National University of Singapore
Department of Mathematics

Semester 1, 2018/2019 **MA1101R Linear Algebra I** **Homework 4**

Instruction

(a) This homework set consists of 2 pages and 7 questions.

(b) Do all the problems and submit on Nov. 5 (Monday) during lecture for SL1 group or before noon, Nov. 7 (Wednesday) for SL2 group in Dr. Ng Kah Loon's office (S17 07–20) or Zhang Lei's office (S17 06–05).

(c) Use A4 size writing paper. Write your full name, student number and tutorial group clearly on the first page of your answer scripts.

(d) Indicate the question numbers clearly (you do not need to copy the questions in your answer sheets).

(e) Show your steps of your working how the answers are derived, unless the questions state otherwise.

(f) Late Submission will not be accepted.

(g) **Warning**: If you are found to have copied answers from your friend(s), both you and your friend(s) will be penalized.

Problem Set (covering Lectures 15–19).

1. For the following matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 3 \\ 3 & 6 & 5 & 2 & 7 \\ 2 & 4 & 1 & -1 & 0 \end{pmatrix}$$

(a) find a basis for the row space and a basis for the column space;

Let us perform the Gauss-Jordan elimination and obtain its reduced row-echelon form:

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 3 \\ 3 & 6 & 5 & 2 & 7 \\ 2 & 4 & 1 & -1 & 0 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 2 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since the elementary row operations preserve the row space, $\{(1,2,0,-1,-1),(0,0,1,1,2)\}$ is a basis of the row space of $\boldsymbol{A}$.

Since the 1st and 3rd columns are the pivot columns of the reduced row-echelon form, the corresponding columns $\{(1,1,3,2)^T,(1,2,5,1)^T\}$ of $\boldsymbol{A}$ are a basis for the column space of $\boldsymbol{A}$.

(b) extend the basis for the row space in Part (a) to a basis for $\mathbb{R}^5$;

Since the 1st and 3rd columns are the pivot columns of the reduced row-echelon form, we need to add three more rows such that all the rest columns are pivot columns. Hence $\{(1,2,0,-1,-1),(0,0,1,1,2),(0,1,0,0,0),(0,0,0,1,0),(0,0,0,0,1),\}$ is a basis of $\mathbb{R}^5$.

(c) extend the basis for the column space in Part (a) to a basis for $\mathbb{R}^4$;

We follow the same method as Part (b). (You can try a different method to extend the basis for the column space. For instance, find two more vectors in $\mathbb{R}^4$ to form a 4×4 whose determinant is nonzero.)
Let us perform the Gauss-Jordan elimination for the transpose $\boldsymbol{A}^T$ and obtain its reduced row-echelon form:

$$\begin{pmatrix} 1 & 1 & 3 & 2 \\ 2 & 2 & 6 & 4 \\ 1 & 2 & 5 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 3 & 7 & 0 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $\{(1,1,3,2)^T, (1,2,5,1)^T, (0,0,1,0)^T, (0,0,0,1)^T\}$ is a basis of the column space.

(d) find a basis for the nullspace;

Following the reduced row-echelon form in Part (a), we may solve the homogeneous linear system $\boldsymbol{Ax} = \boldsymbol{0}$ and have a general solution

$$\begin{cases} x_1 &= -2r + s + t \\ x_2 &= r \\ x_3 &= -s - 2t \\ x_4 &= s \\ x_5 &= t \end{cases}$$

Then

$$\boldsymbol{x} = r\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

Therefore, $\{(-2,1,0,0,0)^T, (1,0,-1,1,0)^T, (1,0,-2,0,1)^T\}$ is a basis of the nullspace.

(e) find the rank and nullity of the matrix and hence verify the Dimension Theorem for $\boldsymbol{A}$;

From Parts (a) and (d), we have $\text{rank}(\boldsymbol{A}) = 2$ and the nullity of $\boldsymbol{A} = 3$. Since the number of the columns is 5, we have $\text{rank}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{A}) = 3 + 2 = 5 =$ the number of the columns.

