DEPARTMENT OF MATHEMATICS, IIT Guwahati

MA102: Mathematics II

Lecture Notes

Note: If a proof of a result or a solution of an example is not written in this summary of lecture notes, then you are advised to find the same in the recommended text book (Linear Algebra: A Modern Introduction by David Poole). Most of the materials in this summary have been extracted from this text book. However, a few results and examples have also been taken from other sources. Proofs/solutions of such results/examples have either been provided or relatively easier to be done by yourself.

1 System of Linear Equations and Matrices

An Example for Motivation:

To solve the system of linear equations: $x - y - z = 2,\ 3x - 3y + 2z = 16,\ 2x - y + z = 9$.
We solve this system by eliminating the variables.

Step 1: Represent the given system of equations in the rectangular array form as follows.

$$\begin{array}{rcrcrcr} x & - & y & - & z & = & 2 \\ 3x & - & 3y & + & 2z & = & 16 \\ 2x & - & y & + & z & = & 9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Step 2: Subtract 3 times the 1st equation from the 2nd equation; and
subtract* 3 *times the 1st row from the 2nd row.

$$\begin{array}{rcrcrcr} x & - & y & - & z & = & 2 \\ & & & & 5z & = & 10 \\ 2x & - & y & + & z & = & 9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Step 3: Subtract 2 times the 1st equation from the 3rd equation; and
subtract* 2 *times the 1st row from the 3rd row.

$$\begin{array}{rcrcrcr} x & - & y & - & z & = & 2 \\ & & & & 5z & = & 10 \\ & & y & + & 3z & = & 5 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right]$$

Step 4: Interchange the 2nd and 3rd equation; and ***interchange the 2nd and 3rd row.***

$$\begin{array}{rcrcrcr} x & - & y & - & z & = & 2 \\ & & y & + & 3z & = & 5 \\ & & & & 5z & = & 10 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right]$$

Now by backward substitution, we find that $z = 2, y = -1, x = 3$ is a solution of the given system of equations.

A **linear system of m equations** in n unknowns $x_1, x_2, \dots, x_n$ is a set of equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad \vdots\ &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{1}$$

where $a_{ij}, b_i \in \mathbb{C}$ for each $1 \leq i \leq m$ and $1 \leq j \leq n$. Letting

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

we can represent the above system of equations as $A\mathbf{x} = \mathbf{b}$.

It is now clear that matrices are quite handy for expressing a system of linear equations in compact form.

Definition:

- A rectangular array of (complex) numbers is called a **matrix**. Formally, an $m \times n$ matrix $A = [a_{ij}]$ is an array of numbers in m rows and n columns as shown below:
$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$
The matrix A is called a matrix of **size** $m \times n$ or a matrix of **order** $m \times n$.

- The number a_{ij} is called the (i, j)-th entry of A.

- A $1 \times n$ matrix is called a **row matrix** (or *row vector*) and an $n \times 1$ matrix is called a **column matrix** (or *column vector*).

- Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are said to be **equal** if they are of same size and $a_{ij} = b_{ij}$ for each $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

- An $m \times n$ matrix is called a **zero matrix** of size $m \times n$, denoted $\mathbf{O}_{m\times n}$ (or simply $\mathbf{O}$), if all the entries are equal to 0.

- If $m = n$, then A is called a **square** matrix.

- If A is a square matrix, then the entries a_{ii} are called the **diagonal** entries of A.

- If A is a square matrix and if $a_{ij} = 0$ for all $i \neq j$, then A is called a **diagonal matrix**.

- If an $n \times n$ diagonal matrix has all diagonal entries equal to 1, then it is called the **identity matrix** of size n, and is denoted by I_n (or simply by I).

- A matrix B is said to be a **sub matrix** of A if B is obtained by deleting some rows and/or columns of A.

- Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices. Then the **sum** $A + B$ is defined to be the matrix $C = [c_{ij}]$, where $c_{ij} = a_{ij} + b_{ij}$. Similarly, the **difference** $A - B$ is defined to be the matrix $D = [d_{ij}]$, where $d_{ij} = a_{ij} - b_{ij}$.

- For a matrix $A = [a_{ij}]$ and $c \in \mathbb{C}$ (**set of complex numbers**), we define cA to be the matrix $[ca_{ij}]$.

- Let $A = [a_{ij}]$ and $B = [b_{jk}]$ be two $m \times n$ and $n \times r$ matrices, respectively. Then the **product** AB is defined to be the $m \times r$ matrix $AB = [c_{ik}]$, where
$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk}.$$

- The **transpose** A^t of an $m \times n$ matrix $A = [a_{ij}]$ is defined to be the $n \times m$ matrix $A^t = [a_{ji}]$, where the i-th row of A^t is the i-th column of A for all $i = 1, 2, \dots, n$.

- The matrix A is said to be **symmetric** if $A^t = A$, and **skew-symmetric** if $A^t = -A$.

- If A is a complex matrix, then $\overline{A} = [\overline{a}_{ij}]$ and $A^* = \overline{A}^t$.

- The matrix A^* is called the **conjugate transpose** of A.

- The (complex) matrix A is said to be **Hermitian** if $A^* = A$, and **skew-Hermitian** if $A^* = -A$.

- A square matrix A is said to be **upper triangular** if $a_{ij} = 0$ for all $i > j$.

- A square matrix A is said to be **lower triangular** if $a_{ij} = 0$ for all $i < j$.

- Let A be an $n \times n$ square matrix. Then we define $A^0 = I_n, A^1 = A$ and $A^2 = AA$.

- In general, if k is a positive integer, we define the power A^k as follows
$$A^k = \underbrace{AA \ldots A}_{k \text{ times}}.$$

It is obvious to see that if A and $\mathbf{O}$ are matrices of the same size, then $A+\mathbf{O} = A = \mathbf{O}+A$, $A-\mathbf{O} = A$, $\mathbf{O}-A = -A$ and $A - A = \mathbf{O}$.

Unless otherwise mentioned, all the matrices will be taken to have complex numbers as entries.

Method of mathematical induction is an useful tool for proving many results of mathematics. We now present two equivalent versions of the method of induction.

Method of Induction: [Version I] Let $P(n)$ be a mathematical statement based on all positive integers n. Suppose that $P(1)$ is true. If $k \geq 1$ and if the assumption that $P(k)$ is true gives that $P(k+1)$ is also true, then the statement $P(n)$ is true for all positive integers.

Method of Induction: [Version II] Let i be an integer and let $P(n)$ be a mathematical statement based on all integers n of the set $\{i, i+1, i+2, \ldots\}$. Suppose that $P(i)$ is true. If $k \geq i$ and if the assumption that $P(k)$ is true gives that $P(k+1)$ is also true, then the statement $P(n)$ is true for all integers of the set $\{i, i+1, i+2, \ldots\}$.

Result 1.1. *Let A be an $m \times n$ matrix, $\mathbf{e}_i$ an $1 \times m$ standard unit row vector, and $\mathbf{e}_j$ an $n \times 1$ standard unit column vector. Then $\mathbf{e}_i A$ is the i-th row of A and $A\mathbf{e}_j$ is the j-th column of A.*

Result 1.2. *Let A be a square matrix and let r and s be non-negative integers. Then $A^r A^s = A^{r+s}$ and $(A^r)^s = A^{rs}$.*

Proof. We have $A^r A^s = (\underbrace{AA\ldots A}_{r \text{ times}})\underbrace{AA\ldots A}_{s \text{ times}} = \underbrace{AA\ldots A}_{r+s \text{ times}} = A^{r+s}$.

Also $(A^r)^s = \underbrace{A^r A^r \ldots A^r}_{s \text{ times}} = \underbrace{AA\ldots A}_{rs \text{ times}} = A^{rs}$. □

Result 1.3. *Let A, B and C be matrices of size $m \times n$, and let $s, r \in \mathbb{C}$. Then*

1. ***Commutative Law**: $A + B = B + A$.*
2. ***Associative Law**: $(A + B) + C = A + (B + C)$.*
3. *$1A = A$, $\quad s(rA) = (sr)A$.*
4. *$s(A + B) = sA + sB$ and $(s + r)A = sA + rA$.*

Proof. Let $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]$ and $s, r \in \mathbb{C}$. We have

1. ij-th entry of $A + B = a_{ij} + b_{ij} = b_{ij} + a_{ij} = ij$-th entry of $B + A$. Hence $A + B = B + A$.
2. ij-th entry of $(A+B)+C = (a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}) = ij$-th entry of $A+(B+C)$. Hence $(A+B)+C = A + (B + C)$.
3. ij-th entry of $1A = 1.a_{ij} = a_{ij} = ij$-th entry of A. Hence $1A = A$.
   Also ij-th entry of $s(rA) = s(ra_{ij}) = (sr)a_{ij} = ij$-th entry of $(sr)A$. Hence $s(rA) = (sr)A$.
4. ij-th entry of $s(A + B) = s(a_{ij} + b_{ij}) = sa_{ij} + sb_{ij} = ij$-th entry of $sA + sB$. Hence $s(A + B) = sA + sB$.
   Also ij-th entry of $(s + r)A = (s + r)a_{ij} = sa_{ij} + ra_{ij} = ij$-th entry of $(sr)A$. Hence $s(rA) = (sr)A$. □

Result 1.4. *Let A, B and C be matrices, and let $s \in \mathbb{C}$. Then*

1. ***Associative Law**: $(AB)C = A(BC)$, if the respective matrix products are defined.*
2. ***Distributive Law**: $A(B + C) = AB + AC$, $\quad (A + B)C = AC + BC$, if the respective matrix sum and matrix products are defined.*
3. *$s(AB) = (sA)B = A(sB)$, if the respective matrix products are defined.*
4. *$I_m A = A = AI_n$, if A is of size $m \times n$.*

Proof. Let $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]$ and $s \in \mathbb{C}$. Let $I_n = [e_{ij}]$ so that $e_{ij} = 1$ if $i = j$ and $e_{ij} = 0$ otherwise.

1. Let the orders of A, B, C be $m \times p, p \times q, q \times n$, respectively. Then the ij-th entry of $(AB)C$ is equal to
   $\sum\limits_{k=1}^{q} \left(\sum\limits_{r=1}^{p} a_{ir} b_{rk} \right) c_{kj} = \sum\limits_{r=1}^{p} a_{ir} \left(\sum\limits_{k=1}^{q} b_{rk} c_{kj} \right) = ij$-th entry of $A(BC)$. Hence $(AB)C = A(BC)$.

2. Let the orders of A, B, C be $m \times p, p \times n, p \times n$, respectively. Then
   ij-th entry of $A(B+C) = \sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj} = ij$-th entry of $AB + AC$. Hence $A(B+C) = AB + AC$.

   Again let the orders of A, B, C be $m \times p, m \times p, p \times n$, respectively. Then
   ij-th entry of $(A+B)C = \sum_{k=1}^{p} (a_{ik} + b_{ik})c_{kj} = \sum_{k=1}^{p} a_{ik}c_{kj} + \sum_{k=1}^{p} b_{ik}c_{kj} = ij$-th entry of $AC + BC$. Hence $(A+B)C = AC + BC$.

3. Let the orders of A, B be $m \times p, p \times n$, respectively. Then
   ij-th entry of $s(AB) = s\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right) = \sum_{k=1}^{p} (sa_{ik})b_{kj} = ij$-th entry of $(sA)B$. Hence $s(AB) = (sA)B$. Again
   ij-th entry of $s(AB) = s\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right) = \sum_{k=1}^{p} a_{ik}(sb_{kj}) = ij$-th entry of $A(sB)$. Hence $s(AB) = A(sB)$.

4. Let the order of A be $m \times n$. Then
   ij-th entry of $I_mA = \sum_{k=1}^{m} e_{ik}a_{kj} = e_{ii}a_{ij} = a_{ij} = ij$-th entry of A. Hence $I_mA = A$. Again
   ij-th entry of $AI_n = \sum_{k=1}^{n} a_{ik}e_{kj} = a_{ij}e_{jj} = a_{ij} = ij$-th entry of A. Hence $AI_n = A$. □

Result 1.5. *Let A and B be two matrices and $k \in \mathbb{C}$. Then*

1. $(A^t)^t = A, \quad (kA)^t = kA^t$.
2. $(A+B)^t = A^t + B^t$ *if A and B are of the same size.*
3. $(AB)^t = B^tA^t$ *if the matrix product AB is defined.*
4. $(A^r)^t = (A^t)^r$ *for any non-negative integer r.*

Proof. Let $A = [a_{ij}], B = [b_{ij}], A^t = [a'_{ij}], B^t = [b'_{ij}]$ and $k \in \mathbb{C}$ so that $a'_{ij} = a_{ji}$ and $b'_{ij} = b_{ji}$.

1. ij-th entry of $(A^t)^t = ji$-th entry of $A^t = ij$-th entry of A. Hence $(A^t)^t = A$. Again
   ij-th entry of $(kA)^t = ji$-th entry of $kA = ka_{ji} = ka'_{ij} = ij$-th entry of kA^t. Hence $(kA)^t = kA^t$.

2. ij-th entry of $(A+B)^t = ji$-th entry of $A + B = a_{ji} + b_{ji} = a'_{ij} + b'_{ij} = ij$-th entry of $A^t + B^t$. Hence $(A+B)^t = A^t + B^t$.

3. ij-th entry of $(AB)^t = ji$-th entry of $AB = \sum_{k=1}^{p} a_{jk}b_{ki} = \sum_{k=1}^{p} b'_{ik}a'_{kj} = ij$-th entry of B^tA^t. Hence $(AB)^t = B^tA^t$.

4. We use induction on r. For $r = 0$, it is clear that $(A^0)^t = I = (A^t)^0$. Now let us assume that $(A^k)^t = (A^t)^k$ for a $k \geq 0$. Then applying this assumption as well as Part 3, we have
   $$(A^{k+1})^t = (A^k.A)^t = A^t(A^k)^t = A^t(A^t)^k = (A^t)^{k+1}.$$
   Thus $(A^r)^t = (A^t)^r$ is also true for $r = k + 1$. Hence by principle of mathematical induction, we have that $(A^r)^t = (A^t)^r$ for any non-negative integer r. □

Partitioned Matrix: By introducing vertical and horizontal lines into a given matrix, we can partition it into some blocks of smaller sub-matrices. A given matrix can have several partitions possible. For example, three partitions of the matrix A are given below:

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 5 \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right] = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ \hline 0 & 0 & 1 & 5 \end{array}\right].$$

- A **block matrix** $A = [A_{ij}]$ is a matrix, where each entry A_{ij} is itself a matrix.
- Thus a partition of a given matrix give us a block matrix.

- For example, the first partition of the previous matrix A is the block matrix $\begin{bmatrix} I & B \\ \mathbf{O} & C \end{bmatrix}$, where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{O} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$.

Result 1.6. *Let A and B be two matrices of sizes $m \times n$ and $n \times r$, respectively.*

1. *If $B = [\mathbf{b}_1 \mid \mathbf{b}_2 \mid \ldots \mid \mathbf{b}_r]$, where $\mathbf{b}_j$ is the j-th column of B then $AB = [A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \ldots \mid A\mathbf{b}_r]$.*

2. *If $A = \begin{bmatrix} \mathbf{a}_1 \\ \hline \mathbf{a}_2 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{bmatrix}$, where $\mathbf{a}_i$ is the i-th row of A then $AB = \begin{bmatrix} \mathbf{a}_1 B \\ \hline \mathbf{a}_2 B \\ \hline \vdots \\ \hline \mathbf{a}_m B \end{bmatrix}$.*

3. *If $m = n = r$ and if $A = [\mathbf{a}_1^t \mid \mathbf{a}_2^t \mid \ldots \mid \mathbf{a}_n^t]$ and $B = \begin{bmatrix} \mathbf{b}_1^t \\ \hline \mathbf{b}_2^t \\ \hline \vdots \\ \hline \mathbf{b}_n^t \end{bmatrix}$, where $\mathbf{a}_k^t$ is the k-th column of A and $\mathbf{b}_k^t$ is the k-th row of B then $AB = \mathbf{a}_1^t\mathbf{b}_1^t + \mathbf{a}_2^t\mathbf{b}_2^t + \ldots + \mathbf{a}_n^t\mathbf{b}_n^t$.*

Proof. Let $A = [a_{ij}], B = [b_{ij}]$ of sizes $m \times n$ and $n \times r$, respectively.

1. Here $\mathbf{b}_j = [b_{1j}, b_{2j}, \ldots, b_{ij}, \ldots, b_{nj}]^t$. We have

$$j\text{-th column of } AB = \begin{bmatrix} \sum\limits_{k=1}^{n} a_{1k}b_{kj} \\ \sum\limits_{k=1}^{n} a_{2k}b_{kj} \\ \vdots \\ \sum\limits_{k=1}^{n} a_{mk}b_{kj} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = A\mathbf{b}_j.$$

   Hence $AB = [A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \ldots \mid A\mathbf{b}_r]$.

2. Here $\mathbf{a}_i = [a_{i1} \;\; a_{i2} \;\; \ldots \;\; a_{ij} \;\; \ldots \;\; a_{in}]$. We have

$$\begin{aligned} i\text{-th row of } AB &= \left[\sum_{k=1}^{n} a_{ik}b_{k1} \quad \sum_{k=1}^{n} a_{ik}b_{k2} \quad \ldots\ldots \quad \sum_{k=1}^{n} a_{ik}b_{kj} \quad \ldots\ldots \quad \sum_{k=1}^{n} a_{ik}b_{kr}\right] \\ &= [a_{i1} \;\; a_{i2} \;\; \ldots \;\; a_{ij} \;\; \ldots \;\; a_{in}] \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1r} \\ b_{21} & b_{22} & \ldots & b_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nr} \end{bmatrix} \\ &= \mathbf{a}_i B. \end{aligned}$$

   Hence $AB = \begin{bmatrix} \mathbf{a}_1 B \\ \hline \mathbf{a}_2 B \\ \hline \vdots \\ \hline \mathbf{a}_m B \end{bmatrix}$.

3. Here $\mathbf{a}_k^t = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}$ and $\mathbf{b}_k^t = [b_{k1} \;\; b_{k2} \;\; \ldots \;\; b_{kj} \;\; \ldots \;\; b_{kr}]$. Now clearly the ij-th entry of $\mathbf{a}_k^t\mathbf{b}_k^t$ is $a_{ik}b_{kj}$. Therefore

$$ij\text{-th entry of } \mathbf{a}_1^t\mathbf{b}_1^t + \mathbf{a}_2^t\mathbf{b}_2^t + \ldots + \mathbf{a}_n^t\mathbf{b}_n^t = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = ij\text{-th entry of } AB.$$

   Hence $AB = \mathbf{a}_1^t\mathbf{b}_1^t + \mathbf{a}_2^t\mathbf{b}_2^t + \ldots + \mathbf{a}_n^t\mathbf{b}_n^t$. □

Practice Problems Set 1

1. Find two different 2×2 real matrices A and B such that $A^2 = \mathbf{O} = B^2$ but $A \neq \mathbf{O}, B \neq \mathbf{O}$.

2. Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -3 & 4 \\ 4 & -1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 2 & -3 \\ 2 & 1 & -1 \end{bmatrix}$. Compute $A^2 + 2AB + B^2$ and $(A+B)^2$. Are they equal? If not, give reasons.

3. Let A and B be two $n \times n$ matrices. If $AB = BA$ then show that $(AB)^m = A^m B^m$ and $(A+B)^m = \sum_{i=0}^{m} \binom{m}{i} A^{m-i} B^i$ for every $m \in \mathbb{N}$. If $AB \neq BA$ then show that these two results need not be true.

4. Show that every square matrix can be written as a sum of a symmetric and an skew-symmetric matrix. Further, show that if A and B are symmetric, then AB is symmetric if and only if $AB = BA$. Give an example to show that if A and B are symmetric then AB need not be symmetric.

5. Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 \\ -4 & 4 \end{bmatrix}$. Is there a matrix C such that $CA = B$?

6. Let $C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. Show that there exist two 2×2 matrices A and B satisfying $C = AB - BA$ if and only if $a + d = 0$.

7. Find conditions on the numbers a, b, c and d such that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ commutes with every 2×2 matrix.

8. Let $A_1, A_2, \ldots, A_n$ be matrices of the same size, where $n \geq 1$. Using mathematical induction, prove that
   (a) $(A_1 + A_2 + \ldots + A_n)^t = A_1^t + A_2^t + \ldots + A_n^t$; and
   (b) $(A_1 A_2 \ldots A_n)^t = A_n^t A_{n-1}^t \ldots A_1^t$.

9. For an $n \times n$ matrix $A = [a_{ij}]$, the trace is defined as $\text{tr}(A) = a_{11} + \ldots + a_{nn}$. If A and B are two $n \times n$ matrices then prove that $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$, $\text{tr}(AB) = \text{tr}(BA)$ and $\text{tr}(kA) = k.\text{tr}(A)$, where $k \in \mathbb{R}$. Also find an expression of $\text{tr}(AA^t)$ in terms of entries of A.

10. Let A and B be two $n \times n$ matrices. If $AB = \mathbf{O}$ then show that $\text{tr}((A+B)^k) = \text{tr}(A^k) + \text{tr}(B^k)$ for any positive integer k.

11. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Show that $\text{tr}(A^k) = \text{tr}(A^{k-1}) + \text{tr}(A^{k-2})$, for any positive integer k.

12. Let $A = [a_{ij}]$ be an $n \times n$ matrix such that $a_{ij} = \begin{cases} 1 & \text{if } i = j+1, \\ 0 & \text{otherwise.} \end{cases}$ Show that $A^n = \mathbf{O}$ but $A^l \neq \mathbf{O}$ for $1 \leq l \leq n-1$.

13. Show that the product of two lower triangular matrices is a lower triangular matrix. (A similar statement also holds for upper triangular matrices.)

14. Let A and B be two skew-symmetric matrices such that $AB = BA$. Is the matrix AB symmetric or skew-symmetric?

15. Let A and B be two $m \times n$ matrices. Prove that if $A\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$ then A is the zero matrix. Further, prove that if $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ then $A = B$.

16. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$. Show that there exist infinitely many matrices B such that $BA = I_2$. Also, show that there does not exist any matrix C such that $AC = I_3$.

17. Show that if A is a complex triangular matrix and $AA^* = A^*A$, then A is a diagonal matrix.

18. Let A be a real matrix such that $AA^t = \mathbf{O}$. Show that $A = \mathbf{O}$. Is the same true if A is a complex matrix?

Hints to Practice Problems Set 1

1. For example, $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$.

2. The reason is $AB \neq BA$.

3. Use induction on m. Find counterexamples for the last part.

4. $A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$. For the last part, take $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 3 & 0 \end{bmatrix}$.

5. Solve $CA = B$, for $C = \begin{bmatrix} a & b & c \\ x & y & z \end{bmatrix}$. One solution is $C = \begin{bmatrix} 1 & 1 & 0 \\ -4 & 0 & 0 \end{bmatrix}$.

6. For the first part, use $\operatorname{tr}(C) = \operatorname{tr}(AB - BA)$. For the second part, if $a \neq 0$ then take $A = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & \frac{b+c}{a} \end{bmatrix}$. If $a = 0$, then solve $C = \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}\begin{bmatrix} p & q \\ r & s \end{bmatrix} - \begin{bmatrix} p & q \\ r & s \end{bmatrix}\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$ to obtain $A = \begin{bmatrix} 1 & 0 \\ 0 & 1+c \end{bmatrix}, B = \begin{bmatrix} 0 & -b/c \\ 1 & 0 \end{bmatrix}$ for $b \neq 0$ and $A = \begin{bmatrix} 1 & 0 \\ 0 & 1-b \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ -c/b & 0 \end{bmatrix}$ for $c \neq 0$.

7. $a = d,\ b = c = 0$

8. Easy.

9. $\operatorname{tr}(AA^t) = \sum\limits_{i,j=1}^{n} a_{ij}^2$.

10. Use the expansion of $(A + B)^k$ and the facts $AB = \mathbf{O},\ \operatorname{tr}(XY) = \operatorname{tr}(YX)$.

11. Use induction on k.

12. Show that A^2 has $n - 2$ non-zero rows, and so on.

13. If $[a_{ij}]$ and $[b_{ij}]$ are lower triangular matrices then show that $\sum\limits_{k=1}^{n} a_{ik}b_{kj} = 0$ for $i < j$.

14. AB is symmetric.

15. Take $\mathbf{x} = \mathbf{e}_i$ to show that the i-th column of A is zero.

16. Take $X = \begin{bmatrix} x & y & z \\ a & b & c \end{bmatrix}$ to solve the system $XA = I_2$ and $AX = I_3$.

17. Compare the (i, i)-entries of AA^* and A^*A.

18. Use Problem 9 or compare the (i, i)-entries of AA^t and $\mathbf{O}$.

2 Solutions of System of Linear Equations

Row Echelon Form: A matrix A is said to be in row echelon form if it satisfies the following properties:

1. All rows consisting entirely of 0's are at the bottom.
2. In each non-zero row, the first non-zero entry (called the **leading entry** or **pivot**) is in a column to the left (**strictly**) of any leading entry below it. [The columns containing a leading entry is called a **leading column**.]

Notice that if a matrix A is in row echelon form, then in each column of A containing a leading entry, the entries below that leading entry are zero. For example, the following matrices are in row echelon form:

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 3 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

However, the following matrices are not in row echelon form:

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 5 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 5 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Elementary Row Operations: The following row operations are called elementary row operation of a matrix:

1. Interchange of two rows R_i and R_j (shorthand notation $R_i \leftrightarrow R_j$).
2. Multiply a row R_i by a non-zero constant c (shorthand notation $R_i \to cR_i$).
3. Add a multiple of a row R_j to another row R_i (shorthand notation $R_i \to R_i + cR_j$).

Any given matrix can be reduced to a row echelon form by applying suitable elementary row operations on the matrix.

Example 2.1. *Transform the following matrix to row echelon form*

$$\begin{bmatrix} 0 & 2 & 3 & 8 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{bmatrix}.$$

Solution. We have

$$\begin{bmatrix} 0 & 2 & 3 & 8 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & -1 & -2 & -5 \\ 2 & 3 & 1 & 5 \\ 0 & 2 & 3 & 8 \end{bmatrix}$$

$$\xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & -1 & -2 & -5 \\ 0 & 5 & 5 & 15 \\ 0 & 2 & 3 & 8 \end{bmatrix}$$

$$\xrightarrow{R_2 \to \frac{1}{5}R_2} \begin{bmatrix} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 2 & 3 & 8 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 2R_2} \begin{bmatrix} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Thus a row echelon form of the given matrix is $\begin{bmatrix} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}$. □

Note that if A is in row echelon form then for any $c \neq 0$, the matrix cA is also in row echelon form. Thus a given matrix can be reduced to several row echelon forms.

Row Equivalent Matrices: Matrices A and B are said to be row equivalent if there is a finite sequence of elementary row operations that converts A into B or B into A. We will see later that if A can be converted into B through a finite sequence of elementary row operations, then B can also be converted into A through a finite sequence of elementary row operations.

Result 2.1. *Matrices A and B are row equivalent if and only if (**iff**) they can be reduced to the same row echelon form.*

Linear System of Equations: Consider the following linear system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad &\;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{2}$$

where $a_{ij}, b_i \in \mathbb{C}$ for each $1 \leq i \leq m$ and $1 \leq j \leq n$. Letting

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

we write the above system of equations as $A\mathbf{x} = \mathbf{b}$.

- The matrix A is called the **coefficient matrix** of the system of equations $A\mathbf{x} = \mathbf{b}$.
- The matrix $[A|\ \mathbf{b}]$, as given below, is called the **augmented matrix** of the system of equations $A\mathbf{x} = \mathbf{b}$.

$$[A|\ \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

  The vertical bar is used in the augmented matrix $[A|\ \mathbf{b}]$ only to separate the column vector $\mathbf{b}$ from the coefficient matrix A.

- If $\mathbf{b} = \mathbf{0} = [0, 0, \ldots, 0]^t$, i.e., if $b_1 = b_2 = \ldots = b_m = 0$, the system $A\mathbf{x} = \mathbf{0}$ is called a **homogeneous** system of equations. Otherwise, if $\mathbf{b} \neq \mathbf{0}$ then $A\mathbf{x} = \mathbf{b}$ is called a **non-homogeneous** system of equations.
- A **solution** of the linear system $A\mathbf{x} = \mathbf{b}$ is a column vector $\mathbf{y} = [y_1, y_2, \ldots, y_n]^t$ such that the linear system (2) is satisfied by substituting y_i in place of x_i. That is, $A\mathbf{y} = \mathbf{b}$ holds true.
- The solution $\mathbf{0}$ of $A\mathbf{x} = \mathbf{0}$ is called the **trivial solution** and any other solution of $A\mathbf{x} = \mathbf{0}$ is called a **non-trivial** solution.
- Two system of linear equations are called equivalent if their augmented matrices are row-equivalent.
- An elementary row operation on $A\mathbf{x} = \mathbf{b}$ is an elementary row operation on the augmented matrix $[A \mid \mathbf{b}]$.

Result 2.2. *Let $C\mathbf{x} = \mathbf{d}$ be the linear system obtained from the linear system $A\mathbf{x} = \mathbf{b}$ by a single elementary row operation. Then the linear systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions.*

Proof. Let the system $A\mathbf{x} = \mathbf{b}$ be given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad &\;\; \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad &\;\; \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n &= b_j \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad &\;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Let $[\alpha_1, \alpha_2, \ldots, \alpha_n]^t$ be a solution of $A\mathbf{x} = \mathbf{b}$. Then we have

$$\begin{aligned} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{j1}\alpha_1 + a_{j2}\alpha_2 + \ldots + a_{jn}\alpha_n &= b_j \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n &= b_m. \end{aligned} \tag{3}$$

Let $C\mathbf{x} = \mathbf{d}$ be obtained from $A\mathbf{x} = \mathbf{b}$ by a single elementary operation. We consider three cases.

<u>Case I</u>: Let $C\mathbf{x} = \mathbf{d}$ be obtained from $A\mathbf{x} = \mathbf{b}$ by applying $R_i \leftrightarrow R_j$. Then the system $C\mathbf{x} = \mathbf{d}$ is given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n &= b_j \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Notice that (3) can also be written as

$$\begin{aligned} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{j1}\alpha_1 + a_{j2}\alpha_2 + \ldots + a_{jn}\alpha_n &= b_j \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n &= b_m, \end{aligned}$$

which shows that $[\alpha_1, \alpha_2, \ldots, \alpha_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$. Since $A\mathbf{x} = \mathbf{b}$ is obtained from $C\mathbf{x} = \mathbf{d}$ by applying again $R_i \leftrightarrow R_j$, we can conclude by the same argument as above that, if $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$ then $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is also a solution of $A\mathbf{x} = \mathbf{b}$. Thus $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions.

<u>Case II</u>: Let $C\mathbf{x} = \mathbf{d}$ be obtained from $A\mathbf{x} = \mathbf{b}$ by applying $R_i \to cR_i, (c \neq 0)$. Then the system $C\mathbf{x} = \mathbf{d}$ is given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ ca_{i1}x_1 + ca_{i2}x_2 + \ldots + ca_{in}x_n &= cb_i \\ \vdots \qquad\quad \vdots \qquad\quad\ \ \vdots\ \ &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Notice that by multiplying the i-th equality in (3) by c, we find

$$\begin{aligned} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ ca_{i1}\alpha_1 + ca_{i2}\alpha_2 + \ldots + ca_{in}\alpha_n &= cb_i \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n &= b_m, \end{aligned}$$

which shows that $[\alpha_1, \alpha_2, \ldots, \alpha_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$. Since $A\mathbf{x} = \mathbf{b}$ is obtained from $C\mathbf{x} = \mathbf{d}$ by applying $R_i \to \frac{1}{c}R_i$, we can conclude by the same argument as above that, if $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$, then $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is also a solution of $A\mathbf{x} = \mathbf{b}$. Thus $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions.

Case III: Let $C\mathbf{x} = \mathbf{d}$ be obtained from $A\mathbf{x} = \mathbf{b}$ by applying $R_j \to R_j + cR_i$. Then the system $C\mathbf{x} = \mathbf{d}$ is given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ (a_{j1} + ca_{i1})x_1 + (a_{j2} + ca_{i2})x_2 + \ldots + (a_{jn} + ca_{in})x_n &= b_j + cb_i \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Notice that replacing the j-th equality in (3) by the result of addition of c times the i-th equality to the j-th equality, we find

$$\begin{aligned} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n &= b_1 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n &= b_i \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ (a_{j1} + ca_{i1})\alpha_1 + (a_{j2} + ca_{i2})\alpha_2 + \ldots + (a_{jn} + ca_{in})\alpha_n &= b_j + cb_i \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \; &\; \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n &= b_m, \end{aligned}$$

which shows that $[\alpha_1, \alpha_2, \ldots, \alpha_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$. Since $A\mathbf{x} = \mathbf{b}$ is obtained from $C\mathbf{x} = \mathbf{d}$ by applying $R_j \to R_j - cR_i$, we can conclude by the same argument as above that, if $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is a solution of $C\mathbf{x} = \mathbf{d}$ then $[\beta_1, \beta_2, \ldots, \beta_n]^t$ is also a solution of $A\mathbf{x} = \mathbf{b}$. Thus $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions. □

Result 2.3. *Two equivalent system of linear equations have the same set of solutions.*

Proof. Let $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ be two equivalent system of linear equations, so that the augmented matrices $[A \mid \mathbf{b}]$ and $[C \mid \mathbf{d}]$ are row equivalent. Then the matrix $[C \mid \mathbf{d}]$ is obtained from an application of finite number of elementary row operations on $[A \mid \mathbf{b}]$. By **Result 2.2**, we know that the solution set does not change by an application of a single elementary row operation on $A\mathbf{x} = \mathbf{b}$. Therefore the solution set does not change by an application of finite number of elementary row operations on $A\mathbf{x} = \mathbf{b}$. Hence the systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions. □

Gaussian Elimination Method: Use the following steps to solve a system of equations $A\mathbf{x} = \mathbf{b}$.

1. Write the augmented matrix $[A \mid \mathbf{b}]$.
2. Use elementary row operations to reduce $[A \mid \mathbf{b}]$ to row echelon form.
3. Use ***back substitution*** to solve the equivalent system that corresponds to the row echelon form.

Example 2.2. *Use Gaussian Elimination method to solve the system:*

(a) $y+z=1, \quad x+y+z=2, \quad x+2y+2z=3$

(b) $y+z=1, \quad x+y+z=2, \quad x+2y+3z=4$

(c) $y+z=1, \quad x+y+z=2, \quad x+2y+2z=4.$

Solution.

(a) The augmented matrix of the given system is

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 2 & 2 & 3 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 2 & 3 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the given system of equation is reduced to $x+y+z=2$ and $y+z=1$. Letting $z=s$, we find $y=1-z=1-s$ and $x=2-y-z=2-(1-s)-s=1$. Thus a solution $[x,y,z]^t$ of the given system is of the form $[1,1-s,s]^t$. Hence the required set of solution is $\{[1,1-s,s]^t : s \in \mathbb{C}\}$.

(b) The augmented matrix of the given system is

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Thus the given system of equation is reduced to $x+y+z=2, y+z=1$ and $z=1$. This gives $y=1-z=1-1=0$ and $x=2-y-z=2-0-1=1$. Thus a solution $[x,y,z]^t$ of the given system is of the form $[1,0,1]^t$. Hence the required set of solution is $\{[1,0,1]^t\}$.

(c) The augmented matrix of the given system is

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 2 & 2 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 2 & 4 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

The equation corresponding to the last row of this equivalent augmented matrix is $0.x+0.y+0.z=1 \Rightarrow 0=1$, which is absurd. Hence the given system of equation **does not** have a solution, that is, the solution set is $\emptyset$. □

- If the system $A\mathbf{x}=\mathbf{b}$ has at least one solution then it is called a **consistent** system. Otherwise, it is called an **inconsistent** system.

Reduced Row Echelon Form: A matrix A is said to be in *reduced row echelon form* (**RREF**) if it satisfies the following properties:

1. A is in row echelon form.
2. The leading entry in each non-zero row is a 1.
3. Each column containing a leading 1 has zeros everywhere else.

For example, the following matrices are in reduced row echelon form.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Applying elementary row operations, any given matrix can be transformed to a reduced row echelon form.

Leading and Free Variable: Consider the linear system $A\mathbf{x}=\mathbf{b}$ in n variables and m equations. Let $[R \mid \mathbf{r}]$ be the reduced row echelon form of the augmented matrix $[A \mid \mathbf{b}]$.

- Then the variables corresponding to the leading columns in the first n columns of $[R \mid \mathbf{r}]$ are called the **leading variables** or **basic variables**.
- The variables which are not leading are called **free variables**.

Remark: If $A=[B \mid C]$, and $RREF(A)=[R_1 \mid R_2]$, then note that R_1 is the RREF of B.

Result 2.4. *Every matrix has a **unique** reduced row echelon form.*

Proof. We will apply induction on number of columns.
Base Case: Consider a matrix with one column, *i.e.,* $A_{m\times 1}$. The possibilities of the RREF being $\mathbf{O}_{m\times 1}$ and e_1, its RREF is unique.
Induction Hypothesis: Let any matrix having $k-1$ columns have a unique RREF.
Inductive Case: We need to show that $A_{m\times k}$ also has unique RREF. Assume A has two RREF, say B and C. Note that because of induction hypothesis the sub-matrices B_1 and C_1 formed by the first $k-1$ columns of B and C, respectively, are identical. In particular, they have same number of nonzero rows.

We have that the systems $A\mathbf{x} = 0$, $B\mathbf{x} = 0$ and $C\mathbf{x} = 0$ have the same set of solutions. Let the i^{th} row of the k^{th} column be different in B and C i.e. $b_{ik} \neq c_{ik} \Rightarrow b_{ik} - c_{ik} \neq 0$.

Let $\mathbf{u}$ be an arbitrary solution of $A\mathbf{x} = 0$.
$\Rightarrow B\mathbf{u} = 0$, $C\mathbf{u} = 0$ and $(B-C)\mathbf{u} = 0$
$\Rightarrow (b_{ik} - c_{ik})u_k = 0 \Rightarrow u_k$ must be 0
$\Rightarrow x_k$ is not a free variable, as x_k can take only one value
$\Rightarrow$ there must be a leading 1 in the k^{th} column of B and C
$\Rightarrow$ the location of 1 is different in the k^{th} column
$\Rightarrow$ numbers of nonzero rows in B_1 and C_1 are different, a contradiction.
This contradiction leads us to conclude that $B = C$.

Hence by the Method of Induction, we conclude that every matrix has a **unique** reduced row echelon form. ☐

Result 2.5. *The matrices A and B are row equivalent iff $RREF(A) = RREF(B)$.*

Method of transforming a given matrix to reduced row echelon form: Let A be an $m\times n$ matrix. Then the following step by step method is used to obtain the reduced row echelon form of the matrix A.

1. Let the i-th column be the left most non-zero column of A. Interchange rows, if necessary, to make the first entry of this column non-zero. Now use elementary row operations to make all the entries below this first entry equal to 0.
2. Forget the first row and first i columns. Start with the lower $(m-1)\times(n-i)$ sub-matrix of the matrix obtained in the first step and proceed as in **Step 1**.
3. Repeat the above steps until we get a row echelon form of A.
4. Now use the leading term in each of the leading column to make (by elementary row operations) all other entries in that column equal to zero. Use this step starting form the rightmost leading column.
5. Scale the leading entries of the matrix obtained in the previous step, by multiplying the rows by suitable constants, to make all the leading entries equal to 1, ending with the unique reduced row echelon form of A.

Gauss-Jordan Elimination Method: Use the following steps to solve a system of equations $A\mathbf{x} = \mathbf{b}$.

1. Write the augmented matrix $[A \mid \mathbf{b}]$.
2. Use elementary row operations to transform $[A \mid \mathbf{b}]$ to reduced row echelon form.
3. Use back substitution to solve the equivalent system that corresponds to the reduced row echelon form. That is, solve for the leading variables in terms of the remaining free variables, if possible.

Example 2.3. *Solve the system $w - x - y + 2z = 1,\ 2w - 2x - y + 3z = 3,\ -w + x - y = -3$ using Gauss-Jordan elimination method.*

Example 2.4. *Solve the following systems using Gauss-Jordan elimination method:*
(a) $2y + 3z = 8, \quad 2x + 3y + z = 5, \quad x - y - 2z = -5$
(b) $x - y + 2z = 3, \quad x + 2y - z = -3, \quad 2y - 2z = 1.$

Rank: The rank of a matrix A, denoted rank(A), is the number of non-zero rows in its reduced row echelon form.

Result 2.6. *Let $A\mathbf{x} = \mathbf{b}$ be a system of equations with n variables. Then number of free variables is equal to $n-$rank(A).*

Proof. By definition of rank, the number of basic variables is equal to the rank of A. Hence the number of free variables is equal to $n - \text{rank}(A)$. ☐

Result 2.7. *Let $A\mathbf{x} = \mathbf{0}$ be a homogeneous system of equations with n variables. If rank$(A) < n$ then the system has infinitely many solutions.*

Proof. Since rank$(A) < n$ so that $n - \text{rank}(A) > 0$, we find that the system has **at least one free variables**. Since free variables can be assigned any real numbers, the system has infinitely many solutions. $\square$

Result 2.8. *Let $A\mathbf{x} = \mathbf{b}$ be a system of equations with n variables.*

1. *If rank$(A) \neq$ rank$([A \mid \mathbf{b}])$ then the system $A\mathbf{x} = \mathbf{b}$ is inconsistent;*
2. *If rank$(A) =$ rank$([A \mid \mathbf{b}]) = n$ then the system $A\mathbf{x} = \mathbf{b}$ has a unique solution; and*
3. *If rank$(A) =$ rank$([A \mid \mathbf{b}]) < n$ then the system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.*

Proof. Given system is $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$.

1. If rank$(A) \neq$ rank$([A \mid \mathbf{b}])$, that is, rank$(A) <$ rank$([A \mid \mathbf{b}])$, then the last non-zero row of reduced row echelon form of $[A \mid \mathbf{b}]$ will be of the form $[0, 0, \ldots, 0 \mid 1]$. The equation corresponding to this row becomes $0 = 1$, which is absurd. Hence the system $A\mathbf{x} = \mathbf{b}$ does not have a solution, that is, it is inconsistent.

2. Let $[R \mid \mathbf{r}]$ be the sub-matrix consisting of the non-zero rows of reduced row echelon form of $[A \mid \mathbf{b}]$ and let $\mathbf{r} = [r_1, r_2, \ldots, r_n]^t$. Then the system $A\mathbf{x} = \mathbf{b}$ is equivalent to the system $R\mathbf{x} = \mathbf{r}$, and R is an identity matrix of order n. Now the equations of the system $R\mathbf{x} = \mathbf{r}$ give $x_1 = r_1, x_2 = r_2, \ldots, x_n = r_n$. Thus the system $R\mathbf{x} = \mathbf{r}$ has the unique solution $\mathbf{r}$, and hence the system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

3. Let $[R \mid \mathbf{r}]$ be the sub-matrix consisting of the non-zero rows of reduced row echelon form of $[A \mid \mathbf{b}]$. Then the system $A\mathbf{x} = \mathbf{b}$ is equivalent to the system $R\mathbf{x} = \mathbf{r}$. Since rank$(A) =$ rank$([A \mid \mathbf{b}]) < n$, the number of rows of $[R \mid \mathbf{r}]$ is less than n. Therefore there is at least one free variable to the system $R\mathbf{x} = \mathbf{r}$. Further, since rank$(A) =$ rank$([A \mid \mathbf{b}])$, no row of $[R \mid \mathbf{r}]$ is of the form $[0, 0, \ldots, 0 \mid 1]$. Therefore the system $R\mathbf{x} = \mathbf{r}$ cannot be inconsistent.

   Since free variables can be assigned any real numbers, we find infinitely many choices for the values of all the variables of the system $R\mathbf{x} = \mathbf{r}$. Therefore the system $R\mathbf{x} = \mathbf{r}$ has infinitely many solutions, and hence the system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions. $\square$

Practice Problems Set 2

1. What are the possible types of reduced row echelon forms of 2×2 and 3×3 matrices?

2. Find the reduced row echelon form of each of the following matrices.

$$\begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 5 & 6 & 2 \\ -1 & 2 & 4 & 3 \\ 1 & 2 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}, \begin{bmatrix} 3 & 4 & 5 & -6 \\ 2 & 3 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 5 & -5 & 5 & 5 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 2 & 2 & 2 & 4 \\ 2 & -2 & 4 & 0 & 8 \\ 4 & 2 & 5 & 6 & 10 \end{bmatrix}.$$

3. Solve the following systems of equations using Gaussian elimination method as well as Gauss-Jordan elimination method:

   (a) $4x + 2y - 5z = 0, \quad 3x + 3y + z = 0, \quad 2x + 8y + 5z = 0;$

   (b) $-x + y + z + w = 0, \quad x - y + z + w = 0, \quad -x + y + 3z + 3w = 0, \quad x - y + 5z + 5w = 0;$

   (c) $x + y + z = 3, \quad x - y - z = -1, \quad 4x + 4y + z = 9;$

   (d) $x + y + 2z = 3, \quad -x - 3y + 4z = 2, \quad -x - 5y + 10z = 11;$

   (e) $x - 3y - 2z = 0, \quad -x + 2y + z = 0, \quad 2x + 4y + 6z = 0;$ and

   (f) $2w + 3x - y + 4z = 0, \quad 3w - x + z = 1, \quad 3w - 4x + y - z = 2.$

4. Consider the system $A\mathbf{x} = \mathbf{0}$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Prove the following:

   (a) If each entry of A is 0 then each vector $\mathbf{x} = [x, y]^t$ is a solution of $A\mathbf{x} = \mathbf{0}$.

   (b) If $ad - bc \neq 0$ then the system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0} = [0, 0]^t$.

