2]^T$$

$$\boldsymbol{CSC}^\top = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

$$(\boldsymbol{CSC}^\top)^{-1} = \begin{bmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{bmatrix}$$

$$(\boldsymbol{C}\vec{x})^\top (\boldsymbol{CSC}^\top)^{-1}(\boldsymbol{C}\vec{x}) = \frac{22}{7}$$

$$\vec{x}^\top \boldsymbol{S}^{-1}\vec{x} = \frac{10}{3}$$

Though the equality in (a) does not hold, it does not contradict with (a) since here $\boldsymbol{C}$ is not a invertible square matrix.