Linear Maps and Matrices

14 avril 2024

1 Linear mappings

The vector mapping $f : E \to F$, (where E and F are vector spaces over $\mathbb{K}$) is said to be linear if the following conditions hold :

1. $\forall u, v \in E : f(u+v) = f(u) + f(v)$.
2. $\forall u \in E, \forall \lambda \in \mathbb{K}, \quad f(\lambda u) = \lambda f(u)$.

This definition is equivalent to :

$$\forall u, v \in E, \forall \lambda, \beta \in \mathbb{K}, \quad f(\lambda u + \beta v) = \lambda f(u) + \beta f(v).$$

Example 1.0.1 *1. The identity map on E , which sends each $u \in E$ to u, is denoted I_E, or just I if the vector space E is clear from context.*
Note that all linear maps (not just the identity) send zero to zero.

Proof 1.0.1 *For any $u \in E$ we have $f(u) = f(u + 0_E) = f(u) + f(0_E)$.*
so by subtracting $f(u)$ from both sides we obtain $f(0_E) = 0_F$.

2. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ a map defined by $f(x, y) = (x + y, x - y)$. f is linear.
Indeed, let $u = (x, y), v = (x_0, y_0) \in \mathbb{R}^2$ and let $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned}
f(\alpha u + \beta v) &= f(\alpha(x, y) + \beta(x_0, y_0)) \\
&= f((\alpha x + \beta x_0, \alpha y + \beta y_0)) \\
&= (\alpha x + \beta x_0 + \alpha y + \beta y_0, \alpha x + \beta x_0 - \alpha y - \beta y_0) \\
&= (\alpha(x + y) + \beta(x_0 + y_0), (\alpha(x - y) + \beta(x_0 - y_0) \\
&= \alpha(x + y, x - y) + \beta(x_0 + y_0, x_0 - y_0) \\
&= \alpha f((x, y)) + \beta f((x_0, y_0))
\end{aligned}$$

3. $f : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $f(x, y) = x^2 - y^2$ is nonlinear because $f((1, 0)) + f((-1, 0)) = 1 + 1 = 2 \neq f((1, 0) + (-1, 0)) = f((0, 0)) = 0$.

4. The mapping $f(x, y) = (x^2 \sin y - \cos(x^2 - 1), x^2 + y^2 + 1)$ is nonlinear. To see this, we computed that
$f((3, 0)) = (-\cos 8, 10) \neq 3f((1, 0)) = (-3, 6)$. Then f is not linear.

Remark 1.0.1 *One useful fact regarding linear maps is that they are uniquely determined by their action on a basis. Let $\{v_1, v_2, \cdots, v_n\}$ be a basis for a vector space E. Then*

$$\forall u \in E : \quad u = \sum_{i=1}^{n} \alpha_i v_i et \quad f(u) = f(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i f(v_i),$$

1.1 Isomorphisms

Observe that the definition of a linear map is suited to reflect the structure of vector spaces, since it preserves vector spaces two main operations, addition and scalar multiplication. In algebraic terms, a linear map is said to be a **homomorphism** of vector spaces. An invertible homomorphism where the inverse is also a homomorphism is called an **isomorphism**. If there exists an isomorphism from E to F, then E and F are said to be **isomorphic**, and we write $E \cong F$. Isomorphic vector spaces are essentially "the same" in terms of their algebraic structure. It is an interesting fact that finite-dimensional vector spaces of the same dimension over the same field are always isomorphic.

Operations on linear mappings

Definition 1.1.1 *The set of linear mappings from E into F is denoted $\mathcal{L}(E,F)$. Let's equip $\mathcal{L}(E,F)$ with the addition of mappings $(f+g)$ and multiplication by a scalar $(\alpha.f)$ as follows :*

1. $\forall x \in E : (f+g)(x) = f(x) + g(x)$,
2. $\forall x \in E, \forall \alpha \in \mathbb{R} : (\alpha.f)(x) = \alpha f(x)$.