(f) determine if $\boldsymbol{A}$ has full rank.

$\boldsymbol{A}$ is not of full rank. To be of full rank, we need $\text{rank}(\boldsymbol{A}) = 4$, which is not the case.

2. Let $\boldsymbol{A}$ be a square matrix of order n and W be subspace of $\mathbb{R}^n$. Referring to Tutorial 8, $W^\perp = \{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u} \text{ is orthogonal to } W\}$, called the orthogonal complement of W.

(a) Show that $W^\perp \cap W = \{\mathbf{0}\}$ and $W + W^\perp = \mathbb{R}^n$.

First, let $\boldsymbol{u} \in W^\perp \cap W$. By definition, $\boldsymbol{u} \cdot \boldsymbol{u} = \mathbf{0}$, which follows $\boldsymbol{u} = \mathbf{0}$. Hence $W^\perp \cap W = \{\mathbf{0}\}$.
Second, following from $W^\perp \cap W = \{\mathbf{0}\}$, by Question 8 in Homework 3, we have $\dim(W + W^\perp) = \dim(W) + \dim(W^\perp)$. Referring to Tutorial 8, $\dim(W) + \dim(W^\perp) = n$. Thus $\dim(W + W^\perp) = \dim(\mathbb{R}^n)$. Since $W + W^\perp$ is a subspace of $\mathbb{R}^n$, $W + W^\perp = \mathbb{R}^n$ following from Question 3(a) in Tutorial 6 or Theorem 3.6.9.

(b) Show that the nullspace of $\boldsymbol{A}$ and the row space of $\boldsymbol{A}$ are orthogonal complements in $\mathbb{R}^n$.

Let

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}$$

where $\boldsymbol{a}_i$ is the i-th row of $\boldsymbol{A}$. For a column vector $\boldsymbol{u}$ in $\mathbb{R}^n$, since

$$\boldsymbol{A}\boldsymbol{u} = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{u} \\ \boldsymbol{a}_2 \cdot \boldsymbol{u} \\ \vdots \\ \boldsymbol{a}_n \cdot \boldsymbol{u} \end{pmatrix},$$

we have $\boldsymbol{A}\boldsymbol{u} = \mathbf{0}$ if and only if $\boldsymbol{a}_i \cdot \boldsymbol{u} = 0$. Therefore, $\boldsymbol{u}$ is a solution, i.e., in the nullspace of $\boldsymbol{A}$ if and only if $\boldsymbol{u}$ is in the orthogonal complement of the row space of $\boldsymbol{A}$. Then the nullspace of $\boldsymbol{A}$ is identical to the orthogonal complement of the row space of $\boldsymbol{A}$. Hence the nullspace of $\boldsymbol{A}$ and the row space of $\boldsymbol{A}$ are orthogonal complements in $\mathbb{R}^n$.

(c) Show that the nullspace of $\boldsymbol{A}^T$ and the column space of $\boldsymbol{A}$ are orthogonal complements in $\mathbb{R}^n$.

By Part (b), the nullspace of $\boldsymbol{A}^T$ and the row space of $\boldsymbol{A}^T$ are orthogonal complements in $\mathbb{R}^n$. Since the row space of $\boldsymbol{A}^T$ and the column space of $\boldsymbol{A}$ are identical. Hence the nullspace of $\boldsymbol{A}^T$ and the column space of $\boldsymbol{A}$ are orthogonal complements in $\mathbb{R}^n$.

3. Let

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 3 & 3 \\ 1 & 1 & 5 \\ -1 & 3 & 7 \end{pmatrix}, \quad \boldsymbol{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 3 \end{pmatrix}, \quad \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}.$$

(a) **(1 point)** Use the Gram-Schmidt Process to transform $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ into an orthonormal basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ for the column space of $\boldsymbol{A}$. (If you change the order of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ when applying the Gram-Schmidt Process, your answer will be much more complicated.)

