Homework 1

Math and Statistics Foundations for Data Science

2022-24790 Sungwoo PARK

Graduate School of Data Science

March 22, 2022

Exercise 2.4

(a)

Answer: Multiplication is impossible.

Let the left matrix be $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix} \in \mathbb{R}^{3\times 2}$ and the right one be $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. For matrix multiplication to be possible, the number of columns of the left matrix must be equal to the number of rows of the right matrix. However, $\boldsymbol{A}$ has two columns while $\boldsymbol{B}$ has three rows, thus unequal. Therefore, the multiplication between $\boldsymbol{A}$ and $\boldsymbol{B}$ is impossible.

(b)

Answer: Possible, and the answer is

$$\begin{pmatrix} 4 & 3 & 5 \\ 10 & 9 & 11 \\ 12 & 9 & 15 \end{pmatrix}$$

(c)

Answer: Possible, and the answer is

$$\begin{pmatrix} 5 & 7 & 9 \\ 11 & 13 & 15 \\ 8 & 10 & 12 \end{pmatrix}$$

Exercise 2.5

(a)

Answer: No solution, *i.e.* $S = \emptyset$, where S denotes the set of solutions.

Following the Gaussian Elimination below, we can figure out that there is no such $\boldsymbol{x}$ satisfying the last row of the row echelon form.

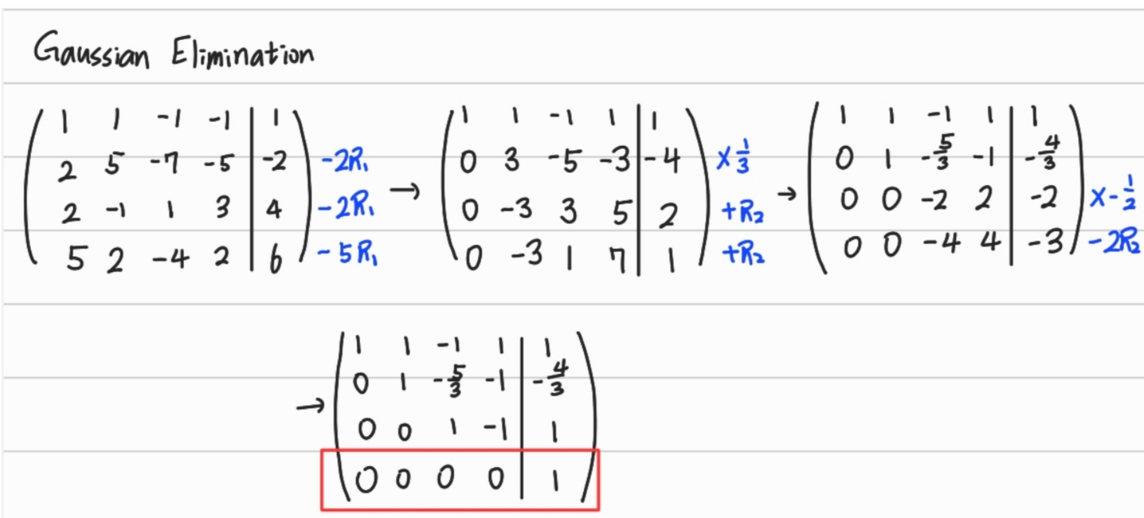

Figure 1: Gaussian Elimination of Augmented Matrix $[\boldsymbol{A} \mid \boldsymbol{b}]$

Exercise 2.8

(a)

Answer: Inverse does not exist.

First, let's obtain the determinant of $\boldsymbol{A}$. Since $\boldsymbol{A} \in \mathbb{R}^{3\times3}$, let's apply the Rule of Sarrus.

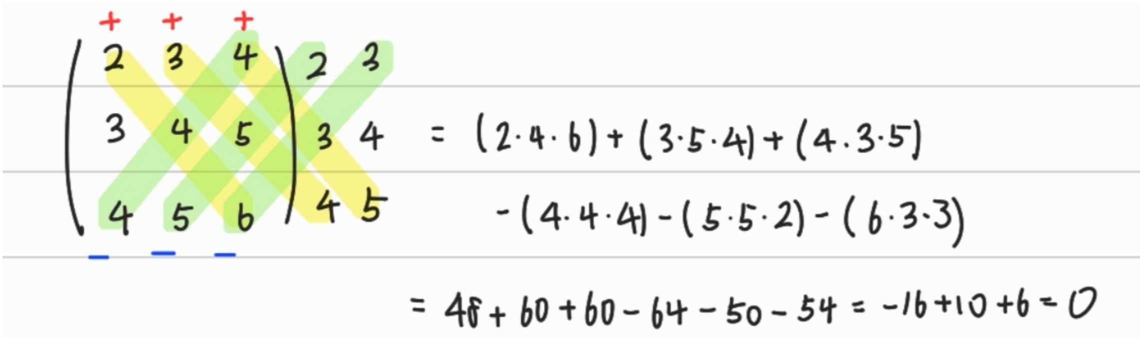

Figure 2: Determinant of $\boldsymbol{A}$ obtained using the Rule of Sarrus

As we can see, since the determinant of $\boldsymbol{A}$ equals to 0, the inverse does not exist.