Proposition 1.1.1 *$\mathcal{L}(E,F)$ with addition and multiplication has a vector space structure on $\mathbb{R}$.*

Linear maps can be added and scaled to produce new linear maps. That is, if f and g are linear maps from E into F, and $\alpha \in \mathbb{K}$, it is straightforward to verify that $f+g$ and $\alpha \cdot f$ are linear maps from E into F. Since addition and scaling of functions satisfy the usual commutativity/ associativity/distributivity rules, the set of linear maps from E into F is also a vector space over $\mathbb{K}$. The additive identity here is the zero map which sends every $u \in E$ to 0_F.

Theoreme 1.1.1 ***(Composition)*** *Let E, F and G be vector spaces over a common field $\mathbb{K}$, and suppose f be a linear map from E to F and and g a linear map from F to G. Then $g \circ f$ is a linear map from E to G.*

Theoreme 1.1.2 *Let f be an isomorphism from E to F. Then f^{-1} is an isomorphism from F to E.*

Proposition 1.1.2 *Let f be an automorphism of E (isomorphism from E to E). Then f^{-1} is an automorphism of E. Let f and g be two automorphisms of E, then $g \circ f$ is an automorphism of E and we have $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Proposition 1.1.3 *If f from E to F is a linear map then :*

1. $\forall u \in E : f(-u) = -f(u)$.
2. $f(0_E) = 0_F$.

1.2 Nullspace and Range

1.2.1 Nullspace

Definition 1.2.1 *If $f : E \to F$ is a linear map, the nullspace of f is the set of all vectors in E that get mapped to zero :*

$$null(f) = \{u \in E, f(u) = 0_F\} = f^{-1}(\{0_F\}).$$

Proposition 1.2.1 *If $f : E \to F$ is a linear map, then $null(f)$ is a subspace of E .*

Proof 1.2.1

1. *We have already seen that $f(0_E) = 0_F$. So $0_E \in null(T)$.*
2. *If $u, v \in null(T)$, then $f(\alpha u + v) = \alpha f(u) + f(v) = 0_F + 0_F = 0_F$. So $\alpha u + v \in null(f)$.*

Proposition 1.2.2 *A linear map $f : E \to F$ is injective if and only if $null(f) = \{0_E\}$.*

Proof 1.2.2 *Assume f is injective. We have seen that $f(0_E) = 0_F$, so $0_E \in null(f)$, and moreover if $f(v) = 0_F = f(0_E)$ the injectivity of f implies $v = 0_E$, so $null(f) = \{0_E\}$.*
Conversely, assume $null(f) = \{0_E\}$, and suppose $f(u) = f(v)$ for some $u, v \in E$.
Then $0 = f(u) - f(v) = f(u - v)$, so $u - v \in null(f)$, which by assumption implies $u - v = 0_E$, i.e. $u = v$. Hence f is injective.

1.2.2 Range

Definition 1.2.2 *The range of f is the set of all possible outputs of f :*

$$range(f) = \{f(u) : \quad u \in E\} = f(E)$$

Proposition 1.2.3 *If $f : E \to F$ is a linear map, then $range(f)$ is a subspace of F.*

Proof 1.2.3

1. *We have already seen that $f(0_E) = 0_F$, so $0_F \in range(f)$.*
2. *If $w, z \in range(f)$, there exist $u, v \in E$ such that $w = f(u)$ and $z = f(v)$. Then $f(u + v) = f(u) + f(v) = w + z$ so $w + z \in range(f)$.*
3. *If $v \in range(f), \alpha \in \mathbb{K}, \exists u \in E$ such that $v = f(u)$. Then $f(\alpha u) = \alpha f(u) = \alpha v$ so $\alpha v \in range(f)$. Thus $range(f)$ is a subspace of F.*

Theoreme 1.2.1 *Let $f : E \to F$ be a linear map. So :*

1. *f is injective $\Leftrightarrow null(f) = \{0_E\}$.*
2. *f is surjective $\Leftrightarrow range(f) = F$.*

Example 1.2.1 *Determine the image and the nullspace of the map f from $\mathbb{R}^3$ to $\mathbb{R}^2$ defined by :*

$$f(x, y, z) = (x - y - z, x + y + z), \quad for \quad all(x, y, z) \in \mathbb{R}^3.$$