Let us perform the Gram-Schmidt Process:

$$\begin{aligned}
\boldsymbol{v}_1 =& \boldsymbol{u}_1 \\
\boldsymbol{v}_2 =& \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 = (1,3,1,3)^T - \frac{-4}{4}(1,-1,1,-1)^T = (2,2,2,2)^T \\
\boldsymbol{v}_3 =& \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 \\
=& (1,3,5,7)^T - \frac{-4}{4}(1,-1,1,-1)^T - \frac{32}{16}(2,2,2,2)^T = (-2,-2,2,2)^T.
\end{aligned}$$

Next, we normalize the vectors:

$$\begin{aligned}
\boldsymbol{w}_1 =& (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^T \\
\boldsymbol{w}_2 =& (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T \\
\boldsymbol{w}_3 =& (-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T.
\end{aligned}$$

(b) **(1 point)** Write each of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ as a linear combination of $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$.

Since $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ is an orthonormal set, for $\boldsymbol{u} \in \text{span}(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$, we have

$$\boldsymbol{u} = (\boldsymbol{u} \cdot \boldsymbol{w}_1)\boldsymbol{w}_1 + (\boldsymbol{u} \cdot \boldsymbol{w}_2)\boldsymbol{w}_2 + (\boldsymbol{u} \cdot \boldsymbol{w}_3)\boldsymbol{w}_3.$$

Thus,

$$\begin{aligned}
\boldsymbol{u}_1 =& 2\boldsymbol{w}_1 \\
\boldsymbol{u}_2 =& -2\boldsymbol{w}_1 + 4\boldsymbol{w}_2 \\
\boldsymbol{u}_3 =& -2\boldsymbol{w}_1 + 8\boldsymbol{w}_2 + 4\boldsymbol{w}_3
\end{aligned}$$

(c) **(1 point)** Hence, or otherwise, write $\boldsymbol{A} = \boldsymbol{QR}$ where $\boldsymbol{Q}$ is a 4×3 matrix with orthonormal columns and $\boldsymbol{R}$ is a 3×3 upper triangular matrix with positive entries along its diagonal. (The process of writing a matrix in the form described in Part (c) is called the QR factorization. It is widely used in computer algorithms for various computations concerning matrices.)

Following Part (a), we may form the matrix $\boldsymbol{Q} = (\boldsymbol{w}_1\ \boldsymbol{w}_2\ \boldsymbol{w}_3)$, which is a 4×3 matrix with orthonormal columns. By Part (b), we have

$$(\boldsymbol{u}_1\ \boldsymbol{u}_2\ \boldsymbol{u}_3) = (\boldsymbol{w}_1\ \boldsymbol{w}_2\ \boldsymbol{w}_3)\begin{pmatrix} 2 & -2 & -2 \\ 0 & 4 & 8 \\ 0 & 0 & 4 \end{pmatrix}.$$

Thus we take

$$\boldsymbol{R} = \begin{pmatrix} 2 & -2 & -2 \\ 0 & 4 & 8 \\ 0 & 0 & 4 \end{pmatrix}$$

which is a 3×3 upper triangular matrix with positive entries along its diagonal. Therefore, $\boldsymbol{A} = \boldsymbol{QR}$, i.e.,

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 3 & 3 \\ 1 & 1 & 5 \\ -1 & 3 & 7 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & -2 & -2 \\ 0 & 4 & 8 \\ 0 & 0 & 4 \end{pmatrix}$$

(d) **(1 point)** Let $\boldsymbol{B}$ be an $m \times n$ matrix, whose columns are linearly independent. Suppose that $\boldsymbol{B}$ has a QR factorization, i.e., $\boldsymbol{B} = \boldsymbol{QR}$ where $\boldsymbol{Q}$ is an $m \times n$ matrix with orthonormal columns and $\boldsymbol{R}$ is an $n \times n$ upper triangular matrix with positive entries along its diagonal, described as in Part (c). Show that if $\boldsymbol{u}$ is a least squares solution of $\boldsymbol{Bx} = \boldsymbol{b}$ then $\boldsymbol{u} = \boldsymbol{R}^{-1}\boldsymbol{Q}^T\boldsymbol{b}$.