Exercise 2.10

(a)

Answer: Not linearly independent, *i.e.* linearly dependent.

Let's make a matrix, putting all three vectors together inside, and denote the matrix by $\boldsymbol{A}$. Then,

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 1 & -3 \\ 3 & -2 & 8 \end{pmatrix} \in \mathbb{R}^{3\times3}$$

Let's make a row echelon form of $\boldsymbol{A}$ using the Gaussian Elimination.

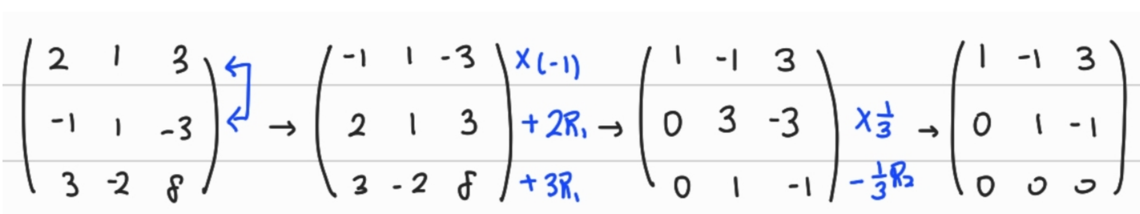

Figure 3: Row echelon form of $\boldsymbol{A}$ using the Gaussian Elimination

As we can see here, the last row of the row echelon form is the zero vector. This implies that the row echelon form is not of full rank, so is $\boldsymbol{A}$. By definition, a square matrix is of full rank if every column vector is linearly independent to one another. Since $\boldsymbol{A}$ is not of full rank, this means that the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ are not linearly independent.

(b)

Answer: Linearly independent.

Like above, make a matrix putting all three vectors together inside, and denote the matrix by $\boldsymbol{B}$. Then,

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{5\times 3}$$

Let's make a row echelon form of $\boldsymbol{B}$ using the Gaussian Elimination.

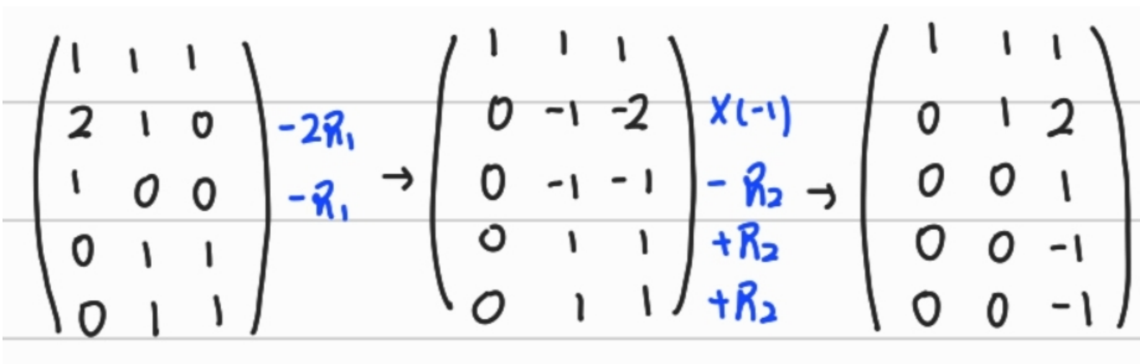

Figure 4: Row echelon form of $\boldsymbol{B}$ using the Gaussian Elimination

As we can see here, the first three rows of the row echelon form are linearly independent. By definition, for a matrix $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ is of full rank if the number of linearly independent column(or row) vectors are equal to $\min\{m, n\}$. Therefore, $\boldsymbol{B}$ is of full rank, which implies that all three vectors are linearly independent.

Exercise 2.11

Answer: $\boldsymbol{y} = -6\boldsymbol{x}_1 + 3\boldsymbol{x}_2, +2\boldsymbol{x}_3$

Let's represent $\boldsymbol{y}$ as linear combination of $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$. Then,

$$\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 + 2\lambda_3 \\ \lambda_1 + 2\lambda_2 - \lambda_3 \\ \lambda_1 + 3\lambda_2 + \lambda_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$$

Thus, this problem is equivalent to solving the system of linear equations. Let's apply the Gaussian Elimination on the augmented matrix.