(c) If $ad-bc=0$ and some entry of A is non-zero, then there is a solution $\mathbf{x}_0=[x_0,y_0]^t \neq \mathbf{0}$ such that $\mathbf{x}=[x,y]^t$ is a solution if and only if $x=kx_0, y=ky_0$ for some constant k.

5. For what values of $c \in \mathbb{R}$ and $k \in \mathbb{R}$, the following systems of equations has (i) no solution, (ii) a unique solution, and (iii) infinitely many solutions?

   (a) $x+y+z=3, \quad x+2y+cz=4, \quad 2x+3y+2cz=k$;

   (b) $x+y+2z=3, \quad x+2y+cz=5, \quad x+2y+4z=k$; and

   (c) $x+2y-z=1, \quad 2x+3y+kz=3, \quad x+ky+3z=2$.

   Also, find the solutions whenever they exist.

6. For what values of $a \in \mathbb{R}$ and $b \in \mathbb{R}$, the following system of equations in unknowns x, y and z, has (i) no solution, (ii) a unique solution, and (iii) infinitely many solutions:

$$ax+by+2z=1, \quad ax+(2b-1)y+3z=1, \quad ax+by+(b+3)z=2b-1 \ \ ?$$

   Also, find the solutions whenever they exist.

7. Solve the following system of equations applying Gaussian elimination method:

$$\begin{aligned}(1-n)x_1+x_2+\ldots+x_n&=0\\ x_1+(1-n)x_2+x_3+\ldots+x_n&=0\\ \vdots\quad \vdots\quad \vdots\quad \vdots\quad \vdots\quad \vdots\quad &\vdots\quad \vdots\\ x_1+x_2+\ldots+x_{n-1}+(1-n)x_n&=0.\end{aligned}$$

8. Prove that if $r<s$, then the r-th and the s-th rows of a matrix can be interchanged by performing $2(s-r)-1$ interchanges of adjacent rows.

9. Let A be an $m \times n$ matrix. Prove that the system of equations $A\mathbf{x}=\mathbf{b}$ is consistent for all $\mathbf{b} \in \mathbb{R}^m$ if and only if each row of the row echelon form of A contains a leading term.

10. Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Prove that the system of equations $A\mathbf{x}=\mathbf{b}$ is inconsistent if and only if there is a leading term in the last column of a row echelon form of its augmented matrix.

11. Let $A=\begin{bmatrix} i & -1-i & 0\\ 1 & -2 & 1\\ 1 & 2i & -1\end{bmatrix}$. Determine the reduced row echelon form of A. Hence solve the system $A\mathbf{x}=\mathbf{0}$.

12. Show that $\mathbf{x}=\mathbf{0}$ is the only solution of the system of equations $A\mathbf{x}=\mathbf{0}$ if and only if the rank of A equals the number of variables.

13. Show that a consistent system of equations $A\mathbf{x}=\mathbf{b}$ has a unique solution if and only if the rank of A equals the number of variables.

14. Prove that if two homogeneous systems of m linear equations in two variables have the same solution, then these two systems are equivalent. Is the same true for more than two variables? Justify.

15. (a) If $\mathbf{x}_1$ is a solution of the non-homogeneous system $A\mathbf{x}=\mathbf{b}$ and if $\mathbf{y}_0$ is a solution of the system $A\mathbf{x}=\mathbf{0}$ then show that $\mathbf{x}_1+\mathbf{y}_0$ is a solution of $A\mathbf{x}=\mathbf{b}$.

    (b) If $\mathbf{x}_0, \mathbf{x}_1$ are solutions of the system $A\mathbf{x}=\mathbf{b}$ then show that there is a solution $\mathbf{y}_0$ of the system $A\mathbf{x}=\mathbf{0}$ such that $\mathbf{x}_0=\mathbf{x}_1+\mathbf{y}_0$.

    [Let S_h and S_{nh} be the solution sets of $A\mathbf{x}=\mathbf{0}$ and $A\mathbf{x}=\mathbf{b}$, respectively. Note that part (a) implies $x_1+S_h \subseteq S_{nh}$ and part (b) implies $S_{nh} \subseteq x_1+S_h$.]

16. Suppose the system $A\mathbf{x}=\mathbf{b}$ is consistent with s_0 as one of the solutions. If S_h is the set of solutions of $A\mathbf{x}=\mathbf{0}$, then show that the set of solutions of $A\mathbf{x}=\mathbf{b}$ is s_0+S_h.

17. Suppose that the non-homogeneous system $A\mathbf{x}=\mathbf{b}$ has solutions $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$. Show that a linear combination $a_1\mathbf{u}_1+a_2\mathbf{u}_2+\ldots+a_k\mathbf{u}_k$ is a solution of $A\mathbf{x}=\mathbf{b}$ if and only if $a_1+a_2+\ldots+a_k=1$. Also, show that $c_1\mathbf{u}_1+c_2\mathbf{u}_2+\ldots+c_k\mathbf{u}_k=\mathbf{0}$ implies $c_1+c_2+\ldots+c_k=0$.

18. (**Not for Exam**) Solve the following systems of equations over the indicated $\mathbb{Z}_p$.

   (a) $x+2y=1,\ x+y=2$ over $\mathbb{Z}_3$.

   (b) $x+y=1,\ y+z=0, x+z=1$ over $\mathbb{Z}_2$.

   (c) $3x+2y=1,\ x+4y=1$ over $\mathbb{Z}_5$.

   (d) $2x+3y=4,\ 4x+3y=2$ over $\mathbb{Z}_6$.

19. (**Not for Exam**) Let A be an $m\times n$ matrix of rank r with entries in $\mathbb{Z}_p$, where p is a prime number. Prove that every consistent system of equations with coefficient matrix A has exactly p^{n-r} solutions over $\mathbb{Z}_p$.

Hints to Practice Problems Set 2

1. There are four and eight possible reduced row echelon forms of a 2×2 and a 3×3 matrices, respectively. These matrices can be obtained by considering the cases based on the ranks and the position of the leading entries.

2. The reduced row echelon form of the given matrices are the following (in order):

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\ \begin{bmatrix}1&0&-1&-2\\0&1&2&3\\0&0&0&0\\0&0&0&0\end{bmatrix},\ \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\ \text{and}\ \begin{bmatrix}1&0&0&1&0\\0&1&0&1&0\\0&0&1&0&0\\0&0&0&0&1\end{bmatrix}.$$

3. (a) $\{[0,0,0]^t\}$, (b) $\{[r,r,-s,s]^t : r,s\in\mathbb{R}\}$, (c) $\{[1,1,1]^t\}$, (d) Inconsistent, (e) $\{[s,s,-s]^t : s\in\mathbb{R}\}$, (f) $\left\{\left[\frac{5}{22}-s,\ \frac{15}{22}-2s,\ \frac{5}{2}-4s,\ s\right]^t :\ s\in\mathbb{R}\right\}$.

4. (b) If $a=0$ or $ac\neq 0$ or $c=0$ then the reduced row echelon form of A is $\begin{bmatrix}1&0\\0&1\end{bmatrix}$.

   (c) If $a=c=0$ then the reduced row echelon form of A is $\begin{bmatrix}0&1\\0&0\end{bmatrix}$. If $(ac\neq 0)$ or $(a=0, c\neq 0)$ or $(a\neq 0, c=0)$, then the reduced row echelon form of A is $\begin{bmatrix}1&\alpha\\0&0\end{bmatrix}$ for some $\alpha\in\mathbb{R}$.

5. (a) If $c=1, k\neq 7$ then no solution. If $c\neq 1$ then unique solution: $\left[k-5-\frac{k-7}{c-1},\ 8-k,\ \frac{k-7}{c-1}\right]^t$.
   If $c=1, k=7$ then infinitely many solutions: $\{[2-s,1,s]^t : s\in\mathbb{R}\}$.

   (b) If $c=4, k\neq 5$ then no solution. If $c\neq 4$ then unique solution: $\left[6-k,\ 2+\frac{(c-2)(k-5)}{c-4},\ \frac{k-5}{4-c}\right]^t$.
   If $c=4, k=5$ then infinitely many solutions: $\{[1,2-2s,s]^t : s\in\mathbb{R}\}$.

   (c) If $k=0$ then no solution. If $k\neq 0$ then unique solution: $\left[\frac{k^2-k+3}{k^2},\ \frac{k-2}{k^2},\ \frac{k-1}{k^2}\right]^t$.

6. The augmented matrix can be reduced to

$$\left[\begin{array}{ccc|c}a&1&1&1\\0&b-1&1&0\\0&0&b+1&2(b-1)\end{array}\right].$$

   If $a=0, b=-1$, then no solution.
   If $a=0, b=1$, then infinitely many solutions: $\{[s,1,0]^t : s\in\mathbb{R}\}$.
   If $a=0, b=5$, then infinitely many solutions: $\{[s,-1/3,4/3]^t : s\in\mathbb{R}\}$.
   If $a=0, b\neq 5,\pm 1$, then no solution.
   If $a\neq 0, b=-1$, then no solution.
   If $a\neq 0, b=1$, then infinitely many solutions: $\{[\frac{1-s}{a}, s, 0]^t : s\in\mathbb{R}\}$.
   If $a\neq 0, b\neq\pm 1$, then unique solution: $\left[\frac{5-b}{a(b+1)}, -\frac{2}{b+1}, \frac{2(b-1)}{b+1}\right]^t$.

7. The coefficient matrix can be transformed to

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1-n \\ 0 & 1 & 0 & 0 & \dots & 0 & -1 \\ 0 & 0 & 1 & 0 & \dots & 0 & -1 \\ 0 & 0 & 0 & 1 & \dots & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

   Hence $x_n = x_{n-1} = \dots\dots = x_2 = x_1$.

8. Apply the following sequence of elementary row operations:

$$R_s \leftrightarrow R_{s-1},\quad R_{s-1} \leftrightarrow R_{s-2},\quad \dots\dots, R_{r+1} \leftrightarrow R_r,\quad R_{r+1} \leftrightarrow R_{r+2},\quad R_{r+2} \leftrightarrow R_{r+3},\quad \dots\dots, R_{s-1} \leftrightarrow R_s.$$

9. For the 'only if' part ($\Rightarrow$), take $\mathbf{b} = \mathbf{e}_i$ for $i = 1, 2, \dots, m$. For the 'if' part ($\Leftarrow$), $m = \text{rank}(A) \le \text{rank}([A \mid \mathbf{b}]) \le m$.

10. Prove the contra-positive for the 'only if' part. That is, assume $A\mathbf{x} = \mathbf{b}$ is consistent and then prove that the last column of a row echelon form of its augmented matrix cannot contain a leading term.

11. The reduced row echelon form is $\begin{bmatrix} 1 & 0 & -i \\ 0 & 1 & -\frac{1}{2}(1-i) \\ 0 & 0 & 0 \end{bmatrix}$. The solution set is $\{t[i, \frac{1}{2}(1-i), 1]^t \;:\; t \in \mathbb{R}\}$.

12. For the 'only if' part ($\Rightarrow$), prove the contra-positive, so that there will be at least one free variable. For the 'if' part ($\Leftarrow$), there will be no free variable.

13. Similar to Problem 12.

14. Similar to a tutorial problem. The statement is also true in general. Indeed, if $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ are two systems of m equations and having equal solution sets (**non-empty**) then $[A \mid \mathbf{b}]$ is row equivalent to $[C \mid \mathbf{d}]$. This has been partially covered in a tutorial problem.

15. Take $\mathbf{y}_0 = \mathbf{x}_0 - \mathbf{x}_1$ for the second part.

16. Use Problem 15.

17. Easy.

18. (a) $\{[0,2]^t\}$, (b) $\{[0,1,1]^t, [1,0,0]^t\}$, (c) Inconsistent, (d) $\{[2,0]^t, [2,2]^t, [2,4]^t\}$.

19. There are p choices to assign values to each of the $n-r$ free variables.

---

3 Vector Space

Definition 3.1. *Let n be a positive integer. Then the space $\mathbb{R}^n$, as defined below, is called the n-dimensional **Euclidean space**:*

$$\mathbb{R}^n = \{[x_1, x_2, \ldots, x_n]^t : x_1, x_2, \ldots, x_n \in \mathbb{R}\}.$$

- Elements of $\mathbb{R}^n$ are called n-vectors or simply **vectors**.
- Note that $[x_1, x_2, \ldots, x_n]^t = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a column vector.
- Sometimes, an element $[x_1, x_2, \ldots, x_n]^t$ of $\mathbb{R}^n$ is also written as a row vector $[x_1, x_2, \ldots, x_n]$ or $(x_1, x_2, \ldots, x_n)$.
- The element $(x_1, x_2, \ldots, x_n)$ is termed as an n-tuple.
- Normally, while discussing a system of linear equations, elements of $\mathbb{R}^n$ are regarded as column vectors. Otherwise, elements of $\mathbb{R}^n$ may also be regarded as row vectors.
- In a similar way, we can define $\mathbb{C}^n$.

Recall that a vector $\vec{x} = a\hat{i} + b\hat{j} + c\hat{k}$ can also be written as (a, b, c) or $[a, b, c]^t$. Thus elements of $\mathbb{R}^3$ are basically the vectors that we had learnt in school. For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^3$, $c, d \in \mathbb{R}$ and for the zero vector $\mathbf{0}$, we learnt in school that

1. $\mathbf{u} + \mathbf{v} \in \mathbb{R}^3$;
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
4. $\mathbf{u} + \mathbf{0} = \mathbf{u}$;
5. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$;
6. $c.\mathbf{u} \in \mathbb{R}^3$;
7. $c.(\mathbf{u} + \mathbf{v}) = c.\mathbf{u} + c.\mathbf{v}$;
8. $(c + d).\mathbf{u} = c.\mathbf{u} + d.\mathbf{u}$;
9. $c.(d.\mathbf{u}) = (cd).\mathbf{u}$; and
10. $1.\mathbf{u} = \mathbf{u}$.

The above properties are sufficient to do vector algebra in $\mathbb{R}^3$. The same properties, exactly in the same way, hold good in $\mathbb{R}^n$ too.

- If $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0} \in \mathbb{C}^n$ and $c, d \in \mathbb{R}$, we get all the previous ten properties.
- If $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0} \in \mathbb{C}^n$ and $c, d \in \mathbb{C}$, we get all the previous ten properties.
- If $A, B, C, \mathbf{O} \in M_2(\mathbb{R})$ (set of all 2×2 real matrices) and $c, d \in \mathbb{R}$, we get all the previous ten properties.
- If $p(x), q(x), r(x), \mathbf{0} \in \mathbb{R}_2[x]$ (set of all polynomials of degree at most two with real coefficients) and $c, d \in \mathbb{R}$, we get all the previous ten properties.
- In our discussion, the symbol $\mathbb{F}$ will be used as a representative of $\mathbb{R}$ or $\mathbb{C}$. That is, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.
- The elements of $\mathbb{F}$ will be termed as **scalars**.

Definition 3.2. *Let V be a non-empty set. For every $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{F}$, let the addition $\mathbf{u} + \mathbf{v}$ (called the vector addition) and the multiplication $c.\mathbf{u}$ (called the scalar multiplication) be defined. Then V is called a **vector space** over $\mathbb{F}$ if for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $c, d \in \mathbb{F}$, the following properties are satisfied:*

1. $\mathbf{u} + \mathbf{v} \in V$;

2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;

4. *There is an element* $\mathbf{0}$, *called a zero, such that* $\mathbf{u} + \mathbf{0} = \mathbf{u}$;

5. *For each* $\mathbf{u} \in V$, *there is an element* $-\mathbf{u}$ *such that* $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$;

6. $c.\mathbf{u} \in V$;

7. $c.(\mathbf{u} + \mathbf{v}) = c.\mathbf{u} + c.\mathbf{v}$;

8. $(c + d).\mathbf{u} = c.\mathbf{u} + d.\mathbf{u}$;

9. $c.(d.\mathbf{u}) = (cd).\mathbf{u}$; *and*

10. $1.\mathbf{u} = \mathbf{u}$.

Example 3.1. *For any $n \geq 1$, the set $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ with respect to usual operations of addition and scalar multiplication. [Show that all the ten defining properties of a vector space are satisfied].*

Example 3.2. *For any $n \geq 1$, the set $\mathbb{C}^n$ is a vector space over $\mathbb{R}$ with respect to usual operations of addition and scalar multiplication. [Show that all the ten defining properties of a vector space are satisfied].*

Example 3.3. *For any $n \geq 1$, the set $\mathbb{C}^n$ is a vector space over $\mathbb{C}$ with respect to usual operations of addition and scalar multiplication. [Show that all the ten defining properties of a vector space are satisfied].*

Example 3.4. *The set $\mathbb{R}^n$ is* ***not*** *a vector space over $\mathbb{C}$ with respect to usual operations of addition and scalar multiplication.*

Solution. For $c = i$ and $\mathbf{u} = [1, 1, \ldots, 1]^t \in \mathbb{R}^n$, we have $c.\mathbf{u} = i.[1, 1, \ldots, 1]^t = [i, i, \ldots, i]^t \notin \mathbb{R}^n$, where $i = \sqrt{-1}$. Hence $\mathbb{R}^n$ is **not** a vector space over $\mathbb{C}$ with respect to usual operations of addition and scalar multiplication [the 6th property is violated]. ☐

Example 3.5. *The set $\mathbb{Z}$ is* ***not*** *a vector space over $\mathbb{R}$ with respect to usual operations of addition and scalar multiplication.*

Solution. For $c = \frac{1}{2} \in \mathbb{R}$ and $\mathbf{u} = 3 \in \mathbb{Z}$, we have $c.\mathbf{u} = \frac{1}{2}.3 = \frac{3}{2} \notin \mathbb{Z}$. Hence $\mathbb{Z}$ is **not** a vector space over $\mathbb{R}$ with respect to usual operations of addition and scalar multiplication [the 6th property is violated]. ☐

Example 3.6. *The set $M_2(\mathbb{C})$ of all 2×2 complex matrices is a vector space over $\mathbb{C}$ with respect to usual operations of matrix addition and matrix scalar multiplication.*

Example 3.7. *The set $\mathbb{R}^2$ is* ***not*** *a vector space over $\mathbb{R}$ with respect to usual operations of addition and the following definition of scalar multiplication:*

$$c.[x, y]^t = [cx, 0]^t \quad \text{for } [x, y]^t \in \mathbb{R}^2, c \in \mathbb{R}.$$

Solution. For $c = 1$ and $\mathbf{u} = [2, 3]^t \in \mathbb{R}^2$, we have $c.\mathbf{u} = 1.[2, 3]^t = [2, 0]^t \neq \mathbf{u}$. Hence $\mathbb{R}^2$ is **not** a vector space over $\mathbb{R}$ with respect to the operations defined in the example [the 10th property is violated]. ☐

Example 3.8. *Let $\mathbb{R}_2[x]$ denote the set of all polynomials of degree at most two with real coefficients. That is,*

$$\mathbb{R}_2[x] = \{a + bx + cx^2 : a, b, c \in \mathbb{R}\}.$$

For $p(x) = a_0 + b_0x + c_0x^2, q(x) = a_1 + b_1x + c_1x^2 \in \mathbb{R}_2[x]$ and $k \in \mathbb{R}$, define

$$p(x) + q(x) = (a_0 + a_1) + (b_0 + b_1)x + (c_0 + c_1)x^2 \text{ and } k.p(x) = (ka_0) + (kb_0)x + (kc_0)x^2.$$

Then $\mathbb{R}_2[x]$ is a vector space over $\mathbb{R}$.

- If V is a vector space, then the elements of V are called **vectors**.

- If there is no confusion, $c.\mathbf{u}$ is simply written as $c\mathbf{u}$.

- In general, we take V to be a vector space over $\mathbb{F}$.
- If $\mathbb{F}$ needs to be specified, then we write $V(\mathbb{F})$ to denote that V is a vector space over $\mathbb{F}$.
- We call V a **real vector space** or **complex vector space** according as $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Result 3.1. *Let V be a vector space over $\mathbb{F}$. Let $\mathbf{u} \in V$ and $c \in \mathbb{F}$.*

1. $0.\mathbf{u} = \mathbf{0}$;
2. $c.\mathbf{0} = \mathbf{0}$;
3. $(-1).\mathbf{u} = -\mathbf{u}$; *and*
4. *If $c.\mathbf{u} = \mathbf{0}$ then either $c = 0$ or $\mathbf{u} = \mathbf{0}$.*

Proof. [Proof of **Part 1** and **Part 3**] We have $0.\mathbf{u} = (0+0).\mathbf{u} = 0.\mathbf{u} + 0.\mathbf{u}$. Therefore adding $-(0.\mathbf{u})$ on both sides,

$$\begin{aligned}[-(0.\mathbf{u}] + 0.\mathbf{u} = [-(0.\mathbf{u})] + (0.\mathbf{u} + 0.\mathbf{u}) &\Rightarrow -(0.\mathbf{u}) + 0.\mathbf{u} = [-(0.\mathbf{u}) + 0.\mathbf{u}] + 0.\mathbf{u} \\ &\Rightarrow \mathbf{0} = \mathbf{0} + 0.\mathbf{u} \\ &\Rightarrow \mathbf{0} = 0.\mathbf{u}.\end{aligned}$$

We have $\mathbf{0} = 0.\mathbf{u} = [(1+(-1)].\mathbf{u} = 1.\mathbf{u} + (-1).\mathbf{u} = \mathbf{u} + (-1).\mathbf{u}$. Therefore adding $-\mathbf{u}$ on both sides, we get

$$\begin{aligned}(-\mathbf{u}) + \mathbf{0} = (-\mathbf{u}) + [\mathbf{u} + (-1).\mathbf{u}] &\Rightarrow -\mathbf{u} = [(-\mathbf{u}) + \mathbf{u}] + (-1).\mathbf{u} \\ &\Rightarrow -\mathbf{u} = \mathbf{0} + (-1).\mathbf{u} \\ &\Rightarrow -\mathbf{u} = (-1).\mathbf{u}.\end{aligned}$$

□

Subspace:
Let V be a vector space and $(\emptyset \neq) W \subseteq V$. Then W is called a **subspace** of V if and only if $a\mathbf{u} + b\mathbf{v} \in W$ for every $\mathbf{u}, \mathbf{v} \in W$ and for every $a, b \in \mathbb{F}$.

- If W is a subspace of a vector space V, then W is also a vector space.
- If W is a subspace of a vector space V then $\mathbf{0} \in W$.
- The sets $\{\mathbf{0}\}$ and V are always subspaces of any vectors space V. These are called the **trivial** subspaces.
- If W is a subspace of V then it is clear that if $\mathbf{u}, \mathbf{v} \in W$ then $a\mathbf{u} = a\mathbf{u} + 0\mathbf{u} \in W$ and $\mathbf{u} + \mathbf{v} = 1\mathbf{u} + 1\mathbf{v} \in W$.
- W is a subspace iff W is closed under addition and scalar multiplication.

Example 3.9. *Examine whether the sets $S = \{[x, y, z]^t \in \mathbb{R}^3 : x = y + 1\}$, $T = \{[x, y, z]^t \in \mathbb{R}^3 : x = z^2\}$ and $U = \{[x, y, z]^t \in \mathbb{R}^3 : x = 5y\}$ are subspaces of $\mathbb{R}^3$.*

Solution. Notice that $[0, 0, 0]^t$ does not satisfy $x = y + 1$ and so $[0, 0, 0]^t \notin S$. Therefore S is not a subspace.

We have $[1, 1, 1]^t \in T$ but $2.[1, 1, 1]^t = [2, 2, 2]^t \notin T$. Therefore T is not a subspace.

For the third set, note that $[0, 0, 0]^t \in U$ and so $U \neq \emptyset$. let $\mathbf{u} = [x_1, y_1, z_1]^t, \mathbf{v} = [x_2, y_2, z_2]^t \in U$ and $a, b \in \mathbb{R}$. Then we have $x_1 = 5y_1$ and $x_2 = 5y_2$. Now

$$a\mathbf{u} + b\mathbf{v} = [ax_1 + bx_2, ay_1 + by_2, az_1 + bz_2]^t \text{ and } ax_1 + bx_2 = a(5y_1) + b(5y_2) = 5(ay_1 + by_2).$$

Therefore $a\mathbf{u} + b\mathbf{v} \in U$, and hence U is a subspace. □

Example 3.10. *Let W be the set of all 2×2 real symmetric matrices. Then W is a subspace of $M_2(\mathbb{R})$.*

Example 3.11. *Let $W = \{[x, y, z]^t \in \mathbb{R}^3 : x + y - z = 0\}$. Then W is a subspace of $\mathbb{R}^3$.*

Solution. Let $\mathbf{u} = [x_1, y_1, z_1]^t, \mathbf{v} = [x_2, y_2, z_2]^t \in W$. Then we have $x_1 + y_1 - z_1 = 0, x_2 + y_2 - z_2 = 0$. Now for $a, b \in \mathbb{R}$, we have $a\mathbf{u} + b\mathbf{v} = [ax_1 + b\mathbf{x}_2, ay_1 + b\mathbf{y}_2, az_1 + bz_2]^t$ and $(ax_1 + b\mathbf{x}_2) + (ay_1 + b\mathbf{y}_2) - (az_1 + bz_2) = a(x_1 + y_1 - z_1) + b(x_2 + y_2 - z_2) = 0$. So $a\mathbf{u} + b\mathbf{v} \in W$ and hence W is a subspace of $\mathbb{R}^3$. □

Example 3.12. *Let A be an $m \times n$ matrix. Then $S = \{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \mathbf{0}\}$ is a subspace, called the **null space** of A.*

Example 3.13. *Let U and W be two subspaces of V. Then $U+W=\{\mathbf{u}+\mathbf{w} : \mathbf{u}\in U, \mathbf{w}\in W\}$ is also a subspace of V.*

Solution. Let $\mathbf{x},\mathbf{y}\in U+W$ and $a,b\in\mathbb{F}$. Then $\mathbf{x}=\mathbf{u}_1+\mathbf{v}_1$ and $\mathbf{y}=\mathbf{u}_2+\mathbf{v}_2$ for some $\mathbf{u}_1,\mathbf{u}_2\in U, \mathbf{v}_1,\mathbf{v}_2\in W$. Now we have $a\mathbf{x}+b\mathbf{y}=a(\mathbf{u}_1+\mathbf{v}_1)+b(\mathbf{u}_2+\mathbf{v}_2)=(a\mathbf{u}_1+b\mathbf{u}_2)+(a\mathbf{v}_1+b\mathbf{v}_2)$. Since U and W are subspaces, we have $a\mathbf{u}_1+b\mathbf{u}_2\in U, a\mathbf{v}_1+b\mathbf{v}_2\in W$. Therefore $a\mathbf{x}+b\mathbf{y}\in U+W$ and hence $U+W$ is a subspace of V. □

- If U and W are subspaces of V such that $U\cap W=\{\mathbf{0}\}$, then $U+W$ is called a **direct sum**, denoted by $U\oplus W$.

Result 3.2. *Let V be a vector space.*

1. *If U and W are subspaces of V, then $U\cap W$ is also a subspace of V.*
2. *If $\{U_i : i\in\Delta\}$ is a non-empty collection of subspaces of V, then $\bigcap\limits_{i\in\Delta} U_i$ is also a subspace of V.*
3. *Let U and W be subspaces of V. Then $U\cup W$ is a subspace of V iff either $U\subseteq W$ or $W\subseteq U$.*

Proof.

1. Clearly $U\cap W\neq\emptyset$, as $\mathbf{0}\in U\cap W$. Let $\mathbf{x},\mathbf{y}\in U\cap W$ and $a,b\in\mathbb{F}$. Then we have $\mathbf{x},\mathbf{y}\in U$ and $\mathbf{x},\mathbf{y}\in W$. Since U and W are subspaces of V, we have
$$a\mathbf{x}+b\mathbf{y}\in U \quad\text{and}\quad a\mathbf{x}+b\mathbf{y}\in W\Rightarrow a\mathbf{x}+b\mathbf{y}\in U\cap W.$$
Thus $U\cap W$ is a subspace of V.
2. Clearly $\bigcap\limits_{i\in\Delta} U_i\neq\emptyset$, as $\mathbf{0}\in U_i$ for all $i\Rightarrow\mathbf{0}\in\bigcap\limits_{i\in\Delta} U_i$. Let $\mathbf{x},\mathbf{y}\in\bigcap\limits_{i\in\Delta} U_i$ and $a,b\in\mathbb{F}$. Then $\mathbf{x},\mathbf{y}\in U_i$ for each i. Since U_i is a subspace for each i, we have
$$a\mathbf{x}+b\mathbf{y}\in U_i \quad\text{for each}\quad i\Rightarrow a\mathbf{x}+b\mathbf{y}\in\bigcap_{i\in\Delta} U_i.$$
Thus $\bigcap\limits_{i\in\Delta} U_i$ is a subspace of V.
3. If $U\subseteq W$ or $W\subseteq U$, then we have $U\cup W=W$ or $U\cup W=U$. So, clearly $U\cup W$ is a subspace of V.

   Conversely, assume that $U\cup W$ is a subspace of V. To show that $U\subseteq W$ or $W\subseteq U$. Let $\mathbf{u}\in U$ and $\mathbf{w}\in W$. Now
$$\begin{aligned}
\mathbf{u},\mathbf{w}\in U\cup W &\Rightarrow \mathbf{u}+\mathbf{w}\in U\cup W\\
&\Rightarrow \mathbf{u}+\mathbf{w}\in U \quad\text{or}\quad \mathbf{u}+\mathbf{w}\in W\\
&\Rightarrow (-1)\mathbf{u}+1.(\mathbf{u}+\mathbf{w})\in U \quad\text{or}\quad 1.(\mathbf{u}+\mathbf{w})+(-1)\mathbf{w}\in W, \quad\text{as } U,W \text{ are subspaces}\\
&\Rightarrow -\mathbf{u}+\mathbf{u}+\mathbf{w}\in U \quad\text{or}\quad \mathbf{u}+\mathbf{w}-\mathbf{w}\in W\\
&\Rightarrow \mathbf{w}\in U \quad\text{or}\quad \mathbf{u}\in W\\
&\Rightarrow W\subseteq U \quad\text{or}\quad U\subseteq W.
\end{aligned}$$

□

Linear Combination: Let V be a vector space. Let $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\in V$ and $c_1,c_2,\ldots,c_k\in\mathbb{F}$. Then the vector $c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_k\mathbf{v}_k$ is called a linear combination of the vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k$.

A vector $\mathbf{v}$ in V is called a linear combination of the vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k$ in V if there exists scalars $c_1,c_2,\ldots,c_k$ such that $\mathbf{v}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_k\mathbf{v}_k$. The numbers $c_1,c_2,\ldots,c_k$ are called the **coefficients** of the linear combination.

Example 3.14. *Is the vector $[1,2,3]^t$ a linear combination of $[1,0,3]^t$ and $[-1,1,-3]^t$?*

Result 3.3. *A system of linear equations with augmented matrix $[A\mid\mathbf{b}]$ is consistent if and only if $\mathbf{b}$ is a linear combination of the columns of A.*

Proof. Let $\mathbf{a}=[\alpha_1,\alpha_2,\ldots,\alpha_n]^t\in\mathbb{F}^n$ and let the system $A\mathbf{x}=\mathbf{b}$ be given by
$$\begin{aligned}
a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1\\
\vdots\qquad\quad \vdots\qquad\qquad \vdots\ \ &\ \ \vdots\\
a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m,
\end{aligned}$$

Now we have, $\quad \mathbf{a}$ is a solution of $A\mathbf{x} = \mathbf{b} \Leftrightarrow$

$$\begin{aligned} &a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = b_1 \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \;\; \vdots \\ &a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n = b_m \\ \Leftrightarrow &\alpha_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \\ \Leftrightarrow &\mathbf{b} \text{ is a linear combination of the columns of } A. \quad \square \end{aligned}$$

Span of a Set: Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a subset of a vector space V. Then the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called the **span** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, and is denoted by $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ or $\text{span}(S)$. That is,

$$\text{span}(S) = \{\mathbf{v} \mid \mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k \text{ for some } c_1, c_2, \ldots, c_k \in \mathbb{F}\}.$$

Let $S \subseteq V$ (may be infinite!) The span of S is defined by

$$\text{span}(S) := \left\{ \sum_{i=1}^{m} \alpha_i \mathbf{v}_i \mid \mathbf{v}_i \in S, \alpha_i \in \mathbb{F}, m \text{ a nonnegative integer} \right\}.$$

- Convention: $\text{span}(\emptyset) = \{\mathbf{0}\}$.
- If $\text{span}(S) = V$, then S is called a s**panning set** for V.
- For example, $\mathbb{R}^2 = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, where $\mathbf{e}_1 = [1, 0]^t$ and $\mathbf{e}_2 = [0, 1]^t$, since for any $[x, y]^t \in \mathbb{R}^2$ we have $[x, y]^t = x\mathbf{e}_1 + y\mathbf{e}_2$.
- $\mathbb{R}_2[x] = \text{span}(1, x, x^2)$.
- $\mathbb{R}[x] = \text{span}(1, x, x^2, \ldots)$ [$\mathbb{R}[x] :=$ set of all polynomials in x].

Clearly $\text{span}(1, x, x^2, \ldots) \subseteq \mathbb{R}[x]$. Also, if $p(x) \in \mathbb{R}[x]$ then $p(x) = a_1x^{m_1} + \ldots + a_kx^{m_k}$ for some $a_1, \ldots, a_k \in \mathbb{R}$ and $m_1, \ldots, m_k \in \mathbb{N} \cup \{0\}$. As $x^{m_1}, \ldots, x^{m_k} \in \{1, x, x^2, \ldots\}$, we have $p(x) \in \text{span}(1, x, x^2, \ldots)$ and consequently $\mathbb{R}[x] \subseteq \text{span}(1, x, x^2, \ldots)$. Thus we get $\mathbb{R}[x] = \text{span}(1, x, x^2, \ldots)$. $\square$

Example 3.15. *Let* $\mathbf{u} = [1, 2, 3]^t$ *and* $\mathbf{v} = [-1, 1, -3]^t$*. Describe span*$(\mathbf{u}, \mathbf{v})$ *in* $\mathbb{R}^3$ *geometrically.*

Solution. Let $\mathbf{w} = [a, b, c]^t \in \text{span}(\mathbf{u}, \mathbf{v})$. Now $\mathbf{w} \in \text{span}(\mathbf{u}, \mathbf{v})$ if and only if there exist $\alpha, \beta \in \mathbb{R}$ such that $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$, which implies that the system of equation $\begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is consistent. We have

$$\left[\begin{array}{rr|r} 1 & -1 & a \\ 2 & 1 & b \\ 3 & -3 & c \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{rr|r} 1 & -1 & a \\ 0 & 3 & b-2a \\ 3 & -3 & c \end{array}\right] \xrightarrow{R_3 \to R_3 - 3R_1} \left[\begin{array}{rr|r} 1 & -1 & a \\ 0 & 3 & b-2a \\ 0 & 0 & c-3a \end{array}\right].$$

Thus the above system of equation is consistent iff $c - 3a = 0$. Hence $\text{span}(\mathbf{u}, \mathbf{v})$ is the plane whose equation is given by $z - 3x = 0$. $\square$

Result 3.4. *Let* S *be a subset of a vector space* V*. Then span*(S) *is a subspace of* V*.*

Proof. If $S = \emptyset$, then $\text{span}(\emptyset) = \{\mathbf{0}\}$ is clearly a subspace of V. For $S \neq \emptyset$, let $\mathbf{u}, \mathbf{v} \in \text{span}(S)$. Then $\mathbf{u} = \alpha_1\mathbf{u}_1 + \ldots + \alpha_k\mathbf{u}_k$ and $\mathbf{v} = \beta_1\mathbf{v}_1 + \ldots + \beta_m\mathbf{v}_m$ for some $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_m \in S$ and $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m \in \mathbb{F}$. Let $a, b \in \mathbb{F}$. We have

$$\begin{aligned} a\mathbf{u} + b\mathbf{v} &= a(\alpha_1\mathbf{u}_1 + \ldots + \alpha_k\mathbf{u}_k) + b(\beta_1\mathbf{v}_1 + \ldots + \beta_m\mathbf{v}_m) \\ &= (a\alpha_1)\mathbf{u}_1 + \ldots + (a\alpha_k)\mathbf{u}_k + (b\beta_1)\mathbf{v}_1 + \ldots + (b\beta_m)\mathbf{v}_m \in \text{span}(S). \end{aligned}$$

Hence $\text{span}(S)$ is a subspace of V. $\square$

Linear Dependence: A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in a vector space V is said to be **linearly dependent** if there are scalars $c_1, c_2, \ldots, c_k$, at least one of them non-zero, such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$.

An **infinite** set $S \subseteq V$ is linearly dependent if there is some **finite** linearly dependent subset of S.

Linear Independence: A set S of vectors in a vector space V is said to be **linearly independent** (LI) if it is **not** linearly dependent. Thus
$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is **linearly independent** (LI) if $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = \mathbf{0} \Rightarrow c_1 = c_2 = \ldots = c_k = 0$.
If S is **infinite** then S is **linearly independent** (LI) if every **finite subset** of S is *linearly independent.*

- Set $\{\mathbf{0}\}$ is linearly dependent as $1.\mathbf{0} = \mathbf{0}$. [A non-trivial linear combination of $\mathbf{0}$ is $\mathbf{0}$.]
- If $\mathbf{0} \in S$, then S is always linearly dependent as S contains a linearly dependent set $\{\mathbf{0}\}$.
- If $\{\mathbf{u}, \mathbf{v}\}$ is linearly dependent, then $a\mathbf{u} + b\mathbf{v} = \mathbf{0}$, where either $a \neq 0$ or $b \neq 0$. If $a \neq 0$, then $\mathbf{u} = \left(-\frac{b}{a}\right)\mathbf{v}$. If $b \neq 0$, then $\mathbf{v} = \left(-\frac{a}{b}\right)\mathbf{u}$. Thus two vectors are linearly dependent iff one of them is a scalar multiple of the other.
- Convention: the set $\emptyset$ is linearly independent.

Example 3.16. *Let $\mathbf{e}_i \in \mathbb{R}^n$ be the i-th column of the identity matrix I_n. Is $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ linearly independent?*

Solution. For $a_1, a_2, \ldots, a_n \in \mathbb{R}$, we have

$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_n\mathbf{e}_n = \mathbf{0} \Rightarrow [a_1, a_2, \ldots, a_n]^t = \mathbf{0} \Rightarrow a_1 = a_2 = \ldots = a_n = 0.$$

Hence $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent. □

Example 3.17. *Examine whether the sets $T = \{[1,2,0]^t, [1,1,-1]^t, [1,4,2]^t\}$ and $S = \{[1,4]^t, [-1,2]^t\}$ are linearly dependent.*

Solution. For $a, b, c \in \mathbb{R}$, we have

$$a[1,2,0]^t + b[1,1,-1]^t + c[1,4,2]^t = \mathbf{0} \Rightarrow a+b+c = 0, \quad 2a+b+4c = 0, \quad -b+2c = 0.$$

Now the augmented matrix of this system of equation is

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_2} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The equations corresponding to the last row echelon form are $a+b+c = 0$ and $-b+2c = 0$. Taking $c = 1$, we have $b = 2$ and $a = -(b+c) = -3$. Therefore

$$(-3)[1,2,0]^t + 2.[1,1,-1]^t + [1,4,2]^t = \mathbf{0}.$$

Hence T is a linearly dependent set.

It is easy to see that $[1,4]^t = a[-1,2]^t, a \in \mathbb{R} \Rightarrow 1 = -a, 4 = 2a$, which is impossible.
Similarly, $[-1,2]^t = b[1,4]^t, b \in \mathbb{R}$ is not possible. Hence S is a linearly independent set. □

Example 3.18. *The set $\{A, B, C\}$ is linearly dependent in $M_2(\mathbb{R})$, where*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \text{ and } C = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

Example 3.19. *The set $\{1, x, x^2, \ldots, x^n\}$ is linearly independent in $\mathbb{R}_n[x]$.*

Example 3.20. *The set $\{1, x, x^2, \ldots\}$ is linearly independent in $\mathbb{R}[x]$.*

Result 3.5. *The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a vector space V are linearly dependent **iff** either $\mathbf{v}_1 = \mathbf{0}$ or there is an integer r such that $\mathbf{v}_r$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}$.*

Linear combinations of rows:

Suppose $A = \begin{bmatrix} \mathbf{a}_1^t \\ \mathbf{a}_2^t \\ \vdots \\ \mathbf{a}_m^t \end{bmatrix}$ is an $m \times n$ matrix. Then

- For $c_i \in \mathbb{R}$, $\mathbf{a} = c_1\mathbf{a}_1^t + \ldots + c_m\mathbf{a}_m^t$ is a linear combination of the rows of A. Note that $\mathbf{a}$ is a $1 \times n$ matrix and $\mathbf{a}^t \in \mathbb{R}^n$.
- Note: $c_1\mathbf{a}_1^t + \ldots + c_m\mathbf{a}_m^t = [c_1, \ldots, c_m]A$. Thus, for any $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{c}^tA$ is a linear combination of rows of A.
- The rows of A are linearly dependent **iff** $\mathbf{c}^tA = c_1\mathbf{a}_1^t + \ldots + c_m\mathbf{a}_m^t = \mathbf{0}^t$ (zero row) for some nonzero $\mathbf{c} \in \mathbb{R}^m$.
- The rows of A are linearly dependent **iff** $\mathbf{a}_1, \ldots, \mathbf{a}_m$ are linearly dependent, i.e., the columns of A^t are linearly dependent.

Result 3.6. *Suppose $R = RREF(A)$ has a zero row. Then the rows of A are linearly dependent.*

Proof. If there is a zero row in A, then clearly the rows of A are linearly dependent. Otherwise, a zero row in the RREF(A) can be obtained from adding multiples of some rows to a fixed row and then by moving the rows (by interchange of rows) so that zero rows are at the bottom. This process basically express the zero vector as a non-trivial linear combination of row vectors of A. Hence the rows of A are linearly dependent.

[The converse of this result is also true. You are asked to prove the converse in a practice problem.] □

Result 3.7. *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be vectors in $\mathbb{R}^n$ and let A be the $n \times m$ matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_m]$ with these vectors as its columns. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly dependent if and only if the homogeneous system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution.*

Result 3.8. *Any set of m vectors in $\mathbb{R}^n$ is linearly dependent if $m > n$.*

Result 3.9. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$ and consider the $m \times n$ matrix $A = \begin{bmatrix} \mathbf{v}_1^t \\ \mathbf{v}_2^t \\ \vdots \\ \mathbf{v}_m^t \end{bmatrix}$. Then S is linearly dependent if and only if $rank(A) < m$.*

Result 3.10. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$ and $A = [\mathbf{v}_1\ \ \mathbf{v}_2 \cdots \mathbf{v}_m]$. Then the following are equivalent.*

1. *S is linearly dependent.*
2. *Columns of A are linearly dependent.*
3. *$A\mathbf{x} = \mathbf{0}$ has a nontrivial solution.*
4. *Rows of A^t are linearly dependent.*
5. *$rank(A^t) < m$.*
6. *$RREF(A^t)$ has a zero row.*

Proof. Follows from previous few results. □

Result 3.11. *Let $S = \{\mathbf{v}_1 \ldots, \mathbf{v}_k\}$ be a linearly independent set in V and $\mathbf{v} \notin span(S)$ Then $S \cup \{\mathbf{v}\}$ is also linearly independent.*

Proof. Let $a\mathbf{u} + a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = \mathbf{0}$ for $a, a_1, \ldots, a_r \in \mathbb{F}$. Now if $a \neq 0$, then we have

$$a\mathbf{u} + a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = \mathbf{0} \Rightarrow \mathbf{u} = \left(-\frac{a_1}{a}\right)\mathbf{v}_1 + \ldots + \left(-\frac{a_r}{a}\right)\mathbf{v}_r \in \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\},$$

a contradiction to the given assumption. Therefore we must have $a = 0$. Now using $a = 0$, we have

$$\begin{aligned} &a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = \mathbf{0} \\ \Rightarrow &a_1 = 0, \ldots, a_r = 0, \text{ as } \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \text{ is linearly independent.} \end{aligned}$$

Thus $a\mathbf{u} + a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = \mathbf{0} \Rightarrow a = 0, a_1 = 0, \ldots, a_r = 0$ and hence $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is linearly independent. □

Basis: A subset B of a vector space V is said to be a **basis** for V if span$(B) = V$ and if B is linearly independent.

- The set $\{1\}$ is a basis for $\mathbb{R}^1$ $(= \mathbb{R})$.
- The set $\{\mathbf{e}_1, \mathbf{e}_2\}$, where $\mathbf{e}_1 = [1, 0]^t$ and $\mathbf{e}_2 = [0, 1]^t$, is a basis for $\mathbb{R}^2$. Similarly, $\{[1, 0]^t, [1, 1]^t\}$ is also a basis for $\mathbb{R}^2$.

- In general, if $\mathbf{e}_i = [0, \ldots, 0, 1, 0, \ldots, 0]^t$, where 1 is the i-th entry and the other entries are zero, then $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$.
- The vectors $\mathbf{e}_i$ (for $i = 1, 2, \ldots, n$) are called the $n \times 1$ standard **unit vectors**.
- If $\mathbf{e}_i$ is written as a row vector, then the vectors $\mathbf{e}_i$ (for $i = 1, 2, \ldots, n$) are called the $1 \times n$ standard **unit vectors**.

Example 3.21. $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ *is a basis for* $\mathbb{F}^n$. *This basis is called the* ***standard basis*** *for* $\mathbb{F}^n$.

Example 3.22. $\{1, x, x^2, \ldots, x^n\}$ *is a basis for* $\mathbb{R}_n[x]$, *known as the* ***standard basis*** *for* $\mathbb{R}_n[x]$.

Example 3.23. $\{1, x, x^2, \ldots\}$ *is a basis for* $\mathbb{R}[x]$, *known as the* ***standard basis*** *for* $\mathbb{R}[x]$.

Example 3.24. $\mathcal{E} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ *is a basis for* $M_2(\mathbb{R})$, *where*

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ and } E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example 3.25. $B = \{1 + x, x + x^2, 1 + x^2\}$ *is a basis for* $\mathbb{R}_2[x]$.