Following Theorem 5.3.10, $\boldsymbol{u}$ is a least squares solution of $\boldsymbol{Bx} = \boldsymbol{b}$ if and only if $\boldsymbol{u}$ is a solution of $\boldsymbol{B}^T\boldsymbol{Bx} = \boldsymbol{B}^T\boldsymbol{b}$. Plugging in $\boldsymbol{B} = \boldsymbol{QR}$,

$$\begin{aligned} (\boldsymbol{B}^T\boldsymbol{B})\boldsymbol{u} &= \boldsymbol{B}^T\boldsymbol{b} \\ (\boldsymbol{R}^T\boldsymbol{Q}^T\boldsymbol{QR})\boldsymbol{u} &= \boldsymbol{R}^T\boldsymbol{Q}^T\boldsymbol{b}. \end{aligned}$$

Since the columns of $\boldsymbol{Q}$ are orthonormal, $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_n$. (Note that the columns of $\boldsymbol{B}$ is linearly independent, which implies $m \geq n$.) Also, as $\boldsymbol{R}$ is upper triangular matrix with positive entries along its diagonal, $\boldsymbol{R}$ is invertible. We have

$$\begin{aligned} (\boldsymbol{R}^T\boldsymbol{R})\boldsymbol{u} &= \boldsymbol{R}^T\boldsymbol{Q}^T\boldsymbol{b} \\ \boldsymbol{u} &= (\boldsymbol{R}^T\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{Q}^T\boldsymbol{b} \\ \boldsymbol{u} &= \boldsymbol{R}^{-1}\boldsymbol{Q}^T\boldsymbol{b}. \end{aligned}$$

(e) Let $\boldsymbol{w} = (5, 1, -2, 0)^T$ and $\boldsymbol{x} = (x_1, x_2, x_3)^T$.

(i) **(1 point)** By using the result in Parts (c) and (d), find a least squares solution to the linear system $\boldsymbol{Ax} = \boldsymbol{w}$.

By Parts (d) and (c), we have

$$\boldsymbol{u} = \boldsymbol{R}^{-1}\boldsymbol{Q}^T\boldsymbol{w} = \begin{pmatrix} 2 & -2 & -2 \\ 0 & 4 & 8 \\ 0 & 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}^T \begin{pmatrix} 5 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{5}{2} \\ -1 \end{pmatrix}$$

(ii) **(1 point)** By the result in Part (i), compute the projection of $\boldsymbol{w}$ onto the column space of $\boldsymbol{A}$.

By Theorem 5.3.8, $\boldsymbol{Au}$ is the projection of $\boldsymbol{w}$ onto the column space of $\boldsymbol{A}$. Thus the projection of $\boldsymbol{w}$ is $(\frac{7}{2}, \frac{5}{2}, -\frac{1}{2}, -\frac{3}{2})$

4. (a) Find the equation of the plane P passing through the point $(2, 1, 0)$ and having $\boldsymbol{n} = (1, 2, -2)$ as a normal vector.

   Since the plane P has the normal vector $\boldsymbol{n} = (1, 2, -2)$, we may assume the equation of plane: $x + 2y - 2z = d$. Since it passes through the point $(2, 1, 0)$, $2 + 2 \times 1 - 2 \times 0 = d$, which follows $d = 4$. Hence the equation of the plane P is $x + 2y - 2z = 4$.

   (b) Find the distance between the point $(3, 1, -2)$ and the plane $x + 2y - 2z = 0$.