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 1 & 2 & -1 & -2 \\ 1 & 3 & 1 & 5 \end{array}\right) \begin{array}{l} \\ -R_1 \\ -R_1 \end{array} \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & -3 & -3 \\ 0 & 2 & -1 & 4 \end{array}\right) \begin{array}{l} \\ \\ -2R_2 \end{array} \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & 5 & 10 \end{array}\right)$$

$$\rightsquigarrow \begin{aligned} \lambda_1 + \lambda_2 + 2\lambda_3 &= 1 \\ \lambda_2 - 3\lambda_3 &= -3 \\ 5\lambda_3 &= 10 \end{aligned} \qquad \therefore \lambda_3 = 2, \ \lambda_2 = 3, \ \lambda_1 = -6$$

Figure 5: Solution of the system of linear equations

Exercise 2.14

$$U_1 = \operatorname{span}\left[\begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 3 \\ 1 \end{pmatrix}\right], \quad U_2 = \operatorname{span}\left[\begin{pmatrix} 3 \\ 1 \\ 7 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \\ -5 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 2 \\ 2 \end{pmatrix}\right]$$

(a)

Answer: $\dim(U_1) = \dim(U_2) = 2$

The dimension of a span equals to the cardinality of a basis, or in other words, the number of linearly independent vectors of a generating set. Since both U_1 and U_2 are spanned by the columns of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively, the dimension of a span equals to the rank of a matrix.

For $\boldsymbol{A}_1$, the third column can be represented as linear combination of the other two columns, where

$$\begin{pmatrix}1\\1\\2\\1\end{pmatrix}+\begin{pmatrix}0\\-2\\1\\0\end{pmatrix}=\begin{pmatrix}1\\-1\\3\\1\end{pmatrix}$$

Thus $\operatorname{rank}(\boldsymbol{A}_1)=2$, so is the dimension of U_1.

Likewise, the third column of $\boldsymbol{A}_2$ can be represented as linear combination of the other columns as well. Therefore, $\operatorname{rank}(\boldsymbol{A}_2)=2$, so is the dimension of U_2.

(b)

Since $\dim(U_1)=\dim(U_2)=2$, we can make a basis by choosing any two column vectors of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively. Let's denote a basis of U_1 and U_2 by $\mathcal{B}_{U_1}$ and $\mathcal{B}_{U_2}$, respectively. Then,

$$\mathcal{B}_{U_1}=\left\{\begin{pmatrix}1\\1\\2\\1\end{pmatrix},\begin{pmatrix}0\\-2\\1\\0\end{pmatrix}\right\}\text{ and }\mathcal{B}_{U_2}=\left\{\begin{pmatrix}3\\1\\7\\3\end{pmatrix},\begin{pmatrix}-3\\2\\-5\\-1\end{pmatrix}\right\}$$

Exercise 2.19

(a)

Answer: $\ker(\Phi)=\{\mathbf{0}\}$ and $\operatorname{Im}(\Phi)=\mathbb{R}^3$

$\ker(\Phi)$ is a subset of the domain, which is $\mathbb{R}^3$ here, such that satisfies $\boldsymbol{A}_\Phi\boldsymbol{x}=\mathbf{0}$. By the Gaussian Elimination, we can obtain the reduced row echelon form of $\boldsymbol{A}_\Phi$, which equals to $\boldsymbol{I}_3$.

$$\begin{pmatrix}1&1&0\\1&-1&0\\1&1&1\end{pmatrix}\begin{matrix}\\-R_1\\-R_1\end{matrix}\rightarrow\begin{pmatrix}1&1&0\\0&-2&0\\0&0&1\end{pmatrix}\begin{matrix}+\frac{1}{2}R_2\\\times(-\frac{1}{2})\\ \\\end{matrix}\rightarrow\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=I_3$$

Figure 6: Reduced row echelon form of $\boldsymbol{A}_\Phi$

Therefore, the equation $\boldsymbol{A}_\Phi\boldsymbol{x}=\mathbf{0}$ has the unique solution: $x_1=x_2=x_3=0$, which makes $\ker(\Phi)=\{\mathbf{0}\}$ and thus $\dim(\ker(\Phi))=0$.

There is a theorem called the dimension theorem, which specifies the relationship between the dimension of the kernel and the dimension of the image of a linear mapping. Specifically, for vector spaces V, W and a linear mapping $\Phi : V \to W$,

$$\dim(V) = \dim(\ker(\Phi)) + \dim(\mathrm{Im}(\Phi))$$

Following this theorem, $\dim(\mathrm{Im}(\Phi)) = \dim(\mathbb{R}^3) - \dim(\ker(\Phi)) = 3 - 0 = 3$. This implies that the set of column vectors of $\boldsymbol{A}_\Phi$ is a basis of $\mathbb{R}^3$. Therefore, $\mathrm{Im}(\Phi) = \mathbb{R}^3$.