Result 3.12. *For a vector space* U, *a subset* $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq U$ *is a basis of* U *iff every element of* U *is a unique linear combination of* $\mathbf{v}_1, \ldots, \mathbf{v}_r$.

Proof. Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be a basis of U. Therefore B spans U and so every element of U can surely be expressed as a linear combination of elements of $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. Let $\mathbf{v} \in U$ be expressed as $\mathbf{v} = a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r$ and $\mathbf{v} = b_1\mathbf{v}_1 + \ldots + b_r\mathbf{v}_r$ for $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{R}$. Then we have

$$\begin{aligned}
& a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = \mathbf{v} = b_1\mathbf{v}_1 + \ldots + b_r\mathbf{v}_r \\
\Rightarrow & (a_1 - b_1)\mathbf{v}_1 + \ldots + (a_r - b_r)\mathbf{v}_r = \mathbf{0} \\
\Rightarrow & a_1 - b_1 = \ldots = a_r - b_r = 0, \quad \text{as } \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \text{ is linearly independent} \\
\Rightarrow & a_1 = b_1, \ldots, a_r = b_r.
\end{aligned}$$

Thus every element of U is a unique linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_r$.

Conversely, suppose that every element of U is a **unique** linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_r$. Then clearly B spans U. Now for $a_1, \ldots, a_r \in \mathbb{R}$, assume that $a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r = 0$. Since $0.\mathbf{v}_1 + \ldots + 0.\mathbf{v}_r = 0$, we find two linear combinations of the element $\mathbf{0}$. By the uniqueness of expressions of elements of U as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_r$, we find that $a_1 = 0, \ldots, a_r = 0$. Thus B is also linearly independent. Hence B is a basis of U. $\square$

Example 3.26. *Find a basis for the subspace* $S = \{\mathbf{x} \in \mathbb{C}^4 : A\mathbf{x} = \mathbf{0}\}$, *where*

$$A = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & -2 & -1 & 3 \\ -1 & 1 & -1 & 0 \end{bmatrix}.$$

Solution. Let $\mathbf{x} = [x, y, z, w]^t$ be a solution of $A\mathbf{x} = \mathbf{0}$. We have

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 0 \\ 2 & -2 & -1 & 3 & 0 \\ -1 & 1 & -1 & 0 & 0 \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ -1 & 1 & -1 & 0 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 + R_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 + 2R_2} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_1 \to R_1 + R_2} \left[\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The last RREF gives $x - y + w = 0$ and $z - w = 0$, where y and w are free variables. Setting $y = s, w = r$, we find $z = r, x = s - r$. So

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s - r \\ s \\ r \\ r \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + r \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Thus $\{[1,1,0,0]^t, [-1,0,1,1]^t\}$ spans S. Also for $a, b \in \mathbb{C}$,

$$a[1,1,0,0]^t + b[-1,0,1,1]^t = 0 \Rightarrow [a-b, a, b, b]^t = \mathbf{0} \Rightarrow a = 0 = b.$$

Thus $\{[1,1,0,0]^t, [-1,0,1,1]^t\}$ is also linearly independent. Hence $\{[1,1,0,0]^t, [-1,0,1,1]^t\}$ is a basis for S. □

Result 3.13. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq \mathbb{R}^n$ and $T \subseteq span(S)$ such that $m = |T| > r$. Then T is linearly dependent.*

Proof. Let $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Write

$$\mathbf{u}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \cdots + a_{ir}\mathbf{v}_r, \ 1 \le i \le m.$$

Let $A = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mr} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{a}_m^t \end{bmatrix}$. So $\mathbf{u}_i = \mathbf{a}_i^t \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_r \end{bmatrix}$.

Since $m > r$, the rows of A are linearly dependent. Suppose $\alpha_1 \mathbf{a}_1^t + \cdots + \alpha_m \mathbf{a}_m^t = \mathbf{0}^t$, where at least one $\alpha_i \neq 0$. Then

$$\sum_{i=1}^m \alpha_i \mathbf{u}_i = \sum_{i=1}^m \alpha_i \mathbf{a}_i^t \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_r \end{bmatrix} = \mathbf{0}^t \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_r \end{bmatrix} = \mathbf{0}^t.$$

□

Let $a, b, c \in \mathbb{R}$. Then

$$\begin{aligned} a(1+x) + b(x+x^2) + c(1+x^2) = 0 &\Rightarrow \begin{bmatrix} 1+x & x+x^2 & 1+x^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \\ &\Rightarrow \begin{bmatrix} 1 & x & x^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \\ &\Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{0}, \text{ as } \{1, x, x^2\} \text{ is LI} \\ &\Rightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{0}, \quad \text{as the rank of } \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ is } 3 \\ &\Rightarrow \{1+x, x+x^2, 1+x^2\} \quad \text{is LI.} \end{aligned}$$

- Note the correspondence $1+x \longleftrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x+x^2 \longleftrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad 1+x^2 \longleftrightarrow \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.
- $\{1+x, x+x^2, 1+x^2\}$ is LI **iff** $\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\}$ is LI.

Coordinate: Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis for a vector space $V(\mathbb{F})$ and let $\mathbf{v} \in V$. Let $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$. Then the scalars $c_1, c_2, \ldots, c_n$ are called the **coordinates of v with respect to** B, and the column vector

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

is called the **coordinate vector of v with respect to** B.
★ Coordinate of a vector is always associated with an **ordered** basis.

Example 3.27. *The coordinate vector $[p(x)]_B$ of $p(x) = 1 - 3x + 4x^2$ with respect to the ordered basis $B = \{1, x, x^2\}$ of $\mathbb{R}_2[x]$ is $[p(x)]_B = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$.*

Result 3.14. *Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis for a vector space V, let $\mathbf{u}, \mathbf{v} \in V$ and let $c \in \mathbb{F}$. Then*

$$[\mathbf{u}+\mathbf{v}]_B = [\mathbf{u}]_B + [\mathbf{v}]_B \qquad \text{and} \qquad [c\mathbf{u}]_B = c[\mathbf{u}]_B.$$

Result 3.15. *Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis for a vector space V, and let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ be vectors in V. Then $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is linearly independent in V if and only if $\{[\mathbf{u}_1]_B, [\mathbf{u}_2]_B, \ldots, [\mathbf{u}_k]_B\}$ is linearly independent in $\mathbb{F}^n$.*

Result 3.16. *Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V.*

1. *Any set of more than n vectors in V must be linearly dependent.*

2. *Any set of fewer than n vectors in V cannot span V.*

Result 3.17. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq V$ and $T \subseteq span(S)$ be such that $m = |T| > r$. Then T is linearly dependent.*

Proof. Similar to Result 3.13. □

Result 3.18. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$. Then S contains a basis of span(S).*

Proof. Let $\mathbf{v}_{i_1} \in S$ and let $\mathbf{v}_{i_1} \neq \mathbf{0}$. If $\text{span}(\{\mathbf{v}_{i_1}\}) \neq \text{span}(S)$, then $\text{span}(\{\mathbf{v}_{i_1}\})$ is a proper subset of $\text{span}(S)$. So, S must have some non-zero elements which do not belong to $\text{span}(\{\mathbf{v}_{i_1}\})$. Let $\mathbf{v}_{i_2} \in S \setminus \text{span}(\{\mathbf{v}_{i_1}\})$. Then clearly, $\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}\}$ is linearly independent. If $\text{span}(\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}\}) \neq \text{span}(S)$, then let $\mathbf{v}_{i_3} \in S \setminus \text{span}\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}\}$. Then clearly, $\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \mathbf{v}_{i_3}\}$ is linearly independent. Continuing in this way, we find a linearly independent set $\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_k}\} \subseteq S$ such that $\text{span}(\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_k}\}) = \text{span}(S)$. Thus $\{\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_k}\}$ is a basis for $\text{span}(S)$. □

- The vector space V is called **finite-dimensional** if V has a finite spanning set. Else, V is called **infinite-dimensional**.

Result 3.19 (The Basis Theorem). *Every finite-dimensional vector space V has a basis. Further, every basis for V has equal number of vectors.*

Dimension: Let V be a vector space.

- The **dimension** of V, denoted dim V, is the number of vectors in a basis for V. We write dim $V = \infty$ if V does not have a finite basis.

- The dimension of the zero space $\{\mathbf{0}\}$ is defined to be 0.

- The vector space V is called **finite-dimensional** if dim V is a finite number.

- The vector space V is called **infinite-dimensional** if there is no **finite** basis for V.

Example 3.28. *$dim(\mathbb{R}^n) = n$, $dim\ \mathbb{C}(\mathbb{C}) = 1$, $dim\ \mathbb{C}(\mathbb{R}) = 2$, $dim\ M_2(\mathbb{R}) = 4$ and $dim\ \mathbb{R}_n[x] = n + 1$.*

Solution. A basis of $\mathbb{R}^n$ over $\mathbb{R}$ is $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. A basis of $\mathbb{C}$ over $\mathbb{C}$ is $\{1\}$. A basis of $\mathbb{C}$ over $\mathbb{R}$ is $\{1, i\}$.

Result 3.20. *Let V be a vector space with dim $V = n$.*

1. *Any linearly independent set in V contains at most n vectors.*

2. *Any spanning set for V contains at least n vectors.*

3. *Any linearly independent set of exactly n vectors in V is a basis for V.*

4. *Any spanning set for V of exactly n vectors is a basis for V.*

5. *Any linearly independent set in V can be extended to a basis for V.*

6. *Any spanning set for V can be reduced to a basis for V.*

Result 3.21. *Let W be a subspace of a finite-dimensional vector space V. Then*

1. *W is also finite-dimensional and dim $W \leq$ dim V; and*

2. *dim W = dim V if and only if $W = V$.*

Proof. [Proof of **Part 2**] Let $\dim(W) = \dim(V) = k$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for W so that $\text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}) = W$. If $W \neq V$, then $\text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ is a proper subset of V and so there exists $\mathbf{v} \in V$ such that $\mathbf{v} \notin \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$. By Result 3.11, the set $\{\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent in V, which gives that $\dim(V) \geq k + 1$. This is a contradiction to the hypothesis that $\dim(W) = \dim(V)$. Hence we must have $W = V$. □

Change of Basis: Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $C = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be ordered bases for a vector space V. The $n \times n$ matrix whose columns are the coordinate vectors $[\mathbf{u}_1]_C, [\mathbf{u}_2]_C, \ldots, [\mathbf{u}_n]_C$ is denoted by $P_{C \leftarrow B}$, and is called the **change of basis matrix** from B to C. That is,

$$P_{C \leftarrow B} = [[\mathbf{u}_1]_C, [\mathbf{u}_2]_C, \ldots, [\mathbf{u}_n]_C].$$

Result 3.22. *Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $C = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be ordered bases for a vector space V and let $P_{C\leftarrow B}$ be the change of basis matrix from B to C. Then*

1. $P_{C\leftarrow B}[\mathbf{x}]_B = [\mathbf{x}]_C$ *for all* $\mathbf{x} \in V$*;*
2. $P_{C\leftarrow B}$ *is the unique matrix P with the property that* $P[\mathbf{x}]_B = [\mathbf{x}]_C$ *for all* $\mathbf{x} \in V$*;*
3. $P_{C\leftarrow B}P_{B\leftarrow C} = I_n = P_{B\leftarrow C}P_{C\leftarrow B}$.

Example 3.29. *Find the change of basis matrix $P_{C\leftarrow B}$ and $P_{B\leftarrow C}$ for the ordered bases $B = \{1, x, x^2\}$ and $C = \{1 + x, x + x^2, 1 + x^2\}$ of $\mathbb{R}_2[x]$. Then find the coordinate vector of $p(x) = 1 + 2x - x^2$ with respect to C.*

Practice Problems Set 3

1. Consider the vector space $\mathbb{R}^2$ over $\mathbb{R}$. Give an example of a subset of $\mathbb{R}^2$ which is
   (a) closed under addition but not closed under scalar multiplication; and
   (b) closed under scalar multiplication but not closed under addition.

2. Let S be a non-empty set and let V be the set of all functions from S to $\mathbb{R}$. Show that V is a vector space with respect to addition $(f + g)(x) = f(x) + g(x)$ for $f, g \in V$ and scalar multiplication $(c.f)(x) = cf(x)$ for $c \in \mathbb{R}, f \in V$.

3. Show that the space of all real (respectively complex) matrices is a vector space over $\mathbb{R}$ (respectively $\mathbb{C}$) with respect to usual addition and scalar multiplication of matrices.

4. A set V and the operation of vector addition and scalar multiplication are given below. Examine whether V is a vector space over $\mathbb{R}$. If not, then find which of the vector space properties it violates.
   (a) $V = \mathbb{R}^2$ and for $[x, y]^t, [z, w]^t \in V, a \in \mathbb{R}$ define $[x, y]^t + [z, w]^t = [x + z, y + w]^t$ and $a.[x, y]^t = [ax, 0]^t$.
   (b) $V = \mathbb{R}^2$ and for $[x, y]^t, [z, w]^t \in V, a \in \mathbb{R}$ define $[x, y]^t + [z, w]^t = [x + z, 0]^t$ and $a.[x, y]^t = [ax, 0]^t$.
   (c) $V = \mathbb{R}$ and for $x, y \in V, a \in \mathbb{R}$ define $x \oplus y = x - y$ and $a \odot x = -ax$.
   (d) $V = \mathbb{R}^2$ and for $[x, y]^t, [z, w]^t \in V, a \in \mathbb{R}$ define $[x, y]^t + [z, w]^t = [x + z, y + w]^t$ and $a.[x, y]^t = [ax, y]^t$.
   (e) $V = \mathbb{R}^2$ and for $[x, y]^t, [z, w]^t \in V, a \in \mathbb{R}$ define $[x, y]^t + [z, w]^t = [x - z, y - w]^t$ and $a.[x, y]^t = [-ax, -ay]^t$.
   (f) $V = \{[x, 3x + 1]^t : x \in \mathbb{R}\}$ with usual addition and scalar multiplication in $\mathbb{R}^2$.
   (g) $V = \{[x, mx + c]^t : x \in \mathbb{R}\}$, where m and c are some fixed real numbers, with usual addition and scalar multiplication in $\mathbb{R}^2$.
   (h) $V = \{f : \mathbb{R} \to \mathbb{C} \mid f \text{ is a function such that } f(-t) = \overline{f(t)} \text{ for all } t \in \mathbb{R}\}$, with usual addition and scalar multiplication of functions (as defined in Problem 2). Also, find a function in V whose range is contained in $\mathbb{R}$.
   (i) $V = M_2(\mathbb{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{R}, ad - bc \neq 0 \right\}$ with usual addition and scalar multiplication of matrices.

5. Let $V = \mathbb{R}^+$ (the set of positive real numbers). This is **not** a vector space under usual operations of addition and scalar multiplication (why?). Define a new vector addition and scalar multiplication as
$$v_1 \oplus v_2 = v_1.v_2 \quad \text{and} \quad \alpha \odot v = v^\alpha,$$
for all $v_1, v_2, v \in \mathbb{R}^+$ and $\alpha \in \mathbb{R}$. Show that $\mathbb{R}^+$ is a vector space over $\mathbb{R}$ with 1 as the additive identity.

6. (Optional) Let $V = \mathbb{R}^2$. Define $[x_1, x_2]^t \oplus [y_1, y_2]^t = [x_1 + y_1 + 1, x_2 + y_2 - 3]^t, \alpha \odot [x_1, x_2]^t = [\alpha x_1 + \alpha - 1, \alpha x_2 - 3\alpha + 3]^t$ for $[x_1, x_2]^t, [y_1, y_2]^t \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$. Verify that the vector $[-1, 3]^t$ is the additive identity, and V is indeed a vector space over $\mathbb{R}$.

7. (Optional) Let V and W be vector spaces over $\mathbb{R}$ with binary operations $(+, \bullet)$ and $(\oplus, \odot)$, respectively. Consider the following operations on the set $V \times W$: for $(x_1, y_1), (x_2, y_2) \in V \times W$ and $\alpha \in \mathbb{R}$, define
$$(x_1, y_1) \ominus (x_2, y_2) = (x_1 + x_2, y_1 \oplus y_2) \quad \text{and} \quad \alpha \circ (x_1, y_1) = (\alpha \bullet x_1, \alpha \odot y_1).$$
With the above definitions, show that $V \times W$ is also a vector space over $\mathbb{R}$.

8. Show that $\mathbb{R}$ is a vector space over $\mathbb{R}$ with additive identity 1 with respect to the addition $\oplus$ and scalar multiplication $\otimes$ defined as follows:
$$x \oplus y = x + y - 1 \text{ for } x, y \in \mathbb{R} \quad \text{and} \quad a \otimes x = a(x-1) + 1 \text{ for } a, x \in \mathbb{R}.$$

9. Let $V = \mathbb{C}[x]$ be the set of all polynomials with complex coefficients. Show that V is a vector space over $\mathbb{C}$ with respect to the following definition of addition and scalar multiplication: If $p(x) = \sum\limits_{i=0}^{n} a_i x^i, q(x) = \sum\limits_{i=0}^{m} b_i x^i \in V$, $m \geq n$, then
$$p(x) + q(x) = \sum_{i=0}^{m} (a_i + b_i)x^i, \text{ where } a_i = 0 \text{ for } i > n \text{ and } c.p(x) = \sum_{i=0}^{n} (ca_i)x^i \text{ for } c \in \mathbb{C}.$$

10. Prove that every vector space has a **unique** zero vector.

11. Prove that for every vector $\mathbf{v}$ in a vector space V, there is a **unique** $\mathbf{v}' \in V$ such that $\mathbf{v} + \mathbf{v}' = \mathbf{0}$.

12. Examine whether the following sets are subspaces of $\mathbb{R}^2$.
$$\{[x,y]^t \in \mathbb{R}^2 : x^2 + y^2 = 0\}, \quad \{[x,y]^t \in \mathbb{R}^2 : x = y\}, \quad \{[x,y]^t \in \mathbb{R}^2 : x, y \in \mathbb{Z}\},$$
$$\{[x,y]^t \in \mathbb{R}^2 : x - y = 1\}, \quad \{[x,y]^t \in \mathbb{R}^2 : xy \geq 0\} \quad \text{and} \quad \{[x,y]^t \in \mathbb{R}^2 : \frac{x}{y} = 1\}.$$

13. Examine whether the following sets are subspaces of $\mathbb{R}^3$.
$$\{[x,y,z]^t \in \mathbb{R}^3 \mid x \geq 0\}, \quad \{[x,y,z]^t \in \mathbb{R}^3 \mid x + y = z\} \quad \text{and} \quad \{[x,y,z]^t \in \mathbb{R}^3 \mid x = y^2\}.$$

14. Examine whether the following sets are subspaces of $\mathbb{R}^n$.
$$S_1 = \{[x_1, \ldots, x_n]^t : x_1 \geq 0\}, \qquad S_2 = \{[x_1, x_2, \ldots, x_n]^t) : x_2 = x_1^2\}$$
$$\text{and} \quad S_3 = \{[x_1, x_2, x_3, x_4, \ldots, x_n]^t : 3x_1 - x_2 + x_3 + 2x_4 = 0\}.$$

15. Determine the subspaces of $\mathbb{R}^3$ spanned by each of the following sets:
$$\{[1,1,1]^t, [0,1,2]^t, [1,0,-1]^t\}, \quad \{[1,2,3]^t, [1,3,5]^t\}, \quad \{[2,1,0]^t, [2,0,-2]^t\},$$
$$\{[1,2,3]^t, [1,3,5]^t, [1,2,4]^t\} \quad \text{and} \quad \{[1,2,0]^t, [1,0,5]^t, [1,2,3]^t\}.$$

16. Show that $S = \{[1,0,0]^t, [1,1,0]^t, [1,1,1]^t\}$ is a linearly independent set in $\mathbb{R}^3$. In general, if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a linearly independent set of vectors in some $\mathbb{R}^n$ then prove that $\{\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$ is also a linearly independent set.

17. Show, without using the following result, that any set of k vectors in $\mathbb{R}^3$ is linearly dependent if $k \geq 4$.

    Result: Any set of m vectors in $\mathbb{R}^n$ is linearly dependent if $m > n$.

18. Let W be a subspace of $\mathbb{R}^3$. Show that $\{[1,0,0]^t, [0,1,0]^t, [0,0,1]^t\} \subseteq W$ if and only if $W = \mathbb{R}^3$. Determine which of the following sets span $\mathbb{R}^3$:
$$\{[0,0,2]^t, [2,2,0]^t, [0,2,2]^t\}, \quad \{[3,3,1]^t, [1,1,0]^t, [0,0,1]^t\} \quad \text{and} \quad \{[-1,2,3]^t, [0,1,2]^t, [3,2,1]^t\}.$$

19. Determine whether $s(x) = 3 - 5x + x^2$ is in $\text{span}(p(x), q(x), r(x))$, where $p(x) = 1 - 2x$, $q(x) = x - x^2$ and $r(x) = -2 + 3x + x^2$.

20. Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a spanning set for a vector space V. Show that if $\mathbf{u}_k \in \text{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{k-1})$ then $\text{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}) = V$.

21. Let V be a vector space over $\mathbb{R}$ (or $\mathbb{C}$) and let $\emptyset \neq S \subseteq V$. Let $\mathcal{A} = \{W \mid S \subseteq W, W \text{ is a subspace of V}\}$. Show that $\text{span}(S) = \bigcap\limits_{W \in \mathcal{A}} W$ (*i.e.*, $\text{span}(S)$ is the smallest subspace of V containing S).

22. Prove or disprove:

    (a) If $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ are linearly dependent vectors in a vector space then each of these vectors is a linear combination of the other vectors.

(b) If any $r-1$ vectors of the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ are linearly independent in a vector space then the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ is also linearly independent.

(c) If $V = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and if every $\mathbf{u}_i$ $(1 \le i \le n)$ is a linear combination of no more than r vectors in $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\} \setminus \{\mathbf{u}_i\}$ then $\dim V \le r$.

23. If $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ are linearly independent vectors in a vector space V and the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{v}$ are linearly dependent, then show that $\mathbf{v}$ can be uniquely expressed as a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$.

24. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for a vector space V, where $n \ge 2$. Show that $\{\mathbf{u}_1, \mathbf{u}_1 + \mathbf{u}_2, \ldots, \mathbf{u}_1 + \ldots + \mathbf{u}_n\}$ is also a basis for V. How about the converse?

25. Discuss the linear independence/linear dependence of the following sets. For those that are linearly dependent, express one of the vectors as a linear combination of the others.

(a) $\{[1,0,0]^t, [1,1,0]^t, [1,1,1]^t\}$ of $\mathbb{R}^3$.

(b) $\{[1,0,0,0]^t, [1,1,0,0]^t, [1,2,0,0]^t, [1,1,1,1]^t\}$ of $\mathbb{R}^4$.

(c) $\{[1,i,0]^t, [1,0,1]^t, [i+2,-1,2]^t\}$ of $\mathbb{C}^3(\mathbb{C})$.

(d) $\{[1,i,0]^t, [1,0,1]^t, [i+2,-1,2]^t\}$ of $\mathbb{C}^3(\mathbb{R})$.

(e) $\{1, i\}$ of $\mathbb{C}(\mathbb{C})$ and $\{1, i\}$ of $\mathbb{C}(\mathbb{R})$.

(f) $\{1+x, 1+x^2, 1-x+x^2\}$ of $\mathbb{R}_2[x]$.

(g) $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} \right\}$ of $M_2(\mathbb{R})$.

26. Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a linearly independent set of vectors in a vector space.

(a) Is $\{\mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w}, \mathbf{w}+\mathbf{u}\}$ linearly independent? Either prove that it is or give a counterexample to show that it is not.

(b) Is $\{\mathbf{u}-\mathbf{v}, \mathbf{v}-\mathbf{w}, \mathbf{u}-\mathbf{w}\}$ linearly independent? Either prove that it is or give a counterexample to show that it is not.

27. Let A be a lower triangular matrix such that none of the diagonal entries are zero. Show that the row (or column) vectors of A are linearly independent.

28. Let V be a vector space and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$. Let $W = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ and $\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 - 2\mathbf{v}_2 \in \text{span}(\mathbf{v}_3, \mathbf{v}_4, \ldots, \mathbf{v}_k)$. Prove that $W = \text{span}(\mathbf{v}_3, \mathbf{v}_4, \ldots, \mathbf{v}_k)$.

29. Examine whether the sets $\{1-x, 1-x^2, x-x^2\}$ and $\{1, 1+2x+3x^2\}$ are bases for $\mathbb{R}_2[x]$.

30. Find all values of a for which the set $\{[a^2,0,1]^t, [0,a,2]^t, [1,0,1]^t\}$ is a basis for $\mathbb{R}^3$.

31. Find a basis for the solution space of the following system of $n+1$ linear equations in $2n$ unknowns:

$$\begin{aligned} x_1 + x_2 + \ldots + x_n &= 0 \\ x_2 + x_3 + \ldots + x_{n+1} &= 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \; &\vdots \\ x_{n+1} + x_{n+2} + \ldots + x_{2n} &= 0. \end{aligned}$$

32. Let $t \in \mathbb{R}$. Discuss the linear independence of the following three vectors over $\mathbb{R}$:

$$\mathbf{u} = [1,1,0]^t, \quad \mathbf{v} = [1,3,-1]^t, \quad \mathbf{w} = [5,3,t]^t.$$

33. Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and $\mathbf{x}$ be four linearly independent vectors in $\mathbb{R}^n$, where $n \ge 4$. Write true or false for each of the following statements, with proper justification:

(a) The vectors $\mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w}, \mathbf{w}+\mathbf{x}$ and $\mathbf{x}+\mathbf{u}$ are linearly independent.

(b) The vectors $\mathbf{u}-\mathbf{v}, \mathbf{v}-\mathbf{w}, \mathbf{w}-\mathbf{x}$ and $\mathbf{x}-\mathbf{u}$ are linearly independent.

(c) The vectors $\mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w}, \mathbf{w}+\mathbf{x}$ and $\mathbf{x}-\mathbf{u}$ are linearly independent.

(d) The vectors $\mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w}, \mathbf{w}-\mathbf{x}$ and $\mathbf{x}-\mathbf{u}$ are linearly independent.

34. Let $S = \{[x_1, x_2, \ldots, x_n]^t \in \mathbb{R}^n : \sum\limits_{i=1}^{n} x_i = 0\}$. Show that S is a subspace of $\mathbb{R}^n$. Find a basis and $\dim(S)$.

35. Prove that the row vectors of a matrix form a linearly dependent set if and only if there is a zero row in any row echelon form of that matrix.

36. Prove that the column vectors of a matrix form a linearly dependent set if and only if there is a column containing no leading term in any row echelon form of that matrix.

37. Prove that the column vectors of a matrix form a linearly independent set if and only if each column contains a leading term in any row echelon form of that matrix.

38. Let V be a vector space over $\mathbb{R}$ and let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ such that $\mathbf{x}+2\mathbf{y}+7\mathbf{z} = \mathbf{0}$. Show that $\operatorname{span}(\mathbf{x}, \mathbf{y}) = \operatorname{span}(\mathbf{y}, \mathbf{z}) = \operatorname{span}(\mathbf{z}, \mathbf{x})$.

39. Show that the set $\{[1, 0]^t, [i, 0]^t\}$ is linearly independent over $\mathbb{R}$ but is linearly dependent over $\mathbb{C}$.

40. Under what conditions on the complex number α are the vectors $[1+\alpha, 1-\alpha]^t$ and $[\alpha-1, 1+\alpha]^t$ in $\mathbb{C}^2(\mathbb{R})$ linearly independent?

41. Let V be a vector space over $\mathbb{C}$ and let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be a linearly independent subset of V. Show that the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, i\mathbf{x}_1, i\mathbf{x}_2, \ldots, i\mathbf{x}_k\}$ is linearly independent over $\mathbb{R}$.

42. Examine whether the following sets are subspaces of each of $\mathbb{C}^3(\mathbb{R})$ and $\mathbb{C}^3(\mathbb{C})$:

$$\{[z_1, z_2, z_3]^t \in \mathbb{C}^3 : z_1 \text{ is real}\}, \quad \{[z_1, z_2, z_3]^t \in \mathbb{C}^3 : z_1 + z_2 = 10z_3\},$$

$$\{[z_1, z_2, z_3]^t \in \mathbb{C}^3 : |z_1| = |z_2|\} \quad \text{and} \quad \{[z_1, z_2, z_3]^t \in \mathbb{C}^3 : z_1 + z_2 = 2\overline{z}_3\}.$$

43. Let U and W be two subspaces of a vector space V. Show that $U \cap W$ is also a subspace of V. Give examples to show that $U \cup W$ need not be a subspace of V.

44. Let U and W be subspaces of a vector space V. Define the *linear sum* of U and W to be

$$U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}.$$

Show that $U+W$ is a subspace of V. If $V = \mathbb{R}^3$, U is the x-axis and W is the y-axis, what is $U+W$ geometrically?

45. Let W be a subspace of a vector space V. Show that $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w} \in W\}$ is a subspace of $V \times V$ and that $\dim \Delta = \dim W$.

46. Let U and V be two finite-dimensional vector spaces. Find a formula for $\dim\,(U \times V)$ in terms of $\dim U$ and $\dim V$.

47. Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for a vector space V. Show that the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, c\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n\}$ is also a basis for V, where c is a non-zero scalar.

48. Find a basis for $\operatorname{span}(1-x, x-x^2, 1-x^2, 1-2x+x^2)$ in $\mathbb{R}_2[x]$.

49. Extend $\{1+x, 1+x+x^2\}$ to a basis for $\mathbb{R}_2[x]$.

50. Extend $\left\{ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$ to a basis for $M_2(\mathbb{R})$.

51. Prove or disprove:

(a) Every lineally independent set of a vector space V is a basis for some subspace of V.

(b) If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a lineally dependent subset of a vector space V, then $\mathbf{x}_n \in \operatorname{span}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1})$.

(c) Let W_1 and W_2 be two subspaces of a vector space V. If $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W_1 and $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$ is a basis for W_2, then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for $W_1 \cap W_2$.

(d) $B = \{x^3, x^3+2, x^2+1, x+1\}$ is a basis for $\mathbb{R}_3[x]$.

52. Find a basis for each of the following subspaces.

   (a) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) : a - d = 0\right\}$.

   (b) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) : 2a - c - d = 0, a + 3b = 0\right\}$.

   (c) $\{a + bx + cx^3 \in \mathbb{R}_3[x] : a - 2b + c = 0\}$.

   (d) $\{A = [a_{ij}] \in M_{m\times n}(\mathbb{R}) : \sum\limits_{j=1}^{n} a_{ij} = 0 \text{ for } i = 1, \ldots, m\}$.

53. Recall that $M_n(\mathbb{R})$ denote the space of all $n \times n$ real matrices. Show that the following sets are subspaces of $M_n(\mathbb{R})$. Also, find a basis and the dimension of each of these subspaces.

   (a) $U_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ is upper triangular}\}$.

   (b) $L_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ is lower triangular}\}$.

   (c) $D_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ is diagonal}\}$.

   (d) $sl_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : \mathrm{tr}(A) = 0\}$.

   (e) $A_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A + A^t = 0\}$.

54. Let $W = \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : a, b, c \in \mathbb{R}\right\}$. Show that W is a subspace of $M_2(\mathbb{R})$ and that the following set form an ordered basis for W:

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}.$$

Find the coordinate of the matrix $\begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix}$ with respect to this ordered basis.

55. Find a basis and the dimension for each of the following vector spaces:

   (a) $M_n(\mathbb{C})$ over $\mathbb{C}$;

   (b) $M_n(\mathbb{C})$ over $\mathbb{R}$;

   (c) $M_n(\mathbb{R})$ over $\mathbb{R}$.

   (d) $H_n(\mathbb{C})$ (the space of all $n \times n$ Hermitian matrices) over $\mathbb{R}$; and

   (e) $S_n(\mathbb{C})$ (the space of all $n \times n$ skew-Hermitian matrices) over $\mathbb{R}$.

56. Let $V = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{C}) : a + d = 0\right\}$ and $W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V : c = -\bar{b}\right\}$. Show that V is a vector space over $\mathbb{R}$ and find a basis for V. Also, show that W is a subspace of V, and find the dimension of W.

57. Show that $W = \{[x, y, z, t]^t \in \mathbb{R}^4 : x + y + z + 2t = 0 = x + y + z\}$ is a subspace of $\mathbb{R}^4$. Find a basis for W, and extend it to a basis for $\mathbb{R}^4$.

58. Show that $W = \{[v_1, \ldots, v_6]^t \in \mathbb{R}^6 \mid v_1 + v_2 + v_3 = 0, v_2 + v_3 + v_4 = 0, v_5 + v_6 = 0\}$ is a subspace of $\mathbb{R}^6$. Find a basis for W, and extend it to a basis for $\mathbb{R}^6$.

59. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for $\mathbb{R}^n$. Show that $\{\mathbf{v}_1, \alpha\mathbf{v}_1 + \mathbf{v}_2, \alpha\mathbf{v}_1 + \mathbf{v}_3, \ldots, \alpha\mathbf{v}_1 + \mathbf{v}_n\}$ is also a basis of $\mathbb{R}^n$ for every $\alpha \in \mathbb{R}$.

60. Show that the set $\{[1, 0, 1]^t, [1, i, 0]^t, [1, 1, 1 - i]^t\}$ is a basis for $\mathbb{C}^3(\mathbb{C})$.

61. Show that $\{[1, 0, 0]^t, [1, 1, 1]^t, [1, 1, -1]^t\}$ is a basis for $\mathbb{C}^3(\mathbb{C})$. Is it a basis for $\mathbb{C}^3(\mathbb{R})$ as well? If not, extend it to a basis for $\mathbb{C}^3(\mathbb{R})$.

62. Show that $\{1, (x - 1), (x - 1)(x - 2)\}$ is a basis for $\mathbb{R}_2[x]$, and that $W = \{p(x) \in \mathbb{R}_2[x] : p(1) = 0\}$ is a subspace of $\mathbb{R}_2[x]$. Also, find $\dim W$.

63. For each of the following statements, write true or false with proper justification:

(a) $\{p(x) \in \mathbb{R}_3[x] : p(x) = ax + b\}$ is a subspace of $\mathbb{R}_3[x]$.

(b) $\{p(x) \in \mathbb{R}_3[x] : p(x) = ax^2\}$ is a subspace of $\mathbb{R}_3[x]$.

(c) $\{p(x) \in \mathbb{R}[x] : p(0) = 1\}$ is a subspace of $\mathbb{R}[x]$.

(d) $\{p(x) \in \mathbb{R}[x] : p(0) = 0\}$ is a subspace of $\mathbb{R}[x]$.

(e) $\{p(x) \in \mathbb{R}[x] : p(x) = p(-x)\}$ is a subspace of $\mathbb{R}[x]$.

64. Let B be a set of vectors in a vector space V with the property that every vector in V can be uniquely expressed as a linear combination of the vectors in B. Prove that B is a basis for V.

65. Let W_1 be the set of all real matrices of the form $\begin{bmatrix} x & -x \\ y & z \end{bmatrix}$ and W_2 be the set of all real matrices of the form $\begin{bmatrix} a & b \\ -a & c \end{bmatrix}$. Show that W_1 and W_2 are subspaces of $M_2(\mathbb{R})$. Find the dimension for each of $W_1, W_2, W_1 + W_2$ and $W_1 \cap W_2$.

66. Let W_1 and W_2 be two subspaces of $\mathbb{R}^8$ and $\dim(W_1) = 6, \dim(W_2) = 5$. What are the possible values for $\dim(W_1 \cap W_2)$?

67. Does there exist subspaces M and N of $\mathbb{R}^7$ such that $\dim(M) = 4 = \dim(N)$ and $M \cap N = \{\mathbf{0}\}$? Justify.

68. Let M be an m-dimensional subspace of an n-dimensional vector space V. Prove that there exists an $(n-m)$-dimensional subspace N of V such that $M + N = V$ and $M \cap N = \{\mathbf{0}\}$.

69. Let $W_1 = \{[x, y, z]^t \in \mathbb{R}^3 : 2x + y + 4z = 0\}$ and $W_2 = \{[x, y, z]^t \in \mathbb{R}^3 : x - y + z = 0\}$. Show that W_1 and W_2 are subspaces of $\mathbb{R}^3$, and find a basis for each of $W_1, W_2, W_1 \cap W_2$ and $W_1 + W_2$.

70. Let $W_1 = \text{span}([1, 1, 0]^t, [-1, 1, 0]^t)$ and $W_2 = \text{span}([1, 0, 2]^t, [-1, 0, 4]^t)$. Show that $W_1 + W_2 = \mathbb{R}^3$. Give an example of a vector $\mathbf{v} \in \mathbb{R}^3$ such that $\mathbf{v}$ can be expressed in two different ways in the form $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2$.

71. Are the vector spaces $M_{6\times4}(\mathbb{R})$ (the space of all 6×4 real matrices) and $M_{3\times8}(\mathbb{R})$ (the space of all 3×8 real matrices) isomorphic? Justify your answer.

72. Let $P = \text{span}([1, 0, 0]^t, [1, 1, 0]^t)$ and $Q = \text{span}([1, 1, 1]^t)$. Show that $\mathbb{R}^3 = P + Q$ and $P \cap Q = \{\mathbf{0}\}$. For an $\mathbf{x} \in \mathbb{R}^3$, find $\mathbf{x}_p \in P$ and $\mathbf{x}_q \in Q$ such that $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_q$. Is the choice of $\mathbf{x}_p$ and $\mathbf{x}_q$ unique? Justify.

73. Let $V = \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ is a function}\}$, $V_e = \{f \in V : f(-x) = f(x) \text{ for all } x \in \mathbb{R}\}$ and $V_o = \{f \in V : f(-x) = -f(x) \text{ for all } x \in \mathbb{R}\}$. Prove that V_e and V_o are subspaces of V, $V_e \cap V_o = \{\mathbf{0}\}$ and $V = V_e + V_o$.

74. Let V_1 and V_2 be two subspaces of a vector space V such that $V = V_1 + V_2$ and $V_1 \cap V_2 = \{\mathbf{0}\}$. Prove that for each vector $\mathbf{v} \in V$, there are unique vectors $\mathbf{v}_1 \in V_1$ and $\mathbf{v}_2 \in V_2$ such that $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. (In such a situation, V is called the direct sum of V_1 and V_2, and is written as $V = V_1 \oplus V_2$.)

75. Find the coordinate of $p(x) = 2 - x + 3x^2$ with respect to the ordered basis $\{1 + x, 1 - x, x^2\}$ of $\mathbb{R}_2[x]$.

76. Find the coordinate of $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ with respect to the ordered basis $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ of $M_2(\mathbb{R})$.

77. Consider the ordered bases $B = \{1 + x + x^2, x + x^2, x^2\}$ and $C = \{1, x, x^2\}$ for $\mathbb{R}_2[x]$, and let $p(x) = 1 + x^2$. Find the coordinate vectors $[p(x)]_B$ and $[p(x)]_C$. Also, find the change of basis matrices $P_{C\leftarrow B}$ and $P_{B\leftarrow C}$. Finally, compute $[p(x)]_B$ and $[p(x)]_C$ using the change of basis matrices and compare the result with the initially computed coordinates of $p(x)$.

78. Show that the set $\mathbb{R}_n[x]$ of all real polynomials of degree at most n is a subspace of the vector space $\mathbb{R}[x]$ of all real polynomials. Find a basis for $\mathbb{R}_n[x]$. Also, show that the set of all real polynomials of degree exactly n is not a subspace of $\mathbb{R}[x]$. Further, show that $\{x + 1, x^2 + x - 1, x^2 - x + 1\}$ is a basis for $\mathbb{R}_2[x]$. Finally, find the coordinates of the elements $2x - 1, x^2 + 1$ and $x^2 + 5x - 1$ of $\mathbb{R}_2[x]$ with respect to the above ordered basis.

79. For $1 \le i \le n$, let $\mathbf{x}_i = [0, \ldots, 0, 1, 1, \ldots, 1]^t \in \mathbb{R}^n$ (*i.e.*, the first $(i-1)$ entries of $\mathbf{x}_i$ are 0 and the rest are 1). Show that $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a basis for $\mathbb{R}^n$. Also, find the coordinates of the vectors $[1, 1, \ldots, 1]^t$, $[1, 2, 3, \ldots, n]^t$ and $[0, \ldots, 0, 1, 0, \ldots, 0]^t$ with respect to the ordered basis B.

80. Consider the ordered bases $B = \{[1, 2, 0]^t, [1, 3, 2]^t, [0, 1, 3]^t\}$ and $C = \{[1, 2, 1]^t, [0, 1, 2]^t, [1, 4, 6]^t\}$ for $\mathbb{R}^3$. Find the change of basis matrix P from B to C. Also, find the change of basis matrix Q from C to B. Verify that $PQ = I_3$.

81. Show that the vectors $\mathbf{u}_1 = [1, 1, 0, 0]^t, \mathbf{u}_2 = [0, 0, 1, 1]^t, \mathbf{u}_3 = [1, 0, 0, 4]^t$ and $\mathbf{u}_4 = [0, 0, 0, 2]^t$ form a basis for $\mathbb{R}^4$. Find the coordinates of each of the standard basis vectors of $\mathbb{R}^4$ with respect to the ordered basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$.

82. Let W be the subspace of $\mathbb{C}^3$ spanned by $\mathbf{u}_1 = [1, 0, i]^t$ and $\mathbf{u}_2 = [1 + i, 1, -1]^t$. Show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for W. Also, show that $\mathbf{v}_1 = [1, 1, 0]^t$ and $\mathbf{v}_2 = [1, i, 1 + i]^t$ are in W, and $\{\mathbf{v}_1, \mathbf{v}_2\}$ form a basis for W. Finally, find the coordinates of $\mathbf{u}_1$ and $\mathbf{u}_2$ with respect to the ordered basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

83. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for an n-dimensional vector space V. Show that $\{a_1\mathbf{u}_1, a_2\mathbf{u}_2, \ldots, a_n\mathbf{u}_n\}$ is also a basis for V, for any non-zero scalars $a_1, a_2, \ldots, a_n$. If the coordinate of a vector $\mathbf{v}$ with respect to the ordered basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$, what is the coordinate of $\mathbf{v}$ with respect to the ordered basis $\{a_1\mathbf{u}_1, a_2\mathbf{u}_2, \ldots, a_n\mathbf{u}_n\}$? What are the coordinates of $\mathbf{w} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n$ with respect to each of the ordered bases $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\{a_1\mathbf{u}_1, a_2\mathbf{u}_2, \ldots, a_n\mathbf{u}_n\}$?

84. Let $A = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be a set of vectors in an n-dimensional vectors space V, and let B be a basis for V. Let $S = \{[\mathbf{u}_1]_B, [\mathbf{u}_2]_B, \ldots, [\mathbf{u}_m]_B\}$ be the set of coordinate vectors of A with respect to the ordered basis B. Prove that $\text{span}(A) = V$ if and only if $\text{span}(S) = \mathbb{R}^n$.

85. Consider two ordered bases B and C for $\mathbb{R}_2[x]$. Find C, if $B = \{x, 1 + x, 1 - x + x^2\}$ and the change of basis matrix from B to C is
$$P_{C\leftarrow B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

86. Let V be an n-dimensional vector space with an ordered basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Let $P = [p_{ij}]$ be an $n \times n$ invertible matrix, and set
$$\mathbf{u}_j = p_{1j}\mathbf{v}_1 + p_{2j}\mathbf{v}_2 + \ldots + p_{nj}\mathbf{v}_n \quad \text{for all } j = 1, 2, \ldots, n.$$
Prove that $C = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an ordered basis for V and $P = P_{B\leftarrow C}$.

Hints to Practice Problems Set 3

1. Take $\{[x, y]^t : x \in \mathbb{R},\ y \in \mathbb{Z}\}$ and $\{[x, y]^t \in \mathbb{R}^2 : 2x + y = 0 \text{ or } x + 2y = 0\}$.

2. Routine check.

3. Routine check.

4. $(a), (b)$ $1[2, 2]^t \neq [2, 2]^t$. (c) Addition is not commutative. (d) $(1 + 1)[2, 2]^t \neq 1[2, 2]^t + 1[2, 2]^t$. (e) Addition is not commutative. $(f), (i)$ $\mathbf{0} \notin V$. (g) V is a vector space iff $c = 0$. (h) Routine Check.

5. Routine check.

6. Routine check.

7. Routine check.

8. Routine check.

9. Routine check.

10. If $\mathbf{0}$ and $\mathbf{0}'$ are two zeros then $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$.

11. $\mathbf{v} + \mathbf{v}' = \mathbf{0} = \mathbf{v} + \mathbf{v}''$. Now $\Rightarrow \mathbf{v}'' + (\mathbf{v} + \mathbf{v}') = \mathbf{v}'' + \mathbf{0} \Rightarrow (\mathbf{v}'' + \mathbf{v}) + \mathbf{v}' = \mathbf{v}'' \Rightarrow \mathbf{0} + \mathbf{v}' = \mathbf{v}''$.

12. Only the first two sets are subspaces.

13. Only the second set is a subspace.

14. Only the S_3 is a subspace.

15. The span of the first three sets is the plane $x - 2y + z = 0$. The last two sets span $\mathbb{R}^3$.

16. Easy.

17. For any $(\emptyset \neq) S \subseteq \mathbb{R}^3$, we have $\dim \text{span}(S) \leq \dim \mathbb{R}^3 = 3$.

18. Only the first and the third sets span $\mathbb{R}^3$.

19. No. Compare coefficients of $s(x) = ap(x) + bq(x) + cr(x)$ and solve for a, b, c.

20. $\text{span}(S) = \text{span}(S \setminus \{\mathbf{u}_k\})$.

21. $\text{span}(S) \in \mathcal{A} \Rightarrow \bigcap\limits_{W \in \mathcal{A}} W \subseteq \text{span}(S)$. Also $S \subseteq W \Rightarrow \text{span}(S) \subseteq W \Rightarrow \text{span}(S) \subseteq \bigcap\limits_{W \in \mathcal{A}} W$.