   Following the proof of Theorem 5.3.2, to find the distance, we need to calculate the projection $\boldsymbol{p}$ from the vector $\boldsymbol{u} = (3, 1, -2)$ to the plane $x + 2y - 2z = 0$. Then the distance $d(\boldsymbol{u}, \boldsymbol{p})$ is the distance between the point and the plane $x + 2y - 2z = 0$. Thus, the goal is to write $\boldsymbol{u} = \boldsymbol{p} + \boldsymbol{v}$ where $\boldsymbol{p}$ is in the plane P and $\boldsymbol{v}$ is orthogonal to P, which follows $d(\boldsymbol{u}, \boldsymbol{p}) = \|\boldsymbol{v}\|$

   Since both $\boldsymbol{v}$ and $\boldsymbol{n}$ are orthogonal to the plane $x + 2y - 2z = 0$, we may assume that $\boldsymbol{v} = t\boldsymbol{n}$ for some $t \in \mathbb{R}$. By $\boldsymbol{u} - \boldsymbol{v}$ in the plane $x + 2y - 2z = 0$, we have $(3 - t, 1 - 2t, -2 + 2t)$ satisfying the equation $x + 2y - 2z = 0$, i.e., $(3 - t) + 2(1 - 2t) - 2(-2 + 2t) = 0$. We solve $t = 1$. Then $d(\boldsymbol{u}, \boldsymbol{p}) = \|\boldsymbol{n}\| = 3$.

   Remark that there are many different ways to solve this problem. The way present here is not one of the easiest ways. You may use the distance between the given vector and the normal vector instead, which is much easier.

   (c) Find the distance between the point $(3, 1, -2)$ and the plane P in Part (a).

   Following from Part (b), the line $(3 - t, 1 - 2t, -2 + 2t)$ passes through the point $(3, 1, -2)$ and is orthogonal to the plane P. The intersection point is determined by $(3 - t) + 2(1 - 2t) - 2(-2 + 2t) = 4$ and solve $t = \frac{5}{9}$. We have the intersection $(\frac{22}{9}, -\frac{1}{9}, -\frac{8}{9})$. The distance between the point $(3, 1, -2)$ and the plane P is given by

$$d((3, 1, -2), (\frac{22}{9}, -\frac{1}{9}, -\frac{8}{9})) = \frac{5}{3}.$$

   (d) Find the distance between the plane P in Part (a) and the plane $2x + 4y - 4z = 18$.

   Following the same method in Part (c), we pick up a point in the plane P and calculate the distance between the point and the pane. We choose a point $(0, 2, 0)$. (Since two planes are parallel, it does not matter which point is chosen.) We have the line $(t, 2t + 2, -2t)$ passing through the point $(0, 2, 0)$ and orthogonal to the plane $2x + 4y - 4z = 18$. Then the intersection point is determined by $2t + 4(2t + 2) - 4(-2t) = 18$ and solve $t = \frac{5}{9}$. We have the intersection $(\frac{5}{9}, \frac{28}{9}, -\frac{10}{9})$. The distance between the point $(3, 1, -2)$ and the plane P is given by

$$d((0, 2, 0), (\frac{5}{9}, \frac{28}{9}, -\frac{10}{9})) = \frac{5}{3}.$$

Remark that $(3,1,-2)$ is also a point in the plane $2x+4y-4z=18$. It directly followw from Part (c) and we have the distance is $\frac{5}{3}$.

5. Let

$$\boldsymbol{A}=\begin{pmatrix}6 & 3 & -8\\ 0 & -2 & 0\\ 1 & 0 & -3\end{pmatrix}.$$

(a) Find the characteristic equation of $\boldsymbol{A}$.

$$\begin{vmatrix}\lambda-6 & -3 & 8\\ 0 & \lambda+2 & 0\\ -1 & 0 & \lambda+3\end{vmatrix}=(\lambda+2)(\lambda^2-3\lambda-10)=\lambda^3-\lambda^2-16\lambda-20.$$

(b) Find all the eigenvalues of $\boldsymbol{A}$.

