CMPUT340 Written Assignment 2

Yiyang Zuo (Student No: 1861800)

Problem 1

Solution:

Given a matrix $\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 4 & 5 & 9 \end{pmatrix}$

(a) *Proof*

Method 1: Compute its determinant

$$\det(\mathbf{X}) = \begin{vmatrix} 1 & 3 \\ 5 & 9 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 \\ 5 & 9 \end{vmatrix} + 4\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} = -6 - 2 \cdot 3 + 4 \cdot 3 = 0$$

For a square matrix with the value of 0 for its determinant, indicating that the matrix $\mathbf{X}$ is singular.

Method 2: Find its rank using Gaussian Elimination

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 4 & 5 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 4 & 5 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, we can obtain that rank($\mathbf{X}$) = 2 < 3, indicating that the matrix $\mathbf{X}$ is singular.

Assume that $\mathbf{X}' = [\mathbf{X}\ \ \mathbf{z}]$.

(b)

Since we are given, $\mathbf{z} \in \operatorname{span}(\mathbf{X})$ and $\mathbf{z} \in \mathbb{R}^3$, indicating that there exists a unique solution, denoted by $\boldsymbol{y}$, to the square linear system $\mathbf{X}\boldsymbol{y} = \mathbf{z}$. In reverse, we can deduct that rank($\mathbf{X}$) = 3.

Assume that **z may** lie in span($\mathbf{X}$) in subproblems (c), (d), and (e).

(c)

Note that $\mathbf{X}'$ is the augmented matrix of the system, after obtaining the value of its rank using row reduction, if rank($\mathbf{X}'$) = 2 < 3, indicating that the last row of augmented matrix is filled with zeros since the pivot in the last column is 0 and the first two entries have been eliminated to zero in previous row reductions. Thus, the last row consists of zero, indicating there exists a free variable and revealing that the system has infinitely many solutions. (Therefore, $\mathbf{z} \notin \operatorname{span}(\mathbf{X})$)

(d)

Likewise, if rank ($\mathbf{X}'$) = 3, indicating that the pivot in each column after elimination is non-zero, indicating there does not exist any free variables. Hence, the system has a unique solution. (Hence, $\mathbf{z} \in \operatorname{span}(\mathbf{X})$)

(e)

From the 2nd method given in subproblem (a), we can know that the base of span($\mathbf{X}$) consist of the first two column vectors in matrix $\mathbf{X}$, denoted by $\mathbf{v}_1$ and $\mathbf{v}_2$.

Thus, we can express the solution of the system as

$$\mathbf{X}\boldsymbol{d} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\begin{pmatrix}1\\2\\4\end{pmatrix} + c_2\begin{pmatrix}2\\1\\5\end{pmatrix} = \begin{pmatrix}c_1+2c_2\\2c_1+c_2\\4c_1+5c_2\end{pmatrix}(c_1, c_2 \in \mathbb{R})$$

According to the known condition, we can obtain that

$$\mathbf{X}\boldsymbol{d} = \begin{pmatrix}c_1+2c_2\\2c_1+c_2\\4c_1+5c_2\end{pmatrix} = \begin{pmatrix}14\\13\\41\end{pmatrix}$$

Hence, we can derive that $c_1 = 4$, $c_2 = 5$, indicating that $\mathbf{X}\boldsymbol{d} \in \operatorname{span}(\mathbf{X})$. Therefore, we can know that the system has infinitely many solutions.

Assume $\boldsymbol{d}$ is not equal to $\boldsymbol{c}$, so we can obtain the augmented matrix $[\mathbf{X}\ \mathbf{X}\boldsymbol{d}]$.

$$[\mathbf{X}\ \mathbf{X}\boldsymbol{d}] = \begin{pmatrix}1&2&3&14\\2&1&3&13\\4&5&9&41\end{pmatrix} \sim \begin{pmatrix}1&2&3&14\\0&-3&-3&-15\\4&5&9&41\end{pmatrix} \sim \begin{pmatrix}1&2&3&14\\0&-3&-3&-15\\0&-3&-3&-15\end{pmatrix} \sim \begin{pmatrix}1&2&3&14\\0&-3&-3&-15\\0&0&0&0\end{pmatrix}$$

Thus, we can obtain that $\boldsymbol{d} = d_1\begin{pmatrix}1\\1\\-1\end{pmatrix} + \begin{pmatrix}0\\1\\4\end{pmatrix}$, when $d_1 = 1$, $\boldsymbol{d} = \boldsymbol{c}$.