22. (a) Consider $\mathbf{u}_1 = [1,0]^t, \mathbf{u}_2 = [2,0]^t, \mathbf{u}_3 = [0,1]^t$. (a) Consider $\mathbf{u}_1 = [1,0,0]^t, \mathbf{u}_2 = [0,1,0]^t, \mathbf{u}_3 = [1,1,0]^t$.
(c) Consider $\mathbf{u}_1 = [1,0]^t, \mathbf{u}_2 = [2,0]^t, \mathbf{u}_3 = [0,1]^t, \mathbf{u}_4 = [0,2]^t$.

23. $a\mathbf{v} + \sum\limits_{i=1}^{k} a_i \mathbf{u}_i = \mathbf{0} \Rightarrow a \neq 0$. Also $\sum\limits_{i=1}^{k} b_i \mathbf{u}_i = \mathbf{v} = \sum\limits_{i=1}^{k} c_i \mathbf{u}_i \Rightarrow b_i = c_i$.

24. $a_1\mathbf{u}_1 + a_2(\mathbf{u}_1 + \mathbf{u}_2) + \ldots + a_n(\mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n) = \mathbf{0} \Rightarrow a_1 + \ldots + a_n = 0, \ldots, a_{n-1} + a_n = 0, a_n = 0$. The converse is also true.

25. $(a), (d), (f), (g)$ Linearly independent, (b) Linearly dependent, $[1,2,0,0]^t = 2[1,1,0,0]^t - [1,0,0,0]^t + 0[1,1,1,1]^t$,
(c) Linearly dependent, $[i+2, -1, 2]^t = i[1, i, 0]^t + 2[1,0,1]^t$,
(e) $\{1, i\}$ is linearly independent in $\mathbb{C}(\mathbb{R})$, but linearly dependent in $\mathbb{C}(\mathbb{C})$, $1 = -i.i$.

26. (a) Yes. (b) $(\mathbf{u} - \mathbf{v}) + (\mathbf{v} - \mathbf{w}) = \mathbf{u} - \mathbf{w}$.

27. Similar to **Problem 24**. Or, the reduced row echelon form of A is I_n.

28. $\mathbf{v}_1 = \frac{1}{3}[2(\mathbf{v}_1 + \mathbf{v}_2) + (\mathbf{v}_1 - 2\mathbf{v}_2)]$, $\mathbf{v}_2 = \frac{1}{3}[(\mathbf{v}_1 + \mathbf{v}_2) - (\mathbf{v}_1 - 2\mathbf{v}_2)]$.

29. None of them are bases. Notice that $x - x^2 = (1 - x^2) - (1 - x)$, and the second set is linearly independent but not spanning.

30. $a \neq 0, 1, -1$.

31. $x_{n+2}, x_{n+3}, \ldots, x_{2n}$ are the free variables. A basis for the solutions space is $\{f_1, f_2, \ldots, f_{n-1}\}$, where

$$f_i = [-1, 0, 0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0, 1, 0, \ldots, 0],\ 1 \leq i \leq n-1.$$

Note that, the first and the $(n+1)$-th entry of each f_i is -1. The two positive entries of f_i is at the $(i+1)$-th position and at the $(n+i+1)$-th position, respectively.

32. The set is linearly independent iff $t \neq 1$.

33. Only part (c) is true.

34. A basis is $\{f_1, f_2, \ldots, f_{n-1}\}$, where

$$f_i = [-1, 0, 0, \ldots, 0, 1, 0, \ldots, 0],\ 1 \leq i \leq n-1.$$

Note that, the first entry of each f_i is -1 and the other positive entry of f_i is at the $(i+1)$-th position.

35. Similar to **Result 3.6**.

36. Take hint from **Result 3.5**.

37. Same as Problem no. 36.

38. $\mathbf{z} = -\frac{1}{7}(\mathbf{x} + 2\mathbf{y}), \mathbf{y} \in \text{span}(\mathbf{x}, \mathbf{y}) \Rightarrow \text{span}(\mathbf{y}, \mathbf{z}) \subseteq \text{span}(\mathbf{x}, \mathbf{y})$.

39. $[1,0]^t + i[i,0]^t = [0,0]^t$.

40. The set is linearly independent for all values of α.

41. $\sum a_k \mathbf{x}_k + \sum b_k (i\mathbf{x}_k) = \mathbf{0} \Rightarrow \sum (a_k + ib_k)\mathbf{x}_k = \mathbf{0}$.

42. The 1st and the 4th sets are subspaces of $\mathbb{C}^3(\mathbb{R})$ but not of $\mathbb{C}^3(\mathbb{C})$. The 2nd set is a subspace of each of $\mathbb{C}^3(\mathbb{R})$ and $\mathbb{C}^3(\mathbb{C})$. The 3rd set is neither a subspace of $\mathbb{C}^3(\mathbb{R})$ nor a subspace of $\mathbb{C}^3(\mathbb{C})$.

43. Take $U = \{[x,0]^t : x \in \mathbb{R}\}$ and $W = \{[0,x]^t : x \in \mathbb{R}\}$.

44. Routine check. $U + W$ will represent the xy-plane in $\mathbb{R}^3$.

45. If $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is a basis for W then $\{(\mathbf{w}_i, \mathbf{w}_i) : i = 1, 2, \ldots, k\}$ is a basis for Δ.

46. $\dim(U \times V) = \dim(U).\dim(V)$.

47. $\sum\limits_{k \neq i} a_k \mathbf{x}_k + a_i(c\mathbf{x}_i) = \mathbf{0} \Rightarrow a_k = \mathbf{0}$ for $k \neq i$ and $ca_i = \mathbf{0}$.

48. $\{1 - x, x - x^2\}$.

49. $\{1, 1 + x, 1 + x + x^2\}$ (some other answers are also possible).

50. $\left\{\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}$ (some other answers are also possible).

51. $(a), (d)$ True statements. (b) Take $\mathbf{x}_1 = [1,0]^t, \mathbf{x}_2 = [2,0]^t, \mathbf{x}_3 = [0,1]^t$. (c) If $W_1 = \text{span}(1, x, x + x^2, x^3)$, $W_2 = \text{span}(1, x, x^2, x^4, x^5)$ then $W_1 \cap W_2 = \text{span}(1, x, x^2) = \text{span}(1, x, x + x^2)$.

52. (a) $\left\{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, (b) $\left\{\begin{bmatrix} 3 & -1 \\ 6 & 0 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 0 & 6 \end{bmatrix}\right\}$, (c) $\{2 + x, 1 - x^3\}$,
(d) Let $A_{lk} = [a_{ij}]$, where $a_{l1} = 1, a_{lk} = -1$ and $a_{ij} = 0$ otherwise. Then $\{A_{lk} : 1 \leq l \leq m, 2 \leq k \leq n\}$ is a basis.

53. (a) $\{A_{lk} : 1 \leq l \leq n, 1 \leq k \leq n, l \leq k\}$, where $A_{lk} = [a_{ij}]$ and $a_{ij} = \begin{cases} 1 & \text{if } i = l, j = k, \\ 0 & \text{otherwise.} \end{cases}$

(b) $\{B_{lk} : 1 \leq l \leq n, 1 \leq k \leq n, l \geq k\}$, where $B_{lk} = [b_{ij}]$ and $b_{ij} = \begin{cases} 1 & \text{if } i = l, j = k, \\ 0 & \text{otherwise.} \end{cases}$

(c) $\{A_l : 1 \leq l \leq n\}$, where $A_l = [a_{ij}]$ and $a_{ij} = \begin{cases} 1 & \text{if } i = j = l, \\ 0 & \text{otherwise.} \end{cases}$

(d) $\{A_{lk} : 1 \leq l \leq n, 1 \leq k \leq n, l \neq k\} \cup \{B_l : 2 \leq l \leq n\}$, where $A_{lk} = [a_{ij}]$, $a_{ij} = \begin{cases} 1 & \text{if } i = l, j = k, \\ 0 & \text{otherwise,} \end{cases}$ and $B_l = [b_{ij}]$, $b_{ij} = \begin{cases} 1 & \text{if } i = j = 1, \\ -1 & \text{if } i = j = l, \\ 0 & \text{otherwise.} \end{cases}$

(e) $\{A_{lk} : 1 \leq l \leq n, 1 \leq k \leq n, l < k\}$, where $A_{lk} = [a_{ij}]$ and $a_{ij} = \begin{cases} 1 & \text{if } i = l, j = k, \\ -1 & \text{if } i = k, j = l, \\ 0 & \text{otherwise.} \end{cases}$

54. The required coordinate is $[1, -2, 3]^t$.

55. $(a), (c), (d), (e)$ n^2, (b) $2n^2$.

56. A basis for V is $\left\{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ i & 0 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}\right\}$.
A basis for W is $\left\{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}\right\}$.

57. $B = \{[1, -1, 0, 0]^t, [1, 0, -1, 0]^t\}$ is a basis for W. An extension is $B \cup \{[1, 0, 0, 0]^t, [0, 0, 0, 1]^t\}$.

58. $B = \{[0, 0, 0, 0, 1, -1]^t, [0, 1, -1, 0, 0, 0]^t, [1, 0, -1, 1, 0, 0]^t\}$ is a basis for W. An extension is $B \cup \{[1, 0, 0, 0, 0, 0]^t, [0, 1, 0, 0, 0, 0]^t, [0, 0, 0, 0, 1, 0]^t\}$.

59. $a_1\mathbf{v}_1 + \sum\limits_{i=2}^{n} a_i(\alpha\mathbf{v}_1 + \mathbf{v}_i) = \mathbf{0} \Rightarrow a_1 + \alpha a_2 + \ldots + \alpha a_n = 0, a_2 = 0, \ldots, a_n = 0.$

60. Examine that the given set is linearly independent.

61. The given set is linearly independent over $\mathbb{C}$. An extension is $\{[1,0,0]^t, [1,1,1]^t, [1,1,-1]^t, [i,0,0]^t, [i,i,i]^t, [i,i,-i]^t\}$.

62. A basis for W is $\{x-1, (x-1)(x-2)\}$.

63. Only part (c) is not a subspace.

64. If $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ then because of the uniqueness, $\sum a_i\mathbf{v}_i = \mathbf{0} \Rightarrow a_i = 0$.

65. $\dim(W_1) = 3, \dim(W_2) = 3, \dim(W_1 + W_2) = 4$ and $\dim(W_1 \cap W_2) = 2$.

66. $\dim(W_1 \cap W_2) \le \dim(W_2) = 5$. Also $\dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 + W_2) \ge 6 + 5 - 8 = 3$.

67. $\dim(M \cap N) = \dim(M) + \dim(N) - \dim(M + N) \ge 4 + 4 - 7 = 1$.

68. If $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a basis for M and $B \cup \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-m}\}$ is an extension of B, then consider $N = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{n-m})$.

69. $\{[-1,2,0]^t, [-2,0,1]^t\}$, $\{[1,1,0]^t, [1,0,-1]^t\}$, $\{[-5,-2,3]^t\}$ and $\{[1,1,0]^t, [1,0,-1]^t, [-1,2,0]^t\}$ are bases for W_1, $W_2, W_1 \cap W_2$ and $W_1 + W_2$, respectively.

70. $\{[1,1,0]^t, [-1,1,0]^t, [1,0,2]^t\}$ is linearly independent. Also $[1,1,1]^t = [0,1,0]^t + [1,0,1]^t = [1,1,0]^t + [0,0,1]^t$.

71. Both the spaces have equal dimension.

72. The set $\{[1,0,0]^t, [1,1,0]^t, [1,1,1]^t\}$ is linearly independent. Also $[x,y,z]^t = [x-z, y-z, 0]^t + [z,z,z]^t$.

73. $f \in V_e \cap V_o \Rightarrow f(x) = f(-x), f(-x) = -f(x)$ for all $x \in \mathbb{R} \Rightarrow f(x) = -f(x)$ for all $x \in \mathbb{R} \Rightarrow f = \mathbf{0}$. Also $f = f_e + f_o$, where $f_e(x) = \frac{f(x)+f(-x)}{2}$ and $f_o(x) = \frac{f(x)-f(-x)}{2}$.

74. $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_1' + \mathbf{v}_2' \Rightarrow \mathbf{v}_1 - \mathbf{v}_1' = \mathbf{v}_2' - \mathbf{v}_2 \in V_1 \cap V_2 = \{\mathbf{0}\}$.

75. $[\frac{1}{2}, \frac{3}{2}, 3]^t$.

76. $[-1,-1,-1,4]^t$.

77. $[p(x)]_B = [1,-1,1]^t$, $[p(x)]_C = [1,0,1]^t$, $P_{C\leftarrow B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, $P_{B\leftarrow C} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

78. $(1 + x^n) + (-x^n) = 1$. If B is the given basis then $[2x-1]_B = [\frac{1}{2}, \frac{3}{4}, -\frac{3}{4}]^t$, $[x^2+1]_B = [\frac{1}{2}, \frac{1}{4}, -\frac{3}{4}]^t$ and $[x^2+5x-1]_B = [2,2,-1]^t$.

79. Let $\mathbf{u}_1 = [1,1,\ldots,1]^t, \mathbf{u}_2 = [1,2,3,\ldots,n]^t$ and $\mathbf{u}_3 = [0,\ldots,0,1,0,\ldots,0]^t$. Then $[\mathbf{u}_1]_B = \mathbf{e}_1, [\mathbf{u}_2]_B = \mathbf{u}_1$ and $[\mathbf{u}_3]_B = [0,\ldots,0,1,-1,0,\ldots,0]^t$, where 1 is at the i-th place.

80. $P_{C\leftarrow B} = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 3 & -1 \\ -1 & -1 & 1 \end{bmatrix}$, $P_{B\leftarrow C} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

81. $[\mathbf{e}_1]_B = [0,0,1,-2]^t, [\mathbf{e}_2]_B = [1,0,-1,2]^t, [\mathbf{e}_3]_B = [0,1,0,-\frac{1}{2}]^t$ and $[\mathbf{e}_4]_B = [0,0,0,\frac{1}{2}]^t$.

82. $\mathbf{v}_1 = -i\mathbf{u}_1 + \mathbf{u}_2$, $\mathbf{v}_2 = (2-i)\mathbf{u}_1 + i\mathbf{u}_2$, $[\mathbf{u}_1]_B = [\frac{1-i}{2}, \frac{1+i}{2}]^t$, $[\mathbf{u}_2]_B = [\frac{i+1}{2}, \frac{i-1}{2}]^t$.

83. $[\mathbf{v}]_C = [\frac{x_1}{a_1}, \ldots, \frac{x_n}{a_n}]^t$, $[\mathbf{w}]_B = [1,1,\ldots,1]^t$ and $[\mathbf{w}]_C = [\frac{1}{a_1}, \ldots, \frac{1}{a_n}]^t$.

84. Let $\operatorname{span}(A) = V$ and $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathbf{x} = [x_1, \ldots, x_n]^t$. Take $\mathbf{v} = \sum\limits_{i=1}^{n} x_i\mathbf{v}_i$.
Then $\mathbf{v} = \sum\limits_{i=1}^{m} y_i\mathbf{u}_i \Rightarrow \mathbf{x} = [\mathbf{v}]_B = y_1[\mathbf{u}_1]_B + \ldots + y_m[\mathbf{u}_m]_B$.

85. $P_{B\leftarrow C} = (P_{C\leftarrow B})^{-1}$ and $C = \{1 - 2x + 2x^2, 2x - x^2, 1 - 3x + 2x^2\}$.

86. $a_1\mathbf{u}_1 + \ldots + a_n\mathbf{u}_n = \mathbf{0} \Rightarrow P\mathbf{a} = \mathbf{0}$, where $\mathbf{a} = [a_1, \ldots, a_n]^t \Rightarrow \mathbf{a} = \mathbf{0}$.

4 The Inverse of a Matrix

Definition 4.1. *An $n \times n$ matrix A is said to be* ***invertible*** *if there exists an $n \times n$ matrix B satisfying $AB = I_n = BA$, and B is called an* ***inverse*** *of A.*

- Note that we can talk of invertibility only for square matrices.
- For example, the matrix $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ is invertible since

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}.$$

- It is easy to see that the zero matrix $\mathbf{O}$ is never invertible.
- The matrix $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ is not invertible.

Result 4.1. *If A is an invertible matrix, then its inverse is unique.*

- We write A^{-1} to denote the inverse of an invertible matrix A.
- That is, if A is invertible then $AA^{-1} = I_n = A^{-1}A$.
- If A is a 1×1 invertible matrix, what is A^{-1}?

Result 4.2. *Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$ then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. If $ad - bc = 0$ then A is not invertible.*

Result 4.3. *Let A and B be two invertible matrices of the same size.*

1. *The matrix A^{-1} is also invertible, and $(A^{-1})^{-1} = A$.*
2. *If $c \neq 0$ then cA is also invertible, and $(cA)^{-1} = \frac{1}{c}A^{-1}$.*
3. *The matrix AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.*
4. *The matrix A^t is invertible, and $(A^t)^{-1} = (A^{-1})^t$.*
5. *For any integer n, the matrix A^n is invertible, and $(A^n)^{-1} = (A^{-1})^n$.*

Proof. [**Proof of Part 2**] We have $(cA)(\frac{1}{c}A^{-1}) = c.\frac{1}{c}AA^{-1} = I$ and $(\frac{1}{c}A^{-1})(cA) = \frac{1}{c}.cA^{-1}A = I$. Hence cA is also invertible, and $(cA)^{-1} = \frac{1}{c}A^{-1}$.

[**Proof of Part 4**] Using the fact $(AB)^t = B^tA^t$, we have

$$A^t(A^{-1})^t = (A^{-1}A)^t = I^t = I \quad \text{and} \quad (A^{-1})^tA^t = (AA^{-1})^t = I^t = I.$$

Hence A^t is invertible, and $(A^t)^{-1} = (A^{-1})^t$. □

Elementary Matrices: An **elementary matrix** is a matrix that can be obtained by performing **an** elementary row operation on the identity matrix.

- Since there are three types of elementary row operations, there are three types of corresponding elementary matrices.
- For example, the following are examples of the three types of elementary matrices of size 3.

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \; E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

- Note that E_1 is obtained by performing $R_2 \leftrightarrow R_3$ on I_3, E_2 is obtained by performing $R_2 \to 5R_2$ on I_3 and E_3 is obtained by performing $R_3 \to R_3 - 2R_1$ on I_3.

- Let A be the 3×3 matrix as given below:
$$A = \begin{bmatrix} a & b & c \\ x & y & z \\ p & q & r \end{bmatrix}.$$
Then we have
$$E_1A = \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix},\ E_2A = \begin{bmatrix} a & b & c \\ 5x & 5y & 5z \\ p & q & r \end{bmatrix} \text{ and } E_3A = \begin{bmatrix} a & b & c \\ x & y & z \\ p-2a & q-2b & r-2c \end{bmatrix}.$$
- Notice that E_1A is the matrix obtained from A by performing the elementary row operation $R_2 \leftrightarrow R_3$.
- The matrix E_2A is the matrix obtained from A by performing the elementary row operation $R_2 \to 5R_2$.
- The matrix E_3A is the matrix obtained from A by performing the elementary row operation $R_3 \to R_3 - 2R_1$.

Result 4.4.

1. *Let E be an elementary matrix obtained by an elementary row operation on I_n. If the same elementary row operation is performed on an $n \times r$ matrix A, then the resulting matrix is equal to EA.*
2. *The matrices B and A are row equivalent iff there are elementary matrices $E_1, \cdots, E_k$ such that $B = E_k \cdots E_1 A$.*

Proof.

1. Let us write $A = [a_{ij}]$ as a row of columns $A = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \ \ldots \ \mid \mathbf{a}_r]$. Let $E = [r_{ij}]$, a column of rows $\begin{bmatrix} \mathbf{r}_1 \\ \hline \mathbf{r}_2 \\ \hline \vdots \\ \hline \mathbf{r}_n \end{bmatrix}$. The ij-th entry of EA is $\sum\limits_{k=1}^{n} r_{ik}a_{kj}$, that is, $\mathbf{r}_i\mathbf{a}_j$. We consider three cases.

   <u>Case I</u>: Let E be the elementary matrix corresponding to the row operation $R_i \leftrightarrow R_j$. Then each row $\mathbf{r}_k$, for $k \neq i, j$, has 1 on the k-th place and zeroes elsewhere. The row $\mathbf{r}_i$ has 1 on the j-th place and zeroes elsewhere. The row $\mathbf{r}_j$ has 1 on the i-th place and zeroes elsewhere. For $k \neq i, j$, we have
$$kl\text{-th entry of } EA = \mathbf{r}_k\mathbf{a}_l = r_{k1}a_{1l} + \ldots + r_{kk}a_{kl} + \ldots + r_{kn}a_{nl} = r_{kk}a_{kl} = a_{kl} = kl\text{-th entry of } A.$$
   Thus all rows of EA except for the i-th and j-th rows coincide with the corresponding rows of A. Again
$$\begin{aligned} il\text{-th entry of } EA =& \mathbf{r}_i\mathbf{a}_l \\ =& r_{i1}a_{1l} + \ldots + r_{ij}a_{jl} + \ldots + r_{in}a_{nl} \\ =& r_{ij}a_{jl}, \text{ as } \mathbf{r}_i \text{ has 1 on the } j\text{-th place and zero elsewhere} \\ =& a_{jl} \\ =& jl\text{-th entry of } A. \end{aligned}$$
   Also
$$\begin{aligned} jl\text{-th entry of } EA =& \mathbf{r}_j\mathbf{a}_l \\ =& r_{j1}a_{1l} + \ldots + r_{ji}a_{il} + \ldots + r_{jn}a_{nl} \\ =& r_{ji}a_{il}, \text{ as } \mathbf{r}_j \text{ has 1 on the } i\text{-th place and zero elsewhere} \\ =& a_{il} \\ =& il\text{-th entry of } A. \end{aligned}$$
   Thus the i-th row of EA is equal to the j-th row of A and the j-th row of EA is equal to the i-th row of A. Thus we see that EA is obtained from A by applying the same row operation used in order to get E from I.

   <u>Case II</u>: Let E be the elementary matrix corresponding to the row operation $R_i \to xR_i, x \neq 0$. Then each row $\mathbf{r}_k$, for $k \neq i$, has 1 on the k-th place and zeroes elsewhere. The row $\mathbf{r}_i$ has x on the i-th place and zeroes elsewhere. For $k \neq i$, we have
$$kl\text{-th entry of } EA = \mathbf{r}_k\mathbf{a}_l = r_{k1}a_{1l} + \ldots + r_{kk}a_{kl} + \ldots + r_{kn}a_{nl} = r_{kk}a_{kl} = a_{kl} = kl\text{-th entry of } A.$$

Thus all rows of EA except for the i-th row coincide with the corresponding rows of A. Again

$$\begin{aligned}
il\text{-th entry of } EA =& \mathbf{r}_i \mathbf{a}_l \\
=& r_{i1}a_{1l} + \ldots + r_{ii}a_{il} + \ldots + r_{in}a_{nl} \\
=& r_{ii}a_{il}, \text{ as } \mathbf{r}_i \text{ has } x \text{ on the } i\text{-th place and zero elsewhere} \\
=& x a_{il}.
\end{aligned}$$

Thus the i-th row of EA is equal to x times the i-th row of A. Thus we see that EA is obtained from A by applying the same operation used in order to get E from I.

Case III: Let E be the elementary matrix corresponding to the row operation $R_i \to R_i + xR_j$. Then each row $\mathbf{r}_k$, for $k \neq i$, has 1 on the k-th place and zeroes elsewhere. The row $\mathbf{r}_i$ has 1 on the i-th place and x on the j-th place and zeroes elsewhere. For $k \neq i$, we have

$$kl\text{-th entry of } EA = \mathbf{r}_k \mathbf{a}_l = r_{k1}a_{1l} + \ldots + r_{kk}a_{kl} + \ldots + r_{kn}a_{nl} = r_{kk}a_{kl} = a_{kl} = kl\text{-th entry of } A.$$

Thus all rows of EA except for the i-th row coincide with the corresponding rows of A. Again

$$\begin{aligned}
& il\text{-th entry of } EA \\
=& \mathbf{r}_i \mathbf{a}_l \\
=& r_{i1}a_{1l} + \ldots + r_{ii}a_{il} + \ldots + r_{ij}a_{jl} + \ldots + r_{in}a_{nl} \\
=& r_{ii}a_{il} + r_{ij}a_{jl}, \text{ as } \mathbf{r}_i \text{ has 1 on the } i\text{-th place and } x \text{ on the } j\text{-th place and zero elsewhere} \\
=& a_{il} + x a_{jl}.
\end{aligned}$$

Thus the i-th row of EA is obtained by adding x times the j-th row of A to the i-th row of A. Thus we see that EA is obtained from A by applying the same operation used in order to get E from I.

2. The matrices B and A are row equivalent **iff** B can be obtained from A by applying a finite sequence of elementary row operations. Since each application of elementary row operations on A is as good as pre-multiplication of A by the corresponding elementary matrix, we find a sequence of elementary matrices $E_1, E_2, \cdots, E_k$ such that $B = E_k \cdots E_2 E_1 A$. ☐

Note that every elementary row operation can be undone or reversed. Applying this fact to the previous elementary matrices E_1, E_2 and E_3, we see that they are invertible.

- Indeed, applying $R_2 \leftrightarrow R_3$ on I_3 we find E_1^{-1}, applying $R_2 \to \frac{1}{5}R_2$ on I_3 we find E_2^{-1}, and applying $R_3 \to R_3 + 2R_1$ on I_3 we find E_3^{-1}.

We have

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = E_1, \; E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

Result 4.5. *Every elementary matrix is invertible, and its inverse is an elementary matrix of the same type.*

Proof. We consider three cases.

Case I: Let E be the elementary matrix corresponding to the elementary row operation $R_i \leftrightarrow R_j$. It is easy to see that two times application of this row operation on any matrix does not change the matrix. Since the result of application of the row operation on A is equal to EA, we find that $EEA = A$. In particular, we have $EE = EEI = I$. Hence E is invertible, and $E = E^{-1}$.

Case II: Let $k \neq 0$, and let E and F be the elementary matrices corresponding to the elementary row operations $R_i \leftrightarrow kR_i$ and $R_i \leftrightarrow \frac{1}{k}R_i$, respectively. For any matrix $A = [a_{ij}]$, using **Result 4.4**, the i-th row of EA becomes $[ka_{i1} \quad ka_{i2} \quad \ldots \quad ka_{ij} \quad \ldots ka_{in}]$ and all other rows remain unchanged. Therefore the i-th row of FEA becomes $[a_{i1} \quad a_{i2} \quad \ldots \quad a_{ij} \quad \ldots a_{in}]$ and all other rows remain unchanged. That is, $FEA = A$. Similarly, $EFA = A$. In particular, for $A = I$, we have $FE = I = EF$. Hence E is invertible, and $F = E^{-1}$.

Case III: Let E and F be the elementary matrices corresponding to the elementary row operations $R_i \leftrightarrow R_i + kR_l$ and $R_i \leftrightarrow R_i - kR_l$, respectively. For any matrix $A = [a_{ij}]$, using **Result 4.4**, the i-th row EA becomes $[a_{i1} + ka_{l1} \quad \ldots \quad a_{ij} + ka_{lj} \quad \ldots a_{in} + ka_{ln}]$ and all other rows remain unchanged. Therefore the i-th row FEA becomes

$$[(a_{i1} + ka_{l1}) - ka_{l1} \quad \ldots \quad (a_{ij} + ka_{lj}) - ka_{lj} \quad \ldots (a_{in} + ka_{ln}) - ka_{ln}] = [a_{i1} \quad \ldots \quad a_{ij} \quad \ldots a_{in}],$$

and all other rows remain unchanged. That is, $FEA = A$. Similarly, $EFA = A$. In particular, for $A = I$, we have $FE = I = EF$. Hence E is invertible, and $F = E^{-1}$.

From the above three cases, we conclude that every elementary matrix is invertible, and its inverse is an elementary matrix of the same type. □

Result 4.6 (The Fundamental Theorem of Invertible Matrices: Version I). *Let A be an $n \times n$ matrix. Then the following statements are equivalent.*

1. *A is invertible.*
2. *A^t is invertible.*
3. *$A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^n$.*
4. *$A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$.*
5. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
6. *The reduced row echelon form of A is I_n.*
7. *The rows of A are linearly independent.*
8. *The columns of A are linearly independent.*
9. *$rank(A) = n$.*
10. *A is a product of elementary matrices.*

Proof. We will prove the equivalence of the statements as follows:

$$\begin{array}{ccccccccccccc}
(1) & \Longrightarrow & (3) & \Longrightarrow & (4) & \Longrightarrow & (5) & \Longrightarrow & (6) & \Longrightarrow & (10) & \Longrightarrow & (1) \\
\Updownarrow & & & & & & \Updownarrow & & \Updownarrow & & & & \\
(2) & & & & & & (8) & & (7),(9) & & & &
\end{array}$$

$\underline{(1) \Longleftrightarrow (2)}$ Follows from Part 4 of **Result 4.3**.

$\underline{(1) \Longrightarrow (3)}$ If A is invertible, then for every $\mathbf{b} \in \mathbb{R}^n$, the vector $\mathbf{y} = A^{-1}\mathbf{b}$ is a solution of $A\mathbf{x} = \mathbf{b}$.

$\underline{(3) \Longrightarrow (4)}$ Assume that $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^n$. We first claim that $\tilde{A}\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^n$, where $\tilde{A} = RREF(A)$. Suppose, if possible, $\tilde{A}\mathbf{x} = \mathbf{b}$ does not have a solution for some $\mathbf{b} \in \mathbb{R}^n$. Apply the elementary row operations on $[\tilde{A} \mid \mathbf{b}]$, that converts $\tilde{A}$ to A, to obtain the matrix $[A \mid \mathbf{b}']$. Since $[A \mid \mathbf{b}']$ and $[\tilde{A} \mid \mathbf{b}]$ are row-equivalent and $\tilde{A}\mathbf{x} = \mathbf{b}$ does not have a solution, we find that the system $A\mathbf{x} = \mathbf{b}'$ does not have a solution. This contradicts our initial assumption. Thus we find that $\tilde{A}\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^n$. In particular, $\tilde{A}\mathbf{x} = \mathbf{e}_n$ has a solution, where $\mathbf{e}_n = [0, 0, 0, \ldots, 0, 1]^t$.

If $\tilde{A}$ has a zero row, then, the last equation of $\tilde{A}\mathbf{x} = \mathbf{e}_n$ gives $0 = 1$, which is absurd. Hence we find that $RREF(A)$ does not have a zero row and therefore all the columns of A are leading columns. Consequently, all the variables of $A\mathbf{x} = \mathbf{b}$ are leading variables. Since there are no free variables, we find that $A\mathbf{x} = \mathbf{b}$ cannot have more than one solution for every $\mathbf{b}$ in $\mathbb{R}^n$.

Aliter

Assume that $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^n$. If $RREF(A)$ has a zero row, then some i-th row of A will be a linear combination of the remaining rows of A. Consequently, the i-th row row of $[A \mid \mathbf{e}_i]$ can be made $[0 \ \ 0 \ \ 0 \ \ \ldots \ \ 0 \mid 1]$ by elementary row operations. This gives that the system $A\mathbf{x} = \mathbf{e}_i$ is in-consistent, which is against the initial assumption. Thus all the rows of $RREF(A)$ are non-zero, and therefore all the columns of A are leading columns. Consequently, all the variables of $A\mathbf{x} = \mathbf{b}$ are leading variables. Since there are no free variables, we find that $A\mathbf{x} = \mathbf{b}$ cannot have more than one solution for every $\mathbf{b}$ in $\mathbb{R}^n$.

$\underline{(4) \Longrightarrow (5)}$ Find it from the text book.

$\underline{(5) \Longrightarrow (6)}$ Find it from the text book.

$\underline{(6) \Longrightarrow (10)}$ Find it from the text book.

$\underline{(10) \implies (1)}$ Find it from the text book.

$\underline{(5) \iff (8)}$ Let the columns of A be $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, respectively. For $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n]^t \in \mathbb{R}^n$, we find that $A\alpha = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n$. Therefore

$$\begin{aligned}
&A\mathbf{x} = \mathbf{0} \text{ has only the trivial solution} \\
\Leftrightarrow &A\alpha = \mathbf{0} \text{ implies } \alpha = \mathbf{0} \\
\Leftrightarrow &\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n = \mathbf{0} \text{ implies } \alpha_1 = 0 = \alpha_2 = \ldots = \alpha_n \\
\Leftrightarrow &\text{ the columns } \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \text{ are linearly independent.}
\end{aligned}$$

$\underline{(6) \iff (7)}$ Notice that, since A is a square matrix, $RREF(A)$ is **not** I_n **iff** $RREF(A)$ contains a zero row. Thus the contra-positive of the equivalence to be proved is: $RREF(A)$ contains a zero row **iff** the rows of A are linearly dependent, which follows from **Result 3.6**.

$\underline{(6) \iff (9)}$ Recall that rank(A) is defined to be the number of non-zero rows in the reduced row echelon form (RREF) of A. Therefore if $RREF(A)$ is I_n, then clearly rank$(A) = n$. Conversely, if rank$(A) = n$, then there **cannot** be a zero row in the $RREF(A)$. Since A is a square matrix and all the rows of $RREF(A)$ are non-zero, we find that all the rows of $RREF(A)$ contains a leading term. Consequently, the columns of $RREF(A)$ are the standard unit vector $\mathbf{e}_i$. Hence the $RREF(A)$ is I_n. □

Result 4.7. *Let A be a square matrix. If B is a square matrix such that either $AB = I$ or $BA = I$, then A is invertible and $B = A^{-1}$.*

Proof. Assume that $BA = I$ and consider the equation $A\mathbf{x} = \mathbf{0}$. Multiplying both sides by B, we have $BA\mathbf{x} = B\mathbf{0} = \mathbf{0} \Rightarrow I\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = 0$. Thus the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, and hence by Fundamental Theorem of Invertible Matrices, the matrix A is invertible, that is, A^{-1} exists and satisfy $AA^{-1} = I = A^{-1}A$. Now $BA = I \Rightarrow (BA)A^{-1} = IA^{-1} \Rightarrow B = A^{-1}$.

Now let us assume that $AB = I$. As in the previous case, we find that B is invertible and $A = B^{-1}$. However, this implies that A is also invertible and that $A^{-1} = (B^{-1})^{-1} = B$. □

Result 4.8. *Let A be a square matrix. If a finite sequence of elementary row operations transforms A to the identity matrix I, then the same sequence of elementary row operations transforms I into A^{-1}.*

Gauss-Jordan Method for Computing Inverse:

Let A be an $n \times n$ matrix.

- Apply elementary row operations on the augmented matrix $[A \mid I_n]$.
- If A is invertible, then $[A \mid I_n]$ will be transformed to $[I_n \mid A^{-1}]$.
- If A is not invertible, then $[A \mid I_n]$ can never be transformed to a matrix of the type $[I_n \mid B]$.

Example 4.1. *Find the inverse of the matrix* $\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, *if it exists.*

Solution. We have

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 1 & 0 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -2 & -1 \end{array}\right]$$

$$\xrightarrow{R_3 \to (-1)R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 2 & 1 \end{array}\right].$$

Hence the required inverse is $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$. □

Practice Problems Set 4

1. Using mathematical induction, prove that if $A_1, A_2, \ldots, A_n$ are invertible matrices of the same size then the product $A_1A_2\ldots A_n$ is also invertible and that $(A_1A_2\ldots A_n)^{-1} = A_n^{-1}A_{n-1}^{-1}\ldots A_1^{-1}$ for all $n \geq 1$.

2. Give a counterexample to show that $(AB)^{-1} \neq A^{-1}B^{-1}$ in general, where A and B are two invertible matrices of the same size. Find a necessary and sufficient condition such that $(AB)^{-1} = A^{-1}B^{-1}$.

3. Give a counterexample to show that $(A+B)^{-1} \neq A^{-1} + B^{-1}$ in general, where A and B are two matrices of the same size such that each of A, B and $A+B$ are invertible.

4. Show that if A is a square matrix that satisfy the equation $A^2 - 2A + I = \mathbf{O}$, then A is invertible and $A^{-1} = 2I - A$.

5. Solve each of the following matrix equations for X:
$$XA^2 = A^{-1},\ (A^{-1}X)^{-1} = A(B^{-2}A)^{-1} \text{ and } ABXA^{-1}B^{-1} = I + A.$$

6. Let A be an invertible matrix. Show that no row or column of A can be entirely zero.

7. Find the inverse of the following matrices, whenever they exist, preferably using Gauss-Jordan method:
$$\begin{bmatrix} 2+i & 6 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} 3 & -1 & 2 \\ -6 & 3 & 1 \\ -7 & -2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ a^2 & a & 1 & 0 \\ a^3 & a^2 & a & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & -6 & -2 \end{bmatrix},$$
where x, y, z are distinct real numbers.

8. Find the inverse of the following matrices, where a, b, c, d are all non-zero real numbers:
$$\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & 0 \\ 0 & c & 0 & 0 \\ d & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & a & 0 & 0 \\ b & 0 & 0 & 0 \\ 0 & 0 & 0 & c \\ 0 & 0 & d & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \end{bmatrix}.$$

9. Find a 3×3 real matrix A such that $A\mathbf{u}_1 = \mathbf{u}_1, A\mathbf{u}_2 = 2\mathbf{u}_2$ and $A\mathbf{u}_3 = 3\mathbf{u}_3$, where
$$\mathbf{u}_1 = [1, 2, 2]^t,\ \mathbf{u}_2 = [2, -2, 1]^t \text{ and } \mathbf{u}_3 = [-2, -1, 2]^t.$$

10. Let A, B, C, D be $n \times n$ matrices such that $ABCD = I$. Show that
$$ABCD = DABC = CDAB = BCDA = I.$$

11. Let A and B be two $n \times n$ matrices. Show that if $AB = A \pm B$ then $AB = BA$.

12. (a) Let A be the 3×3 matrix all of whose main diagonal entries are 0, and elsewhere 1, *i.e.*, $a_{ii} = 0$ for $1 \leq i \leq 3$ and $a_{ij} = 1$ for $i \neq j$. Show that A is invertible and find A^{-1}.
    (b) Let A be the $n \times n$ matrix all of whose main diagonal entries are 0, and elsewhere 1, *i.e.*, $a_{ii} = 0$ for $1 \leq i \leq n$ and $a_{ij} = 1$ for $i \neq j$. Show that A is invertible and find A^{-1}.

13. Let $A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$.
    (a) Find elementary matrices E_1 and E_2 such that $E_2E_1A = I_2$.
    (b) Write A and A^{-1} as a product of elementary matrices.

14. Let A and B be two $n \times n$ matrices and let B be invertible. If $\mathbf{b} \in \mathbb{R}^n$ then show that the system of equations $A\mathbf{x} = \mathbf{b}$ and $BA\mathbf{x} = B\mathbf{b}$ are equivalent.

15. Let $A = [a_{ij}]$ be a 5×5 invertible matrix such that $\sum\limits_{j=1}^{5} a_{ij} = 1$ for $i = 1, 2, 3, 4, 5$. Show that the sum of all the entries of A^{-1} is 5.

16. Prove that if a matrix A is row equivalent to B, then there exists a invertible matrix P such that $B = PA$. Further, show that

    (a) if A is $n \times n$ and invertible then P is unique; and

    (b) if A is $n \times n$ and non-invertible then P need not be unique.

17. Find invertible matrices P and Q such that $B = PAQ$, where

$$A = \begin{bmatrix} 2 & 4 & 8 \\ 1 & 3 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

18. Show that the inverse of an invertible Hermitian matrix (*i.e.*, $A = A^*$) is Hermitian. Also, show that the product of two Hermitian matrices is Hermitian if and only if they commute.

19. Find the inverse of the matrix

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2 & 3 & \dots & n \end{bmatrix}.$$

20. Let A be an $m \times n$ matrix. Show that, by means of a finite number of elementary row and/or column operations, one can transform A to a matrix $R = [r_{ij}]$ which is both 'reduced row echelon form' and 'reduced column echelon form', i.e., $r_{ij} = 0$ if $i \neq j$, and there is a $k \in \{1, 2, \dots, n\}$ such that $r_{ii} = 1$ if $1 \leq i \leq k$ and $r_{ii} = 0$ if $i > k$. Show that $R = PAQ$, where P and Q are invertible matrices of sizes $m \times m$ and $n \times n$, respectively.

Hints to Practice Problems Set 4

1. Use induction on n.

2. Take two invertible matrices which do not commute.

3. Take $A = B = I_2$.

4. The given condition gives $I = (2I - A)A$.

5. The solutions are (in order): $X = A^{-3}$, $X = AB^{-2}$ and $X = B^{-1}A^{-1}BA + A$.

6. Show that $AB = I$ is never possible if A has a zero row.

7. For each matrix A if you get the answer as B, then to be sure verify that $AB = I$.

8. For each matrix A if you get the answer as B, then to be sure verify that $AB = I$.

9. $A[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = [\mathbf{u}_1, 2\mathbf{u}_2, 3\mathbf{u}_3]$.

10. For square matrices, $XY = I \Leftrightarrow YX = I$.

11. $AB = A - B \Rightarrow (A + I)(I - B) = I \Rightarrow (A + I)(I - B) = I = (I - B)(A + I) \Rightarrow AB = BA$.

12. Here $A = J - I$, where each entry of J is 1. We have $(J - I)\left(\frac{1}{n-1}J - I\right) = I$.

13. $E_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$ and $A^{-1} = E_2E_1$, $\quad A = E_1^{-1}E_2^{-1}$.

14. Use the fact that B is a product of elementary matrices, **OR** directly show that both the systems have the same solution sets.

15. If $B = [b_{ij}]$ is the inverse of A, then $\sum\limits_{j=1}^{5}\sum\limits_{i=1}^{5}\left(\sum\limits_{k=1}^{5} b_{ik}a_{kj}\right) = 5 \Rightarrow \sum\limits_{k=1}^{5}\left[\sum\limits_{i=1}^{5} b_{ik}\left(\sum\limits_{j=1}^{5} a_{kj}\right)\right] = 5.$

    Or note that $A[1, 1, \dots, 1]^t = [1, 1, \dots, 1]^t$ which gives $[1, 1, \dots, 1]^t = A^{-1}[1, 1, \dots, 1]^t$.

16. Each elementary row operation corresponds to an elementary matrix. (a) $PA = B = QA \Rightarrow P = Q$.
(b) Take $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Then $P = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $Q = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

17. Find P and Q from the multiplication of elementary matrices corresponding to the respective elementary row and column operations. $P = \begin{bmatrix} 3/2 & -2 \\ -1/2 & 1 \end{bmatrix}$ and $Q = \begin{bmatrix} 1 & 0 & -8 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ are two such matrices. There could be other such matrices also.

18. Easy.

19. $\mathbf{x}_i$ is the solution of $A\mathbf{x} = \mathbf{e}_i$, where $\mathbf{x}_1 = [2, -1, 0, \ldots, 0]^t, \mathbf{x}_n = [0, \ldots, 0, -1, 1]^t$ and $\mathbf{x}_i = [0, \ldots, 0, -1, 2, -1, 0, \ldots, 0]^t$ for $1 < i < n$. Then $A^{-1} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$.

20. Apply sufficient numbers of elementary row and column operations on A.

5 Determinant

Definition 5.1. *Let $A = [a_{ij}]$ be an $n \times n$ matrix.*

- *For a 1×1 matrix $A = [a]$, we define the **determinant** of A as $det(A) = a$.*
- *If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then we define $det(A) = ad - bc$.*
- *In general, if A_{ij} is the submatrix of A obtained by deleting the i-th row and the j-th column of A, then $det(A)$ is defined recursively as follows:*

$$det(A) = a_{11}\,det(A_{11}) - a_{12}\,det(A_{12}) + \ldots + (-1)^{1+n}a_{1n}\,det(A_{1n}) = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\,det(A_{1j}).$$

- *Sometimes $det(A)$ is also denoted by $|A|$.*
- *We define $det(A_{ij})$ to be the (i, j)**-minor** of A.*
- *The number $C_{ij} = (-1)^{i+j}det(A_{ij})$ is called the (i, j)**-cofactor** of A.*
- *Thus we can write*

$$det(A) = |A| = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n} = \sum_{j=1}^{n} a_{1j}C_{1j}.$$

Result 5.1 (Properties of Determinants). *Let $A = [a_{ij}]$ be an $n \times n$ matrix. Then*

$$a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n} = det(A) = a_{11}C_{11} + a_{21}C_{21} + \ldots + a_{n1}C_{n1}.$$

Result 5.2 (Properties of Determinants). *Let A be an $n \times n$ matrix. If B is obtained by interchanging any two rows of A, then $det(B) = -det(A)$.*

Result 5.3 (Laplace Expansion Theorem). *The determinant of an $n \times n$ matrix $A = [a_{ij}]$, where $n \geq 2$, can be computed as*

$$det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in} = \sum_{j=1}^{n} a_{ij}C_{ij},$$

*(this is the **cofactor expansion along the i-th row**), and also as*

$$det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj} = \sum_{i=1}^{n} a_{ij}C_{ij},$$

*(this is the **cofactor expansion along the j-th column**).*

Result 5.4.

1. *For any square matrix A, $det(A^t) = det(A)$.*
2. *The determinant of a triangular matrix is the product of the diagonal entries. That is, if $A = [a_{ij}]$ is an $n \times n$ triangular matrix then $det(A) = a_{11}a_{22}\ldots a_{nn}$.*

Proof.