$\det(\lambda I_3-\boldsymbol{A})=0$ if and only if $(\lambda+2)(\lambda^2-3\lambda-10)=0$. Solve for λ and we have $\lambda=-2$ or $\lambda=5$.

(c) **(2 points)** Find a basis for the eigenspace associated with each eigenvalue of $\boldsymbol{A}$.

For $\lambda=5$, the linear system $(5I_3-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ is

$$\begin{pmatrix}-1 & -3 & 8\\ 0 & 7 & 0\\ -1 & 0 & 8\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}.$$

Solving the system, we have

$$\begin{pmatrix}x\\ y\\ z\end{pmatrix}=t\begin{pmatrix}8\\ 0\\ 1\end{pmatrix}.$$

For $\lambda=-2$, the linear system $(-2I_3-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ is

$$\begin{pmatrix}-8 & -3 & 8\\ 0 & 0 & 0\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}.$$

Solving the system, we have

$$\begin{pmatrix}x\\ y\\ z\end{pmatrix}=t\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}.$$

Remark that even the multiplicity of the root $\lambda=-2$ is 2 but the dimension of the eigenspace is 1.

6. **(1 point)** Let $\boldsymbol{x}$ be an $n \times 1$ matrix and

$$\boldsymbol{A} = \boldsymbol{I}_n - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T$$

Show that $\boldsymbol{A}$ is both orthogonal and symmetric.

First, we show that $\boldsymbol{A}$ is symmetric, i.e., $\boldsymbol{A} = \boldsymbol{A}^T$:

$$\begin{aligned}
\boldsymbol{A}^T =& (\boldsymbol{I}_n - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T)^T \\
=& \boldsymbol{I}_n^T - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}(\boldsymbol{x}\boldsymbol{x}^T)^T \\
& \text{(because } \boldsymbol{I}_n^T = \boldsymbol{I}_n \text{ and } (\boldsymbol{X}\boldsymbol{Y})^T = \boldsymbol{Y}^T\boldsymbol{X}^T \text{ for any square matrices } \boldsymbol{X} \text{ and } \boldsymbol{Y}) \\
=& \boldsymbol{I}_n - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}(\boldsymbol{x}^T)^T\boldsymbol{x}^T \\
=& \boldsymbol{I}_n - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T \text{ because } (\boldsymbol{x}^T)^T = \boldsymbol{x} \\
=& \boldsymbol{A}.
\end{aligned}$$

Second, we show that $\boldsymbol{A}$ is orthogonal, i.e., $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_n$. By Part (a), it is equivalent to show that $\boldsymbol{A}^2 = \boldsymbol{I}_n$:

$$\begin{aligned}
\boldsymbol{A}^2 =& (\boldsymbol{I}_n - \frac{2}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T)^2 \\
& \text{(because } (I - \boldsymbol{X})^2 = \boldsymbol{I} - 2\boldsymbol{X} + \boldsymbol{X}^2) \\
=& \boldsymbol{I}_n - \frac{4}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T + (\frac{2}{\boldsymbol{x}^T\boldsymbol{x}})^2(\boldsymbol{x}\boldsymbol{x}^T)^2 \\
=& \boldsymbol{I}_n - \frac{4}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T + \frac{4}{(\boldsymbol{x}^T\boldsymbol{x})^2}(\boldsymbol{x}(\boldsymbol{x}^T\boldsymbol{x})\boldsymbol{x}^T) \\
& \text{(because } \boldsymbol{x}^T\boldsymbol{x} \text{ is a scalar, i.e., } \boldsymbol{x}^T\boldsymbol{x} \in \mathbb{R}) \\
=& \boldsymbol{I}_n - \frac{4}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T + \frac{4}{\boldsymbol{x}^T\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}^T \\
=& \boldsymbol{I}_n.
\end{aligned}$$

7. Let $\boldsymbol{A}$ be a square matrix of order n and $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$ be its characteristic polynomial.