Deduce that f is not injective and that f is surjective. We have :

$$\begin{aligned} null(f) &= \left\{(x, y, z) \in \mathbb{R}^3, f((x, y, z)) = (0, 0)\right\} \\ &= \left\{(x, y, z) \in \mathbb{R}^3, (x - y - z, x + y + z) = (0, 0)\right\} \\ &= \left\{(x, y, z) \in \mathbb{R}^3, x = 0 et \quad z = -y\right\} \\ &= \left\{(0, y, -y) \in \mathbb{R}^3, y \in \mathbb{R}\right\} \\ &= span\{(0, 1, -1)\} \end{aligned}$$

null$(f) \neq \{(0,0,0)\}$, *therefore* f *is not injective.*

$$\begin{aligned} range(f) &= \left\{f(x,y,z) : (x,y,z) \in \mathbb{R}^3\right\} \\ &= \{(x-y-z, x+y+z) : x,y,z \in \mathbb{R}\} \\ &= \{x(1,1) + y(-1,1) + z(-1,1) : x,y,z \in \mathbb{R}\} \\ &= span\{(1,1),(-1,1),(-1,1)\} \\ &= span\{(1,1),(-1,1)\} \end{aligned}$$

$range(f) = span\{(1,1),(-1,1)\} = \mathbb{R}^2$, *so* f *is surjective.*

1.3 Linear mappings in finite dimension

Proposition 1.3.1 *Let* f *be a linear map of* E *into* F *and* $B = \{e_1, e_2, \cdots, e_n\}$ *be a base of* E. *Then the image by* f *of* B *is a generating family of* $range(f)$, *that is to say*

$$range(f) = span\{f(e_1), f(e_2), \cdots, f(e_n)\}$$

Remark 1.3.1 *The family* $\{f(e_1), f(e_2), \cdots, f(e_n)\}$ *is not necessarily free, so the family* $\{f(e_1), f(e_2), \cdots, f(e_n)\}$ *is therefore not a basis of* $range(f)$

Theoreme 1.3.1 *Let* $f : E \to F$ *be a linear map with* $dim(E) = n$ *(finite). So :*

$$dim(E) = dim\,(null(f)) + dim\,(range(f))\,.$$

Proposition 1.3.2 *Let* $f : E \to F$ *be a linear map, if* $dim(E) = dim(F) = n$, *then the following properties are equivalent :*

1. f *is bijective.*
2. f *is injective.*
3. f *is surjective.*

1.4 Rank of a linear map

Definition 1.4.1 *Let* $f : E \to F$ *be a linear map. We call* ***rank*** *of* f *the dimension of* $range(f)$, $rank(f) = dim\,(Im(f)) = dim\,(f(E))$.

$$dimE = rank(f) + nullity(f)$$

such that $nullity(f)$=*dimension of its nullspace Null(f).*

Remark 1.4.1 *If* B *is a basis of* E *then the rank of* f *is the number of linearly independent vectors in* $f(B)$, *the image by* f *of the basis* B.

Exercice 1.4.1 *Let* n *be a natural number and let* f *be a linear map defined by* $f(P) = P', \quad \forall P \in \mathbb{R}_n[X]\,.$

1. *Show that* f *is an endomorphism of* $\mathbb{R}_n[X]\,.$
2. *Determine the nullspace of* f.
3. *Deduce the dimension of* $range(f)$.
4. *Check that* $range(f) = \mathbb{R}_{n-1}[X]\,.$

2 Matrices

Definition 2.0.1 *A matrix is a two-dimensional array of numbers. An $m-$rows-by-n-columns (abbreviated $m \times n$) matrix A is represented with a block of numbers :*

$$A = (a_{ij})_{m\times n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

where a_{ij} is the entry in the $i-th$ row and the $j-th$ column of A.