1. The rows of A are the columns of A^t. So, the proof follows from **Laplace Expansion Theorem**.
2. Let A be a lower triangular matrix of size n. We use induction on the size n. For $n = 1$, it is clear that $\det(A) = a_{11}$. Let us assume that for any $k \times k$ lower triangular matrix, the determinant is equal to the product of the diagonal entries. Now let $A = [a_{ij}]$ be a $(k + 1) \times (k + 1)$ lower triangular matrix. Then we have

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n} = a_{11}C_{11} + 0C_{12} + \ldots + 0C_{1n} = a_{11}C_{11}.$$

Since C_{11} is the determinant of a lower triangular matrix of size k with diagonal entries $a_{22}, \ldots, a_{k+1,k+1}$, by induction hypothesis, we have $C_{11} = a_{22}.\ldots.a_{k+1,k+1}$. Therefore $\det(A) = a_{11}a_{22}.\ldots.a_{k+1,k+1}$. Hence by principle of mathematical induction, we conclude that for any lower triangular matrix, the determinant is equal to the product of the diagonal entries.

Now the proof for upper triangular matrix follows from the fact that A^t is lower triangular, whenever A is upper triangular and that $\det(A) = \det(A^t)$. [Induction can also be used separately through column-wise expansion.] ☐

Result 5.5. ***[Properties of Determinants]***

1. *If A has a zero row then $det(A) = 0$.*

2. *If A has two identical rows then $det(A) = 0$.*

3. *If B is obtained by multiplying a row of A by k, then $det(B) = k.det(A)$.*

4. *If the matrices A, B and C are identical except that the i-th row of C is the sum of the i-th rows of A and B, then $det(C) = det(A) + det(B)$.*

5. *If C is obtained by adding a multiple of one row of A to another row, then $det(C) = det(A)$.*

6. *If $A = [a_{ij}]$, where $n \geq 2$, then $a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn} = 0 \quad$ for $i \neq j$.*

Proof. [Proof of Part 1] Expanding through the zero row (by Laplace Expansion Theorem), we find $\det(A) = 0$.

[Proof of Part 5] Let C be obtained by adding k times of the j-th row of A to the i-th row of A. Let B be the matrix whose i-th row is equal to k times of the j-th row of A, and all other rows of B are identical with that of A. Application of **Part 2** and **Part 3** give that $\det(B) = 0$.

Now we see that the matrices A, B and C are identical except that the i-th row of C is the sum of the i-th rows of A and B. Therefore $\det(C) = \det(A) + \det(B) = \det(A) + 0 = \det(A)$.

[Proof of Part 6] Follows from Part 2. □

Result 5.6 (Determinants of Elementary Matrices). *Let E be an $n \times n$ elementary matrix and A be any $n \times n$ matrix. Then*

1. *$det(E) = -1, k$ or 1.*

2. *$det(EA) = det(E)det(A)$.*

Proof.

1. Since $\det(I_n) = 1$, **Result 5.2** and **Result 5.5** (Part 3 and Part 5) give the desired result.

2. Let E_1 be the elementary matrix corresponding to the elementary row operation $R_i \longleftrightarrow R_j$. Then $\det(E_1) = -1$ and E_1A is the matrix obtained by interchanging the i-th and j-th rows of A. Therefore $\det(E_1A) = -\det(A)$ and also $\det(E_1)\det(A) = -\det(A)$. Thus $\det(E_1A) = \det(E_1)\det(A)$.

   Let E_2 be the elementary matrix corresponding to the elementary row operation $R_i \longrightarrow kR_i, k \neq 0$. Then $\det(E_2) = k$ and E_2A is the matrix obtained by multiplying the i-th row of A by k. Therefore $\det(E_2A) = k\det(A)$ and also $\det(E_2)\det(A) = k\det(A)$. Thus $\det(E_2A) = \det(E_2)\det(A)$.

   Let E_3 be the elementary matrix corresponding to the elementary row operation $R_i \longleftrightarrow R_i + kR_j$. Then $\det(E_3) = 1$ and E_3A is the matrix obtained by adding k-times of the j-th row of A to the i-th row of A. Therefore $\det(E_3A) = \det(A)$ and also $\det(E_3)\det(A) = 1.\det(A)$. Thus $\det(E_3A) = \det(E_3)\det(A)$. □

Result 5.7.

1. *A square matrix A is invertible if and only if $det(A) \neq 0$.*

2. *Let A be an $n \times n$ matrix. Then $det(kA) = k^n det(A)$.*

3. *Let A and B be two $n \times n$ matrices. Then $det(AB) = det(A)det(B)$.*

4. *If the matrix A is invertible then $det(A^{-1}) = \frac{1}{det(A)}$.*

Proof. [Proof of Part 2] If $k = 0$, then clearly $\det(kA) = 0 = k^n\det(A)$. For $k \neq 0$, let E_i be the elementary matrix corresponding to the elementary row operation $R_i \longrightarrow kR_i$ for $i = 1, 2, \ldots, n$. Then we see that $kA = E_1E_2 \ldots E_nA$. Therefore

$$\det(kA) = \det(E_1E_2 \ldots E_nA) = \det(E_1)\det(E_2) \ldots \det(E_n)\det(A) = k^n\det(A).$$ □

★ A matrix A is said to be **singular** or **non-singular** according as $\det(A) = 0$ or $\det(A) \neq 0$.

★ As a determinant can be expanded column-wise, all the previous results based on row-wise expansion of determinant are also valid for column-wise expansion.

Result 5.8.

1. *If B is obtained by interchanging any two columns of A, then $det(B) = -det(A)$.*
2. *If A has a zero column then $det(A) = 0$.*
3. *If A has two identical columns then $det(A) = 0$.*
4. *If B is obtained by multiplying a column of A by k, then $det(B) = k det(A)$.*
5. *If A, B and C are identical except that the i-th column of C is the sum of the i-th columns of A and B, then $det(C) = det(A) + det(B)$.*
6. *If C is obtained by adding a multiple of one column of A to another column, then $det(C) = det(A)$.*
7. *If $A = [a_{ij}]$, where $n \geq 2$, then $a_{1i}C_{1j} + a_{2i}C_{2j} + \ldots + a_{ni}C_{nj} = 0 \quad for\ i \neq j$.*

Proof. Follows from **Result 5.2**, **Result 5.5** and $\det(A^t) = \det(A)$. ☐

Result 5.9. *Let the matrices* $\begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix}$, *$P$ and Q be square matrices. Then* $det \begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix} = det(P)det(Q)$.

Proof. We use induction on the size of P. Let $A = \begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix} = [a_{ij}]$. Let C_{ij} be the co-factor of a_{ij} in A. If $P = [a_{11}]$ is 1×1, then $C_{11} = \det(Q)$. Expanding through the first column, we have

$$\det \begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix} = a_{11}C_{11} + 0.C_{21} + \ldots + 0.C_{n1} = a_{11}C_{11} = \det(P)\det(Q).$$

Let us now assume that the result is true whenever the size of P is at most $k-1$. Now let the size of P be k and consider

$$A = \det \begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix} = \left[\begin{array}{cccc|ccc} a_{11} & a_{12} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2k} & a_{2,k+1} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ \hline 0 & 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{nn} \end{array}\right].$$

Then we have

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + \ldots + a_{k1}C_{k1}.$$

For $i = 1, \ldots, k$, let C'_{i1} denote the co-factor of a_{i1} in P. Notice that, for each $i = 1, 2, \ldots, k$, the sub-matrix A_{i1} of A obtained by deleting the i-th row and the first column is also a block matrix of the form $\begin{bmatrix} P_i & R_i \\ \mathbf{O} & Q \end{bmatrix}$, where the size of P_i is $k-1$. Further, P_i is the sub-matrix of P obtained by deleting the i-th row and the first column from P. Therefore by induction hypothesis, we have $C_{i1} = (-1)^{i+1}\det(A_{i1}) = (-1)^{i+1}\det \begin{bmatrix} P_i & R_i \\ \mathbf{O} & Q \end{bmatrix} = (-1)^{i+1}\det(P_i)\det(Q) = C'_{i1}.\det(Q)$. Therefore

$$\begin{aligned} \det(A) =& a_{11}C_{11} + a_{21}C_{21} + \ldots + a_{k1}C_{k1} \\ =& a_{11}C'_{11}.\det(Q) + a_{21}C'_{21}.\det(Q) + \ldots + a_{k1}C'_{k1}.\det(Q) \\ =& (a_{11}C'_{11} + a_{21}C'_{21} + \ldots + a_{k1}C'_{k1})\det(Q) \\ =& \det(P)\det(Q). \end{aligned}$$

Thus we see that the result is also true whenever the size of P is k. Hence by the principle of mathematical induction, we conclude that $\det \begin{bmatrix} P & R \\ \mathbf{O} & Q \end{bmatrix} = \det(P)\det(Q)$. ☐

Definition 5.2. *Let A be an $n \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^n$. Then $A_i(\mathbf{b})$ denotes the matrix obtained by replacing the i-th column of A by $\mathbf{b}$. That is, if $A = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_n]$, then $A_i(\mathbf{b}) = [\mathbf{a}_1\ \mathbf{a}_2\ \dots \mathbf{a}_{i-1}\ \mathbf{b}\ \mathbf{a}_{i+1}\ \dots\ \mathbf{a}_n]$.*

Result 5.10 (Cramer's Rule). *Let A be an $n \times n$ invertible matrix and let $\mathbf{b} \in \mathbb{R}^n$. Then the unique solution $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ of the system $A\mathbf{x} = \mathbf{b}$ is given by*

$$x_i = \frac{det\ (A_i(\mathbf{b}))}{det(A)} \quad for\ \ i = 1, 2, \dots, n.$$

The Adjoint of a Matrix: Let $A = [a_{ij}]$ be an $n \times n$ matrix and let C_{ij} be the (i, j)-cofactor of A. Then the **adjoint** of A, denoted $\mathrm{adj}(A)$, is defined as

$$\mathrm{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} = [C_{ij}]^t.$$

Result 5.11. *Let A be an $n \times n$ invertible matrix. Then $A^{-1} = \frac{1}{det(A)} . adj(A)$.*

Remark: The proof of Result 5.11 can be obtained using Part 6 of Result 5.5 or as given in the text book.

Example 5.1. *Use the adjoint method to find the inverse of the matrix*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}.$$

Practice Problems Set 5

1. Find the inverse of the following matrices using adjoint method:
$$\begin{bmatrix} 1 & 1 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},$$
where x, y, z are distinct real numbers.

2. If A is an idempotent matrix ($i.e.,$ $A^2 = A$), then find all possible values of $\det(A)$.

3. Let $A = [a_{ij}]$ be an $n \times n$ matrix such that $a_{ij} = \max\{i, j\}$. Find $\det(A)$.

4. Let $A = [a_{ij}]$ be an $n \times n$ matrix such that $a_{ij} = j^{i-1}$. Show that
$$\det(A) = (n-1)(n-2)^2(n-3)^3 \dots 3^{n-3}2^{n-2}.$$

5. Let $A = [\mathbf{a}, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4]$ and $B = [\mathbf{b}, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4]$ be two 4×4 matrices, where $\mathbf{a}, \mathbf{b}, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4 \in \mathbb{R}^4$. If $\det(A) = 4$ and $\det(B) = 1$, find $\det(A + B)$. What is $\det(C)$, where $C = [\mathbf{r}_4, \mathbf{r}_3, \mathbf{r}_2, \mathbf{a} + \mathbf{b}]$?

6. Let A be an $n \times n$ matrix.
   (a) Show that if $A^2 + I = \mathbf{O}$ then n must be an even integer.
   (b) Does (a) remain true for complex matrices?

7. Let A be an $n \times n$ matrix. If $AA^t = I$ and $\det(A) < 0$, then find $\det(A + I)$.

8. If A is a matrix satisfying $A^3 = 2I$ then show that the matrix B is invertible, where $B = A^2 - 2A + 2I$.

9. Show that if $a \neq b$ then $\det(A) = \frac{a^{n+1} - b^{n+1}}{a-b}$, where A is an $n \times n$ matrix given as follows:
$$A = \begin{bmatrix} a+b & ab & 0 & \dots & 0 & 0 \\ 1 & a+b & ab & \dots & 0 & 0 \\ 0 & 1 & a+b & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a+b & ab \\ 0 & 0 & 0 & \dots & 1 & a+b \end{bmatrix}.$$
If $a = b$ then what happens to the value of $\det(A)$?

10. The position of the (i, j)-th entry a_{ij} of an $n \times n$ matrix $A = [a_{ij}]$ is called even or odd according as $i + j$ is even or odd.

   (a) Let B be the matrix obtained from multiplying all the entries of A in odd positions by -1. That is, if $B = [b_{ij}]$ then $b_{ij} = a_{ij}$ or $b_{ij} = -a_{ij}$ according as $i + j$ is even or odd. Show that $\det(B) = \det(A)$.

   (b) Let $C = [c_{ij}]$ be the matrix such that $c_{ij} = -a_{ij}$ or $c_{ij} = a_{ij}$ according as $i + j$ is even or odd. Show that $\det(C) = \det(A)$ or $\det(C) = -\det(A)$ according as n is even or odd.

11. Let $\lambda \neq 0$. Determine $|\lambda I - A|$ for the 10×10 matrix

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 10^{10} & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

12. Show that an upper triangular (square) matrix is invertible if and only if every entry on its main diagonal is non-zero.

13. Let A and B be two non-singular matrices of the same size. Are $A + B, A - B$ and $-A$ non-singular? Justify.

14. Each of the numbers $1375, 1287, 4191$ and 5731 is divisible by 11. Show, without calculating the actual value, that the determinant of the matrix $\begin{bmatrix} 1 & 1 & 4 & 5 \\ 3 & 2 & 1 & 7 \\ 7 & 8 & 9 & 3 \\ 5 & 7 & 1 & 1 \end{bmatrix}$ is also divisible by 11.

15. Let $A = [a_{ij}]$ be an $n \times n$ matrix, where $a_{ij} = \frac{1}{i+j}$ for all i, j. Show that A is invertible.

16. Let A be an $n \times n$ matrix. Show that there exists an $n \times n$ non-zero matrix B such that $AB = \mathbf{O}$ if and only if $\det(A) = 0$.

17. Let a_{ij} be integers, where $1 \leq i, j \leq n$. If for any set of integers $b_1, b_2, \dots, b_n$, the system of linear equations

$$\sum_{j=1}^{n} a_{ij} x_j = b_j \quad \text{for} \quad i = 1, 2, \dots, n,$$

   has integer solution $[x_1, x_2, ..., x_n]^t$ then show that $\det(A) = \pm 1$, where $A = [a_{ij}]$.

18. Show that $\det(\overline{A}) = \det(A^*) = \overline{\det(A)}$. Hence show that if A is Hermitian (*i.e.*, $A^* = A$) then $\det(A)$ is a real number.

19. A matrix A is said to be orthogonal if $AA^t = I = A^t A$. Show that if A is orthogonal then $\det(A) = \pm 1$.

20. Let $A = [a_{ij}]$ be an $n \times n$ complex matrix. Show that

   (a) if A is skew-Hermitian (*i.e.*, $\overline{A}^t = -A$) and n is even then $\det(A)$ is real;

   (b) if A is skew-symmetric (*i.e.*, $A^t = -A$) and n is odd then $\det(A) = 0$;

   (c) if A is invertible and $A^{-1} = A^t$ then $\det(A) = \pm 1$; and

   (d) if A is invertible and $A^{-1} = \overline{A}^t$ then $|\det(A)| = 1$.

21. Suppose A is a 2×1 matrix and B is an 1×2 matrix. Prove that the matrix AB is not invertible. When is the matrix BA invertible?

22. Prove or disprove: The matrix A, as given below, is invertible and all entries of A^{-1} are integers.

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{bmatrix}.$$

23. Find the determinant and the inverse of the $n \times n$ matrix

$$\begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}.$$

24. Assuming that all matrix inverses involved below exist, show that

$$(A - B)^{-1} = A^{-1} + A^{-1}(B^{-1} - A^{-1})^{-1}A^{-1}.$$

In particular, show that

$$(I + A)^{-1} = I - (A^{-1} + I)^{-1} \text{ and } |(I + A)^{-1} + (I + A^{-1})^{-1}| = 1.$$

25. Let S be the backward identity matrix; that is,

$$S = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Show that $S^{-1} = S^t = S$. Find $\det(S)$ and SAS for $n \times n$ matrix $A = [a_{ij}]$.

Hints to Practice Problems Set 5

1. For each matrix A if you get the answer as B, then to be sure verify that $AB = I$.
2. $A^2 = A \Rightarrow (\det A)^2 = \det A$.
3. Apply the row operations $R_n \to R_n - R_{n-1}, R_{n-1} \to R_{n-1} - R_{n-2}, \dots, R_2 \to R_2 - R_1$ and then expand through the last column. We get $\det A = (-1)^{n+1}n$.
4. Consider the Vandermonde matrix, where $x_i = i$ for $i = 1, 2, \dots, n$.
5. Use **Result 5.1** repeatedly to find $\det(A + B) = 40$ an $\det C = 5$.
6. $A^2 = -I \Rightarrow (\det A)^2 = (-1)^n$. For the 2nd part, using the fact $i^2 = -1$ try to find a counterexample for $n = 3$.
7. $\det A = -1$ and $\det(A + I) = \det(A + AA^t)$. We find $\det(A + I) = 0$.
8. $B = A(A+2I)(A-I)$. Also $A^3 = 2I \Rightarrow I = (A-I)(A^2+A+I)$ and $A^3+8I = 10I \Rightarrow (A+2I)(A^2-2A+4I) = 10I$. Thus $\det(A - I) \neq 0$ and $\det(A + 2I) \neq 0$, and hence $\det B \neq 0$.
9. Use induction on n. Induction hypothesis: Assume the result to be true for matrices of size n, where $n \leq k$. Now prove for a matrix of size $k + 1$. [**This is another version of method of induction**]

   [Note that we need to prove $\det(A) = \frac{a^{n+1}-b^{n+1}}{a-b}$.] Take limit of $\frac{a^{n+1}-b^{n+1}}{a-b}$ as a tends to b, for the second part.
10. (a). The numbers $i+j$ and $i-j$ are either both even or both odd. Thus, multiplying the entries at odd positions of A by -1 is same as multiplying each entry of A by $(-1)^{i-j}$. Now use **Problem** 32 of Tutorial Sheet.

    (b). Multiplying the entries at even positions of A by -1 is same as multiplying each entry of B by -1.
11. Expand through the last row. The answer is $\lambda^{10} - 10^{10}$.
12. What is the determinant of a triangular matrix?
13. $A + B$ and $A - B$ need not be non-singular (find counterexapmles.) Also, $\det(-A) = (-1)^n \det(A)$.
14. Apply suitable elementary row operations.

15. To show that $\det(A) \neq 0$. Subtract the last row of A from the $n-1$ preceding rows. Now take the factors $\frac{1}{n+1}, \frac{1}{n+2}, \ldots, \frac{1}{2n-1}, \frac{1}{2n}$ common from the respective columns and also take the factors $n-1, n-2, \ldots, 2$ common from the respective rows. Now in the remaining determinant, subtract the last column from each of the preceding columns and take suitable factors common from the rows as well as from the columns. Finally, expand the remaining determinant through a suitable row (column) to get a similar determinant of size $n-1$. Now use mathematical induction on n.

16. $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution iff $\det(A) = 0$.

17. For each i, there is a column vector $\mathbf{c}_i$, each of whose entries are integers, such that $A\mathbf{c}_i = \mathbf{e}_i$. Set $C = [\mathbf{c}_1, \ldots, \mathbf{c}_n]$ so that $AC = I$.

18. Easy.

19. $\det(AA^t) = 1$.

20. (a) $\det(\overline{A}^t) = \det(-A) \Rightarrow \overline{\det(A)} = (-1)^n \det(A)$. (b) Similar to first part. (c) $AA^t = I$. (d) $AA^* = I$.

21. $\det(AB) = 0$.

22. The given statement is correct. The proof is similar to **Problem** 15.

23. Apply the row operations $R_1 \leftrightarrow R_n,\ R_2 \to R_2 - R_1,\ R_3 \to R_3 - R_1, \ldots, R_n \to R_n - R_{n-1}$ and then use induction to find $\det(A) = (-1)^{n-1}(n-1)$. Find A^{-1} as in Problem 12 of **Section 4 OR** solve $A\mathbf{x} = \mathbf{e}_i$ for each $i = 1, 2, \ldots, n$.

24. Check that $(A-B)\left[A^{-1} + A^{-1}(B^{-1} - A^{-1})^{-1}A^{-1}\right] = I$.

25. Direct computation gives $S^2 = I$. Use induction to show that $\det(S) = (-1)^{\frac{n^2+3n}{2}}$. The (i,j)-th entry of SAS is $a_{n-i+1,n-j+1}$. (Notice that if $S = [u_{ij}]$ then $u_{ij} = 1$ iff $i + j = n + 1$.)

6 Subspaces Associated with Matrices

Definition 6.1. *Let A be an $m \times n$ matrix.*

1. *The* ***null space*** *of A, denoted* **null**(A)*, is the subspace of $\mathbb{C}^n$ consisting of the solutions of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$. In other words, null$(A) = \{\mathbf{x} \in \mathbb{C}^n \mid A\mathbf{x} = \mathbf{0}\}$.*

2. *The* ***column space*** *of A, denoted* **col**(A)*, is the subspace of $\mathbb{C}^m$ spanned by the columns of A. In other words, col$(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n\}$.*

3. *The* ***row space*** *of A, denoted* **row**(A)*, is the subspace of $\mathbb{C}^n$ spanned by the rows of A.*
   In other words, row$(A) = \{\mathbf{x}^T A \mid \mathbf{x} \in \mathbb{C}^m\}$.

[Here, elements of row(A) are row vectors. How can they be elements of $\mathbb{C}^n$? In strict sense, row$(A) :=$ col(A^T).]

Result 6.1. *Let the matrices B and A be row equivalent. Then row$(B) =$ row(A).*

Corollary 6.1. *For any A, row$(A) =$ row$(RREF(A))$.*

Proof. Since A and $RREF(A)$ are row equivalent, we find that $\text{row}(A) = \text{row}(\text{RREF}(A))$. □

Corollary 6.2. *For any A, the non-zero rows of RREF(A) form a basis of row(A).*

Proof. Since $\text{row}(A) = \text{row}(\text{RREF}(A))$, the non-zero rows of RREF(A) span row(A). Let $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ be the set of non-zero rows of RREF(A). We claim that $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ is linearly independent to prove that the non-zero rows of RREF(A) form a basis of row(A).

Let $j_1, j_2, \ldots, j_r$ be the position of the first non-zero entry of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, respectively. Since $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ are the non-zero rows of an RREF of a matrix, we find that $j_1 < j_2 < \ldots < j_r$. Therefore for $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$, the j_k-th entry of $\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_r\mathbf{a}_r$ is α_k for $k = 1, 2, \ldots, r$. So

$$\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_r\mathbf{a}_r = \mathbf{0} \Rightarrow \alpha_1 = 0 = \alpha_2 = \ldots = \alpha_r.$$

Hence $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ is linearly independent. □

Suppose A and B are row-equivalent. Are col(A) and col(B) equal? No. Take $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

Suppose A and B are row-equivalent. Do col(A) and col(B) have same dimension? Yes. We will see soon. Indeed, rank$(A) = \dim(\text{row(A)}) = \dim(\text{col(A)})$.

Method of Finding Bases of the Row Space, the Null Space and the Column Space of a Matrix:

Let A be a given matrix and let R be the reduced row echelon form of A.

1. Use the non-zero rows of R to form a basis for row(A).

2. Solve the leading variables of $R\mathbf{x} = \mathbf{0}$ in terms of the free variables, set the free variables equal to parameters, substitute back into $\mathbf{x}$, write the result as a linear combination of k vectors (where k is the number of free variables). These k vectors form a basis for null(A).

3. A basis for row(A^t) will also be a basis for col(A). **Or**, Use the columns of A that correspond to the columns of R containing the leading 1's to form a basis for col(A).

Example 6.1. *Find bases for the row space, column space and the null space of the following matrix:*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 4 & 6 & 2 \end{bmatrix}.$$

Solution. We have

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 4 & 6 & 2 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 6 \\ 4 & 6 & 2 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 4R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 6 \\ 0 & -2 & 6 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 - R_2} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 \to -\frac{1}{2}R_2} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} = R, \quad \text{say.}$$

A basis for row(A) is $\{[1, 0, 5]^t, [0, 1, -3]^t\}$. A basis for col$(A)$ is $\{[1, 2, 4]^t, [2, 2, 6]^t\}$. Also for $\mathbf{x} = [x, y, z]^t$, we have

$$R\mathbf{x} = \mathbf{0} \Rightarrow x + 5z = 0, y - 3z = 0 \Rightarrow x = -5z, y = 3z.$$

Take $z = s$, so that $x = -5s, y = 3s$ and $\mathbf{x} = [-5s, 3s, s]^t = s[-5, 3, 1]^t$. Hence a basis for null$(A)$ is $\{[-5, 3, 1]^t\}$. ☐

Result 6.2. *Let $R = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n]$ be the reduced row echelon form of a matrix $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n]$ of rank r. Let $\mathbf{b}_{j_1}, \mathbf{b}_{j_2}, \ldots, \mathbf{b}_{j_r}$ be the columns of R such that $\mathbf{b}_{j_k} = \mathbf{e}_k$ for $k = 1, \ldots, r$. Then $\{\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}\}$ is a basis for col(A).*

Proof. Since R is the RREF of A, we have $R = E_1E_2 \ldots E_kA$ for some elementary matrices $E_1, E_2, \ldots, E_k$. Taking $E_1E_2 \ldots E_k = M$, we find that M is an invertible matrix and $R = MA$. Since the i-th column of MA is $M\mathbf{a}_i$, we have $\mathbf{b}_i = M\mathbf{a}_i$ for $i = 1, 2, \ldots, n$. Now for $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}^n$, we have

$$\begin{aligned}
&\alpha_1\mathbf{a}_{j_1} + \alpha_2\mathbf{a}_{j_2} + \ldots + \alpha_r\mathbf{a}_{j_r} = \mathbf{0} \\
\Rightarrow &M(\alpha_1\mathbf{a}_{j_1} + \alpha_2\mathbf{a}_{j_2} + \ldots + \alpha_r\mathbf{a}_{j_r}) = M\mathbf{0} = \mathbf{0} \\
\Rightarrow &\alpha_1(M\mathbf{a}_{j_1}) + \alpha_2(M\mathbf{a}_{j_2}) + \ldots + \alpha_r(M\mathbf{a}_{j_r}) = \mathbf{0} \\
\Rightarrow &\alpha_1\mathbf{b}_{j_1} + \alpha_2\mathbf{b}_{j_2} + \ldots + \alpha_r\mathbf{b}_{j_r} = \mathbf{0} \\
\Rightarrow &\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_r\mathbf{e}_r = \mathbf{0} \\
\Rightarrow &\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0.
\end{aligned}$$

Therefore $\{\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}\}$ is linearly independent. Now let $\mathbf{x} \in$ col(A). Then there exist $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}^n$ such that $\mathbf{x} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n$. Notice that each $\mathbf{b}_j$ is of the form $\mathbf{b}_j = [b_{1j}, b_{2j}, \ldots, b_{rj}, 0, 0, \ldots, 0]^t$. Now we have

$$\begin{aligned}
&\mathbf{x} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n \\
\Rightarrow &M\mathbf{x} = \alpha_1(M\mathbf{a}_1) + \alpha_2(M\mathbf{a}_2) + \ldots + (M\alpha_r\mathbf{a}_n) = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \ldots + \alpha_n\mathbf{b}_n \\
\Rightarrow &M\mathbf{x} = \alpha_1(b_{11}\mathbf{e}_1 + \ldots + b_{r1}\mathbf{e}_r) + \alpha_2(b_{12}\mathbf{e}_1 + \ldots + b_{r2}\mathbf{e}_r) + \ldots + \alpha_n(b_{1n}\mathbf{e}_1 + \ldots + b_{rn}\mathbf{e}_r) \\
\Rightarrow &M\mathbf{x} = (\alpha_1b_{11} + \ldots + \alpha_nb_{1n})\mathbf{e}_1 + \ldots + (\alpha_1b_{r1} + \ldots + \alpha_nb_{rn})\mathbf{e}_r \\
\Rightarrow &M\mathbf{x} = (\alpha_1b_{11} + \ldots + \alpha_nb_{1n})\mathbf{b}_{j_1} + \ldots + (\alpha_1b_{r1} + \ldots + \alpha_nb_{rn})\mathbf{b}_{j_r} \\
\Rightarrow &\mathbf{x} = (\alpha_1b_{11} + \ldots + \alpha_nb_{1n})(M^{-1}\mathbf{b}_{j_1}) + \ldots + (\alpha_1b_{r1} + \ldots + \alpha_nb_{rn})(M^{-1}\mathbf{b}_{j_r}) \\
\Rightarrow &\mathbf{x} = (\alpha_1b_{11} + \ldots + \alpha_nb_{1n})\mathbf{a}_{j_1} + \ldots + (\alpha_1b_{r1} + \ldots + \alpha_nb_{rn})\mathbf{a}_{j_r}.
\end{aligned}$$

Thus $\mathbf{x}$ is a linear combination of $\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}$, and so $\{\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}\}$ spans col(A).

Finally, $\{\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_r}\}$ is linearly independent and spans col(A), and hence is a basis for col(A). ☐

Remark: Note that if $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a linearly independent subset of $\mathbb{R}^n$ and P is an $n \times n$ invertible matrix, then it is immediate from the previous proof that $\{P\mathbf{x}_1, P\mathbf{x}_2, \ldots, P\mathbf{x}_k\}$ is also linearly independent. Also, if $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a basis of a subspace of $\mathbb{R}^n$ and P is an $n \times n$ invertible matrix, then it is immediate from the previous proof that $\{P\mathbf{x}_1, P\mathbf{x}_2, \ldots, P\mathbf{x}_k\}$ is also a basis of the same subspace.

Result 6.3. *The row space and the column space of a matrix A have the same dimension, and dim(row(A)) = dim(col(A)) = rank(A).*

Proof. By **Corollary 6.2**, the non-zero rows of $RREF(A)$ form a basis for row(A). Therefore $\dim(\text{row}(A)) = \text{rank}(A)$. Again, from **Result 6.2**, $\dim(\text{col}(A)) = \text{rank}(A)$. $\square$

Result 6.4. *For any matrix A, we have $rank(A^t) = rank(A)$.*

Nullity: The **nullity** of a matrix A is the dimension of its null space, and is denoted by **nullity**(A).

Note that nullity$(A) =$ the number of free variables of the system $A\mathbf{x} = \mathbf{0}$.

Result 6.5 (**Rank Nullity Theorem**). *Let A be an $m \times n$ matrix. Then*

$$rank(A) + nullity(A) = n.$$

Result 6.6 (**The Fundamental Theorem of Invertible Matrices: Version II**). *Let A be an $n \times n$ matrix. Then the following statements are equivalent.*

1. *A is invertible.*
2. *A^t is invertible.*
3. *$A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^n$.*
4. *$A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$.*
5. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
6. *The reduced row echelon form of A is I_n.*
7. *The rows of A are linearly independent.*
8. *The columns of A are linearly independent.*
9. *$rank(A) = n$.*
10. *A is a product of elementary matrices.*
11. *$nullity(A) = 0$.*
12. *The column vectors of A span $\mathbb{R}^n$.*
13. *The column vectors of A form a basis for $\mathbb{R}^n$.*
14. *The row vectors of A span $\mathbb{R}^n$.*
15. *The row vectors of A form a basis for $\mathbb{R}^n$.*

Proof. The equivalence of the statements (1) to (10) are proved in **Result 4.6**. The remaining equivalence will be proved as follows:

$$(9) \Longleftrightarrow (11), \qquad \begin{array}{ccc} (12) & \Longleftrightarrow & (13) \\ & & \Updownarrow \\ & & (8) \end{array} \quad \text{and} \quad \begin{array}{ccc} (14) & \Longleftrightarrow & (15) \\ & & \Updownarrow \\ & & (7) \end{array}.$$

$(9) \Longleftrightarrow (11)$ From $\text{rank}(A) + \text{nullity}(A) = n$, we find that $\text{rank}(A) = n \Longleftrightarrow \text{nullity}(A) = 0$.

$(12) \Longleftrightarrow (13)$ Suppose the column vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ of A spans $\mathbb{R}^n$. If $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is linearly dependent, then some column $\mathbf{a}_i$ is a linear combination of the remaining columns. Consequently, the $n-1$ columns $\{\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n\}$ also span $\mathbb{R}^n$, contradicting the fact that $\dim(\mathbb{R}^n) = n$. Therefore $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ is linearly independent, which along with the fact $\text{span}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}) = \mathbb{R}^n$ give that $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ is a basis for $\mathbb{R}^n$.

Now $(13) \Longrightarrow (12)$ follows directly from the definition of basis.

$(13) \Longleftrightarrow (8)$ Suppose the column vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ are linearly independent. If $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ does not form a basis for $\mathbb{R}^n$, then $n < \dim(\mathbb{R}^n) = n$, a contradiction. Hence column vectors of A form a basis for $\mathbb{R}^n$.

Now $(13) \Longrightarrow (8)$ follows directly from the definition of basis.

(14) $\iff$ (15) Suppose the row vectors $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ of A spans $\mathbb{R}^n$. If $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ is linearly dependent, then some row $\mathbf{r}_i$ is a linear combination of the remaining rows. Consequently, the $n-1$ rows $\{\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n\}$ also span $\mathbb{R}^n$, contradicting the fact that $\dim(\mathbb{R}^n) = n$. Therefore $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ is linearly independent, which along with the fact $\text{span}(\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}) = \mathbb{R}^n$ give that $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ is a basis for $\mathbb{R}^n$.

Now (15) $\Longrightarrow$ (14) follows directly from the definition of basis.

(15) $\iff$ (7) Suppose the row vectors $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ are linearly independent. If $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ does not form a basis for $\mathbb{R}^n$, then $n < \dim(\mathbb{R}^n) = n$, a contradiction. Hence row vectors of A form a basis for $\mathbb{R}^n$.

Now (15) $\Longrightarrow$ (7) follows directly from the definition of basis. $\square$

Example 6.2. *Show that the vectors* $[1, 2, 3]^t, [-1, 0, 1]^t$ *and* $[4, 9, 7]^t$ *form a basis for* $\mathbb{R}^3$.

Result 6.7. *Let A be an $m \times n$ matrix. Then*

1. *rank*$(A^*A) =$ *rank*(A).
2. *The $n \times n$ matrix A^*A is invertible if and only if rank$(A) = n$.*

Note that $\text{rank}(A^tA) \neq \text{rank}(A)$ in case of complex matrices. Take for example, $A = \begin{bmatrix} 1 & 0 \\ i & 0 \end{bmatrix}$.

Result 6.8. *Let A, B, T, S be matrices, where T are S are invertible.*

1. *If TA and AS are defined, then rank$(TA) =$ rank$(A) =$ rank(AS).*
2. *If AB is defined then rank$(AB) \leq \min\{$rank$(A),$ rank$(B)\}$.*
3. *If $A+B$ is defined then rank$(A+B) \leq$ rank$(A) +$ rank(B).*

Proof. We provide the proof of **Part 1**. See Tutorial Sheet for the proof of the other two parts.

If T is invertible, then $T = E_1E_2 \ldots E_k$ for some elementary matrices $E_1, E_2, \ldots, E_k$. Thus $TA = E_1E_2 \ldots E_kA$. Since the effect of pre-multiplication of a given matrix by an elementary matrix is same as the effect of application of the corresponding elementary row operation, we conclude that TA and A are row equivalent. Therefore $\text{row}(TA) = \text{row}(A)$, and hence $\text{rank}(TA) = \text{rank}(A)$.

Now $\text{rank}(AS) = \text{rank}((AS)^t) = \text{rank}(S^tA^t) = \text{rank}(A^t) = \text{rank}(A)$. $\square$

Result 6.9. *Let A be an $n \times n$ matrix of rank r, where $1 \leq r < n$. Then there exist elementary matrices $E_1, \ldots, E_p$ and $F_1, \ldots, F_q$ such that* $E_1 \ldots E_pAF_1 \ldots F_q = \begin{bmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$.

Proof. See Tutorial Sheet. $\square$

Practice Problems Set 6

1. Determine the reduced row echelon form for each of the following matrices. Hence, find a basis for each of the corresponding row spaces, column spaces and the null spaces.

$$\begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 5 & 6 & 2 \\ -1 & 2 & 4 & 3 \\ 1 & 2 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}, \begin{bmatrix} 3 & 4 & 5 & -6 \\ 2 & 3 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 5 & -5 & 5 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 2 & 2 & 2 & 4 \\ 2 & -2 & 4 & 0 & 8 \\ 4 & 2 & 5 & 6 & 10 \end{bmatrix}.$$

2. Prove that if R is an echelon form of a matrix A, then the non-zero rows of R form a basis of $\text{row}(A)$.

3. Give examples to show that the column space of two row equivalent matrices need not be the same.

4. Find a matrix whose row space contains the vector $[1, 2, 1]^t$ and whose null space contains the vector $[1, -2, 1]^t$, or prove that there is no such matrix.

5. If a matrix A has rank r then prove that A can be written as the sum of r matrices, each of which has rank 1.

6. If the rank of the matrix $\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 0 & t & 0 \\ 0 & -4 & 5 & -2 \end{bmatrix}$ is 2, find all the possible values of t.

7. For what values of t is the rank of the matrix $\begin{bmatrix} t & 1 & 1 & 1 \\ 1 & t & 1 & 1 \\ 1 & 1 & t & 1 \\ 1 & 1 & 1 & t \end{bmatrix}$ equal to 3?

8. Let A and B be two $n \times n$ matrices. Show that if $AB = \mathbf{O}$ then $\text{rank}(A) + \text{rank}(B) \leq n$.

9. Let A be an $n \times n$ matrix such that $A^2 = A^3$ and $\text{rank}(A) = \text{rank}(A^2)$. Show that $A = A^2$. Also, show that the condition $\text{rank}(A) = \text{rank}(A^2)$ cannot be dropped even for a 2×2 matrix.

10. If B is a sub matrix of a matrix A obtained by deleting s rows and t columns from A, then show that $\text{rank}(A) \leq s + t + \text{rank}(B)$.

11. Let A be an $n \times n$ matrix. Show that $A^2 = A$ if and only if $\text{rank}(A) + \text{rank}(A - I) = n$.

12. Find the values of $\lambda \in \mathbb{R}$ for which $\beta = [0, \lambda, \lambda^2]^t$ belongs to the column space of A, where

$$A = \begin{bmatrix} 1+\lambda & 1 & 1 \\ 1 & 1+\lambda & 1 \\ 1 & 1 & 1+\lambda \end{bmatrix}.$$

13. Let A be a square matrix. If $\text{rank}(A) = \text{rank}(A^2)$, show that the linear systems of equations $A\mathbf{x} = \mathbf{0}$ and $A^2\mathbf{x} = \mathbf{0}$ have the same solution space.

14. Let A be a $p \times n$ matrix and B be a $q \times n$ matrix. If $\text{rank}(A) + \text{rank}(B) < n$ then show that there exists an $\mathbf{x}\ (\neq \mathbf{0}) \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$.

15. Let A and B be $n \times n$ matrices such that $\text{nullity}(A) = l$ and $\text{nullity}(B) = m$. Show that $\text{nullity}(AB) \geq \max(l, m)$.

16. Let A be an $m \times n$ matrix of rank r. Show that A can be expressed as $A = BC$, where each of B and C have rank r, B is a matrix of size $m \times r$ and C is a matrix of size $r \times n$.

17. Let A and B be two matrices such that AB is defined and $\text{rank}(A) = \text{rank}(AB)$. Show that $A = ABX$ for some matrix X. Similarly, if BA is defined and $\text{rank}(A) = \text{rank}(BA)$ then show that $A = YBA$ for some matrix Y.

18. Let A be an $m \times n$ matrix with complex entries. Show that the system $A^*A\mathbf{x} = A^*\mathbf{b}$ is consistent for each $\mathbf{b} \in \mathbb{C}^n$.

Hints to Practice Problems Set 6

1. The reduced row echelon forms are given in **Problem** 2 of **Section 2**. Bases for the row space, column space and the null space are:

   First matrix: $\{[1,0,0,0]^t, [0,1,0,0]^t, [0,0,1,0]^t, [0,0,0,1]^t\}$, $\{[1,0,-1,1]^t, [-1,5,2,2]^t, [2,6,4,-1]^t, [3,2,3,2]^t\}$ and $\emptyset$, respectively.

   Second matrix: $\{[1,0,-1,-2]^t, [0,1,2,3]^t\}$, $\{[1,5,9,13]^t, [2,6,10,14]^t\}$ and $\{[1,-2,1,0]^t, [2,-3,0,1]^t\}$, respectively.

   Third matrix: $\{[1,0,0,0]^t, [0,1,0,0]^t, [0,0,1,0]^t, [0,0,0,1]^t\}$, $\{[3,2,0,5]^t, [4,3,2,-5]^t, [5,1,0,5]^t, [-6,1,0,5]^t\}$ and $\emptyset$, respectively.

   Fourth matrix: $\{[1,0,0,1,0]^t, [0,1,0,1,0]^t, [0,0,1,0,0]^t, [0,0,0,0,1]^t\}$, $\{[1,0,2,4]^t, [2,2,-2,2]^t, [1,2,4,5]^t, [2,4,8,10]^t\}$ and $\{[-1,-1,0,1,0]^t\}$, respectively.

2. Show that the non-zero rows of R are linearly independent.

3. One example is given in this lecture note. Find other examples of your own.

4. Not possible since $[1,2,1][1,-2,1]^t \neq 0$.

5. There are invertible matrices P and Q such that $PAQ = \begin{bmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$.

6. $t = 3$.

7. $t = -3$.

8. If $AB = \mathbf{O}$ then $\text{col}(B) \subseteq \text{null}(A)$.

9. $I = (I - A) + A \Rightarrow n \leq \text{rank}(I - A) + \text{rank}(A)$. Also $A^3 = A^2 \Rightarrow \text{rank}(I - A) \leq \text{nullity}(A^2) = n - \text{rank}(A^2)$. Now use **Problem** 11. Take $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ as a counterexample.

10. If A_s is the matrix obtained by deleting s rows of A, then $\text{rank}(A) - s \leq \text{rank}(A_s)$.

11. $n \leq \text{rank}(A - I) + \text{rank}(A)$. Also $A^2 = A \Rightarrow \text{rank}(A - I) \leq \text{nullity}(A) = n - \text{rank}(A)$. Conversely, $n = \text{rank}(A - I) + \text{rank}(A) \Rightarrow \text{rank}(A - I) = \text{nullity}(A)$. Then $\mathbf{z} \in \text{null}(A) \Rightarrow \mathbf{z} = -(A - I)\mathbf{z} \in \text{col}(A - I)$. Thus $\text{null}(A) \subseteq \text{col}(A - I) \Rightarrow \text{null}(A) = \text{col}(A - I)$, and so $A(A - I)\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$. Hence $A^2 = A$.

12. $\lambda \neq -3$ (transform the matrix to row echelon form).

13. $\text{null}(A) \subseteq \text{null}(A^2)$. Also $\text{rank}(A) = \text{rank}(A^2) \Rightarrow \text{nullity}(A) = \text{nullity}(A^2)$.

14. The formula $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$ will be discussed in vector space.

$$\begin{aligned} \dim[\text{null}(A) \cap \text{null}(B)] &= \dim[\text{null}(A)] + \dim[\text{null}(B)] - \dim[\text{null}(A) + \text{null}(B)] \\ &= n - \text{rank}(A) + n - \text{rank}(B) - \dim[\text{null}(A) + \text{null}(B)] > n - \dim[\text{null}(A) + \text{null}(B)] \geq 0. \end{aligned}$$

15. $\text{nullity}(AB) = n - \text{rank}(AB) \geq n - \max\{n - l, n - m\} = \max\{l, m\}$.

16. Let $A = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$ and $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$ a basis for $\text{col}(A)$. If $\mathbf{a}_i = \sum_{j=1}^{r} \alpha_{ij}\mathbf{b}_j$ then take $B = [\mathbf{b}_1, \ldots, \mathbf{b}_r]$ and $C = [\alpha_{ji}]$.

17. $\text{col}(AB) \subseteq \text{col}(A), \text{rank}(A) = \text{rank}(AB) \Rightarrow \text{col}(AB) = \text{col}(A)$. Now if $A = [\mathbf{a}_1, \ldots, \mathbf{a}_n], AB = [\mathbf{b}_1, \ldots, \mathbf{b}_k]$ and $\mathbf{a}_i = \sum_{j=1}^{k} \alpha_{ij}\mathbf{b}_j$ then take $X = [\alpha_{ji}]$.

18. Show that $\text{rank}(A^*A) \leq \text{rank}([A^*A \mid A^*\mathbf{b}]) \leq \text{rank}(A^*[A \mid \mathbf{b}]) \leq \text{rank}(A^*)$ and $\text{rank}(A^*A) = \text{rank}(A^*)$. Hence $\text{rank}(A^*A) = \text{rank}([A^*A \mid A^*\mathbf{b}])$.

---

7 Linear Transformation

- Suppose $A \in \mathcal{M}_{m\times n}$. Take $\mathbf{v} \in \mathbb{R}^n$. Then $A\mathbf{v} \in \mathbb{R}^m$. Thus, we have a map (function) $F : \mathbb{R}^n \to \mathbb{R}^m$ given by $F(\mathbf{v}) = A\mathbf{v}$.
- Take $F : \mathbb{R}[x] \to \mathbb{R}[x]$ given by $F(p(x)) = p'(x)$.
- Take $F : \mathbb{R}[x] \to \mathbb{R}$ given by $F(p(x)) = p(3)$.

What is common in all of these? Well, they are maps (functions) with domains and co-domains as vector space's over $\mathbb{R}$. Further,

$$F(\mathbf{u}+\mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v}), \quad F(\alpha\mathbf{v}) = \alpha F(\mathbf{v}),$$

or, equivalently, $F(\alpha\mathbf{u}+\beta\mathbf{v}) = \alpha F(\mathbf{u}) + \beta F(\mathbf{v})$. Such functions are called linear transformations (LT).