   (a) Show that $\det(\boldsymbol{A}) = (-1)^n a_n$.

   Recall that $p(\lambda) = \det(\lambda I_n - \boldsymbol{A})$. Since $p(0) = a_n$ and $p(0) = \det(-\boldsymbol{A})$, we have

$$\det(-\boldsymbol{A}) = (-1)^n \det(\boldsymbol{A}) = a_n \iff \det(\boldsymbol{A}) = (-1)^n a_n.$$

   (b) Show that $p(\lambda)$ is also the characteristic polynomial of $\boldsymbol{A}^T$ and $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ for any invertible matrix $\boldsymbol{P}$.

The characteristic polynomials of $\boldsymbol{A}^T$ and $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ are given by $\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}^T)$ and $\det(\lambda \boldsymbol{I}_n - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})$:

$$\begin{aligned}
&\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}^T) = \det((\lambda \boldsymbol{I}_n - \boldsymbol{A}^T)^T) = \det(\lambda \boldsymbol{I}_n^T - \boldsymbol{A}) = \det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = p(\lambda) \\
&\det(\lambda \boldsymbol{I}_n - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}) = \det(\boldsymbol{P}^{-1}(\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{P}) = \det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = p(\lambda).
\end{aligned}$$

(c) **(1 point)** Let

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}.$$

Can we find an invertible matrix $\boldsymbol{P}$ with all real entries diagonalizing $\boldsymbol{A}$? Please justify.

Remark. We have not learned the diagonalizable algorithm until Lecture 20. Homework 4 is supposed to cover Lectures 15–19. So I will present a more elementary way to solve this problem. It is OK to use any other method if you can get the correct answer.

Yes. The reason is the following. Assume

$$\boldsymbol{P} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}.$$

We need to solve for $\boldsymbol{P}$ satisfying

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \iff \boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Then we have a matrix equation:

$$\begin{pmatrix} 2x+z & w+2y \\ -3x-2z & -2w-3y \end{pmatrix} = \begin{pmatrix} \lambda_1 x & \lambda_2 y \\ \lambda_1 z & \lambda_2 w \end{pmatrix},$$

which gives us a system of linear equations:

$$\begin{cases} 2x + z = \lambda_1 x \\ w + 2y = \lambda_2 y \\ -3x - 2z = \lambda_1 z \\ -2w - 3y = \lambda_2 w \end{cases}$$

Next, we have the following two linear systems:

$$\begin{cases} (2-\lambda_1)x + z = 0 \\ -3x - (2+\lambda_1)z = 0 \end{cases} \tag{1}$$

$$\begin{cases} w + (2-\lambda_2)y = 0 \\ -(2+\lambda_2)w - 3y = 0 \end{cases} \tag{2}$$

Since P is invertible, the homogeneous linear systems (1) and (2) have non-trivial solutions, which implies

$$\begin{vmatrix} (2-\lambda_1) & 1 \\ -3 & -(2+\lambda_1) \end{vmatrix} = \begin{vmatrix} 1 & (2-\lambda_2) \\ -(2+\lambda_2) & -3 \end{vmatrix} = 0.$$

Hence, $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$.
Obviously, $\lambda_1 \neq \lambda_2$. Otherwise,

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \pm \boldsymbol{I}_2 \iff \boldsymbol{A} = \pm \boldsymbol{I}, \text{ a contradiction.}$$

Let us take $\lambda_1 = 1$ and $\lambda_2 = -1$. Under this assumption, we solve the homogeneous linear systems (1) and (2)

$$\begin{cases} x &= -t \\ z &= t \end{cases} \qquad \begin{cases} y &= -s/3 \\ w &= s \end{cases}$$

And

$$\boldsymbol{P} = \begin{pmatrix} -t & -s/3 \\ t & s \end{pmatrix}.$$

Now, we only need to choose t and s such that $\boldsymbol{P}$ is invertible. For example

$$\boldsymbol{P} = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}.$$