Example 2.0.1 $n = 2$, $p = 3$:

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

$a_{12} = 2, \quad a_{23} = 5$

Definition 2.0.2
1. *Two matrices are equal when they have the same size and the corresponding coefficients are equal.*
2. *The set of n row and p column matrices with coefficients in $\mathbb{K}$ is denoted $M_{n,p}(\mathbb{K})$. The elements of $M_{n,p}(\mathbb{R})$ are called real matrices.*

2.1 Types of matrices

Different types of matrices are given below :

1. **Row Matrix :** A Matrix having only one row is called a **Row Matrix**, E.g. $(a_{11}, a_{12}, \cdots, a_{1p})$.
2. **Column Matrix :** A Matrix having only one Column is called a **Column Matrix**,E.g.
$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$
3. **Null Matrix :** $A = (a_{ij})_{m\times n}$such that $a_{ij} = 0, \forall i$ and j. Then A is called a **Zero Matrix** and it is denoted by $0_{m\times n}$.
4. **Rectangular Matrix :**If $A = (a_{ij})_{m\times n}$, and $m \neq n$ then the matrix A is called a **Rectangular Matrix**.
5. **Square Matrix :**If $A = (a_{ij})_{m\times n}$, and $m = n$ then the matrix A is called a **Square Matrix**.
6. **Identity Matrix or Unit matrix**
$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$
n is the dimension of the matrix.

7. **Lower triangular matrix,L**

$$L = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix}$$

8. **Upper triangular matrix,U**

$$U = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix}$$

9. **Diagonal matrix** : A Square Matrix is said to be diagonal matrix, if $a_{ij} = 0$ for $i \neq j$ i.e. all the elements except the principal diagonal elements are zeros.
Note :

a/ Diagonal matrix is both lower and upper triangular.

b/ If $d_1, d_2, \cdots, d_n$ are the diagonal elements in a diagonal matrix it can be represented as $diag\,(d_1, d_2, \cdots, d_n)$. E.g. $diag(3, 2, -1) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

10. A square matrix A is said to be symmetrical if : $a_{ij} = a_{ji}$ for all i different from j

3 Operations on matrices

3.1 Transpose of a Matrix

The transpose of an $m \times n$ matrix $A = (a_{ij})$ is defined as the $n \times m$ matrix $B = (b_{ij})$, with $b_{ij} = a_{ji}$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. The transpose of A is denoted by A^T.

Definition 3.1.1 *That is, by the transpose of an $m \times n$ matrix A, we mean a matrix of order $n \times m$ having the rows of A as its columns and the columns of A as its rows.*

For example, if $A = \begin{pmatrix} 3 & -2 & 5 \\ 1 & 7 & 0 \end{pmatrix}$ then $A^T = \begin{pmatrix} 3 & 1 \\ -2 & 7 \\ 5 & 0 \end{pmatrix}$ Thus, the transpose of a row vector is a column vector and vice-versa.

Theoreme 3.1.1 *For any matrix A, we have $(A^T)^T = A$.*

3.2 Addition of Matrices

Definition 3.2.1 *Let A and B be two matrices of the same size $n \times p$. Then the sum $A + B$ is defined to be the matrix $C = A + B$ of size $n \times p$ with $c_{ij} = a_{ij} + b_{ij}$.*
In other words, we sum coefficient by coefficient.

Example 3.2.1 *If*

$$A = \begin{pmatrix} 3 & -2 \\ 1 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 5 \\ 2 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} -5 \\ 2 \end{pmatrix}.$$

Then

$$A + B = \begin{pmatrix} 3+0 & -2+5=3 \\ 1+2=3 & 7-1=6 \end{pmatrix}.$$

$A + C$ *is not defined.*

3.3 Multiplying a Scalar to a Matrix

Let $A = (a_{ij})$ of $M_{np}(\mathbb{K})$, for a scalar α, we define αA by the matrix (αa_{ij}).

Example 3.3.1 *If*

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

and $\alpha = -3$. *Then*

$$\alpha A = -3A = \begin{pmatrix} -9 & -3 & -6 \\ 0 & 12 & 6 \end{pmatrix}$$

The matrix $(-1)A$ *is the opposite of* A *and is denoted* $-A$. *The difference* $A - B$ *is defined by* $A + (-B)$.