Definition 7.1. *A **linear transformation** (or a linear mapping) from a vector space $V(\mathbb{F})$ into a vector space $W(\mathbb{F})$ is a mapping $T : V \to W$ such that for all $\mathbf{u}, \mathbf{v} \in V$ and for all $a \in \mathbb{F}$*

$$T(a\mathbf{u}+\mathbf{v}) = aT(\mathbf{u}) + T(\mathbf{v}).$$

Equivalently, $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(a\mathbf{v}) = aT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V,\ a \in \mathbb{F}$.

Example 7.1. *Let A be an $m \times n$ matrix. Define $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then T is a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$.*

Example 7.2. *The map $T : \mathbb{R} \to \mathbb{R}$, defined by $T(x) = x + 1$ for all $x \in \mathbb{R}$, is **not** a linear transformation.*

Example 7.3. *The map $T : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $T([x, y]^t) = [2x, x+y]^t$ for all $[x, y]^t \in \mathbb{R}^2$, is a linear transformation.*

Solution. Let $\mathbf{u} = [x_1, y_1]^t, \mathbf{v} = [x_2, y_2]^t \in \mathbb{R}^2$ and $a \in \mathbb{R}$. We have $a\mathbf{u} + \mathbf{v} = [ax_1 + x_2, ay_1 + y_2]^t$. So

$$\begin{aligned} T(a\mathbf{u}+\mathbf{v}) =& T([ax_1 + x_2, ay_1 + y_2]^t) \\ =& [2(ax_1 + x_2), (ax_1 + x_2) + (ay_1 + y_2)]^t \\ =& a[2x_1, x_1 + y_1]^t + [2x_2, x_2 + y_2]^t \\ =& aT(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

Hence T is a linear transformation. □

Example 7.4. *Let V and W be two vector spaces. The map $T_0 : V \to W$, defined by $T_0(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v} \in V$, is a linear transformation. The map T_0 is called the **zero transformation**.*

Example 7.5. *Let V be a vector space. The map $I : V \to V$, defined by $I(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$, is a linear transformation. The map I is called the **identity transformation**.*

Result 7.1. *Let $T : V \to W$ be a linear transformation. Then*

1. *$T(\mathbf{0}) = \mathbf{0}$;*
2. *$T(-\mathbf{v}) = -T(\mathbf{v})$ for all $\mathbf{v} \in V$; and*
3. *$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$.*

Proof. For the second part, we have $T(-\mathbf{v}) = T((-1)\mathbf{v}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$. □

Example 7.6. *Suppose $T : \mathbb{R}^2 \to \mathbb{R}_2[x]$ is a linear transformation such that $T[1, 0]^t = 2 - 3x + x^2$ and $T[0, 1]^t = 1 - x^2$. Find $T[2, 3]^t$ and $T[a, b]^t$.*

Result 7.2. *Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V (dim$(V) = n$). Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be arbitrarily chosen in W. Then there is a unique linear transformation $T : V \longrightarrow W$ such that $T(\mathbf{v}_i) = \mathbf{u}_i$.*

Proof. Let $\mathbf{v} \in V$. Then $\mathbf{v}$ equals a unique linear combination $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Define $T(\mathbf{v}) = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n \in W$. The resulting map T is the linear transformation we are looking for. □

Remark: To define (know) a linear transformation, it is enough to define (know) the images of vectors in any basis of the domain.

Composition of Linear Transformation: Let $T : U \to V$ and $S : V \to W$ be two linear transformations. Then the composition of S with T is the mapping $S \circ T : U \to W$ defined by

$$(S \circ T)(\mathbf{u}) = S(T(\mathbf{u})) \quad \text{for all } \mathbf{u} \in U.$$

Result 7.3. *Let $T : U \to V$ and $S : V \to W$ be two linear transformations. Then the composition $S \circ T$ is also a linear transformation.*

Inverse of a Function: A function $f : X \to Y$ is said to be **invertible** if there is another function $g : Y \to X$ such that

$$g \circ f = I_X \quad \text{and} \quad f \circ g = I_Y.$$

- If f is invertible, the the function g satisfying $g \circ f = I_X, \quad f \circ g = I_Y$ is called inverse of f.
- Inverse of a function, if it exists, is **unique**.
- The symbol f^{-1} is used to denote the inverse of f.
- Inverse of a linear transformation is linear.

Example 7.7. *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by*

$$T[x, y]^t = [x - y, -3x + 4y]^t \text{ and } S[x, y]^t = [4x + y, 3x + y]^t \text{ for all } [x, y]^t \in \mathbb{R}^2.$$

Then S is the inverse of T.

Solution. We have

$$(T \circ S)(x, y) = T(S(x, y)) = T([4x + y, 3x + y]^t) = [(4x + y) - (3x + y), -3(4x + y) + 4(3x + y)]^t = [x, y]^t = I([x, y]^t),$$

$$(S \circ T)(x, y) = S(T(x, y)) = S([x - y, -3x + 4y]^t) = [4(x - y) + (-3x + 4y), 3(x - y) + (-3x + 4y)]^t = [x, y]^t = I([x, y]^t).$$

Thus $T \circ S = I$ and $S \circ T = I$, and therefore $T^{-1} = S$. □

Kernel and Range: Let $T : V \to W$ be a linear transformation. Then the **kernel** of T (null space of T), denoted $\ker(T)$, and the **range** of T, denoted range(T), are defined as

$$\ker(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}, \text{ and}$$

$$\text{range}(T) = \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}.$$

Result 7.4. *Let $T : V \to W$ be a linear transformation and let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a spanning set for V. Then $T(B) = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_k)\}$ spans the range of T.*

Example 7.8. *Let A be an $m \times n$ matrix. Define $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then ker$(T)$ = null(A) and range(T) = col(A).*

Result 7.5. *Let $T : V \to W$ be a linear transformation. Then ker(T) is a subspace of V and range(T) is a subspace of W.*

Definition 7.2. *Let $T : V \to W$ be a linear transformation. Then we define*

- ***rank**(T) = dimension of range(T); and*
- ***nullity**(T) = dimension of ker(T).*

Example 7.9. *Let $D : \mathbb{R}_3[x] \to \mathbb{R}_2[x]$ be defined by $D(p(x)) = \frac{d}{dx}p(x)$. Then rank$(D) = 3$ and nullity$(D) = 1$.*

Result 7.6 (The Rank-Nullity Theorem). *Let $T : V \to W$ be a linear transformation from a finite dimensional vector space V into a vector space W. Then*

$$\mathit{rank}(T) + \mathit{nullity}(T) = \mathit{dim}(V).$$

Definition 7.3. *Let $T : V \to W$ be a linear transformation. Then*

1. *T is called **one-one** if T maps distinct vectors in V into distinct vectors in W.*
2. *T is called **onto** if range$(T) = W$.*

- For all $\mathbf{u}, \mathbf{v} \in V$, if $\mathbf{u} \neq \mathbf{v}$ implies that $T(\mathbf{u}) \neq T(\mathbf{v})$, then T is one-one.
- For all $\mathbf{u}, \mathbf{v} \in V$, if $T(\mathbf{u}) = T(\mathbf{v})$ implies that $\mathbf{u} = \mathbf{v}$, then T is one-one.
- For all $\mathbf{w} \in W$, if there is at least one $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$, then T is onto.

Example 7.10. *Some examples of one-one and onto linear transformation.*

- $T : \mathbb{R} \to \mathbb{R}^2$ *defined by* $T(x) = [x, 0]^t, x \in \mathbb{R}$ *is one-one but not onto.*
- $T : \mathbb{R}^2 \to \mathbb{R}$ *defined by* $T[x, y]^t = x$*, for* $[x, y]^t \in \mathbb{R}^2$ *is onto but not one-one.*
- $T : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by* $T[x, y]^t = [-x, -y]^t$*, for* $[x, y]^t \in \mathbb{R}^2$ *is one-one and onto.*

Solution.

- For $x, y \in \mathbb{R}$, we have $T(x) = T(y) \Rightarrow [x, 0]^t = [y, 0]^t \Rightarrow x = y$. So T is one-one. However, $[1, 1]^t$ does not have a pre-image, and so T is not onto.
- We have $[1, 1]^t, [1, 2]^t \in \mathbb{R}^2$ such that $[1, 1]^t \neq [1, 2]^t$ but $T([1, 1]^t) = T([1, 2]^t)$. So, T is not one-one. However, for every $x \in \mathbb{R}$, we have $[x, 1]^t \in \mathbb{R}^2$ such that $T([x, 1]^t) = x$. So, T is onto.
- For $[x, y]^t, [u, v]^t \in \mathbb{R}^2$, we have $T([x, y]^t) = T([u, v]^t) \Rightarrow [-x, -y]^t = [-u, -v]^t \Rightarrow -x = -u, -y = -v \Rightarrow [x, y]^t = [u, v]^t$. So T is one-one. Also, for every $[x, y]^t \in \mathbb{R}^2$, we have $[-x, -y]^t \in \mathbb{R}^2$ such that $T([-x, -y]^t) = [x, y]^t$. So, T is onto. □

Result 7.7. *A linear transformation* $T : V \to W$ *is one-one iff ker*$(T) = \{\mathbf{0}\}$*.*

Result 7.8. *Let dim*$(V) =$ *dim*(W)*. Then a linear transformation* $T : V \to W$ *is one-one iff* T *is onto.*

Example 7.11. *Let* $T : V \longrightarrow W$ *be a linear transformation and* $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ *such that* $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$ *are linearly independent. Can* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *be linearly dependent? Justify?*

Solution. No, because if $a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k = \mathbf{0}$ then $a_1T(\mathbf{v}_1) + \ldots + a_kT(\mathbf{v}_k) = \mathbf{0}$. □

Result 7.9. *Let* $T : V \to W$ *be a one-one linear transformation. If* $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *is a linearly independent set in* V *then* $T(S) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ *is a linearly independent set in* W*.*

Remark: Find an example to see that if T is **not** one-one then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent does not necessarily imply that $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ is linearly independent.

Result 7.10. *Let dim*$(V) =$ *dim*(W)*. Then a one-one linear transformation* $T : V \to W$ *maps a basis for* V *onto a basis for* W*.*

Isomorphism:

- A linear transformation $T : V \to W$ is called an **isomorphism** if it is one-one and onto.
- If $T : V \to W$ is an isomorphism then we say that V and W are isomorphic, and we write $V \cong W$.

Example 7.12. *The vector spaces* $\mathbb{R}^3$ *and* $\mathbb{R}_2[x]$ *are isomorphic.*

Result 7.11. *Let* $V(\mathbb{F})$ *and* $W(\mathbb{F})$ *be two finite dimensional vector spaces. Then* $V \cong W$ *iff dim*$(V) =$ *dim*(W)*.*

Example 7.13. *The vector spaces* $\mathbb{R}^n$ *and* $\mathbb{R}_n[x]$ *are not isomorphic.*

Matrix of a Linear Transformation:

Result 7.12. *Let* V *and* W *be two finite-dimensional vector spaces with ordered bases* B *and* C *respectively, where* $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *and* $C = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$*. If* $T : V \to W$ *is a linear transformation, then the* $m \times n$ *matrix* A *defined by*

$$A = [[T(\mathbf{v}_1)]_C, [T(\mathbf{v}_2)]_C, \ldots, [T(\mathbf{v}_n)]_C]$$

satisfies

$$A[\mathbf{v}]_B = [T(\mathbf{v})]_C \qquad \text{for all } \mathbf{v} \in V.$$

- The matrix A in Result 7.12 is called the **matrix of** T **with respect to the ordered bases** B **and** C.
- The matrix A is also written as $[T]_{C \leftarrow B}$.
- If $B = C$, then $[T]_{C \leftarrow B}$ is written as $[T]_B$.

Remark 7.1. *Result 7.12 means:*
Suppose we know $[T]_{C\leftarrow B}$ with respect to given ordered bases B and C. Then we know T in the following sense:

If $\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{v}_i$ *and* $[T]_{C\leftarrow B}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, *then* $T(\mathbf{v}) = \sum_{j=1}^{m} b_j\mathbf{u}_j$.

The above expression can be represented by the following diagram.

$$A[\mathbf{v}]_B = [T(\mathbf{v})]_C \quad, i.e., \quad \left\{ \begin{array}{ccc} \mathbf{v} \in V & \xrightarrow{T} & T(\mathbf{v}) \in W \\ \downarrow & & \downarrow \\ [\mathbf{v}]_B \in \mathbb{F}^n & \xrightarrow{T_A} & [T(\mathbf{v})]_C = A[\mathbf{v}]_B \in \mathbb{F}^m \end{array} \right\}.$$

Here $T_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ is the linear transformation given by $T_A(\mathbf{x}) = A\mathbf{x}$.

Example 7.14. *Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by*

$$T([x, y, z]^t) = [x - 2y, x + y - 3z]^t \quad \textit{for } [x, y, z]^t \in \mathbb{R}^3.$$

Let $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $C = \{\mathbf{e}_2, \mathbf{e}_1\}$ be ordered bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Find $[T]_{C\leftarrow B}$ and verify Result 7.12 for $\mathbf{v} = [1, 3, -2]^t$.

Example 7.15. *Consider $D : \mathbb{R}_3[x] \to \mathbb{R}_3[x]$ defined by $D(p(x)) = p'(x)$. Find the matrix $[D]_B$ with respect to the ordered basis $B = \{1, x, x^2, x^3\}$ of $\mathbb{R}_3[x]$. Also verify that $[D]_B[p(x)]_B = [D(p(x))]_B$.*

Solution. Since $D(1) = 0, D(x) = 1, D(x^2) = 2x, D(x^3) = 3x^2$, we get $[D]_B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Consider $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$. Then $D(p(x)) = a_1 + 2a_2x + 3a_3x^2$.

We see that $[p(x)]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$, $[D]_B[p(x)]_B = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{bmatrix} = [D(p(x))]_B$. □

Result 7.13. *Let U, V and W be three finite-dimensional vector spaces with ordered bases B, C and D, respectively. Let $T : U \to V$ and $S : V \to W$ be linear transformations. Then*

$$[S \circ T]_{D\leftarrow B} = [S]_{D\leftarrow C}[T]_{C\leftarrow B}.$$

Result 7.14. *Let $T : V \to W$ be a linear transformation between two n-dimensional vector spaces V and W with ordered bases B and C, respectively. Then T is invertible if and only if the matrix $[T]_{C\leftarrow B}$ is invertible. In this case,*

$$([T]_{C\leftarrow B})^{-1} = [T^{-1}]_{B\leftarrow C}.$$

Example 7.16. *Let $T : \mathbb{R}^2 \to \mathbb{R}_1[x]$ be defined by $T([a, b]^t) = a + (a + b)x$ for $[a, b]^t \in \mathbb{R}^2$. Show that T is invertible, and hence find T^{-1}.*

Practice Problems Set 7

1. Examine whether the following maps $T : V \to W$ are linear transformations.

   (a) $V = W = \mathbb{R}^3$ and $T[x, y, z]^t = [3x + y, z, |x|]^t$ for all $[x, y, z]^t \in \mathbb{R}^3$.

   (b) $V = W = M_2(\mathbb{R})$ and for every $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$, define **(i)** $T(A) = A^t$, **(ii)** $T(A) = A + I_2$, **(iii)** $T(A) = A^2$, **(iv)** $T(A) = \det A$, **(v)** $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$, **(vi)** $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}$.

   (c) $V = W = M_n(\mathbb{R})$ and for every $A = [a_{ij}] \in M_n(\mathbb{R})$, define **(i)** $T(A) = \text{tr}(A)$, **(ii)** $T(A) = \text{rank}(A)$ and **(iii)** $T(A) = a_{11}a_{22}\dots a_{nn}$.

   (d) $V = W = \mathbb{R}_2[x]$ and $T(a + bx + cx^2) = a + b(x + 1) + c(x + 1)^2$ for all $a + bx + cx^2 \in \mathbb{R}_2[x]$.

2. Let $T : \mathbb{R}^2 \to \mathbb{R}_2[x]$ be a linear transformation for which $T([1,1]^t) = 1 - 2x$ and $T([3,-1]^t) = x + 2x^2$. Find $T([-7,9]^t)$ and $T([a,b]^t)$ for $[a,b]^t \in \mathbb{R}_2[x]$.

3. Let $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ be a linear transformation for which $T(1+x) = 1+x^2$, $T(x+x^2) = x - x^2$ and $T(1+x^2) = 1+x+x^2$. Find $T(4 - x + 3x^2)$ and $T(a+bx+cx^2)$ for $a+bx+cx^2 \in \mathbb{R}_2[x]$.

4. Consider the linear transformations $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T[x,y]^t = [0,x]^t$ and $S[x,y]^t = [y,x]^t$ for all $[x,y]^t \in \mathbb{R}^2$. Compute $T \circ S$ and $S \circ T$. What do you observe?

5. Let $T : \mathbb{C} \to \mathbb{C}$ be the map defined by $T(z) = \overline{z}$ for all $z \in \mathbb{C}$. Show that T is a linear transformation on $\mathbb{C}(\mathbb{R})$ but not a linear transformation on $\mathbb{C}(\mathbb{C})$.

6. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for a vector space V and $T : V \to V$ be a linear transformation. Prove that if $T(\mathbf{u}_i) = \mathbf{u}_i$ for all $i = 1, 2, \ldots, n$, then T is the identity transformation on V.

7. Let V be a vector space over $\mathbb{R}$ (or $\mathbb{C}$) of dimension n and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V. If W is another vector space over $\mathbb{R}$ (or $\mathbb{C}$) and $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in W$, then show that there exists a unique linear transformation $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for all $i = 1, 2, \ldots, n$.

8. Examine the linearity of the following maps. Also, find bases for their range spaces and null spaces, whenever they are linear.

   (a) $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ for $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$.

   (b) $T : \mathbb{R}_2[x] \to \mathbb{R}^2$ defined by $T(a+bx+cx^2) = [a-b, b+c]^t$ for $a+bx+cx^2 \in \mathbb{R}_2[x]$.

   (c) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T[x,y]^t = [x, x+y, x-y]^t$ for all $[x,y]^t \in \mathbb{R}^2$.

   (d) $T : \mathbb{R}_2[x] \to \mathbb{R}_3[x]$ defined by $T(p(x)) = x.p(x)$ for all $p(x) \in \mathbb{R}_2[x]$.

   (e) $T : M_2(\mathbb{R}) \to \mathbb{R}^2$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = [a-b, c-d]^t$ for $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$.

9. Examine whether the following linear transformations are one-one and onto.

   (a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T[x,y]^t = [2x-y, x+2y]^t$ for all $[x,y]^t \in \mathbb{R}^2$.

   (b) $T : \mathbb{R}_2[x] \to \mathbb{R}^3$ defined by $T(a+bx+cx^2) = [2a-b, a+b-3c, c-a]^t$ for all $a+bx+cx^2 \in \mathbb{R}_2[x]$.

10. Find a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that

    (a) $\text{range}(T) = \text{span}([1,1,1]^t)$ $\qquad$ (b) $\text{range}(T) = \text{span}([1,2,3]^t, [1,3,2]^t)$.

11. Let $S : V \to W$ and $T : U \to V$ be two linear transformations.

    (a) If $S \circ T$ is one-one then prove that T is also one-one.

    (b) If $S \circ T$ is onto then prove that S is also onto.

12. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation such that $T[1,0,0]^t = [1,0,0]^t$, $T[1,1,0]^t = [1,1,1]^t$ and $T[1,1,1]^t = [1,1,0]^t$. Find $T[x,y,z]^t$, $\text{nullity}(T)$ and $\text{rank}(T)$, where $[x,y,z]^t \in \mathbb{R}^3$. Also, show that $T^2 = T$.

13. Let $z_1, z_2, \ldots, z_k$ be k distinct complex numbers. Let $T : \mathbb{C}_n[z] \to \mathbb{C}^k$ be defined by $T(f(z)) = [f(z_1), f(z_2), \ldots, f(z_k)]^k$ for all $f(z) \in \mathbb{C}_n[z]$. Show that T is a linear transformation, and find the dimension of range(T).

14. Show that each of the following linear transformations is an isomorphism.

    (a) $T : \mathbb{R}_3[x] \to \mathbb{R}^4$ defined by $T(a+bx+cx^2+dx^3) = [a,b,c,d]^t$ for all $a+bx+cx^2+dx^3 \in \mathbb{R}_3[x]$.

    (b) $T : M_2(\mathbb{C}) \to \mathbb{C}^4$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = [a,b,c,d]^t$ for all $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{C})$.

    (c) $T : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ defined by $T(X) = A^{-1}XA$ for all $X \in M_n(\mathbb{R})$, where A is a given $n \times n$ invertible matrix.

    (d) $T : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ defined by $T(p(x)) = p(x) + \frac{d}{dx}(p(x))$ for all $p(x) \in \mathbb{R}_n[x]$.

15. Examine whether the following vector spaces V and W are isomorphic. Whenever they are isomorphic, find an explicit isomorphism $T : V \to W$.

    (a) $V = \mathbb{C}$ and $W = \mathbb{R}^2$.

    (b) $V = \{A \in M_2(\mathbb{R}) : \text{tr}(A) = 0\}$ and $W = \mathbb{R}^2$.

    (c) $V =$ the vector space of all 3×3 diagonal matrices and $W = \mathbb{R}^3$.

    (d) $V =$ the vector space of all 3×3 symmetric matrices and $W =$ the vector space of all 3×3 skew-symmetric matrices.

Result 7.15. *Let $T : V \to W$ be a linear transformation. If V is finite-dimensional then ker(T) and range(T) are also finite-dimensional.*

16. Let A be an $m \times n$ real matrix. Using the above result, show that if $m < n$ then the system of equations $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, and if $m > n$ then there exists a non-zero vector $\mathbf{b} \in \mathbb{R}^m$ such that the system of equations $A\mathbf{x} = \mathbf{b}$ does not have any solution.

17. Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ be a basis for a vector space V of dimension 4. Define a linear transformation $T : V \to V$ such that
$$T(\mathbf{u}_1) = T(\mathbf{u}_2) = T(\mathbf{u}_3) = \mathbf{u}_1, T(\mathbf{u}_4) = \mathbf{u}_2.$$
Describe each of the spaces $\ker(T)$, $\text{range}(T)$, $\ker(T) \cap \text{range}(T)$ and $\ker(T) + \text{range}(T)$.

18. Let V be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation. If $\text{rank}(T) = \text{rank}(T^2)$ then prove that $\text{range}(T) \cap \ker(T) = \{\mathbf{0}\}$.

19. Let U and W be subspaces of a finite-dimensional vector space V. Define $T : U \times W \to V$ by $T(\mathbf{u}, \mathbf{w}) = \mathbf{u} - \mathbf{w}$ for all $(\mathbf{u}, \mathbf{w}) \in U \times W$.

    (a) Show that T is a linear transformation.

    (b) Show that $\text{range}(T) = U + W$.

    (c) Show that $\ker(T) \cong U \cap W$.

    (d) Prove **Grassmann's identity**: $\dim\,(U + W) = \dim U + \dim W - \dim\,(U \cap W)$.

20. Find the matrix $[T]_{C \leftarrow B}$ for each of the following linear transformations $T : V \to W$ with respect to the given ordered bases B and C for V and W, respectively.

    (a) $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T[x, y, z]^t = [x-y+z, y-z]^t$ for all $[x, y, z]^t \in \mathbb{R}^3$, and $B = \{[1, 1, 1]^t, [1, 1, 0]^t, [1, 0, 0]^t\}$, $C = \{[1, 1]^t, [1, -1]^t\}$.

    (b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T[x, y, z]^t = [x+y, y+z, z+x]^t$ for all $[x, y, z]^t \in \mathbb{R}^3$, and $B = C = \{[1, 1, 0]^t, [0, 1, 1]^t, [1, 0, 1]^t\}$.

    (c) $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T[x_1, x_2, x_3, \ldots, x_n]^t = [x_2, x_3, \ldots, x_n, 0]^t$ for all $[x_1, x_2, x_3, \ldots, x_n]^t \in \mathbb{R}^n$, and $B = C =$ the standard basis for $\mathbb{R}^n$.

    (d) $T : \mathbb{R}_3[x] \to \mathbb{R}_4[x]$ defined by $T(p(x)) = x.p(x)$ for all $p(x) \in \mathbb{R}_3[x]$, and $B = \{1, x, x^2, x^3\}$, $C = \{1, x, x^2, x^3, x^4\}$.

    (e) $T : \mathbb{C}^2 \to \mathbb{C}^2$ defined by $T[z_1, z_2]^t = [z_1 + z_2, iz_2]^t$ for all $[z_1, z_2]^t \in \mathbb{C}^2$, and $B =$ the standard basis for $\mathbb{C}^2$, $C = \{[1, 1]^t, [1, 0]^t\}$.

    (f) $T : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ defined by $T(A) = A + iA^t$ for all $A \in M_2(\mathbb{C})$, and
    $B = C = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$.

21. Let A be an $m \times n$ matrix and let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Show that the matrix of T with respect to the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ is A.

22. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation whose matrix with respect to the standard basis B for $\mathbb{R}^3$ is given by
$$A = \begin{bmatrix} 1 & 2 & -4 \\ 2 & -3 & 5 \\ 1 & 0 & 1 \end{bmatrix}.$$
Let $[x, y, z]^t \in \mathbb{R}^3$. Determine $T[x, y, z]^t$ and show that T is invertible. Also, find $T^{-1}[x, y, z]^t$.

23. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation whose matrix with respect to the basis $B = \{[1,1,1]^t, [1,0,1]^t, [1,-1,-1]^t\}$ for $\mathbb{R}^3$ is given by

$$A = \begin{bmatrix} 1 & 2 & -4 \\ 2 & -3 & 5 \\ 1 & 0 & 1 \end{bmatrix}.$$

Determine $T[x,y,z]^t$, where $[x,y,z]^t \in \mathbb{R}^3$.

24. Consider the bases $B = \{[1,1,1]^t, [0,1,1]^t, [0,0,1]^t\}$ and $C = \{1-t, 1+t\}$ for $\mathbb{R}^3$ and $\mathbb{R}_1[t]$, respectively. If $[T]_{C\leftarrow B} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}$ is the matrix of a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}_1[t]$, determine $T[x,y,z]^t$, where $[x,y,z]^t \in \mathbb{R}^3$.

25. Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ be a basis for a vector space V of dimension 4, and let T be a linear transformation on V whose matrix representation with respect to this basis is given by

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{bmatrix}.$$

   (a) Describe each of the spaces $\ker(T)$ and $\text{range}(T)$.

   (b) Find a basis for $\ker(T)$, extend it to a basis for V, and then find the matrix representation of T with respect to this basis.

26. Let T and S be two linear transformations on $\mathbb{R}^2$. Suppose the matrix representation of T with respect to the basis $\{\mathbf{u}_1 = [1,2]^t, \mathbf{u}_2 = [2,1]^t\}$ is $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, and the matrix representation of S with respect to the basis $\{\mathbf{v}_1 = [1,1]^t, \mathbf{v}_2 = [1,2]^t\}$ is $\begin{bmatrix} 3 & 3 \\ 2 & 4 \end{bmatrix}$. Let $\mathbf{u} = [3,3]^t \in \mathbb{R}^2$.

   (a) Find the matrix of $T + S$ with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

   (b) Find the matrix of $T \circ S$ with respect to the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$.

   (c) Find the coordinate of $T(\mathbf{u})$ with respect to the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$.

   (d) Find the coordinate of $S(\mathbf{u})$ with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

27. Let $\dim(V) = n$ and $T : V \to V$ be a linear transformation. If B and C are two bases of V, then show that the matrices $[T]_B$ and $[T]_C$ are similar, that is, there is an invertible matrix P such that $[T]_B = P^{-1}[T]_C P$.

28. Let V and W be two vector spaces of dimensions n and m, respectively, over the same $\mathbb{F}$. For two linear transformations $T : V \to W$ and $S : V \to W$, and for $a \in \mathbb{F}$, define the maps $T + S : V \to W$ and $aT : V \to W$ by

$$(T+S)(v) = T(v) + S(v) \quad \text{and} \quad (aT)(v) = aT(v) \quad \text{for all } v \in V.$$

   (a) Show that $S + T$ and aT are also linear transformations.

   (b) Let $\mathcal{L}(V,W)$ be the set of all linear transformations from V into W. Show that $\mathcal{L}(V,W)$ is a vector space, under the vector addition and scalar multiplication, as defined above.

   (c) Let B and C be bases for V and W, respectively. Define $\mathcal{F} : \mathcal{L}(V,W) \longrightarrow \mathcal{M}_{m\times n}(\mathbb{F})$ by

$$\mathcal{F}(T) = [T]_{C\leftarrow B} \qquad \text{for all } T \in \mathcal{L}(V,W).$$

   Show that $\mathcal{F}$ is an isomorphism.

   (d) Show, using Part (c) or otherwise, that $\dim\mathcal{L}(V,W) = mn$.

Hints to Practice Problems Set 7

1. Yes: $(b)i, (b)v, (b)vi, (c)i, (d)$. No: $(a), (b)ii, (b)iii, (b)iv, (c)ii, (c)iii$.

2. $T[a,b]^t = \frac{1}{4}[(a+3b) - (a+7b)x + 2(a-b)x^2]$.

3. $T(a+bx+cx^2) = a + cx + \frac{(3a-b-c)}{2}x^2$.

4. $(T \circ S)[x,y]^t = [0,y]^t, \quad (S \circ T)[x,y]^t = [x,0]^t$.

5. $T(i.1) \neq iT(1)$.

6. $T(\mathbf{v}) = T(\sum \mathbf{a}_i\mathbf{u}_i) = \sum \mathbf{a}_i T(\mathbf{u}_i) = \sum \mathbf{a}_i\mathbf{u}_i = \mathbf{v}$ for all $\mathbf{v} \in V$.

7. For $\mathbf{v} = \sum \mathbf{a}_i\mathbf{v}_i$, define $T(\mathbf{v}) = \sum \mathbf{a}_i\mathbf{w}_i$.

8. (a) $\ker(T) = \left\{ \begin{bmatrix} 0 & x \\ y & 0 \end{bmatrix} : x,y \in \mathbb{R} \right\}, \quad \text{range}(T) = \left\{ \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} : x,y \in \mathbb{R} \right\}.$
   (b) $\ker(T) = \text{span}(1+x-x^2)$, $\text{range}(T) = \mathbb{R}^2$, (c) $\ker(T) = \{\mathbf{0}\}$, $\text{range}(T) = \{[x,y,z]^t \in \mathbb{R}^3 : 2x - y - z = 0\}$,
   (d) $\ker(T) = \{\mathbf{0}\}$, $\text{range}(T) = \text{span}(x, x^2, x^3)$, (e) $\ker(T) = \left\{ \begin{bmatrix} a & a \\ c & c \end{bmatrix} : a,c \in \mathbb{R} \right\}$, $\text{range}(T) = \mathbb{R}^2$.

9. (a) $\ker(T) = \{\mathbf{0}\}$, $\quad (b)$ $\ker(T) \neq \{\mathbf{0}\}$.

10. (a) Define $T(\mathbf{e}_1) = [1,1,1]^t, T(\mathbf{e}_2) = T(\mathbf{e}_3) = \mathbf{0}$. $\quad (b)$ Define $T(\mathbf{e}_1) = [1,2,3]^t, T(\mathbf{e}_2) = [1,3,2]^t, T(\mathbf{e}_3) = \mathbf{0}$.

11. (a) $T(\mathbf{x}) = \mathbf{0} \Rightarrow (S \circ T)(\mathbf{x}) = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$. $\quad (b)$ For $\mathbf{w} \in W$, $S(T\mathbf{u}) = \mathbf{w}$ where $\mathbf{u} \in U$.

12. $T[x,y,z]^t = [x,y,y-z]^t$, $\text{nullity}(T) = 0$, $\text{rank}(T) = 3$, $T(T[x,y,z]^t) = [x,y,z]^t$.

13. If $k \geq n+1$ then $f \in \ker(T) \Rightarrow f$ will have at least $n+1$ roots $\Rightarrow \ker(T) = \{\mathbf{0}\}$.
    If $k = n$ then $\ker(T) = \text{span}(f(z))$, where $f(z) = (z-z_1)(z-z_2)\ldots(z-z_n)$.
    If $k < n$ then $\ker(T) = \{(z-z_1)(z-z_2)\ldots(z-z_k)q(z) : \ q(z)$ is a polynomial of degree at most $n-k\}$. Now use $\text{rank}(T) + \text{nullity}(T) = n+1$.

14. All are isomorphisms. One method may be by showing that $\ker(T) = \{\mathbf{0}\}$.

15. (a) $T(a+ib) = [a,b]^t$, $\quad (b)$ $\dim(V) \neq \dim(\mathbb{R}^2)$, $\quad (c)$ $T(\text{diag}[x,y,z]) = [x,y,z]^t$, $\quad (d)$ $\dim(V) \neq \dim(W)$.

16. Consider $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$. Now $m < n \Rightarrow T$ is not one-one, and $m > n \Rightarrow T$ is not onto.

17. $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for $\text{range}(T)$, $\{\mathbf{u}_1 - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_3\}$ is a basis for $\ker(T)$, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a basis for $\ker(T) + \text{range}(T)$ and $\{\mathbf{u}_1 - \mathbf{u}_2\}$ is a basis for $\ker(T) \cap \text{range}(T)$.

18. $\text{rank}(T) = \text{rank}(T^2) \Rightarrow \text{nullity}(T) = \text{nullity}(T^2)$. Also $\ker(T) \subseteq \ker(T^2)$. Hence $\ker(T) = \ker(T^2)$. Now $\mathbf{x} \in \ker(T) \cap \text{range}(T) \Rightarrow T\mathbf{x} = \mathbf{0}, \mathbf{x} = T\mathbf{y}$ for some $\mathbf{y}$ $\Rightarrow T^2\mathbf{y} = \mathbf{0} \Rightarrow T\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$.

19. (c) Define $S : \ker(T) \to U \cap W$ by $S(\mathbf{w}, \mathbf{w}) = \mathbf{w}$. Then S is an isomorphism. (d) Use Rank-Nullity Theorem.

20. (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$, (e) $\begin{bmatrix} 0 & i \\ 1 & 1-i \end{bmatrix}$, $\quad (c)$ $[\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}]$, $\quad (d)$ $[\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5]$,
    (b) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $\quad (f)$ $\begin{bmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & 0 & 1+i \end{bmatrix}$.

21. If $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n]$ then $T(\mathbf{e}_i) = \mathbf{a}_i$.

22. $T[x,y,z]^t = A[x,y,z]^t = [x+2y-4z, \ 2x-3y+5z, \ x+z]^t$. Also A is invertible $\Rightarrow T$ is invertible, and $T^{-1}[x,y,z]^t = A^{-1}[x,y,z]^t$.

23. Use $A[\mathbf{x}]_B = [T(\mathbf{x})]_B$. We have $T[x,y,z]^t = \left[3x + 5y - 4z, \ \frac{-5x-4y+9z}{2}, \ x + 3y - 2z\right]$.

24. Use $A[\mathbf{x}]_B = [T(\mathbf{x})]_C$. We have $T[x,y,z]^t = (-2x+2y) + (-4y+2z)t$.

25. (a) A basis for $\text{null}(A)$ is $\{\mathbf{x}, \mathbf{y}\}$, where $\mathbf{x} = [-2, -\frac{3}{2}, 1, 0]^t, \mathbf{y} = [-1,-2,0,1]^t$. We have $\ker(T) = \text{span}(\mathbf{v}, \mathbf{w})$, where $\mathbf{v} = -2\mathbf{u}_1 - \frac{3}{2}\mathbf{u}_2 + \mathbf{u}_3$ and $\mathbf{w} = -\mathbf{u}_1 - 2\mathbf{u}_2 + \mathbf{u}_4$.
    Now $\text{col}(A) = \text{span}(\mathbf{a}_1, \mathbf{a}_2)$, where $\mathbf{a}_1, \mathbf{a}_2$ are the 1st and the 2nd column of A. Therefore $\text{range}(T) = \text{span}(T\mathbf{u}_1, T\mathbf{u}_2)$,

since $\mathbf{a}_1 = [T\mathbf{u}_1]_B,\ \mathbf{a}_2 = [T\mathbf{u}_2]_B$.

(b) $C = \{\mathbf{v}, \mathbf{w}, \mathbf{u}_1, \mathbf{u}_2\}$ will be a basis for V. Also $[T]_C = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 9/2 & 1 \end{bmatrix}$.

26. (a) $\begin{bmatrix} 8 & 9 \\ 4/3 & 3 \end{bmatrix}$, (b) $\begin{bmatrix} 7 & 8 \\ 13 & 14 \end{bmatrix}$, (c) $[3,5]^t$, (d) $[9,6]^t$.

27. Take $P = [I]_{C \leftarrow B}$ and use $[I \circ T]_{C \leftarrow B} = [T \circ I]_{C \leftarrow B} \Rightarrow [I]_{C \leftarrow B}[T]_B = [T]_C[I]_{C \leftarrow B}$.

28. (a) Routine Work.

(b) Routine Work.

(c) For injectiveness, show $\text{Ker}(\mathcal{F}) = \{\mathbf{0}\}$. For surjectiveness, if $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{F})$, $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $C = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$, then take $T : V \to W$ such that $T(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m$ to have $\mathcal{F}(T) = A$.

(d) $\dim\mathcal{L}(V, W) = \dim\left(\mathcal{M}_{m\times n}(\mathbb{F})\right)$.

8 Eigenvalue, Eigenvector and Diagonalizability

Recall that like the space $\mathbb{R}^n$, we also defined the space $\mathbb{C}^n$. Indeed,

$$\mathbb{C}^n = \{[x_1, x_2, \ldots, x_n]^t : x_1, x_2, \ldots, x_n \in \mathbb{C}\}.$$

Definition 8.1. *Let A be an $n \times n$ matrix. A complex number λ is called an **eigenvalue** of A if there is a vector $\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Such a vector $\mathbf{x}$ is called an **eigenvector** of A corresponding to λ.*

Example 8.1. *The numbers $4, -2$ are eigenvalues of $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ with corresponding eigenvectors $[1, 1]^t$ and $[1, -1]^t$, respectively.*

Definition 8.2. *Let λ be an eigenvalue of a matrix A. Then the collection of all eigenvectors of A corresponding to λ, together with the zero vector, is called the eigenspace of λ, and is denoted by E_λ.*

Result 8.1. *Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. Then*

1. *λ is an eigenvalue of A iff det$(A - \lambda I) = 0$.*
2. *λ is an eigenvalue of A iff $A - \lambda I$ is not invertible.*
3. *0 is an eigenvalue of A iff A is not invertible.*
4. *$E_\lambda = $ null$(A - \lambda I)$ is a non-trivial subspace of $\mathbb{C}^n$.*
5. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to λ and $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k \neq \mathbf{0}$. Then $\mathbf{v}$ is also an eigenvector of A corresponding to λ.*
6. *Eigenvalues of a **triangular** matrix are its **diagonal** entries.*
7. *Eigenvalues of $\left[\begin{array}{c|c} A_p & C \\ \hline O & B_q \end{array}\right]$ are the eigenvalues of A_p and B_q.*

Proof. Let λ be an eigenvalue of A with corresponding eigenvector $\mathbf{x}$, so that $A\mathbf{x} = \lambda\mathbf{x}$ and $\mathbf{x} \neq \mathbf{0}$.

1. We have

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\Leftrightarrow A\mathbf{x} = \lambda\mathbf{x}, \mathbf{x} \neq \mathbf{0} \\ &\Leftrightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}, \mathbf{x} \neq \mathbf{0} \\ &\Leftrightarrow (A - \lambda I)\mathbf{y} = \mathbf{0} \text{ has a non-trivial solution} \\ &\Leftrightarrow A - \lambda I \text{ is not invertible} \\ &\Leftrightarrow \det(A - \lambda I) = 0. \end{aligned}$$

2. Use Part 1.
3. Take $\lambda = 0$ and use Part 2.
4. $E_\lambda = \{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \lambda\mathbf{x}\} = \{\mathbf{x} \in \mathbb{C}^n : (A - \lambda I)\mathbf{x} = \mathbf{0}\} = \text{null}(A - \lambda I)$.
5. Given that $A\mathbf{v}_i = \lambda\mathbf{v}_i$ for $i = 1, 2, \ldots, k$ and $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k \neq \mathbf{0}$. Now

$$A\mathbf{v} = A(\alpha_1\mathbf{v}_1) + A(\alpha_2\mathbf{v}_2) + \ldots + A(\alpha_k\mathbf{v}_k) = \alpha_1(\lambda\mathbf{v}_1) + \alpha_2(\lambda\mathbf{v}_2) + \ldots + \alpha_k(\lambda\mathbf{v}_k) = \lambda\mathbf{v}.$$

   Hence $\mathbf{v}$ is also an eigenvector of A corresponding to λ.

6. Let A be a triangular matrix with diagonal entries a_{11}, a_{22} and a_{nn}, respectively. Then we have

$$\det(A - \lambda I) = 0 \Leftrightarrow (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0 \Leftrightarrow \lambda = a_{11}, a_{22}, \ldots, a_{nn}.$$

7. We have

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } \left[\begin{array}{c|c} A_p & C \\ \hline O & B_q \end{array}\right] &\Leftrightarrow \det\left(\left[\begin{array}{c|c} A_p & C \\ \hline O & B_q \end{array}\right] - \lambda I\right) = 0 \\
&\Leftrightarrow \det\left[\begin{array}{c|c} A_p - \lambda I & C \\ \hline O & B_q - \lambda I \end{array}\right] = 0 \\
&\Leftrightarrow \det(A_p - \lambda I).\det(B_q - \lambda I) = 0 \\
&\Leftrightarrow \lambda \text{ is an eigenvalue of } A_p \text{ or } B_q.
\end{aligned}$$

□

Definition 8.3. *Let A be an $n \times n$ matrix. Then*

- $p_A(x) = det(A - xI)$ *is called* ***characteristic polynomial*** *of A.*
- $p_A(x) = 0$ *is called* ***characteristic equation*** *of A.*

Method of Finding Eigenvalues and Bases for Corresponding Eigenspaces:

Let A be an $n \times n$ matrix.

1. Compute the characteristic polynomial $\det(A - \lambda I)$.
2. Find the eigenvalues of A by solving the characteristic equation $\det(A - \lambda I) = 0$ for λ.
3. For each eigenvalue λ, find a basis of $\text{null}(A - \lambda I)$. This null space is the required eigenspace, *i.e.*, $E_\lambda = \text{null}(A - \lambda I)$.

Example 8.2. *Find the eigenvalues and the corresponding eigenspaces of the following matrices:*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}.$$

Example 8.3. *A real matrix may have complex eigenvalues and complex eigenvectors. No wonder, every real matrix is also a complex matrix. Consider* $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. *The characteristic polynomial of A is $\lambda^2 + 1$, eigenvalues are $\pm\, i$. Corresponding to the eigenvalues $\pm\, i$, it has eigenvectors $[1, i]^T$ and $[i, 1]^T$, respectively. Eigenspaces E_i and $E_{(-i)}$ are one dimensional.* ***However, most of our examples will be of real matrices with real eigenvalues.***

Result 8.2 (The Fundamental Theorem of Invertible Matrices: Version III). *Let A be an $n \times n$ matrix. Then the following statements are equivalent.*

1. *A is invertible.*
2. *A^t is invertible.*
3. *$A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{C}^n$.*
4. *$A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{C}^n$.*
5. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
6. *The reduced row echelon form of A is I_n.*
7. *The rows of A are linearly independent.*
8. *The columns of A are linearly independent.*
9. *$\text{rank}(A) = n$.*
10. *A is a product of elementary matrices.*
11. *$\text{nullity}(A) = 0$.*
12. *The column vectors of A span $\mathbb{C}^n$.*

13. *The column vectors of A form a basis for $\mathbb{C}^n$.*

14. *The row vectors of A span $\mathbb{C}^n$.*

15. *The row vectors of A form a basis for $\mathbb{C}^n$.*

16. ***det** $A \neq 0$.*

17. *0 **is not an eigenvalue of** A.*

Proof. The equivalence of the statements (1) to (15) are proved in **Result 6.6**. The equivalence (1) $\Longleftrightarrow$ (16) is proved in **Part 1** of **Result 5.7**. The equivalence (1) $\Longleftrightarrow$ (17) is proved in **Part 3** of **Result** 8.1. ☐

Result 8.3. *Let A be a matrix with eigenvalue λ and corresponding eigenvector $\mathbf{x}$.*

1. *For any positive integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$.*

2. *If A is invertible, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with corresponding eigenvector $\mathbf{x}$.*

3. *If A is invertible then for any integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$.*

Proof. Given that $A\mathbf{x} = \lambda\mathbf{x}$ and $\mathbf{x} \neq \mathbf{0}$.

2. If A is invertible, then $\lambda \neq 0$. Now

$$A\mathbf{x} = \lambda\mathbf{x} \Rightarrow A^{-1}(A\mathbf{x}) = A^{-1}(\lambda\mathbf{x}) \Rightarrow \mathbf{x} = \lambda(A^{-1}\mathbf{x}) \Rightarrow A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}.$$

Thus $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with corresponding eigenvector $\mathbf{x}$.

3. If A is invertible, then $\lambda \neq 0$. For $n > 0$, Part 1 gives that λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$. For $n = 0$, clearly $\lambda^0 = 1$ is an eigenvalue of $A^0 = I$ with corresponding eigenvector $\mathbf{x}$. Now let $n = -m < 0$, where $m > 0$. We have

$$A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x} \Rightarrow (A^{-1})^m\mathbf{x} = \frac{1}{\lambda^m}\mathbf{x} \Rightarrow A^{-m}\mathbf{x} = \lambda^{-m}\mathbf{x} \Rightarrow A^n\mathbf{x} = \lambda^n\mathbf{x}.$$

Thus λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$. ☐

Result 8.4. *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be eigenvectors of a matrix A with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, respectively. Let $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m$. Then for any positive integer k,*

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \ldots + c_m\lambda_m^k\mathbf{v}_m.$$

Proof. Applying **Part 1** of **Result** 8.3, we have $A^k\mathbf{v}_i = \lambda_i^k\mathbf{v}_i$ for $i = 1, 2, \ldots, k$. Therefore

$$\begin{aligned} A^k\mathbf{x} =& A^k(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m) \\ =& c_1(A^k\mathbf{v}_1) + c_2(A^k\mathbf{v}_2) + \ldots + c_m(A^k\mathbf{v}_m) \\ =& c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \ldots + c_m\lambda_m^k\mathbf{v}_m. \quad \square \end{aligned}$$

Result 8.5. *Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be distinct eigenvalues of a matrix A with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$, respectively. Then the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is linearly independent.*

Result 8.6 (Cayley-Hamilton Theorem)**.** *Let $p(\lambda)$ be the characteristic polynomial of a matrix A. Then $p(A) = O$, the zero matrix.*

This is a beautiful and useful theorem. We will see a simple proof of this theorem in the last section of the course.