Plug $d_1 = 2$ in $\boldsymbol{d}$, we can obtain that $\boldsymbol{d} = 2\begin{pmatrix}1\\1\\-1\end{pmatrix} + \begin{pmatrix}0\\1\\4\end{pmatrix} = \begin{pmatrix}2\\3\\2\end{pmatrix}$.

Problem 2

Solution:

Proof: Since the matrix $\mathbf{C}$ is not square, it has pseudoinverse only. From the definition, we can obtain that

$$\mathbf{C}^{+} = (\mathbf{C}^{\mathrm{T}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}}$$

We have assumed that $\mathbf{C}$ is a m by n matrix, which can be decomposed using singular values. Thus, we can obtain the expression of pseudoinverse as following:

$$\begin{aligned}
\mathbf{C}^{+} &= (\mathbf{C}^{\mathrm{T}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}} = ((\mathbf{U}\sum\mathbf{V}^{\mathrm{T}})^{\mathrm{T}}\mathbf{U}\sum\mathbf{V}^{\mathrm{T}})^{-1}(\mathbf{U}\sum\mathbf{V}^{\mathrm{T}})^{\mathrm{T}} \\
&= (\mathbf{V}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\sum\mathbf{V}^{\mathrm{T}})^{-1}\mathbf{V}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}} \\
&= (\mathbf{V}\sum{}^{2}\mathbf{V}^{\mathrm{T}})^{-1}\mathbf{V}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}} \\
&= \mathbf{V}\sum{}^{-2}\mathbf{V}^{\mathrm{T}}\mathbf{V}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}} \\
&= \mathbf{V}\sum{}^{-2}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}
\end{aligned}$$

Note that matrix $\sum$ is non-square matrix, hence $\sum{}^{-2}\sum{}^{T} = \sum{}^{+}$

Thus, we can prove that $\mathbf{C}^{+} = \mathbf{V}\sum{}^{+}\mathbf{U}^{\mathrm{T}}$.

Apply the definition of pseudoinverse again, we can obtain pseudoinverse of $\mathbf{C}$.

$$\sum{}^{+} = (\sum{}^{T}\sum\ )^{-1}\sum{}^{T} = \left((\sum{}_{1}\ \ 0)\begin{pmatrix}\sum{}_{1}\\ 0\end{pmatrix}\right)^{-1}(\sum{}_{1}\ \ 0)$$

Before computing the intermediate expression, we have to determine the size of each matrix. $\sum{}_{1}$ is n by n, 0-partitioned matrix in sigma transpose is m by (m-n), and 0-partitioned matrix in sigma is (m-n) by n. Hence, we know that this expression is multipliable.

Hence, we can derive that

$$\sum{}^{+} = \left((\sum{}_{1}\ \ 0)\begin{pmatrix}\sum{}_{1}\\ 0\end{pmatrix}\right)^{-1}(\sum{}_{1}\ \ 0) = \sum{}^{-2}{}_{1}(\sum{}_{1}\ \ 0) = (\sum{}^{-1}{}_{1}\ \ 0)_{n\times m}$$

Problem 3

(a) *Proof*

We know that eigenvalues are produced from the linear equation $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, indicating that all corresponding entries in two matrices are equal. Thus, we can obtain that

$$\|\mathbf{A}\boldsymbol{x}\| = \|\lambda\boldsymbol{x}\|$$

We can express $|\lambda| = \dfrac{\|\mathbf{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$ (Since $\|\lambda\boldsymbol{x}\| = |\lambda|\,\|\boldsymbol{x}\|$).

We can express $\rho(\mathbf{A})$ based on what we have derived and the definition of spectral radius

$$\rho(\mathbf{A}) = \max_{\boldsymbol{x}\neq\vec{0}} \frac{\|\mathbf{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \max_{\boldsymbol{x}\neq\vec{0}} \frac{\|\mathbf{A}\|\,\|\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\boldsymbol{x}\neq\vec{0}} \|\mathbf{A}\| = \|\mathbf{A}\|$$

Hence, we can prove that $\rho(\mathbf{A}) \leqslant \|\mathbf{A}\|$.