Below are some basic algebraic properties of matrix addition/scalar multiplication.

Theoreme 3.3.1 *Let* A, B, C *be matrices of the same size and let* α, β *be scalars. Then*

(a) $A + B = B + A$,

(d) $\alpha(A + B) = \alpha A + \alpha B$

(b) $(A + B) + C = A + (B + C)$,

(e) $(\alpha + \beta)A = \alpha A + \beta A$,

(c) $A + 0 = A$,

(f) $\alpha(\beta A) = (\alpha\beta)A$

3.4 Multiplication of matrices / Product

Matrices can be multiplied only if the number of columns of the left matrix equals the number of rows of the right matrix. In other words, an n-by-p matrix on the left can only be multiplied by an p-by-q matrix on the right. The resulting matrix will be n-by-q.
In general, an element in the resulting product matrix, say in row i and column j, is obtained by multiplying and summing the elements in row i of the left matrix with the elements in column j of the right matrix.

Definition 3.4.1 ***Product of two matrices*** *Let* $A = (a_{ij})$ *be a matrix* $n \times p$ *and* $B = (b_{ij})$ *be a matrix* $p \times q$. *Then the product* $C = AB$ *is a* $n \times q$ *matrix whose coefficients* c_{ij} *are defined by :*

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

The coefficient can be written in a more elaborate way :

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} + \cdots + a_{ip}b_{pj}.$$

It is convenient to arrange the calculations as follows :

$$\begin{array}{cc} & \begin{pmatrix} & & b_{1j} & \\ & & \times & \\ & & \times & \\ & & \times & \end{pmatrix} \leftarrow B \\ A \rightarrow \begin{pmatrix} & & & \\ & & & \\ a_{i1}\times & \times & \times & \\ & & & \end{pmatrix} & \begin{pmatrix} & & \times & \\ & & \times & \\ \times & \times & c_{ij} & \\ & & & \end{pmatrix} \leftarrow AB \end{array}$$

Example 3.4.1 *1.*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix}$$

A is an $2-by-3$ matrix, B is an $3-by-2$ matrix, then AB is an $2-by-2$ matrix and we have :

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \leftarrow B \\ A \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \leftarrow AB \end{array}$$

$$c_{11} = 1 \times 1 + 2 \times (-1) + 3 \times 1 = 2, \quad c_{12} = 1 \times 2 + 2 \times 1 + 3 \times 1 = 7$$

$$c_{21} = 2 \times 1 + 3 \times (-1) + 4 \times 1 = 3, \quad c_{22} = 2 \times 2 + 3 \times 1 + 4 \times 1 = 11.$$

Then

$$AB = \begin{pmatrix} 2 & 7 \\ 3 & 11 \end{pmatrix}$$

2. $u = (a_1 a_2 \cdots a_n), \quad v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

uv is a matrix of size 1×1 whose single coefficient is $a_1b_1 + a_2b_2 + \cdots + a_nb_n$. This number is called the scalar product of vectors u and v.

Remark 3.4.1 *1. $AB = 0$ does not imply $A = 0$ or $B = 0$. Let*

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix}$$

Calculate AB.

2. *$AB = AC$ does not imply $B = C$. Let*

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -1 \\ 5 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 5 \\ 5 & 4 \end{pmatrix}$$

Calculate AB and AC

Proposition 3.4.1 *Let $A, B, C \in M_n(\mathbb{K})$, be square matrices of order n, then*

1. *$A(BC) = (AB)C$ (associativity)*
2. *$A(B + C) = AB + AC$ (left distributivity) ;*
3. *$(B + C)A = BA + CA$(right-hand distributivity)*
4. *$\exists I_n \in M_n$ such that $AI = IA = A$*
5. *Matrix multiplication is generally not commutative.*

Warning ! If A and B don't switch, i.e. if $AB \neq BA$

$$(A + B)^2 = (A + B)(A + B) = A^2 + BA + AB + B^2 \neq A^2 + 2AB + B^2$$

4 Row echelon form

4.1 Elementary row operations

Elementary row operations are used to transform a system of linear equations into a new system that has the same solutions as the original one (i.e., into an equivalent system).
A system of n linear equations in m unknowns is written in matrix form as $Ax = b$,where A is the $n \times m$ matrix of coefficients ; x is the $m \times 1$ vector of unknowns and b is the $n \times 1$ vector of constants.
Our goal is to begin with the matrix A and apply operations that respect row equivalence until we have a matrix in Reduced Row Echelon Form (RREF). The three elementary row operations are :

1. (**Row Swap**) Exchange any two rows.
2. (**Scalar Multiplication**) Multiply any row by a constant.
3. (**Row Sum**) Add a multiple of one row to another row.