Example 8.4. *Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.*

- *Verify that characteristic polynomial of A is $\lambda^2 - 2\lambda - 3$.*
- *Verify that $A^2 - 2A - 3I = O$.*
- *Use the fact $A^2 = 2A + 3I$ to compute A^5 without computing any matrix multiplication.*

- *Argue that A is invertible using its characteristic polynomial. Since $3I = A^2 - 2A$, we have $A^{-1} = \frac{1}{3}[A - 2I]$. Verify that the inverse you found here is the same obtained by the Gauss-Jordan method.*

Similar Matrices: Let A and B be two $n \times n$ matrices. Then A is said to be **similar** to B if there is an $n \times n$ invertible matrix T such that $T^{-1}AT = B$. Then note that B is also similar to A, since $T^{-1}AT = B \Rightarrow A = TBT^{-1}$.

- If A is similar to B, we write $A \approx B$.
- If $A \approx B$, we can equivalently write that $A = TBT^{-1}$ or $AT = TB$.

Example 8.5. *Let* $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$ *and* $T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. *Then $A \approx B$ since $AT = TB$.*

Result 8.7. *Let A, B and C be $n \times n$ matrices. Then*

1. *$A \approx A$.*
2. *If $A \approx B$ then $B \approx A$.*
3. *If $A \approx B$ and $B \approx C$ then $A \approx C$.*

Proof. [Proof of **Part 3**] Since $A \approx B$ and $B \approx C$, we find invertible matrices T and S such that $A = TBT^{-1}$ and $B = SCS^{-1}$. Now TS is also invertible and $A = TBT^{-1} = A = T(SCS^{-1})T^{-1} = (TS)C(TS)^{-1}$. Hence $A \approx C$. ☐

Result 8.8. *Let A, B and T be three matrices such that $T^{-1}AT = B$, that is, $A \approx B$. Then*

1. *det $A =$ det B.*
2. *A is invertible iff B is invertible.*
3. *A and B have the same rank.*
4. *A and B have the same characteristic polynomial.*
5. *A and B have the same set of eigenvalues.*
6. *λ is an eigenvalue of B with eigenvector $\mathbf{v}$ **iff** λ is an eigenvalue of A with eigenvector $T\mathbf{v}$.*
7. *The dim(E_λ) for A is same as dim(E_λ) for B.*
8. *$A^m \approx B^m$ for all integers $m \geq 0$.*
9. *If A is invertible, then $A^m \approx B^m$ for all integers m.*

Proof. Given that $A \approx B$ so that we find invertible matrix T such that $A = TBT^{-1}$.

1. det $A = \det(TBT^{-1}) = \det(T).\det(B).\det(T^{-1}) = \det(T).\det(B).\frac{1}{\det(T)} = \det B$.
2. A is invertible $\Leftrightarrow \det A \neq 0 \Leftrightarrow \det B \neq 0 \Leftrightarrow B$ is invertible.
3. Recall that pre or post multiplication by an invertible matrix do not alter the rank. Hence A and B have the same rank.
4. $\det(A - \lambda I) = \det(TBT^{-1} - \lambda I) = \det[T(B - \lambda I)T^{-1}] = \det(B - \lambda I)$.
5. Since A and B have the same characteristic polynomial, by Part 4, they must have the same set of eigenvalues.
6. We have

$$\begin{aligned}
&\lambda \text{ is an eigenvalue of } B \text{ with corresponding eigenvector } \mathbf{v} \\
\Leftrightarrow &B\mathbf{v} = \lambda\mathbf{v} \\
\Leftrightarrow &(T^{-1}AT)\mathbf{v} = \lambda\mathbf{v} \\
\Leftrightarrow &A(T\mathbf{v}) = \lambda(T\mathbf{v}) \\
\Leftrightarrow &\lambda \text{ is an eigenvalue of } A \text{ with corresponding eigenvector } T\mathbf{v}.
\end{aligned}$$

7. Let $\dim(E_\lambda) = k$ for B and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis of E_λ for B. We claim that $\{T\mathbf{v}_1, \ldots, T\mathbf{v}_k\}$ is a basis of E_λ for A. We have for $a_1, \ldots, a_k \in \mathbb{C}$,

$$\begin{aligned} a_1(T\mathbf{v}_1) + \ldots + a_k(T\mathbf{v}_k) = 0 \Rightarrow & T(a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k) = 0 \\ \Rightarrow & a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k = 0, \text{ since } T \text{ is invertible} \\ \Rightarrow & a_1 = \ldots = a_k = 0, \text{ since } S \text{ is linearly independent} \\ \Rightarrow & \{T\mathbf{v}_1, \ldots, T\mathbf{v}_k\} \text{ is linearly independent.} \end{aligned}$$

Again if $\mathbf{v}$ is an eigenvector of A, then by Part 6, $\mathbf{v} = T\mathbf{u}$ for some eigenvector $\mathbf{u}$ of B. Now $\mathbf{u} = a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k$ for some $a_1, \ldots, a_k \in \mathbb{C}$ and hence

$$\mathbf{v} = T\mathbf{u} = T(a_1\mathbf{v}_1 + \ldots + a_k\mathbf{v}_k) = a_1(T\mathbf{v}_1) + \ldots + a_k(T\mathbf{v}_k).$$

Thus $\{T\mathbf{v}_1, \ldots, T\mathbf{v}_k\}$ is linearly independent and spans E_λ for A. Therefore $\{T\mathbf{v}_1, \ldots, T\mathbf{v}_k\}$ is a basis of E_λ for A, and hence the $\dim(E_\lambda)$ for A is same as $\dim(E_\lambda)$ for B.

8. Let $m \geq 0$. We use induction to show that $A^m = TB^mT^{-1}$. Clearly $A^m = TB^mT^{-1}$ is true for $m = 0$. Now assume that $A^k = TB^kT^{-1}$ for $k \geq 0$. Now we have

$$A^{k+1} = A^k.A = (TB^kT^{-1}).(TBT^{-1}) = TB^k.BT^{-1} = TB^{k+1}T^{-1}.$$

Thus whenever $A^k = TB^kT^{-1}$, we find $A^{k+1} = TB^{k+1}T^{-1}$. Hence by principle of mathematical induction, $A^m = TB^mT^{-1}$ for all $m \geq 0$. That is, $A^m \approx B^m$ for all integers $m \geq 0$.

9. Let $m = -n$ so that $n > 0$. If A is invertible, then

$$\begin{aligned} A = TBT^{-1} \Rightarrow A^n = TB^nT^{-1} \Rightarrow (A^n)^{-1} = (TB^nT^{-1})^{-1} \Rightarrow & A^{-n} = (T^{-1})^{-1}(B^n)^{-1}T^{-1} = TB^{-n}T^{-1} \\ \Rightarrow & A^m = TB^mT^{-1}. \end{aligned}$$

Thus $A^m = TB^mT^{-1}$, and hence $A^m \approx B^m$ for all integers m. □

Diagonalizable Matrix: A matrix A is said to be **diagonalizable** if there is a diagonal matrix D such that $A \approx D$, that is, if there is an invertible matrix T and a diagonal matrix D such that $AT = TD$.

Example 8.6. *The matrix* $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ *is diagonalizable, since if* $D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$ *and* $T = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix}$ *then* $AT = TD$.

Result 8.9. *Let A be an $n \times n$ matrix. Then A is diagonalizable iff A has n linearly independent eigenvectors.*

- Let A be an $n \times n$ matrix. Then there exists an invertible matrix T and a diagonal matrix D satisfying $T^{-1}AT = D$ **iff** the columns of T are n linearly independent eigenvectors of A and the diagonal entries of D are the eigenvalues of A corresponding to the columns (eigenvectors of A) of T in the same order.

Example 8.7. *Check for the diagonalizablity of the following matrices. If they are diagonalizable, find invertible matrices T that diagonalizes them:*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}.$$

Result 8.10. *If A is an $n \times n$ matrix with n distinct eigenvalues then A is diagonalizable.*

Result 8.11. *Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of a matrix A. If $\mathcal{B}_i$ is a basis for the eigenspace E_{λ_i}, then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$ is a linearly independent set.*

Definition 8.4. *Let λ be an eigenvalue of a matrix A.*

- *The **algebraic multiplicity** of λ is the multiplicity of λ as a root of the characteristic polynomial of A.*
- *The **geometric multiplicity** of λ is the dimension of E_λ.*

Result 8.12. *The geometric multiplicity of each eigenvalue of a matrix is less than or equal to its algebraic multiplicity.*

Result 8.13 (The Diagonalization Theorem). *Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then the following statements are equivalent:*

1. *A is diagonalizable.*
2. *The union $\mathcal{B}$ of the bases of the eigenspaces of A contains n vectors.*
3. *The algebraic multiplicity of each eigenvalue equals its geometric multiplicity.*
4. *A has n linearly independent eigenvectors.*

Practice Problems Set 8

1. Let A be an $n \times n$ matrix and S be an $n \times n$ invertible matrix. Show that the eigenvalues of A and $S^{-1}AS$ are the same. Are their corresponding eigenvectors same?

2. Find the eigenvalues and the corresponding eigenvectors of the following matrices:
$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & i \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 5 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -3 & -3 \\ 2 & 4 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ 3 & 2 & -2 \\ 0 & 3 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

3. Show that the following matrices A, B and C are diagonalizable. Also, find invertible matrices S_1, S_2 and S_3 such that $S_1^{-1}AS_1, S_2^{-1}BS_2$ and $S_3^{-1}CS_3$ are all diagonal matrices.
$$A = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & -2 \\ 0 & 3 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

4. Prove that the following matrices A and B are similar by showing that they are similar to the same diagonal matrix. Also, find an invertible matrix P such that $P^{-1}AP = B$.
   (a) $A = \begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$
   (b) $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 3 & 2 & -5 \\ 1 & 2 & -1 \\ 2 & 2 & -4 \end{bmatrix}.$

5. Let A and B be invertible matrices of the same size. Show that the matrices AB and BA are similar.

6. Let A and B be two similar matrices and let λ be an eigenvalue of A and B. Prove that the algebraic (geometric) multiplicity of λ in A is equal to the algebraic (geometric) multiplicity of λ in B.

7. Show that the matrices $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & i \\ i & 0 \end{bmatrix}$ are not diagonalizable.

8. Let A be a symmetric matrix. Show that the eigenvalues of A are real numbers.

9. Let A be an skew-symmetric matrix of odd order. Prove that 0 is an eigenvalue of A.

10. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Show that $\det(A) = \lambda_1.\lambda_2.\ldots.\lambda_n$ and $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$. Further, show that A is invertible if and only if none of the eigenvalues of A are zero.

11. Let A and B be two $n \times n$ matrices. Prove that the sum of all the eigenvalues of $A+B$ is the sum of all the eigenvalues of A and B individually. Also, prove that the product of all the eigenvalues of AB is the product of all the eigenvalues of A and B individually.

12. Prove or disprove: If $A = \begin{bmatrix} 5 & 4 & 14 & 0 \\ 4 & 13 & 14 & 0 \\ 14 & 14 & 49 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ then there exists a symmetric matrix B such that $A = B^{52}$.

13. Let A and B be two $n \times n$ matrices with eigenvalues λ and μ, respectively.

    (a) Give an example to show that $\lambda + \mu$ need not be an eigenvalue of $A + B$.

    (b) Give an example to show that $\lambda\mu$ need not be an eigenvalue of AB.

    (c) Suppose that λ and μ correspond to the same eigenvector $\mathbf{x}$. Show that $\lambda + \mu$ is an eigenvalue of $A + B$ and $\lambda\mu$ is an eigenvalue of AB.

14. Let A be an $n \times n$ matrix and let c $(\neq 0)$ be a constant. Show that λ is an eigenvalue of A if and only if $c\lambda$ is an eigenvalue of cA.

15. Let A be an $n \times n$ matrix. Show that A and A^t have the same eigenvalues. Are their corresponding eigenspaces same?

16. Let A be a $n \times n$ matrix. Show that the eigenvalues of A are either real numbers or complex conjugates occurring in pairs. Further, show that if the order of A is odd then A has at least one real eigenvalue.

17. Let A be an $n \times n$ complex matrix. Show that

    (a) if A is Hermitian (*i.e.*, $A^* = \overline{A}^t = A$) then all eigenvalues of A are real numbers; and

    (b) if A is skew-Hermitian (*i.e.*, $A^* = \overline{A}^t = -A$) and $A \neq \mathbf{O}$ then all the eigenvalues of A are purely imaginary numbers.

18. Let A be an $n \times n$ matrix. Show that

    (a) if A is idempotent (*i.e.*, $A^2 = A$) then all the eigenvalues of A are either 0 or 1; and

    (b) if A is nilpotent (*i.e.*, $A^m = \mathbf{0}$ for some $m \geq 1$) then all the eigenvalues of A are 0.

19. Let A be an $n \times n$ matrix. Prove that

    (a) if λ $(\neq 0)$ is an eigenvalue of A then $\frac{1}{\lambda}\det(A)$ is an eigenvalue of $\mathrm{adj}(A)$; and

    (b) if $\mathbf{v}$ is an eigenvector of A then $\mathbf{v}$ is also an eigenvector of $\mathrm{adj}(A)$.

20. Let A and B be two $n \times n$ matrices and let A be invertible. Show that the matrices BA^{-1} and $A^{-1}B$ have the same eigenvalues.

21. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (which are not necessarily real). Denote $\lambda_k = x_k + iy_k$ for $k = 1, 2, \ldots, n$. Show that

    (a) $y_1 + y_2 + \ldots + y_n = 0$;

    (b) $x_1y_1 + x_2y_2 + \ldots + x_ny_n = 0$; and

    (c) $\mathrm{tr}(A^2) = (x_1^2 + x_2^2 + \ldots + x_n^2) - (y_1^2 + y_2^2 + \ldots + y_n^2)$.

22. For each of the following matrices, find the eigenvalues and the corresponding eigenspaces over $\mathbb{C}$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}, \begin{bmatrix} i & 1+i \\ -1+i & i \end{bmatrix}, \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ and } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

23. Using **Cayley Hamilton Theorem**, find the inverse of the following matrices, whenever they exist:

$$\begin{bmatrix} 1 & i \\ i & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

24. Find all real values of k for which the following matrices are diagonalizable.

$$\begin{bmatrix} 1 & 1 \\ 0 & k \end{bmatrix}, \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} k & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & k \\ 1 & 1 & k \\ 1 & 1 & k \end{bmatrix}.$$

25. Prove that if A and B are similar matrices then $\mathrm{tr}(A) = \mathrm{tr}(B)$.

26. For any real numbers a, b and c, show that the matrices

$$A = \begin{bmatrix} b & c & a \\ c & a & b \\ a & b & c \end{bmatrix}, \quad B = \begin{bmatrix} c & a & b \\ a & b & c \\ b & c & a \end{bmatrix} \text{ and } C = \begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix}$$

are similar to each other. Moreover, if $BC = CB$ then show that A has two zero eigenvalues.

27. Prove that if the matrix A is similar to B, then A^t is similar to B^t.

28. Prove that if the matrix A is diagonalizable, then A^t is also diagonalizable.

29. Let A be an invertible matrix. Prove that if A is diagonalizable, then A^{-1} is also diagonalizable.

30. Let A be a diagonalizable matrix and let $A^m = \mathbf{O}$ for some $m \geq 1$. Show that $A = \mathbf{O}$.

31. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   (a) Prove that A is diagonalizable if $(a-d)^2 + 4bc > 0$.

   (b) Find two examples to demonstrate that if $(a-d)^2 + 4bc = 0$, then A may or may not be diagonalizable.

32. Let A be a 2×2 matrix with eigenvectors $v_1 = [1, -1]^t$ and $v_2 = [1, 1]^t$ and corresponding eigenvalues $\frac{1}{2}$ and 2, respectively. Find $A^{10}\mathbf{x}$ and $A^k\mathbf{x}$ for $k \geq 1$, where $\mathbf{x} = [5, 1]^t$. What happens if k becomes large (i.e., $k \to \infty$.)

   Definition: Let $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ be a polynomial. Then the companion matrix of $p(x)$ is the $n \times n$ matrix

$$C(p) = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \ldots & -a_1 & -a_0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.$$

33. Show that the companion matrix $C(p)$ of $p(x) = x^2 + ax + b$ has characteristic polynomial $\lambda^2 + a\lambda + b$. Also, show that if λ is an eigenvalue of this companion matrix then $[\lambda, 1]^t$ is an eigenvector of $C(p)$ corresponding to the eigenvalue λ.

34. Show that the companion matrix $C(p)$ of $p(x) = x^3 + ax^2 + bx + c$ has characteristic polynomial $-(\lambda^3 + a\lambda^2 + b\lambda + c)$. Further, show that if λ is an eigenvalue of this companion matrix then $[\lambda^2, \lambda, 1]^t$ is an eigenvector of $C(p)$ corresponding to λ.

35. Use mathematical induction to show that for $n \geq 2$, the companion matrix $C(p)$ of $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ has characteristic polynomial $(-1)^n p(\lambda)$. Further, show that if λ is an eigenvalue of this companion matrix then $[\lambda^{n-1}, \lambda^{n-2}, \ldots, \lambda, 1]^t$ is an eigenvector of $C(p)$ corresponding to λ.

36. Let A be an $n \times n$ non-singular matrix and let $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ be its characteristic polynomial. Show that $A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 I)$.

37. Let A and B be two 2×2 real matrices for which $\det(A) = \det(B)$ and $\mathrm{tr}(A) = \mathrm{tr}(B)$.

   (a) Do A and B have the same set of eigenvalues?

   (b) Give examples to show that the matrices A and B need not be similar.

Hints to Practice Problems Set 8

1. $\det(S^{-1}AS - \lambda I) = \det[S^{-1}(A - \lambda I)S]$. Take the counterexample $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ for the second part.

2. For each of the matrices, the ordered pairs given below consist of an eigenvalue and a basis for the corresponding eigenspace:
1st matrix: $(0, \{[0,1]^t\})$, $(1, \{[1,0]^t\})$, **2nd matrix:** $(1+\sqrt{2}, \{[\sqrt{2}, 1-i]^t\})$, $(1-\sqrt{2}, \{[-\sqrt{2}, 1-i]^t\})$,
3rd matrix: $(2+\sqrt{3}, \{[2, 1+\sqrt{3}]^t\})$, $(2-\sqrt{3}, \{[2, 1-\sqrt{3}]^t\})$, **5th matrix:** $(1, \{[-3,5]^t\})$, $(4, \{[0,1]^t\})$,
4th matrix: $(1+\sqrt{2i}, \{[1, -i\sqrt{2i}]^t\})$, $(1-\sqrt{2i}, \{[1, i\sqrt{2i}]^t\})$,
6th matrix: $(\frac{5}{2} + \frac{\sqrt{33}}{2}, \{[3-\sqrt{33}, -6]^t\})$, $(\frac{5}{2} - \frac{\sqrt{33}}{2}, \{[3+\sqrt{33}, -6]^t\})$,
7th matrix: $(0, \{[0,1,-1]^t\})$, $(2, \{[1,-2,3]^t\})$,
8th matrix: $(1, \{[0,0,1]^t, [1,1,0]^t\})$, $(-1, \{[1,-1,0]^t\})$,
9th matrix: $(3, \{[1,1,1]^t\})$, $(i\sqrt{2}, \{[\frac{2(2-i\sqrt{2})}{9}, \frac{i\sqrt{2}}{3}, 1]^t\})$, $(-i\sqrt{2}, \{[\frac{2(2+i\sqrt{2})}{9}, -\frac{i\sqrt{2}}{3}, 1]^t\})$,
10th matrix: $(1, \{[1,-2,1]^t\})$, $(\frac{3+\sqrt{5}}{2}, \{[-1-\sqrt{5}, -1-\sqrt{5}, 1-\sqrt{5}]^t\})$, $(\frac{3-\sqrt{5}}{2}, \{[-1+\sqrt{5}, -1+\sqrt{5}, 1+\sqrt{5}]^t\})$,
11th matrix: $(0, \{[-1,1,0]^t, [-1,0,1]^t\})$, $(3, \{[1,1,1]^t\})$.

3. A has 3 distinct eigenvalues and $S_1 = \begin{bmatrix} 2 & \frac{-8-2i\sqrt{11}}{9} & \frac{8-2i\sqrt{11}}{9} \\ 2 & \frac{1}{3}(5-i\sqrt{11}) & -\frac{1}{3}(5+i\sqrt{11}) \\ 3 & -1-i\sqrt{11} & 1-i\sqrt{11} \end{bmatrix}$.
B has 3 distinct eigenvalues and $S_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.
C has eigenvalues 0 and 3, but E_0 has dimension 2, and $S_3 = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

4. (a) Both the matrices are similar to $D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$. If $U^{-1}AU = D$ and $V^{-1}AV = D$ then take $P = UV^{-1}$.
(b) Both the matrices are similar to $\text{diag}[2, -2, 1]$.

5. $A^{-1}(AB)A = BA$.

6. $B = P^{-1}AP \Rightarrow A$ and B have the same characteristic polynomial. Also any eigenvector of B must be of the form $P^{-1}\mathbf{v}$ for some eigenvector $\mathbf{v}$ of A.

7. Find the algebraic and geometric multiplicities of the eigenvalues.

8. $A\mathbf{u} = \lambda\mathbf{u} \Rightarrow \overline{\mathbf{u}}^t A = \overline{\lambda}\overline{\mathbf{u}}^t \Rightarrow \overline{\mathbf{u}}^t A\mathbf{u} = \overline{\lambda}\overline{\mathbf{u}}^t\mathbf{u}$.

9. $A^t = -A \Rightarrow \det(A) = (-1)^n \det(A)$.

10. $\det(A - \lambda I) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$. Compare the constant term and the coefficient of λ^{n-1} on both the sides.

11. $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$ and $\det(AB) = \det(A)\det(B)$.

12. -1 is an eigenvalue of A. So, $A = B^{52} \Rightarrow \lambda^{52} = -1$ for some eigenvalue λ of B.

13. (a) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, (b) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}$, (c) Easy.

14. Easy.

15. $\det(A - \lambda I) = \det(A^t - \lambda I)$. Take the counterexample $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ for the second part.

16. Complex roots of a polynomial equation occur in pairs.

17. Similar to Problem 8.

18. (a) $A\mathbf{x} = \lambda\mathbf{x}, \mathbf{x} \neq \mathbf{0} \Rightarrow \lambda^2\mathbf{x} = \lambda\mathbf{x}$. (b) $A\mathbf{x} = \lambda\mathbf{x} \Rightarrow \mathbf{0} = \lambda^m\mathbf{x}$.

19. (a) $A\mathbf{v} = \lambda\mathbf{v}$ and $\text{adj}(A)A = \det(A)I$.
(b) If $A\mathbf{v} = \lambda\mathbf{v}$ and $\lambda \neq 0$, then use Part (a). Now let $A\mathbf{v} = \mathbf{0}$ *i.e.,* $\mathbf{v} \in \text{null}(A)$. If $\text{rank}(A) \leq n-2$ then $\text{adj}(A) = \mathbf{O}$. If $\text{rank}(A) = n-1$ then $\text{null}(A) = \{\alpha\mathbf{v} : \alpha \in \mathbb{R}\}$ and also $A.\text{adj}(A)\mathbf{v} = \det(A)\mathbf{v} = \mathbf{0} \Rightarrow \text{adj}(A)\mathbf{v} \in \text{null}(A)$.

20. $A^{-1}(BA^{-1})A = A^{-1}B$.

21. $\text{tr}(A)$ and $\text{tr}(A^2)$ are real numbers. Also $\text{tr}(A^2) = \sum\limits_{k=1}^{n} \lambda_k^2$.

22. **1st matrix:** $E_0 = \{[0, \alpha]^t : \alpha \in \mathbb{C}\},\ E_1 = \{[\alpha, 0]^t : \alpha \in \mathbb{C}\}$,
2nd matrix: $E_{1+\sqrt{2}} = \{\alpha[\sqrt{2},\ 1-i]^t : \alpha \in \mathbb{C}\},\ E_{1-\sqrt{2}} = \{\alpha[-\sqrt{2},\ 1-i]^t : \alpha \in \mathbb{C}\}$,
3rd matrix: $E_{i+i\sqrt{2}} = \{\alpha[i\sqrt{2},\ i-1]^t : \alpha \in \mathbb{C}\},\ E_{i-i\sqrt{2}} = \{\alpha[-i\sqrt{2},\ i-1]^t : \alpha \in \mathbb{C}\}$,
4th matrix: If $\theta \neq n\pi$ then $E_{e^{i\theta}} = \{\alpha[i, 1]^t : \alpha \in \mathbb{C}\}$ and $E_{e^{-i\theta}} = \{\alpha[-i, 1]^t : \alpha \in \mathbb{C}\}$. If $\theta = 2n\pi$ then $E_1 = \mathbb{R}^2$. If $\theta = (2n+1)\pi$ then $E_{-1} = \mathbb{R}^2$.
5th matrix: Consider each of the cases $\theta \neq n\pi, n\pi + \frac{\pi}{2},\ \theta = 2n\pi, \theta = (2n+1)\pi, \theta = 2n\pi + \frac{\pi}{2}$ and $\theta = (2n+1)\pi + \frac{\pi}{2}$ one by one.

23. If the given matrices are A, B, C, D (in order), then $A^{-1} = \frac{1}{3}(3I - A),\ B^{-1} = \frac{1}{3}(B - 2I),\ C^{-1} = \frac{1}{8}(-C^2 + 5C - 2I)$ and $D^{-1} = -D^2 + D + I$.

24. **1st matrix:** $k \neq 1$, **2nd matrix:** $k = 0$, **3rd matrix:** $k \in \mathbb{R}$, **4th matrix:** $k = 0$, **5th matrix:** $k \neq -2$.

25. A and B have the same set of eigenvalues.

26. Take $P = [\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2]$ and $Q = [\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1]$. Then $AP = PB$ and $AQ = QC$. Also $BC = CB \Rightarrow a^3 + b^3 + c^3 = 3abc$. Use this equation in computing $\det(A - \lambda I)$.

27. Take transpose of both sides of $P^{-1}AP = B$.

28. Use Problem 27.

29. Take inverse of both sides of $P^{-1}AP = B$.

30. $P^{-1}AP = D \Rightarrow P^{-1}A^mP = D^m$.

31. (a) When does the equation $\det(A - \lambda I) = 0$ have distinct real roots? (b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$.

32. $\mathbf{x} = 2\mathbf{v}_1 + 3\mathbf{v}_2,\quad A^k\mathbf{x} = [2^{1-k} + 3.2^k,\ 3.2^k - 2^{1-k}]^t$.

33. Easy.

34. Easy.

35. Expand $\det(C(p) - \lambda I)$ through the last column and use induction.

36. $a_0 = \det(A) \neq 0$. Use Cayley-Hamilton Theorem.

37. (a) $xy = ab, x + y = a + b \Rightarrow \{x, y\} = \{a, b\}$. (b) $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

---

9 Inner Product

Let $\mathbb{F}$ denote $\mathbb{C}$ or $\mathbb{R}$. Let $\mathbf{u} = [u_1, u_2, \dots, u_n]^t$, $\mathbf{v} = [v_1, v_2, \dots, v_n]^t \in \mathbb{F}^n$.

- The **dot product** or the standard **inner product** $\mathbf{u}.\mathbf{v}$ of $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$\mathbf{u}.\mathbf{v} = \overline{u_1}v_1 + \overline{u_2}v_2 + \ldots + \overline{u_n}v_n.$$

  Notice that

$$\mathbf{u}\cdot\mathbf{v} = \begin{bmatrix} \overline{u_1} & \cdots & \overline{u_n} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \overline{u_1}v_1 + \overline{u_2}v_2 + \ldots + \overline{u_n}v_n = \sum \overline{u_i}v_i = \mathbf{u}^*\mathbf{v}.$$

- Sometimes, the notation $\langle \mathbf{u}, \mathbf{v} \rangle$ is also used to denote the inner product of $\mathbf{u}$ and $\mathbf{v}$.

- The **length** or **norm** $\|\mathbf{u}\|$ of $\mathbf{u}$ is defined by

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}.\mathbf{u}} = \sqrt{|u_1|^2 + |u_2|^2 + \ldots + |u_n|^2}.$$

- Observe that $\|\alpha\mathbf{u}\| = |\alpha|\,\|\mathbf{u}\|$.

- The vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthogonal** if $\mathbf{u}.\mathbf{v} = 0$.

- A vector $\mathbf{u}$ is called an **unit** vector if $\|\mathbf{u}\| = 1$.

Result 9.1. *Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{F}^n$ and $c \in \mathbb{F}$. Then*

1. $\mathbf{u}\cdot\mathbf{v} = \overline{\mathbf{v}\cdot\mathbf{u}}$;
2. $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$;
3. $(c\mathbf{u})\cdot\mathbf{v} = \overline{c}(\mathbf{u}\cdot\mathbf{v})$;
4. $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$;
5. $\mathbf{u}.\mathbf{u} \geq 0$ *and* $\mathbf{u}.\mathbf{u} = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$.

Proof. Let $\mathbf{u} = [u_1, u_2, \dots, u_n]^t$, $\mathbf{v} = [v_1, v_2, \dots, v_n]^t$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]^t$.

1. $\mathbf{u}\cdot\mathbf{v} = \sum \overline{u_i}v_i = \overline{\sum u_i\overline{v_i}} = \overline{\mathbf{v}\cdot\mathbf{u}}$.
2. $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = \sum \overline{u_i}(v_i + w_i) = \sum \overline{u_i}v_i + \sum \overline{u_i}w_i = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$.
3. $(c\mathbf{u})\cdot\mathbf{v} = \sum \overline{cu_i}v_i = \overline{c}\sum \overline{u_i}v_i = \overline{c}(\mathbf{u}\cdot\mathbf{v})$.
4. $\mathbf{u}\cdot(c\mathbf{v}) = \sum \overline{u_i}(cv_i) = c\sum \overline{u_i}v_i = c(\mathbf{u}\cdot\mathbf{v})$.
5. $\mathbf{u}.\mathbf{u} = \sum \overline{u_i}u_i = \sum |u_i|^2 \geq 0$ and so $\mathbf{u}.\mathbf{u} = 0 \Leftrightarrow \sum |u_i|^2 = 0 \Leftrightarrow |u_i|^2 = 0$ for all $i \Leftrightarrow u_i = 0$ for all $i \Leftrightarrow \mathbf{u} = \mathbf{0}$. ☐

Orthogonal Set: A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in $\mathbb{F}^n$ is said to be an **orthogonal set** if all pairs of distinct vectors in the set are orthogonal, that is, if

$$\mathbf{v}_i.\mathbf{v}_j = 0 \text{ whenever } i \neq j \quad \text{for } i, j = 1, 2, \dots, k.$$

- The set $\{[2, 1, -1]^t, [0, 1, 1]^t, [1, -1, 1]^t\}$ is an orthogonal set in $\mathbb{R}^3$.

- An orthogonal set can contain a zero vector.

Result 9.2. *If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is an orthogonal set of non-zero vectors, then S is linearly independent.*

Result 9.3. *Let A be a Hermitian (or real symmetric) matrix. Then*

- *all eigenvalues of A are real.*

- *eigenvectors corresponding to distinct eigenvalues of A are orthogonal.*

Proof. Suppose A is a Hermitian matrix, i.e., $A^* = A$. Let λ be an eigenvalue of A. Then for an eigenvector $\mathbf{v}$ of A, we have

$$\begin{aligned}\lambda\|\mathbf{v}\|^2 = \lambda(\mathbf{v}\cdot\mathbf{v}) = \mathbf{v}\cdot(\lambda\mathbf{v}) = \mathbf{v}\cdot A\mathbf{v} = \mathbf{v}^*(A\mathbf{v}) = (\overline{\mathbf{v}}^T A)\mathbf{v} =& (A^T\overline{\mathbf{v}})^T\mathbf{v}\\ =& (\overline{A^*\mathbf{v}})^T\mathbf{v}\\ =& (\overline{A\mathbf{v}})^T\mathbf{v}\\ =& (A\mathbf{v})^*\mathbf{v}\\ =& (\lambda\mathbf{v})\cdot\mathbf{v}\\ =& \overline{\lambda}(\mathbf{v}\cdot\mathbf{v})\\ =& \overline{\lambda}\|\mathbf{v}\|^2.\end{aligned}$$

Since $\mathbf{v}\neq 0$, we have $\|\mathbf{v}\|^2 \neq 0$. Thus, $\lambda = \overline{\lambda}$, i.e., λ is real.

Next, suppose λ, μ are two distinct eigenvalues of A, and $\mathbf{u}$ and $\mathbf{v}$ are corresponding eigenvectors. Then

$$\mu(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(\mu\mathbf{v}) = \mathbf{u}\cdot(A\mathbf{v}) = (\mathbf{u}^*A)\mathbf{v} = (A^*\mathbf{u})^*\mathbf{v} = (A\mathbf{u})^*\mathbf{v} = (\lambda\mathbf{u})\cdot\mathbf{v} = \overline{\lambda}(\mathbf{u}\cdot\mathbf{v}).$$

Since λ is real we get $(\lambda-\mu)(\mathbf{u}\cdot\mathbf{v}) = 0$, i.e., $\mathbf{u}\cdot\mathbf{v} = 0$. ☐

Orthogonal Basis: An **orthogonal basis** for a subspace W is a basis for W that is an orthogonal set.

Example 9.1. *Notice that* $\mathbf{u} = [2,1,-1]^t, \mathbf{v} = [0,1,1]^t, \mathbf{w} = [1,-1,1]^t\}$ *form an orthogonal basis for* $\mathbb{R}^3$. *Take* $\mathbf{x} = [1,1,1]^t \in \mathbb{R}^3$. *Find* a, b, c *such that* $[1,1,1]^t = a[2,1,-1]^t + b[0,1,1]^t + c[1,-1,1]^t$.

Solution. We have $a = \frac{\mathbf{u}\cdot\mathbf{x}}{\|\mathbf{u}\|^2} = \frac{1}{3}$, $b = \frac{\mathbf{v}\cdot\mathbf{x}}{\|\mathbf{v}\|^2} = 1$ and $c = \frac{\mathbf{w}\cdot\mathbf{x}}{\|\mathbf{w}\|^2} = \frac{1}{3}$. ☐

Example 9.2. $\mathbf{u} = \begin{bmatrix}2\\1\\-1\end{bmatrix}, \mathbf{v} = \begin{bmatrix}0\\1\\1\end{bmatrix}, \mathbf{w} = \begin{bmatrix}1\\-1\\1\end{bmatrix}$ *form an orthogonal basis for* $\mathbb{C}^3$. *Take* $\mathbf{x} = \begin{bmatrix}1\\1\\i\end{bmatrix} \in \mathbb{C}^3$. *Find* a, b, c *such that* $\mathbf{x} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$.

Solution. We have $a = \frac{\mathbf{u}\cdot\mathbf{x}}{\|\mathbf{u}\|^2} = \frac{3-i}{6}$, $b = \frac{\mathbf{v}\cdot\mathbf{x}}{\|\mathbf{v}\|^2} = \frac{1+i}{2}$ and $c = \frac{\mathbf{w}\cdot\mathbf{x}}{\|\mathbf{w}\|^2} = \frac{i}{3}$. ☐

Result 9.4. *Let* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *be an orthogonal basis for a subspace* W *and let* $\mathbf{w} \in W$. *Then the unique scalars* $c_1, c_2, \ldots, c_k$ *such that* $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$ *are given by*

$$c_i = \frac{\mathbf{v}_i.\mathbf{w}}{\mathbf{v}_i.\mathbf{v}_i} \quad \textit{for} \;\; i = 1, 2, \ldots, k.$$

Orthonormal Set: A set of vectors is said to be an **orthonormal set** if it is an orthogonal set of unit vectors.

Orthonormal Basis: An **orthonormal basis** of a subspace W is a basis of W that is an orthonormal set.

Result 9.5. *Let* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ *be an orthonormal basis for a subspace* W *and let* $\mathbf{w} \in W$. *Then*

$$\mathbf{w} = (\mathbf{u}_1.\mathbf{w})\mathbf{u}_1 + (\mathbf{u}_2.\mathbf{w})\mathbf{u}_2 + \ldots + (\mathbf{u}_k.\mathbf{w})\mathbf{u}_k,$$

and this representation is unique.

Orthogonal Complement: Let W be a subset of $\mathbb{F}^3$.

- A vector $\mathbf{v}$ is said to be **orthogonal** to W if $\mathbf{v}$ is orthogonal to every vector in W.
- The **orthogonal complement** of W, denoted $W^\perp$, is defined as

$$W^\perp = \{\mathbf{v} \in \mathbb{F}^n : \mathbf{v}.\mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

- The vector $\mathbf{0}$ is orthogonal to W, for any subset W.
- In $\mathbb{R}^3$, take $W = \{\mathbf{e}_1\}$. Then $W^\perp = \;\; yz$-plane.
- In $\mathbb{R}^3$, take $W = \{[1,1,1]^t\}$. Then $W^\perp = \{[x,y,z]^t : \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 0\} = \{[x,y,z]^t : x+y+z = 0\}$.
- In $\mathbb{R}^3$, take $W = \{[1,1,1]^t, [1,2,3]^t\}$. Then $W^\perp = \{[x,y,z]^t : \begin{bmatrix}1 & 1 & 1\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 0\}$.
- In $\mathbb{C}^3$, take $W = \text{span}\{[1,1,i]^t, [1,2i,3]^t\}$. Then $W^\perp = \{\begin{bmatrix}x\\y\\z\end{bmatrix} : \begin{bmatrix}1 & 1 & -i\\1 & -2i & 3\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 0\}$.

Result 9.6. *Let W be a subspace of $\mathbb{F}^n$. Then*

1. *$W^\perp$ is also a subspace of $\mathbb{F}^n$.*
2. *$W \cap W^\perp = \{\mathbf{0}\}$.*
3. *If $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ are linearly independent sets in W and $W^\perp$, respectively, then the union $\{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots$ is also linearly independent.*
4. *If $B = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is a basis for W, then $\mathbf{v} \in W^\perp$ if and only if $\mathbf{v}.\mathbf{w}_i = 0$ for all $i = 1, 2, \ldots, k$.*
5. *Let $A = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k]$, where $\mathbf{w}_i$'s are as in Part 4. Let T be an invertible matrix such that $TA = \left[\frac{I_k}{\mathbf{O}}\right]$. Then the last $n-k$ columns of T^* will form a basis for $W^\perp$.*

Proof.

1. Clearly $W^\perp$ is non-empty since $\mathbf{0} \in W^\perp$. Let $\mathbf{x}, \mathbf{y} \in W^\perp$ and $a, b \in \mathbb{F}$. Then $\mathbf{x}.\mathbf{w} = 0, \mathbf{y}.\mathbf{w} = 0$ for all $\mathbf{w} \in W$. Now $(a\mathbf{x} + b\mathbf{y}).\mathbf{w} = a(\mathbf{x}.\mathbf{w}) + b(\mathbf{y}.\mathbf{w}) = 0$, and so $a\mathbf{x} + b\mathbf{y} \in W^\perp$. Hence $W^\perp$ is a subspace of $\mathbb{F}^n$.
2. If $\mathbf{v} \in W \cap W^\perp$, then $\mathbf{v} \cdot \mathbf{v} = 0 \Rightarrow \mathbf{v} = \mathbf{0}$. Hence $W \cap W^\perp = \{\mathbf{0}\}$.
3. Let $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s \in \mathbb{F}$ be such that $\alpha_1\mathbf{u}_1 + \ldots + \alpha_r\mathbf{u}_r + \beta_1\mathbf{v}_1 + \ldots + \beta_s\mathbf{v}_s = \mathbf{0}$. Then

$$\begin{aligned}
&\alpha_1\mathbf{u}_1 + \ldots + \alpha_r\mathbf{u}_r = -(\beta_1\mathbf{v}_1 + \ldots + \beta_s\mathbf{v}_s) \in W \cap W^\perp = \{\mathbf{0}\} \\
\Rightarrow &\alpha_1\mathbf{u}_1 + \ldots + \alpha_r\mathbf{u}_r = \mathbf{0} = \beta_1\mathbf{v}_1 + \ldots + \beta_s\mathbf{v}_s \\
\Rightarrow &\alpha_1 = 0 = \ldots = \alpha_r, \beta_1 = 0 = \ldots = \beta_s.
\end{aligned}$$

   Hence $\{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots, \mathbf{v}_s\}$ is a linearly independent set.

4. Suppose that $\mathbf{v} \cdot \mathbf{w}_i = 0$, for each i. Let $\mathbf{w} \in W$. Then $\mathbf{w} = \sum\limits_{i=1}^{k} a_i\mathbf{w}_i$ for some $a_1, \ldots, a_k \in \mathbb{F}$.

   Now $\mathbf{v} \cdot \mathbf{w} = \sum a_i(\mathbf{v} \cdot \mathbf{w}_i) = 0$ and so $\mathbf{v} \in W^\perp$.

   Conversely, if $\mathbf{v} \in W^\perp$, that is, if $\mathbf{v} \cdot \mathbf{w} = 0$ for each $\mathbf{w} \in W$, then in particular, $\mathbf{v} \cdot \mathbf{w}_i = 0$ for each i.

5. Note that rank$(A) = k$ and so nullity$(A) = 0$. Also $TA = \left[\frac{I_k}{\mathbf{O}}\right] = RREF(A)$. Consider the last $n-k$ rows of T. These rows are linearly independent, as they are part of an invertible matrix. Notice that the matrix product of such a row with the vectors $\mathbf{w}_i$ is 0. Therefore by Part 4, the last $n-k$ columns of T^* will belong to $W^\perp$. Now by Part 3, $\dim(W^\perp) = n-k$. Thus the last $n-k$ columns (linearly independent) of T^* form a basis for $W^\perp$. □

Corollary 9.1. *Let W be a subspace of $\mathbb{F}^n$. Then dim W + dim $W^\perp = n$.*

Proof. Follows from **Part 5** of **Result** 9.6. □

Result 9.7 (Orthogonal Decomposition Theorem). *Let W be a subspace of $\mathbb{F}^n$ and let $\mathbf{v} \in \mathbb{F}^n$. Then there are unique vectors $\mathbf{w}$ in W and $\mathbf{w}^\perp$ in $W^\perp$ such that $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$. That is, $W \oplus W^\perp = \mathbb{F}^n$.*

Proof. We already know that if the dimension of W is k then the dimension of $W^\perp$ is $n-k$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for W and $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ be a basis for $W^\perp$. Then by Part 3 of Result 9.6, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, and hence it is a basis for $\mathbb{F}^n$.

Thus for any $\mathbf{v} \in \mathbb{F}^n$, we have $a_1, \ldots, a_n \in \mathbb{F}$ such that $\mathbf{v} = \sum_{i=1}^{k} a_i\mathbf{v}_i = \mathbf{w} + \mathbf{w}'$, where $\mathbf{w} = \sum_{i=1}^{k} a_i\mathbf{v}_i$ and $\mathbf{w}' = \sum_{i=k+1}^{n} a_i\mathbf{v}_i$.

Now if $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_1' = \mathbf{w}_2 + \mathbf{w}_2'$, where $\mathbf{w}_1, \mathbf{w}_2 \in W, \mathbf{w}_1', \mathbf{w}_2' \in W^\perp$ then

$$\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{w}_2' - \mathbf{w}_1' \in W \cap W^\perp = \{\mathbf{0}\} \Rightarrow \mathbf{w}_1 = \mathbf{w}_2, \mathbf{w}_1' = \mathbf{w}_2'.$$

Thus there are unique vectors $\mathbf{w}$ in W and $\mathbf{w}^\perp$ in $W^\perp$ such that $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$. That is, $W \oplus W^\perp = \mathbb{F}^n$. □

Result 9.8. *Let W be a subspace of $\mathbb{F}^n$. Then $(W^\perp)^\perp = W$.*

Proof. Note that $(W^\perp)^\perp$ contains those elements of $\mathbb{F}^n$ which are orthogonal to $W^\perp$. In particular, as each element w of W is orthogonal to $W^\perp$, we see that $w \in (W^\perp)^\perp$. Thus $W \subseteq (W^\perp)^\perp$.