(b) *Proof*

We can know that **A** is square and assume $\mathbf{A}\in\mathbb{R}^{n\times n}$ and exists $\boldsymbol{x}\in\mathbb{R}^{n}$.

Since **A** is positive-definite $\Rightarrow \boldsymbol{x}^T\mathbf{A}\boldsymbol{x} > 0$

From the characteristic equation $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, we can derive that

$$\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}^T\boldsymbol{x} = \lambda\|\boldsymbol{x}\|^2 > 0$$

Hence, we can know that $\lambda > 0$, indicating that all eigenvalues of matrix **A** are positive.

Denote the eigenvalue of matrix ($\mathbf{I} - \alpha\mathbf{A}$) by μ.

Based on the characteristic equation, we can express

$$(\mathbf{I} - \alpha\mathbf{A})\boldsymbol{y} = \mu\boldsymbol{y} \Rightarrow \alpha\mathbf{A}\boldsymbol{y} = (1-\mu)\boldsymbol{y}$$

Since $\mathbf{A}\boldsymbol{y} = \lambda_i\boldsymbol{y}$ (i = 1,2,3, …, n)(Since the transformation $(\mathbf{I} - \alpha\mathbf{A})\boldsymbol{y}$ is linear.), we can derive that $\mu = 1 - \alpha\lambda_i$.

Apply the definition of spectral radius, we can find that $\rho(\mathbf{I} - \alpha\mathbf{A}) = \max_{\mathbf{i}} |1 - \alpha\lambda_i|$

Hence, we can prove that $\rho(\mathbf{I}-\alpha\mathbf{A}) = \max_{\mathbf{i}} |1-\alpha\lambda_i| < 1$ when $\alpha > 0$.

Problem 4

(a) *Proof*

Denote the eigenvalue of matrix **A** and (**I** - **A**) by λ_i, and $\mu_i (1 \leqslant i \leqslant n)$. Assume that the solution to characteristic equation regarding **A** is ***x.***

Thus, we can derive the characteristic equation as following:

$$(\mathbf{I}-\mathbf{A})\boldsymbol{x} = \mu_i \boldsymbol{x}$$

Solve that $\mu_i = 1-\lambda_i$.

We are given that $\lambda_i < 1$, hence $\mu_i \neq 0$, indicating that matrix (**I** - **A**) is nonsingular.

(b) *Proof*

From given conditions, we know that matrix **A** is diagonalizable. Thus, we can express A as the following:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

where the matrix **D** is diagonal, and **P** is invertible. Denote $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$.

Thus, we can derive that $\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}$. Take the limit in terms of $\mathbf{A}^k$, we can obtain

$$\lim_{k\to\infty} \mathbf{A}^k = \lim_{k\to\infty} (\mathbf{P}\mathbf{D}^k\mathbf{P}^{-1})$$

Note that we have $\lim_{k\to\infty} \mathbf{D}^k = \lim_{k\to\infty} \begin{pmatrix} \lambda^k{}_1 & 0 & \cdots & 0 \\ 0 & \lambda^k{}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda^k{}_n \end{pmatrix} = \vec{0}$

due to $\lambda_i < 1 \Rightarrow \lambda^k{}_i \to 0$ when $k \to \infty$.

Hence, we can obtain that $\lim_{k\to\infty} \mathbf{A}^k = \lim_{k\to\infty} (\mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}) = \mathbf{P}\vec{0}\mathbf{P}^{-1} = \vec{0}$.

(c) *Proof*

We know that $\sum_{k=0}^{\infty} \mathbf{A}^k = \lim_{n\to\infty} \sum_{k=0}^{n} \mathbf{A}^k$.

Expand the summation, we can obtain that

$$\sum_{k=0}^{n} \mathbf{A}^k = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^n$$

Denote $S_n = \sum_{k=0}^{n} \mathbf{A}^k$.