4.1.1 Row echelon form

Definition 4.1.1 *A matrix $A \in M_{m,n}(\mathbb{K})$, of order $m \times n$ and with coefficients in a field $\mathbb{K}$, is said to be in the row echelon form if the number of zero coefficients starting each row increases as we pass from a row R_i to a row R_j, for $i < j$.*
*The first non-zero coefficient in a row of a matrix is called **pivot**.*

Example 4.1.1 *Consider the following matrices*

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 2 & 3 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 2 & 3 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 & 5 \end{pmatrix}$$

and

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

1. *In matrix A, row R_2 begins with 0, but row R_3 begins with -1. So, going from row R_2 to row R_3, the number of zeros is not increasing. Thus, the matrix A is not in (REF).*
2. *In matrix B, rows R_4 and R_5 begin with the same number of 0, which equals to 4. So, as we move from line R_4 to line R_5, the number of zeros is constant, i.e. it is not increasing. Thus, the matrix B is not in (REF).*
3. *The number of zero coefficients starting the rows of matrix C is increasing, from row to row. Thus, matrix C is in (REF).*

4.1.2 Row Reduced Form of a Matrix

Definition 4.1.2 *A matrix A is said to be in the row reduced form if*

1. *the first non-zero entry in each row of A is 1 ;*
2. *the column containing this 1 has all its other entries zero.*
   A matrix in the row reduced form is also called a row reduced matrix.

Example 4.1.2

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

1. *In the matrix M, the pivot of row 4 is $m_{4,6} = 1$, but it is not the only non-zero coefficient in its column, since $m_{3,6} = 1$. So matrix M is not reduced.*
2. *The matrix N is reduced, since all the pivots are the only non-zero coefficients in their respective columns.*

Every matrix can be put in row echelon form by applying a sequence of elementary row operations.

4.1.3 Method to get the row-reduced echelon form of a given matrix

Let A be an $m \times n$ matrix. Then the following method is used to obtain the row-reduced echelon form of the matrix A.

1. **Step 1 :** Consider the first column of the matrix A. If all the entries in the first column are zero, move to the second column.
   Else, find a row, say ith row, which contains a non-zero entry in the first column. Now,

interchange the first row with the ith row. Suppose the non-zero entry in the $(1,1)$-position is $\alpha \neq 0$. Divide the whole row by α so that the $(1,1)$-entry of the new matrix is 1. Now, use the 1 to make all the entries below this 1 equal to 0.

2. **Step 2 :** Ignore the first row and first column. Start with the lower $(m-1)\times(n-1)$ submatrix of the matrix obtained in the first step and proceed as in step 1.
3. **Step 3 :** Keep repeating this process till we obtain an equivalent where all the entries below a particular row, say r, are zero.
   The integer r is the largest integer such that $a_{rr} \neq 0$ and $a_{ij} = 0$ for $i \geq r+1$.
   The final matrix is the **row-reduced echelon form of the matrix A.**

Example 4.1.3 *Let*

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Let's proceed with some elementary operations to produce matrix in the row echelon form

$$\begin{array}{lcc} R_2 \leftrightarrow R_1 & A \sim & \begin{pmatrix} 1 & 0 & -1 & -1 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 \end{pmatrix} \\ R_2 - 2R_1 & \sim & \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 4 \\ 0 & 1 & 1 & 0 \end{pmatrix} \\ R_3 - R_2 & \sim & \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & -2 & -4 \end{pmatrix} \\ 2R_2, \quad 3R_3 & \sim & \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 6 & 8 \\ 0 & 0 & -6 & -12 \end{pmatrix} \\ R_3 + R_2 & \sim A_5 = & \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 6 & 8 \\ 0 & 0 & 0 & -4 \end{pmatrix} \end{array}$$