As $\dim(W) = k$, we know that $\dim(W^\perp) = n-k$. Hence $\dim((W^\perp)^\perp) = n-(n-k) = k$. Now $W \subseteq (W^\perp)^\perp$ and both have the same dimension, and hence they must be equal. □

Result 9.9. *Let A be an $m \times n$ matrix. Then $(col(A))^{\perp} = null(A^*), (row(\overline{A}))^{\perp} = null(A)$ and $row(\overline{A}) = (null(A))^{\perp}$.*

Proof. Let $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n]$, where $\left\{\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}\right\}$. Notice that if $\mathbf{w} \in \text{col}(A)$ then we have $\mathbf{w} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Therefore $\mathbf{v}.\mathbf{a}_i = 0$ for all $i = 1, 2, \ldots, n$ implies that $\mathbf{v}.\mathbf{w} = 0$. Conversely, it is clear that if $\mathbf{v}.\mathbf{w} = 0$ for all $\mathbf{w} \in \text{col}(A)$, then $\mathbf{v}.\mathbf{a}_i = 0$ for all $i = 1, 2, \ldots, n$. Now we have

$$\begin{aligned}
(\text{col}(A))^{\perp} =& \{\mathbf{v} \ : \ \mathbf{v} \cdot \mathbf{w} = 0 \text{ for each } \mathbf{w} \in \text{col}(A)\} \\
=& \{\mathbf{v} \ : \ \mathbf{v} \cdot \mathbf{a}_i = 0 \text{ for each } i = 1, \ldots, n\} \\
=& \{\mathbf{v} \ : \ \mathbf{a}_i \cdot \mathbf{v} = 0 \text{ for each } i = 1, \ldots, n\} \\
=& \{\mathbf{v} \ : \ \mathbf{a}_i{}^*\mathbf{v} = 0 \text{ for each } i = 1, \ldots, n\} \\
=& \left\{\mathbf{v} \ : \ \begin{bmatrix} \overline{a_{11}} & \cdots & \overline{a_{m1}} \\ & \vdots & \\ \overline{a_{1n}} & \cdots & \overline{a_{mn}} \end{bmatrix} \mathbf{v} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}\right\} \\
=& \text{null}(A^*).
\end{aligned}$$

Now $\text{row}(\overline{A})^{\perp} = (\text{col}(A^*))^{\perp} = \text{null}(A)$, as $\text{row}(A) = \text{col}(A^t)$. Again we have $\text{row}(\overline{A}) = ((row(\overline{A}))^{\perp})^{\perp} = (\text{null}(A))^{\perp}$. $\square$

Corollary 9.2. *Let A be an $m \times n$ matrix. Then $rank(A) + nullity(A) = n$.*

Proof. We first prove that $\text{nullity}(A) = \text{nullity}(\overline{A})$. It is clear that $A\mathbf{v} = \mathbf{0} \Leftrightarrow \overline{A}\overline{\mathbf{v}} = \mathbf{0}$. Let $\{\mathbf{v}_1, \ldots \mathbf{v}_k\}$ be a basis for $\text{null}(A)$. Now for $a_1, \ldots, a_k \in \mathbb{F}$ we have

$$a_1\overline{\mathbf{v}}_1 + \ldots + a_k\overline{\mathbf{v}}_k = \mathbf{0} \Rightarrow \overline{a}_1\mathbf{v}_1 + \ldots + \overline{a}_k\mathbf{v}_k = \mathbf{0} \Rightarrow \overline{a}_1 = 0, \ldots, \overline{a}_k = 0 \Rightarrow a_1 = 0, \ldots, a_k = 0.$$

Thus $\{\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_k\}$ is linearly independent. Also if $\mathbf{v} \in \text{null}(\overline{A})$, then $\overline{\mathbf{v}} \in \text{null}(A)$. Accordingly we find $b_1, \ldots, b_k \in \mathbb{F}$ such that $\overline{\mathbf{v}} = b_1\mathbf{v}_1 + \ldots + b_k\mathbf{v}_k$. Then $\mathbf{v} = \overline{b}_1\overline{\mathbf{v}}_1 + \ldots + \overline{b}_k\overline{\mathbf{v}}_k$. Thus $\{\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_k\}$ spans $\text{null}(\overline{A})$. Finally we conclude that $\{\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_k\}$ is a basis for $\text{null}(\overline{A})$. Hence $\text{nullity}(A) = \text{nullity}(\overline{A})$.

Now we have $\text{nullity}(A) = \text{nullity}(\overline{A}) = \dim(\text{null}(\overline{A})) = \dim(\text{row}(A)^{\perp})$ and $\text{rank}(A) = \dim(\text{row}(A))$. Therefore $n = \dim(\text{row}(A)) + \dim(\text{row}(A)^{\perp}) = \text{rank}(A) + \text{nullity}(A)$. $\square$

Example 9.3. *Let S be a subspace of $\mathbb{F}^n$ and let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ form a basis for $S^{\perp}$. Consider the $k \times n$ matrix A whose i-th row is $\mathbf{v}_i^*$. Show that $S = null(A)$.*

Solution. Note that $\mathbf{v}_i{}^*\mathbf{w} = 0$ for each $\mathbf{w} \in S$. Thus $A\mathbf{w} = \mathbf{0}$ for each $\mathbf{w} \in S$. Hence $S \subseteq \text{null}(A)$. Furthermore, $\text{rank}(A) = k = \dim(S^{\perp})$. Hence $\dim(S) = n - k = \dim(\text{null}(A))$, so that $S = \text{null}(A)$. $\square$

Example 9.4. *Let A and B be two $m \times n$ matrices and let the linear systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution space. Show that the matrix A is row equivalent to B.*

Solution. We have $\text{null}(A) = \text{null}(B)$. Hence $(\text{null}(A))^{\perp} = (\text{null}(B))^{\perp}$, that is, $\text{row}(\overline{A}) = \text{row}(\overline{B})$. Hence $\text{row}(A) = \text{row}(B)$. Let $k = \dim(\text{row}(A))$. Let A' be the matrix obtained by taking the first k rows of RREF of A and B' be the matrix obtained by taking the first k rows of RREF of B. As rows of A' are linear combinations of rows of B', we have $A' = SB'$ for some $k \times k$ matrix S. Note that $k = \text{rank}(A') = \text{rank}(SB') \leq \text{rank}(S)$. That is, S is invertible. Hence A' is row equivalent to B'. Now A is row equivalent to A',A' is row equivalent to B' and B' is row equivalent to B altogether give that A is row equivalent to B. [Indeed, $A' = B'$ as RREF of A is row equivalent to RREF of B and that RREF of a matrix is unique.] $\square$

Example 9.5. *Let A and B be two $m \times n$ matrices and let the* ***consistent*** *linear systems $A\mathbf{x} = \mathbf{c}$ and $B\mathbf{x} = \mathbf{d}$ have the same solution set. Show that the matrix A is row equivalent to B.*

Solution. If $A\mathbf{x} = \mathbf{c}$ and $B\mathbf{x} = \mathbf{d}$ have the same solution set, then $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution set. Now proceed as in **Example** 9.4. $\square$

Example 9.6. *Let W be the subspace of $\mathbb{R}^5$ spanned by the vectors $\mathbf{w}_1 = [1, -3, 5, 0, 5]^t, \mathbf{w}_2 = [-1, 1, 2, -2, 3]^t$ and $\mathbf{w}_3 = [0, -1, 4, -1, 5]^t$. Find a basis for $W^{\perp}$.*

Example 9.7. *Consider $W = \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x + y + z = 0\}$. Find bases for W and $W^{\perp}$. For $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, find (somehow) $\mathbf{w} \in W$ and $\mathbf{w}' \in W^{\perp}$ such that $\mathbf{v} = \mathbf{w} + \mathbf{w}'$.*

Solution. We have $x+y+z=0 \Rightarrow x=-y-z$ so that $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Thus $\{[-1,1,0]^t, [-1,0,1]^t\}$ is a basis for W. Now $[x,y,z]^t \in W^{\perp}$ iff $[x,y,z]^t.[-1,1,0]^t = 0$ and $[x,y,z]^t.[-1,0,1]^t = 0$. That is, $-x+y=0$ and $-x+z=0$ which imply $x=y=z$. Thus $\{[1,1,1]^t\}$ is a basis for $W^{\perp}$.

Solving $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = a\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ for a, b, c, we have $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, so that $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \in W$ and $\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \in W^{\perp}$. $\square$

Orthogonal Projection: Let W be a subspace of $\mathbb{F}^n$ and let $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be an orthogonal basis for W. If $\mathbf{v} \in \mathbb{F}^n$, then the **orthogonal projection of v onto** W is defined as

$$\text{proj}_W(\mathbf{v}) = (\mathbf{w}_1.\mathbf{v})\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|^2} + (\mathbf{w}_2.\mathbf{v})\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|^2} + \ldots + (\mathbf{w}_k.\mathbf{v})\frac{\mathbf{w}_k.\mathbf{v}}{\|\mathbf{w}_k\|^2}.$$

The **component of v orthogonal to** W is the vector

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v}).$$

- Note that if W_i is the subspace spanned by the vector $\mathbf{w}_i$, then $\text{proj}_W(\mathbf{v}) = \text{proj}_{W_1}(\mathbf{v}) + \text{proj}_{W_2}(\mathbf{v}) + \ldots + \text{proj}_{W_k}(\mathbf{v})$.
- If $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is an orthonormal basis for W, then

$$\text{proj}_W(\mathbf{v}) = (\mathbf{w}_1.\mathbf{v})\mathbf{w}_1 + \ldots + (\mathbf{w}_k.\mathbf{v})\mathbf{w}_k.$$

- Assume that $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is an **orthonormal** basis for $\mathbb{F}^n$ such that $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is a basis for W and $\{\mathbf{w}_{k+1}, \ldots, \mathbf{w}_n\}$ is a basis for $W^{\perp}$. Then for $\mathbf{v} \in \mathbb{F}^n$

$$\begin{aligned}\mathbf{v} =& (\mathbf{w}_1.\mathbf{v})\mathbf{w}_1 + \ldots + (\mathbf{w}_n.\mathbf{v})\mathbf{w}_n \\ =& [(\mathbf{w}_1.\mathbf{v})\mathbf{w}_1 + \ldots + (\mathbf{w}_k.\mathbf{v})\mathbf{w}_k] + [(\mathbf{w}_{k+1}.\mathbf{v})\mathbf{w}_{k+1} + \ldots + (\mathbf{w}_n.\mathbf{v})\mathbf{w}_n] \\ =& \mathbf{w} + \mathbf{w}'.\end{aligned}$$

  Notice that $\mathbf{w} = \text{proj}_W(\mathbf{v})$ and $\mathbf{w}' = \text{perp}_W(\mathbf{v})$.

- The **orthogonal projection** of $\mathbf{v}$ on W is the unique vector $\mathbf{w} \in W$ such that $\mathbf{v} = \mathbf{w} + \mathbf{w}'$, for some $\mathbf{w}' \in W^{\perp}$. In other words, it is the unique vector $\mathbf{w} \in W$ such that $\mathbf{v} - \mathbf{w} \in W^{\perp}$. Again in words, it is the unique vector $\mathbf{w} \in W$ such that $(\mathbf{v} - \mathbf{w})$ is orthogonal to $\mathbf{w}$.
- For **Example** 9.7, we have $\{[-1,2,-1]^t, [1,0,-1]^t\}$ is an orthogonal basis for W. Hence

$$\text{proj}_W([1,2,3]^t) = \frac{[-1,2,-1]^t.[1,2,3]^t}{6}[-1,2,-1]^t + \frac{[1,0,-1]^t.[1,2,3]^t}{2}[1,0,-1]^t = [-1,0,1]^t.$$

  Thus,

$$\text{perp}_W([1,2,3]^t) = [1,2,3]^t - [-1,0,1]^t = [2,2,2]^t.$$

Result 9.10 (**The Gram-Schmidt Process**). *Let* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ *be an ordered basis for a subspace* W *of* $\mathbb{F}^n$ *and define*

$$\begin{aligned}
&\mathbf{v}_1 = \mathbf{x}_1, && W_1 = \mathit{span}(\mathbf{x}_1);\\
&\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1.\mathbf{x}_2}{\mathbf{v}_1.\mathbf{v}_1}\right)\mathbf{v}_1, && W_2 = \mathit{span}(\mathbf{x}_1, \mathbf{x}_2);\\
&\mathbf{v}_3 = \mathbf{x}_3 - \left(\frac{\mathbf{v}_1.\mathbf{x}_3}{\mathbf{v}_1.\mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2.\mathbf{x}_3}{\mathbf{v}_2.\mathbf{v}_2}\right)\mathbf{v}_2, && W_3 = \mathit{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3);\\
&\vdots\ \vdots && \vdots\ \vdots\\
&\vdots\ \vdots && \vdots\ \vdots\\
&\mathbf{v}_k = \mathbf{x}_k - \left(\frac{\mathbf{v}_1.\mathbf{x}_k}{\mathbf{v}_1.\mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2.\mathbf{x}_k}{\mathbf{v}_2.\mathbf{v}_2}\right)\mathbf{v}_2 - \ldots - \left(\frac{\mathbf{v}_{k-1}.\mathbf{x}_k}{\mathbf{v}_{k-1}.\mathbf{v}_{k-1}}\right)\mathbf{v}_{k-1}, && W_k = \mathit{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k).
\end{aligned}$$

Then for each $i = 1, 2, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ *is an orthogonal basis for* W_i. *In particular,* $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for* W.

Example 9.8. *Apply the Gram-Schmidt process to find an* ***orthonormal*** *basis of the subspace spanned by* $\mathbf{u} = [1,-1,1]^t$, $\mathbf{v} = [0,3,-3]^t$ *and* $\mathbf{w} = [3,2,2]^t$.

Solution. We have

$$\begin{aligned}
\mathbf{v}_1 =& \mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \\
\mathbf{v}_2 =& \mathbf{x}_2 - \left(\frac{\mathbf{v}_1.\mathbf{x}_2}{\mathbf{v}_1.\mathbf{v}_1}\right)\mathbf{v}_1 = \mathbf{x}_2 - \left(\frac{-6}{3}\right)\mathbf{v}_1 = \begin{bmatrix} 0 \\ 3 \\ -3 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \\
\mathbf{v}_3 =& \mathbf{x}_3 - \left(\frac{\mathbf{v}_1.\mathbf{x}_3}{\mathbf{v}_1.\mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2.\mathbf{x}_3}{\mathbf{v}_2.\mathbf{v}_2}\right)\mathbf{v}_2 = \mathbf{x}_3 - \left(\frac{3}{3}\right)\mathbf{v}_1 - \left(\frac{6}{6}\right)\mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.
\end{aligned}$$

Therefore $\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$, $\frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$ and $\frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

Hence an orthonormal basis is $\left\{ \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$. □

- Given a set of vectors S, we can use Gram-Schmidt process to check its linear dependency.
- We can find an orthonormal basis B for span(S).
- The vectors in S corresponding to the elements of B are linearly independent.
- The **angle** θ between $\mathbf{u}$ and $\mathbf{v}$, $(\mathbf{u}, \mathbf{v} \in \mathbb{R}^n)$, is defined by

$$\cos\theta = \frac{\mathbf{u}.\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}, \theta \in [0, \pi].$$

Orthogonal Matrix: An $n \times n$ matrix Q whose columns form an orthonormal set (*i.e.*, $QQ^t = I = Q^tQ$) is called an **orthogonal matrix**.

Practice Problems Set 9

1. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}^n$ and let $a, b \in \mathbb{R}$. Show that $(a\mathbf{x} + b\mathbf{y}).\mathbf{z} = \overline{a}(\mathbf{x}.\mathbf{z}) + \overline{b}(\mathbf{y}.\mathbf{z})$ and $\mathbf{x}.(a\mathbf{y} + b\mathbf{z}) = a(\mathbf{x}.\mathbf{y}) + b(\mathbf{x}.\mathbf{z})$.

2. For $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, prove the following:
   (a) $\mathbf{x}.\mathbf{y} = 0$ if and only if $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$;
   (b) $\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$;
   (c) $\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4(\mathbf{x}.\mathbf{y})$;
   (d) $\|\mathbf{x}\| = \|\mathbf{y}\|$ if and only if $(\mathbf{x} + \mathbf{y}).(\mathbf{x} - \mathbf{y}) = 0$; and
   (e) $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ if and only if $\|s\mathbf{x} + t\mathbf{y}\| = s\|\mathbf{x}\| + t\|\mathbf{y}\|$ for all $s, t \geq 0$.

3. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal set of vectors in $\mathbb{F}^n$. Prove that
$$\left\| \sum_{i=1}^{k} \mathbf{v}_i \right\|^2 = \sum_{i=1}^{k} \|\mathbf{v}_i\|^2.$$

4. Find a basis for each of the orthogonal complements $S^\perp, M^\perp$ and $W^\perp$, where
$$S = \{[x, y, z]^t \in \mathbb{R}^3 : 2x - y + 3z = 0\}, \quad M = \{[x, y, z]^t \in \mathbb{R}^3 : x = 2t = y, z = -t, t \in \mathbb{R}\}$$
$$\text{and} \quad W = \{[x, y, z]^t \in \mathbb{R}^3 : x = t, y = -t, z = 3t, t \in \mathbb{R}\}.$$
   Also, write the orthogonal complements $S^\perp, M^\perp$ and $W^\perp$.

5. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{F}^n$. Prove that the matrix $A = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ is invertible, and compute its inverse.

6. Let M and N be two subspaces of $\mathbb{F}^n$. Show that $(M + N)^\perp = M^\perp \cap N^\perp$ and $(M \cap N)^\perp = M^\perp + N^\perp$.

7. Let A be a symmetric matrix. Show that the eigenvectors corresponding to distinct eigenvalues of A are orthogonal to each other.

8. If Q is an orthogonal matrix, prove that any matrix obtained by rearranging the rows of Q is also orthogonal.

9. Prove that the columns of an $m \times n$ matrix Q form an orthonormal set if and only if $Q^tQ = I_n$.

10. Show that a square matrix Q is orthogonal if and only if $Q^{-1} = Q^t$.

11. Prove that if an upper triangular matrix is orthogonal, then it must be a diagonal matrix.

12. Let Q be an $n \times n$ matrix. Prove that the following statements are equivalent:

    (a) Q is orthogonal.

    (b) $(Q\mathbf{x}).(Q\mathbf{y}) = \mathbf{x}.\mathbf{y}$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

    (c) $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for every $\mathbf{x} \in \mathbb{R}^n$.

13. Prove that if $n > m$, then there is no $m \times n$ matrix A such that $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.

14. Apply Gram-Schmidt process to the following sets to obtain an orthonormal set in the spaces spanned by the corresponding vectors:

    (a) $\{[1, 1, 0]^t, [3, 4, 2]^t\}$ in $\mathbb{R}^3$;

    (b) $\{[-1, 0, 1]^t, [1, -1, 0]^t, [0, 0, 1]^t\}$ in $\mathbb{R}^3$;

    (c) $\{[2, -1, 1, 2]^t, [3, -1, 0, 4]^t\}$ in $\mathbb{R}^4$; and

    (d) $\{[1, 1, 1, 1]^t, [0, 2, 0, 2]^t, [-1, 1, 3, -1]^t\}$ in $\mathbb{R}^4$.

15. Use Gram-Schmidt process to find an orthogonal basis for the column space of each of the following matrices:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 1 & 0 \\ 1 & 5 & 1 \end{bmatrix}.$$

16. Find an orthogonal basis for $\mathbb{R}^3$ containing the vector $[3, 1, 5]^t$.

17. Find an orthogonal basis for $\mathbb{R}^4$ containing the vectors $[2, 1, 0, -1]^t$ and $[1, 0, 3, 2]^t$.

18. Find an orthogonal basis for the subspace spanned by $[1, 1, 0, 1]^t, [-1, 1, 1, -1]^t$, $[0, 2, 1, 0]^t$ and $[1, 0, 0, 0]^t$.

19. Find the orthogonal projection of $\mathbf{v}$ onto the subspace W spanned by the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, where

$$\mathbf{v} = [1, 2, 3]^t, \ \mathbf{u}_1 = [2, -2, 1]^t \text{ and } \mathbf{u}_2 = [-1, 1, 4]^t.$$

20. Find the orthogonal projection of $\mathbf{v}$ onto the subspace W spanned by the vectors $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$, where

$$\mathbf{v} = [3, -2, 4, -3]^t, \ \mathbf{u}_1 = [1, 1, 0, 0]^t, \mathbf{u}_2 = [1, -1, -1, 1]^t \text{ and } \mathbf{u}_3 = [0, 0, 1, 1]^t.$$

21. Let M be a subspace of $\mathbb{R}^m$ and let $\dim(M) = k$. How many linearly independent vectors can be orthogonal to M? Justify your answer.

22. Let A be an $n \times n$ orthogonal matrix. Show that the rows of A form an orthonormal basis for $\mathbb{R}^n$. Similarly, the columns of A also form an orthonormal basis for $\mathbb{R}^n$.

23. Let $\mathbf{u} = [1, 0, 0, 0]^t$ and $\mathbf{v} = [0, \frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}]^t$. Find vectors $\mathbf{w}$ and $\mathbf{x}$ in $\mathbb{R}^4$ such that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}\}$ form an orthonormal basis for $\mathbb{R}^4$.

24. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthogonal basis for $\mathbb{F}^n$ and let $W = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ for some $1 \le k < n$. Is it necessarily true that $W^\perp = \operatorname{span}(\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$? Either prove that it is true or find a counterexample.

25. Let W be a subspace of $\mathbb{F}^n$ and let $\mathbf{x} \in \mathbb{F}^n$. Prove that

    (a) $\mathbf{x} \in W$ if and only if $\operatorname{proj}_W(\mathbf{x}) = \mathbf{x}$;

    (b) $\mathbf{x}$ is orthogonal to W if and only if $\operatorname{proj}_W(\mathbf{x}) = \mathbf{0}$; and

    (c) $\operatorname{proj}_W(\operatorname{proj}_W(\mathbf{x})) = \operatorname{proj}_W(\mathbf{x})$.

Hints to Practice Problems Set 9

1. Easy.

2. For the first four parts, applly distributive and commutative laws in $\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y}).(\mathbf{x}+\mathbf{y})$.
   (e) $\|s\mathbf{x}+t\mathbf{y}\| \leq s\|\mathbf{x}\| + t\|\mathbf{y}\|$. Also for $t \geq s$, $\|s\mathbf{x}+t\mathbf{y}\| = \|t(\mathbf{x}+\mathbf{y}) - (t-s)\mathbf{x}\| \geq t\|\mathbf{x}+\mathbf{y}\| - (t-s)\|\mathbf{x}\|$. Similarly, for $t < s$, $\|s\mathbf{x}+t\mathbf{y}\| \geq s\|\mathbf{x}\| + t\|\mathbf{y}\|$.

3. Use $\mathbf{v}_i.\mathbf{v}_j = 0$ in the expansion of the RHS of $\left\|\sum\limits_{i=1}^{n} \mathbf{v}_i\right\|^2 = (\sum\limits_{i=1}^{n} \mathbf{v}_i).(\sum\limits_{i=1}^{n} \mathbf{v}_i)$.

4. $S^\perp = \{[u,v,w]^t \in \mathbb{R}^3 : u+2v=0, 3u-2w=0\}$ and a basis for $S^\perp$ is $\{[2,-1,3]^t\}$.
   $M^\perp = \{[u,v,w]^t \in \mathbb{R}^3 : 2u+2v-w=0\}$ and a basis for $M^\perp$ is $\{[-1,1,0]^t, [1,0,2]^t\}$.
   $W^\perp = \{[u,v,w]^t \in \mathbb{R}^3 : u-v+3w=0\}$ and a basis for $W^\perp$ is $\{[1,1,0]^t, [3,0,-1]^t\}$.

5. $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, and $A^{-1} = A^t$.

6. Show that $(M+N)^\perp \subseteq M^\perp \cap N^\perp$ and $M^\perp \cap N^\perp \subseteq (M+N)^\perp$. For the 2nd part, replace M by $M^\perp$ and N by $N^\perp$ in the 1st part.

7. $A\mathbf{x} = \lambda\mathbf{x}, A\mathbf{y} = \mu\mathbf{y} \Rightarrow \lambda\mathbf{x}^t\mathbf{y} = \mathbf{x}^tA\mathbf{y} = \mu\mathbf{x}^t\mathbf{y}$.

8. Rearranging the columns of Q^t will not change the orthogonality of its columns.

9. See **Theorem 5.4** of the text book.

10. Use **Problem 9**.

11. If $A = [a_{ij}]$ is orthogonal and upper triangular, then we have $a_{11}^2 = a_{11}^2 + a_{12}^2 + \ldots + a_{1n}^2$.

12. See **Theorem 5.6** of the text book.

13. Use **Problem 12**.

14. (a) $\{[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]^t, [-\frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{2\sqrt{2}}{3}]^t\}$, (b) Find and check yourself.
    (c) $\{[\frac{2}{\sqrt{10}}, -\frac{1}{\sqrt{10}}, \frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}]^t, [0, \frac{1}{\sqrt{14}}, -\frac{3}{\sqrt{14}}, \frac{\sqrt{2}}{\sqrt{7}}]^t\}$,
    (d) $\{[\frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}]^t, [-\frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}, -\frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}]^t, [-\frac{2}{\sqrt{10}}, \frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}, -\frac{1}{\sqrt{10}}]^t\}$.

15. $\{[0,1,1]^t, [1,-\frac{1}{2},\frac{1}{2}]^t, [\frac{2}{3},\frac{2}{3},-\frac{2}{3}]^t\}$ and $\{[1,1,-1,1]^t, [0,-2,2,4]^t, [0,1,1,0]^t\}$.

16. Apply Gram-Schmidt Process to $\{[3,1,5]^t\} \cup B$, where $\{[3,1,5]^t\} \cup B$ is a basis for $\mathbb{R}^3$.

17. Similar to **Problem 16**.

18. Apply Gram-Schmidt Process to a basis for the subspace.

19. $\text{proj}_W(\mathbf{v}) = [-\frac{1}{2}, \frac{1}{2}, 3]^t$.

20. $\text{proj}_W(\mathbf{v}) = [0,1,1,0]^t$.

21. $m-k$.

22. $AA^t = I = A^tA$.

23. Similar to **Problem 17**.

24. The given statement is true.

25. (a) Direct use of the definition of $\text{proj}_W(\mathbf{v})$. (c) Use part (a).

---

10 Spectral Theorem

In this section, we shall study those matrices which are diagonalizable in a very nice manner, in the sense that $P^{-1}AP$ is diagonal, where P is unitary or orthogonal in some cases. Note that if P is unitary, then $P^{-1} = P^*$, and if P is orthogonal, then $P^{-1} = P^t$.

Example 10.1. *Examine if there is an orthogonal matrix P such that P^tAP is diagonal, where $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.*

Definition 10.1. *Recall the definitions of transpose and conjugate transpose of a matrix, symmetric and Hermitian matrices.*

- *A matrix A is said to be normal if $AA^* = A^*A$.*
- *A **real** matrix A is said to be orthogonal if $AA^t = I = A^tA$.*
- *A matrix A is said to be unitary if $AA^* = I = A^*A$.*
- *Note that symmetric and Hermitian matrices are normal.*
- *A **real** unitary matrix is orthogonal.*
- *A real Hermitian matrix is symmetric.*

Definition 10.2.

- *A **real** matrix A is said to be orthogonally diagonalizable if there is an orthogonal matrix Q such that Q^tAQ is a diagonal matrix.*
- *A matrix A is said to be unitarily diagonalizable if there is an unitary matrix U such that U^*AU is a diagonal matrix.*

Result 10.1 (Schur Unitary Triangularization Theorem)**.** *For a square matrix A, there is an unitary matrix U such that U^*AU is upper triangular.*

Proof. Let A be an $n \times n$ matrix. We use induction on n. The case $n = 1$ is trivial. Let $n > 1$, and assume that the result is true for every $(n-1) \times (n-1)$ matrices.

Let λ_1 be an eigenvalue of A, with eigenvector $\mathbf{w}_1$ such that $\|\mathbf{w}_1\| = 1$. Consider and orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ of $\mathbb{C}^n$, and $W = [\mathbf{w}_1 \cdots \mathbf{w}_n]$. Then $W^*AW = \begin{bmatrix} \lambda_1 & * \\ 0 & A' \end{bmatrix}$, where A' is an $(n-1) \times (n-1)$ matrix.

By induction hypothesis, $U'^*A'U' = T'$ is upper-triangular for some unitary matrix U'. Take $U = W \begin{bmatrix} 1 & 0 \\ 0 & U' \end{bmatrix}$ so that U is unitary. We have

$$\begin{aligned} U^*AU = \begin{bmatrix} 1 & 0 \\ 0 & U'^* \end{bmatrix} W^*AW \begin{bmatrix} 1 & 0 \\ 0 & U' \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & U'^* \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ 0 & A' \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U' \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & * \\ 0 & U'^*A'U' \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & * \\ 0 & T' \end{bmatrix}, \text{ an upper-triangular matrix.} \end{aligned}$$

Hence, we conclude by principle of mathematical induction that the result is true for all n. ☐

Corollary 10.1. *If a square matrix A and its eigenvalues are real, then there is a (real) orthogonal matrix Q such that Q^tAQ is upper triangular.*

Proof. If a square matrix A and its eigenvalues are real, then $A\mathbf{x} = \lambda\mathbf{x}$ can be solved in $\mathbb{R}^n$. Thus if the proof of Schur Unitary Triangularization Theorem is applied for A, every numbers, vectors and matrices can be taken to be real. ☐

Remark 10.1. *The diagonal entries in U^*AU or Q^tAQ of the previous results can be obtained (taken) in any prescribed order, accordingly the U or Q will change. They are the eigenvalues of A.*

Result 10.2.

- *If A is orthogonally diagonalizable then A is symmetric.*

- *If A is unitarily diagonalizable then A is normal.*

Proof.

- Let $P^tAP = D$ be diagonal, where $P^tP = I = PP^t$. We have $A = PDP^t$ and so $A^t = (PDP^t)^t = PDP^t = A$. Hence A is symmetric.
- Let $P^*AP = D$ be diagonal, where $P^*P = I = PP^*$. We have $A = PDP^*$ so that $A^* = PD^*P^*$. Now

$$A^*A = (PD^*P^*)(PDP^*) = PD^*DP^* = PDD^*P^* = (PDP^*)(PD^*P^*) = AA^*.$$

  Hence A is normal.

□

The following is **Result** 9.3 stated again.

Result 10.3. *If A is a Hermitian matrix, then its eigenvalues are real.*

The second part of the following result is a part of **Result** 9.3.

Result 10.4.

- *If A is a normal matrix, then any two eigenvectors corresponding to distinct eigenvalues of A are orthogonal.*
- *If A is a real symmetric matrix, then any two eigenvectors corresponding to distinct eigenvalues of A are orthogonal.*

Proof. Let $A^*A = AA^*$, $A\mathbf{u} = \lambda\mathbf{u}$ and $A\mathbf{v} = \mu\mathbf{v}$, where $\mathbf{u} \neq \mathbf{0}, \mathbf{v} \neq \mathbf{0}$ and $\lambda \neq \mu$. To show that $\mathbf{v}^*\mathbf{u} = 0$.

We notice that

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*(A^*A)\mathbf{x} = \mathbf{x}^*(AA^*)\mathbf{x} = (A^*\mathbf{x})^*(A^*\mathbf{x}) = \|A^*\mathbf{x}\|^2.$$

Therefore

$$A\mathbf{v} = \mu\mathbf{v} \Leftrightarrow \|(A-\mu I)\mathbf{v}\| = 0 \Leftrightarrow \|(A-\mu I)^*\mathbf{v}\| = 0 \Leftrightarrow \|(A^* - \overline{\mu}I)\mathbf{v}\| = 0 \Leftrightarrow A^*\mathbf{v} = \overline{\mu}\mathbf{v} \Leftrightarrow \mathbf{v}^*A = \mu\mathbf{v}^*.$$

Therefore

$$\lambda(\mathbf{v}^*\mathbf{u}) = \mathbf{v}^*(\lambda\mathbf{u}) = \mathbf{v}^*(A\mathbf{u}) = (\mathbf{v}^*A)\mathbf{u} = \mu\mathbf{v}^*\mathbf{u} \Rightarrow (\lambda-\mu)\mathbf{v}^*\mathbf{u} = 0 \Rightarrow \mathbf{v}^*\mathbf{u} = 0, \quad \text{as } \lambda \neq \mu.$$

□

Result 10.5 (Spectral Theorem for Normal Matrices)**.** *A matrix A is normal iff there is an unitary matrix U such that U^*AU is a diagonal matrix.*

Proof. Let A be normal. By Schur's Triangulation Theorem, there exists an unitary matrix U such that $U^*AU = T$ is upper-triangular, where $T = [t_{ij}]$. From $AA^* = A^*A$, we get $TT^* = T^*T$. Now

$$(TT^*)_{11} = (T^*T)_{11} \Rightarrow \sum_{i=1}^{n} |t_{1i}|^2 = |t_{11}|^2 \Rightarrow t_{1i} = 0 \text{ for } i = 2, 3, \ldots, n.$$

Now from $(TT^*)_{22} = (T^*T)_{22}$, we get $t_{2i} = 0$ for $i = 3, 4, \ldots, n$. Repeating this process, we find that T is diagonal.

The other part of the proof is easy, and was done in Result 10.2. □

Corollary 10.2. *A matrix A is Hermitian iff there is an unitary matrix U such that U^*AU is a* ***real*** *diagonal matrix.*

Proof. Let $A = A^*$ so that A is normal as well. By Spectral Theorem, there is an unitary matrix U such that $U^*AU = D$ is a diagonal matrix. Now $A = A^*$ gives that

$$D^* = (U^*AU)^* = U^*AU = D \Rightarrow D \text{ is a real diagonal matrix.}$$

Thus U^*AU is a **real** diagonal matrix. The other part of the proof is easy. □

Corollary 10.3 (Spectral Theorem for Real Symmetric Matrices)**.** *A real matrix A is symmetric iff there is an orthogonal matrix Q such that Q^tAQ is a diagonal matrix.*

Proof. Let A be a real symmetric matrix. Then all eigenvalues of A are real. Therefore by Corollary 10.1, there is a (real) orthogonal matrix Q such that $Q^tAQ = D$ is upper triangular. Now $A = A^t$ and $Q^tAQ = D$ give that $D = D^t$, and hence D is a diagonal matrix.

The other part of the proof is easy, and was done in Result 10.2. $\square$

Result 10.6 (Cayley-Hamilton Theorem)**.** *Every matrix satisfies its characteristic equation.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a square matrix A, so that its characteristic polynomial is $p(x) = \prod\limits_{i=1}^{n}(x - \lambda_i)$. By Schur's Triangulation Theorem, there exists an unitary matrix U such that $U^*AU = T$ is upper-triangular, where $T = [t_{ij}]$ with $t_{ii} = \lambda_i$ for each i. We have

$$p(A) = \prod_{i=1}^{n}(A - \lambda_i I) = \prod_{i=1}^{n}(UTU^* - \lambda_i UIU^*) = U\left[(T - \lambda_1 I)\ldots(T - \lambda_n I)\right]U^* = UOU^* = O.$$

Note that $T - \lambda_1 I$ has the first column zero, $(T - \lambda_1 I)(T - \lambda_2 I)$ has the first two columns zero, and so on. $\square$

Practice Problems Set 10

1. Orthogonally diagonalize the following matrices:

$$(a)\ \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix},\ (b)\ \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}, (c)\ \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix}, (d)\ \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix},\ (e)\ \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{bmatrix},\ (f)\ \begin{bmatrix} 2 & 3 & 0 \\ 3 & 2 & 4 \\ 0 & 4 & 2 \end{bmatrix},$$

$$(g)\ \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix},\quad (h)\ \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},\quad (i)\ \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},\quad (j)\ \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

2. If $b \neq 0$, then orthogonally diagonalize the real matrices $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ and $B = \begin{bmatrix} a & 0 & b \\ 0 & a & 0 \\ b & 0 & a \end{bmatrix}$.

3. Let A and B be orthogonally diagonalizable real matrices of same size and let $c \in \mathbb{R}, k \in \mathbb{N}$. Show that the matrices $A + B, cA, A^2$ and A^k are also orthogonally diagonalizable.

4. If the matrix A is invertible and orthogonally diagonalizable, then show that A^{-1} is also orthogonally diagonalizable.

5. If A and B be orthogonally diagonalizable and $AB = BA$, then show that AB is also orthogonally diagonalizable.

6. If A is a symmetric matrix, then show that every eigenvalue of A is non-negative iff $A = B^2$ for some symmetric matrix B.

7. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal set in $\mathbb{R}^n$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Show that the matrix $\sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^t$ is symmetric.

8. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal set in $\mathbb{C}^n$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Show that the matrix $\sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^*$ is normal.

9. Let A be a real symmetric matrix. Show that there is an orthonormal set $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ in $\mathbb{R}^n$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that $A = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^t$.

10. Let A be a normal matrix. Show that there is an orthonormal set $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ in $\mathbb{C}^n$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that $A = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^*$.

11. Let A be a symmetric matrix. Show that there are distinct real numbers $\alpha_1, \ldots, \alpha_k$ and real matrices $E_1, \ldots, E_k$ such that $A = \alpha_1 E_1 + \ldots + \alpha_k E_k$, where $E_1 + \ldots + E_k = I,\ E_iE_j = O$ for $i \neq j$ and $E_i^2 = E_i$ for $i = 1, \ldots, k$.

12. Find spectral decomposition of the matrices given in Excercise 1.

13. Find symmetric matrices of appropriate sizes with the given eigenvalues and the corresponding eigenvectors for the following problems:

(a) $\lambda_1 = -1, \lambda_2 = 2$ and $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$;

(b) $\lambda_1 = 3, \lambda_2 = -2$ and $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$;

(c) $\lambda_1 = -3, \lambda_2 = -3$ and $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$;

(d) $\lambda_1 = 1, \lambda_2 = -4, \lambda_3 = -4$ and $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 5 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

14. Let A be a real nilpotent matrix. Prove that there is an orthogonal matrix Q such that Q^tAQ is upper-triangular with zeros on its diagonal.

15. Let $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ be an orthonormal set of vectors in $\mathbb{R}^n$ and $W = \text{span}(\mathbf{q}_1, \ldots, \mathbf{q}_k)$.

(a) Show that the matrix of the orthogonal projection onto W is given by

$$P = \mathbf{q}_1\mathbf{q}_1^t + \ldots + \mathbf{q}_k\mathbf{q}_k^t.$$

(b) Show that the matrix in Part (a) above is symmetric and satisfies $P^2 = P$.

(c) Show that $P = QQ^t$, and deduce that $\text{rank}(P) = k$, where $Q = [\mathbf{q}_1 \cdots \mathbf{q}_k]$.

Hints to Practice Problems Set 10

1. (a) $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$.

(b) The eigenvalues of the matrix are 2 and -4 with corresponding eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, respectively.

Normalize the eigenvectors to get $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$.

(c) $Q = \begin{bmatrix} 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{3} & -2/\sqrt{6} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$.

(d) $Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \quad D = \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}$.

(e) $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.

(f) $Q = \frac{1}{5}\begin{bmatrix} 3/\sqrt{2} & 3/\sqrt{2} & 4 \\ 5/\sqrt{2} & -5/\sqrt{2} & 0 \\ 4/\sqrt{2} & 4/\sqrt{2} & -3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(g) $Q = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

(h) $Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{3} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{2} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

(i) $Q = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(j) $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

2. For A: $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} a+b & 0 \\ 0 & a-b \end{bmatrix}$;

   For B: $Q = \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} a & 0 & 0 \\ 0 & a+b & 0 \\ 0 & 0 & a-b \end{bmatrix}$;

3. Since A and B be orthogonally diagonalizable real matrices, they are all symmetric. Therefore the matrices $A+B, cA, A^2$ and A^k are all symmetric. **OR**, if Q is orthogonal and $Q^tAQ = D$ is diagonal, then $Q^t(cA)Q = cD, Q^tA^2Q = D^2$ and $Q^tA^kQ = D^k$.

4. If A is symmetric, then A^{-1} is symmetric. **OR**, if Q is orthogonal and $Q^tAQ = D$, then $Q^tA^{-1}Q = D^{-1}$.

5. If A, B are symmetric and $AB = BA$, then AB is symmetric.

6. If $Q^tAQ = D$, then take $B = Q^tD_1Q$, where each entry of D_1 is the positive square root of the corresponding entry in D, for one part. For the other part, if $B = Q^tD_1Q$, then $A = B^2 = Q^tD_1^2Q \Rightarrow Q^tAQ = D_1^2$.

7. $(\alpha_i \mathbf{u}_i \mathbf{u}_i^t)^t = \alpha_i (\mathbf{u}_i^t)^t \mathbf{u}_i^t = \alpha_i \mathbf{u}_i \mathbf{u}_i^t$.

8. $(\alpha_i \mathbf{u}_i \mathbf{u}_i^*)^*(\alpha_i \mathbf{u}_i \mathbf{u}_i^*) = |\alpha_i|^2 \mathbf{u}_i \mathbf{u}_i^* = (\alpha_i \mathbf{u}_i \mathbf{u}_i^*)(\alpha_i \mathbf{u}_i \mathbf{u}_i^*)^*$. For $i \neq j$, $(\alpha_i \mathbf{u}_i \mathbf{u}_i^*)^*(\alpha_j \mathbf{u}_j \mathbf{u}_j^*) = 0 = (\alpha_i \mathbf{u}_i \mathbf{u}_i^*)(\alpha_j \mathbf{u}_j \mathbf{u}_j^*)^*$.

9. For $Q^tAQ = D$, let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be the columns of Q and $\alpha_1, \ldots, \alpha_n$ be the diagonal entries of D.
   Then $QDQ^t = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^t$.

10. For $Q^*AQ = D$, let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be the columns of Q and $\alpha_1, \ldots, \alpha_n$ be the diagonal entries of D.
    Then $QDQ^* = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i \mathbf{u}_i^*$.

11. Use Problem 9 and proceed as in Spectral Decomposition Theorem.

12. (a) $5\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + 3\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

    (b) $2\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} - 4\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

    (c) $2\begin{bmatrix} \frac{2}{3} & \frac{2}{\sqrt{18}} \\ \frac{2}{\sqrt{18}} & \frac{1}{3} \end{bmatrix} - \begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{18}} \\ -\frac{2}{\sqrt{18}} & \frac{2}{3} \end{bmatrix}$.

    (d) $10\begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} + 5\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$.

    (e) $5\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 4\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} - 2\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

    (f) $7.\begin{bmatrix} \frac{9}{50} & \frac{15}{50} & \frac{12}{50} \\ \frac{15}{50} & \frac{25}{50} & \frac{20}{50} \\ \frac{12}{50} & \frac{20}{50} & \frac{16}{50} \end{bmatrix} - 3.\begin{bmatrix} \frac{9}{50} & -\frac{15}{50} & \frac{12}{50} \\ -\frac{15}{50} & \frac{25}{50} & -\frac{20}{50} \\ \frac{12}{50} & -\frac{20}{50} & \frac{16}{50} \end{bmatrix} + 2.\begin{bmatrix} \frac{16}{25} & 0 & -\frac{12}{25} \\ 0 & 0 & 0 \\ -\frac{12}{25} & 0 & \frac{9}{25} \end{bmatrix}$

    (g) $2.\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + 1.\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 0.\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} = 2\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

    (h) $5.\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} - 1.\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} - 1.\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

    (i) $2\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

(j) $3.\begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix} + 1.\begin{bmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix} + 1.\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 1.\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$

13. (a) $A = \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix}.$

(b) Normalize the given vectors using Gram-Schmidt Process to obtain $\mathbf{u}_1 = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$. Now $A = 3\mathbf{u}_1\mathbf{u}_1^t - 2\mathbf{u}_2\mathbf{u}_2^t = \begin{bmatrix} -\frac{1}{5} & \frac{12}{5} \\ \frac{12}{5} & \frac{6}{5} \end{bmatrix}.$

(c) Normalize the given vectors using Gram-Schmidt Process to obtain $\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$. Now $A = -3\mathbf{u}_1\mathbf{u}_1^t - 3\mathbf{u}_2\mathbf{u}_2^t = \begin{bmatrix} -3 & 0 \\ 0 & -3 \end{bmatrix}.$

(d) Normalize the given vectors using Gram-Schmidt Process to obtain $\mathbf{u}_1 = \frac{1}{\sqrt{42}}\begin{bmatrix} 4 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{u}_3 = \frac{1}{\sqrt{14}}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$. Now $A = 1.\mathbf{u}_1\mathbf{u}_1^t - 4.\mathbf{u}_2\mathbf{u}_2^t - 4.\mathbf{u}_3\mathbf{u}_3^t = \frac{1}{42}\begin{bmatrix} -88 & 100 & -20 \\ 100 & -43 & -25 \\ -20 & -25 & -163 \end{bmatrix}.$

14. 0 is the only eigenvalue of a nilpotent matrix.

15. $\text{proj}_W(\mathbf{v}) = (\mathbf{q}_1^t\mathbf{v})\mathbf{q}_1 + \ldots + (\mathbf{q}_k^t\mathbf{v})\mathbf{q}_k = (\mathbf{q}_1\mathbf{q}_1^t + \ldots + \mathbf{q}_k\mathbf{q}_k^t)\mathbf{v}.$

(a) $T(\mathbf{v}) = A\mathbf{v}$, where $A = \mathbf{q}_1\mathbf{q}_1^t + \ldots + \mathbf{q}_k\mathbf{q}_k^t$.

(b) $(\mathbf{q}_i\mathbf{q}_i^t)^t = \mathbf{q}_i\mathbf{q}_i^t,\;\; (\mathbf{q}_i\mathbf{q}_i^t)(\mathbf{q}_i\mathbf{q}_i^t) = \mathbf{q}_i\mathbf{q}_i^t$ and $(\mathbf{q}_i\mathbf{q}_i^t)(\mathbf{q}_j\mathbf{q}_j^t) = \mathbf{q}_i(\mathbf{q}_i^t\mathbf{q}_j)\mathbf{q}_j^t = O$ for $i \neq j$.

(c) Clearly, $[\mathbf{q}_1 \cdots \mathbf{q}_k]\begin{bmatrix} \mathbf{q}_1^t \\ \vdots \\ \mathbf{q}_k^t \end{bmatrix} = \mathbf{q}_1\mathbf{q}_1^t + \ldots + \mathbf{q}_k\mathbf{q}_k^t$. Also $\text{rank}(QQ^t) \leq \text{rank}(Q) = k$. Further

$$P\mathbf{q}_i = Q\begin{bmatrix} \mathbf{q}_1^t \\ \vdots \\ \mathbf{q}_k^t \end{bmatrix}\mathbf{q}_i = Q\begin{bmatrix} \mathbf{q}_1^t\mathbf{q}_i \\ \vdots \\ \mathbf{q}_i^t\mathbf{q}_i \\ \vdots \\ \mathbf{q}_n^t\mathbf{q}_i \end{bmatrix} = Q\mathbf{e}_i = \mathbf{q}_i \Rightarrow \mathbf{q}_i \in \text{range}(P).$$

End of Note