Exercise 3.5

(a)

First, let's determine the dimension of U by applying the Gaussian Elimination on the matrix composed of the vectors in U.

$$\begin{pmatrix} 0 & 1 & -3 & -1 \\ -1 & -3 & 4 & -3 \\ 2 & 1 & 1 & 5 \\ 0 & -1 & 2 & 0 \\ 2 & 2 & 1 & 7 \end{pmatrix} \to \begin{array}{c} \begin{pmatrix} -1 & -3 & 4 & -3 \\ 0 & 1 & -3 & -1 \\ 2 & 1 & 1 & 5 \\ 0 & -1 & 2 & 0 \\ 2 & 2 & 1 & 7 \end{pmatrix} \end{array} \begin{array}{l} \times(-1) \\ \\ +2R_1 \\ \\ +2R_1 \end{array} \to \begin{pmatrix} 1 & 3 & -4 & 3 \\ 0 & 1 & -3 & -1 \\ 0 & -5 & 9 & -1 \\ 0 & -1 & 2 & 0 \\ 0 & -4 & 9 & 1 \end{pmatrix} \begin{array}{l} \\ \\ +5R_2 \\ +R_2 \\ +4R_2 \end{array}$$

$$\to \begin{pmatrix} 1 & 3 & -4 & 3 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & -6 & -6 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -3 & -3 \end{pmatrix} \begin{array}{l} \\ \\ \times -\frac{1}{6} \\ -\frac{1}{6}R_3 \\ -\frac{1}{2}R_3 \end{array} \to \begin{pmatrix} 1 & 3 & -4 & 3 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 7: Gaussian Elimination on the matrix $\boldsymbol{U}$

As we can see here, $\mathrm{rank}(\boldsymbol{U}) = 3$, thus $\dim(U) = 3$ as well. Now, let $\mathcal{B}_{\mathcal{U}}$ be a basis of U, then the orthogonal projection $\pi_U(\boldsymbol{x})$ can be represented as linear combination of the basis.

$$\mathcal{B}_U = \left\{ \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 5 \\ 0 \\ 7 \end{pmatrix} \right\} \text{ and } \pi_U(\boldsymbol{x}) = \lambda_1 \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ -3 \\ 5 \\ 0 \\ 7 \end{pmatrix}$$

Then let $\boldsymbol{B}$ denote a matrix putting all elements of $\mathcal{B}_U$. Since the inner product is defined as the dot product here, the vector of lambda, λ, can be obtained by the formula $\left(\boldsymbol{B}^T\boldsymbol{B}\right)^{-1}\boldsymbol{B}^T\boldsymbol{x}$. The below part is a Python code snippet to obtain the vector λ.

```python
import numpy as np

## define B and X
B = np.array([[0, 1, -1],
              [-1, -3, -3],
              [2, 1, 5],
              [0, -1, 0],
              [2, 2, 7]], dtype=np.float64)
X = np.array([[-1, -9, -1, 4, 1]], dtype=np.float64).T

## obtain the inverse part of the projection formula
inv_part = np.linalg.inv(B.T.dot(B))
mat_part = inv_part.dot(B.T) # equivalent to (B^T*B)^(-1)*B^T
lbda = mat_part.dot(X)
print(lbda) # -4, 2, 1

## obtain the projection
pi = np.zeros(shape=(B.shape[0], ))
for i in range(len(lbda)):
    print(f"lambda:{lbda[i][0]}, vector:{B[:, i]}")
    pi += lbda[i][0] * B[:, i]

pi = pi.reshape(X.shape)
print(pi) # 1, -5, -1, -2, 3
```

Figure 8: Screenshot of the Python code snippet

Thus the vector $\lambda = \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix}$, and finally the projection $\pi_U\left(\boldsymbol{x}\right) = \begin{pmatrix} 1 \\ -5 \\ -1 \\ -2 \\ 3 \end{pmatrix}$

(b)

Since $\pi_U(\boldsymbol{x})$ is the orthogonal projection onto U, the distance between $\boldsymbol{x}$ and U equals to the length(or sqaure root of L_2 norm) of the vector $\boldsymbol{x} - \pi_U(\boldsymbol{x}) = \begin{pmatrix} -2 \\ -4 \\ 0 \\ 6 \\ -2 \end{pmatrix}$. Continuing to the code snippet above, we can obtain the length of $\boldsymbol{x} - \pi_U(\boldsymbol{x})$ as below.

```python
def l2_norm(X:np.ndarray) -> float:
    X = np.array(X, dtype=np.float64)
    self_dot = X.T.dot(X).flatten() # returns 1d array
    return self_dot[0]

diff = X - pi
print(l2_norm(diff))
```

Figure 9: Screenshot of the Python code snippet

Since the L_2 norm of $\boldsymbol{x} - \pi_U(\boldsymbol{x})$ equals to 60, the length of the vector, or the distance between $\boldsymbol{x}$ and U, equals to $\sqrt{60} \approx 7.746$.