Note that the reduced form of a matrix in a row echelon form is obtained using the following steps :

1. *Multiply rows R_i of non-zero pivots a_i by $\lambda = \frac{1}{a_i}$, giving pivots all equal to 1.*
2. *We proceed with elementary operations on the rows, starting from the bottom of the matrix, to eliminate the coefficients in the column of each pivot.*

1. *To obtain pivots equal to 1, we perform the elementary operation $-\frac{1}{4}R_3$ on A_5.*

$$A_5 = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 6 & 8 \\ 0 & 0 & 0 & -4 \end{pmatrix} \overset{-\frac{1}{4}R_3}{\Longrightarrow} A_6 = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 6 & 8 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2. *Elimination of coefficients above the pivot in the pivot column. We then perform the following elementary operations :* $R_2 - 8R_3$ *and* $R_1 + R_3$.

$$A_6 = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 6 & 8 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{R_1+R_3}{\overset{R_2-8R_3}{\Longrightarrow}} A_7 = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We then obtain the matrix A_7*, which is the* ***row reduced form*** *of the matrix* A.

4.2 Rank of a Matrix

Definition 4.2.1 ***Row rank of a Matrix***
The number of non-zero rows in the row reduced form of a matrix is called the row-rank of the matrix.
For a matrix A*,we write* '$row - rank(A)'$ *to denote the row-rank of* A.

Example 4.2.1 *Find the rank of the matrix* $A = \begin{pmatrix} 2 & 3 & 7 \\ 3 & -2 & 4 \\ 1 & -3 & -1 \end{pmatrix}$ *by reducing it to Echelon form.*
Solution : *Applying row transformations on* A.

$$R_1 \leftrightarrow R_3 \qquad A \sim \begin{pmatrix} 1 & -3 & -1 \\ 3 & -2 & 4 \\ 2 & 3 & 7 \end{pmatrix}$$

$$R_2 = R_2 - 3R_1, \quad R_3 = R_3 - 2R_1 \qquad A \sim \begin{pmatrix} 1 & -3 & -1 \\ 0 & 7 & 7 \\ 0 & 9 & 9 \end{pmatrix}$$

$$R_2 = R_2/7, \quad R_3 = R_3/9 \qquad \sim \begin{pmatrix} 1 & -3 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$R_3 = R_3 - R_2 \qquad \sim \begin{pmatrix} 1 & -3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

This is the Echelon form of matrix A. *The rank of a matrix* A*=Number of non-zero rows =2.*

Note
In the previous example we have $A \sim A_7$,Consequently, the rank of the matrix A is equal to 3.

5 The inverse of a matrix

An $n \times n$ matrix A is invertible if there is a matrix B such that $AB = BA = I_n$.
In that case, B is the inverse of A and we write $A^{-1} = B$.

Theoreme 5.0.1 *Suppose* A *and* B *are invertible. Then :*

1. A^{-1} *is invertible and* $(A^{-1})^{-1} = A$.
2. A^T *is invertible and* $(A^T)^{-1} = (A^{-1})^T$.
3. AB *is invertible and* $(AB)^{-1} = B^{-1}A^{-1}$.

5.1 Solving systems using matrix inverse

1. To solve $Ax = b$, we do row reduction on $[A \mid b]$.
2. To solve $AX = I$, we do row reduction on $[A \mid I]$.
3. To compute A^{-1}

a/ Form the augmented matrix $[A \mid I]$.

b/ Compute the reduced echelon form.

c/ If A is invertible, the result is of the form $[I \mid A^{-1}]$.