Pre-multiply the summation by (**I**-**A**), we can obtain that

$$(\mathbf{I} - \mathbf{A})S_n = (\mathbf{I} - \mathbf{A}) \sum_{k=0}^{n} \mathbf{A}^k = \sum_{k=0}^{n} (\mathbf{A}^k - \mathbf{A}^{k+1})$$

Expand the summation, we can obtain that

$$\sum_{k=0}^{n} (\mathbf{A}^k - \mathbf{A}^{k+1}) = \mathbf{I} - \mathbf{A}^{n+1}$$

Take the limit on both sides of the equation, we can derive that

$$\lim_{n\to\infty} \sum_{k=0}^{n} (\mathbf{A}^k - \mathbf{A}^{k+1}) = \lim_{n\to\infty} (\mathbf{I} - \mathbf{A}^{n+1}) = \mathbf{I} = (\mathbf{I} - \mathbf{A}) \lim_{n\to\infty} S_n$$

Hence, we can derive that $\lim_{n\to\infty} S_n = \sum_{k=0}^{\infty} \mathbf{A}^k = (\mathbf{I} - \mathbf{A})^{-1}$

Problem 5

(a) *Proof*

Based on the given condition, we can assume that there are eigenvalues for a matrix $\mathbf{A}$ and size of $\mathbf{A}$ is n by n. Denote the eigenvalues by $\lambda_i (1 \leqslant i \leqslant n)$

We can derive the characteristic equation in terms of $\mathbf{A}$

$$\mathbf{A}\boldsymbol{x} = \lambda_i \boldsymbol{x}$$

Pre-multiply both side of the equation by $\boldsymbol{x}^T$, we can obtain that

$$\boldsymbol{x}^T \mathbf{A}\boldsymbol{x} = \boldsymbol{x}^T \lambda_i \boldsymbol{x} = \lambda_i \boldsymbol{x}^T \boldsymbol{x} = \lambda_i \|\boldsymbol{x}\|^2$$

Since $\boldsymbol{x}^T \mathbf{A}\boldsymbol{x} = \lambda_i \|\boldsymbol{x}\|^2 > 0$, we can prove that $\lambda_i > 0$, indicating that the eigenvalues of a positive definite matrix $\mathbf{A}$ are all positive.

(b) *Proof*

Considering that the algorithm is an iteration, we can prove the conclusion with Mathematical Induction.

Base case: k=1

$$\mathbf{A}_0 = \mathbf{Q}_1\mathbf{R}_1 = \mathbf{A} \Rightarrow \mathbf{A}_1 = \mathbf{R}_1\mathbf{Q}_1 = \mathbf{Q}^H{}_1\mathbf{Q}_1\mathbf{R}_1\mathbf{Q}_1 = \mathbf{Q}^H{}_1\mathbf{A}\mathbf{Q}_1$$

Inducive case: Assume for $m \geqslant 1$, we have:

$$\mathbf{A}_m = \hat{\mathbf{Q}}^H{}_m \mathbf{A}\hat{\mathbf{Q}}_m$$

where $\hat{\mathbf{Q}}_m = \prod_{i=1}^{m} \mathbf{Q}_i$.

Thus, we must show that this also holds for the (m + 1) th iteration.

For the next iteration, we can obtain that

$$\mathbf{A}_m = \mathbf{Q}_{m+1}\mathbf{R}_{m+1} \Rightarrow \mathbf{A}_{m+1} = \mathbf{R}_{m+1}\mathbf{Q}_{m+1} = \mathbf{Q}^H{}_{m+1}\mathbf{Q}_{m+1}\mathbf{R}_{m+1}\mathbf{Q}_{m+1}$$

Thus, we can substitute $\mathbf{Q}_{m+1}\mathbf{R}_{m+1}$ to $\mathbf{A}_m$

$$\mathbf{A}_{m+1} = \mathbf{Q}^H{}_{m+1}\mathbf{A}_m\mathbf{Q}_{m+1}$$

Likewise, we can substitute $\mathbf{A}_m$

$$\mathbf{A}_{m+1} = \mathbf{Q}^H{}_{m+1}\mathbf{Q}^H{}_m\mathbf{A}_m\mathbf{Q}_m\mathbf{Q}_{m+1}$$

Therefore, make continuous substitutions, we can finally obtain that

$$\mathbf{A}_{m+1} = \mathbf{Q}_{m+1}^H \mathbf{Q}_m^H \cdots \mathbf{Q}_1^H \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m+1} = (\mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m+1})^H \mathbf{A} (\mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m+1})$$

Hence, we can prove that $\mathbf{A}_{m+1} = \hat{\mathbf{Q}}^H{}_{m+1} \mathbf{A} \hat{\mathbf{Q}}_{m+1}$ also holds for the (m + 1) th iteration.

Hence, we can prove that $\mathbf{A}_k = \hat{\mathbf{Q}}^H{}_k \mathbf{A} \hat{\mathbf{Q}}_k$.