Example 5.1.1 *Let the system :*

$$\begin{aligned} 3x_3 &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ \frac{1}{3}x_1 + 2x_2 &= 3 \end{aligned}$$

First we write the system as an augmented matrix :

$$(A \mid b) = \begin{pmatrix} 0 & 0 & 3 & \mid 9 \\ 1 & 5 & -2 & \mid 2 \\ \frac{1}{3} & 2 & 0 & \mid 3 \end{pmatrix}$$

$$R_1 \leftrightarrow R_3 \qquad (A \mid b) \sim \begin{pmatrix} \frac{1}{3} & 2 & 0 & \mid 3 \\ 1 & 5 & -2 & \mid 2 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$3R_1 \qquad \sim \begin{pmatrix} 1 & 6 & 0 & \mid 9 \\ 1 & 5 & -2 & \mid 2 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$R_2 = R_2 - R_1 \qquad \sim \begin{pmatrix} 1 & 6 & 0 & \mid 9 \\ 0 & -1 & -2 & \mid -7 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$-R_2 \qquad \sim \begin{pmatrix} 1 & 6 & 0 & \mid 9 \\ 0 & 1 & 2 & \mid 7 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$R_1 = R_1 - 6R_2 \qquad \sim \begin{pmatrix} 1 & 0 & -12 & \mid -33 \\ 0 & 1 & 2 & \mid 7 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$\frac{1}{3}R_3 \qquad \sim \begin{pmatrix} 1 & 0 & -12 & \mid -33 \\ 0 & 1 & 2 & \mid 7 \\ 0 & 0 & 1 & \mid 3 \end{pmatrix}$$

$$R_1 = R_1 + 12R_3 \qquad \sim \begin{pmatrix} 1 & 0 & 0 & \mid 3 \\ 0 & 1 & 2 & \mid 7 \\ 0 & 0 & 3 & \mid 9 \end{pmatrix}$$

$$R_2 = R_2 - 2R_3 \qquad \sim \begin{pmatrix} 1 & 0 & 0 & \mid 3 \\ 0 & 1 & 0 & \mid 1 \\ 0 & 0 & 1 & \mid 3 \end{pmatrix}$$

Now we're in RREF and can see that the solution to the system is given by $x_1 = 3, \quad x_2 = 1$ *and* $x_3 = 1$ *; it happens to be a unique solution. Notice that we kept track of the steps we were taking ; this is important for checking work !*

Example 5.1.2 *Find the inverse of* $\begin{pmatrix} 2 & 0 & 1 \\ 0 & -1 & -3 \\ -2 & 4 & 3 \end{pmatrix}$

Solution

$$(A \mid I_3) = \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -3 & 0 & 1 & 0 \\ -2 & 4 & 3 & 0 & 0 & 1 \end{array}\right)$$

$$R_3 = R_3 + R_1 \qquad (A \mid I_3) \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -3 & 0 & 1 & 0 \\ 0 & 4 & 4 & 1 & 0 & 1 \end{array}\right)$$

$$R_3 = R_3 + 4R_2 \qquad \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -3 & 0 & 1 & 0 \\ 0 & 0 & -8 & 1 & 4 & 1 \end{array}\right)$$

$$R_2 = 8R_2 - 3R_3 \qquad \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -8 & 0 & -3 & -4 & -3 \\ 0 & 0 & -8 & 1 & 4 & 1 \end{array}\right)$$

$$R_1 = 8R_1 + R_3 \qquad \sim \left(\begin{array}{ccc|ccc} 16 & 0 & 8 & 9 & 4 & 1 \\ 0 & -8 & 0 & -3 & -4 & -3 \\ 0 & 0 & -8 & 1 & 4 & 1 \end{array}\right)$$

$$R_1 = \frac{1}{16}8R_1, \quad R_2 = \frac{-1}{8}R_2, \quad R_3 = \frac{-1}{8}R_3 \qquad \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 8 & 9/16 & 1/4 & 1/16 \\ 0 & 1 & 0 & 3/8 & 1/2 & 3/8 \\ 0 & 0 & 1 & -1/8 & -1/2 & -1/8 \end{array}\right)$$

The inverse is the right side.

$$\begin{pmatrix} 9/16 & 1/4 & 1/16 \\ 3/8 & 1/2 & 3/8 \\ -1/8 & -1/2 & -1/8 \end{pmatrix}$$