Chapter 4

Matrices

Matrices are perhaps the most important concept in linear algebra and are used extensively in many areas of science and engineering.

4.1 Matrix Representation of Linear Mappings

Consider the linear mapping $L : V_n(F) \to W_m(F)$ and suppose we select a basis $\{v_1, v_2, \cdots, v_n\}$ for $V_n(F)$ and a basis $\{w_1, w_2, \cdots, w_m\}$ for $W_m(F)$.

We have already seen how to represent a vector $v \in V_n(F)$ in the basis $\{v_1, v_2, \cdots, v_n\}$ as a column vector, say $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ representing the linear combination $v = \sum_{i=1}^{n} x_i v_i$. If w is the image of v under L i.e. $w = L(v)$, we can also find a representation of w in the basis $\{w_1, w_2, \cdots, w_m\}$ in the form of a column vector, say $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ representing the linear combination $w = \sum_{j=1}^{m} y_j w_j$ so that X represents v and Y represents $w = L(v)$. In this section, we want to find a representation of the mapping L describing the relationship between X and Y. Since $\{L(v_1), L(v_2), \cdots, L(v_n)\}$ are vectors of $W_m(F)$, we can express uniquely each one of them in the basis $\{w_1, w_2, \cdots, w_m\}$ as:

$L(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m$
$L(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m$
$\vdots$
$L(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m$

This set of equations describes the effect of L on the basis of the domain and can be summarized in a structure called **matrix** as:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & a_{11} \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

This matrix A allows to represent $w = L(v)$ in the form $Y = AX$ where Y represents w, A represents L and X represents v. It allows to describe the effect of L on any vector of the domain.

Notice that A is organized in **rows** and **columns**. The number of rows is the dimension of the codomain (m) and the number of columns is the dimension of the domain (n) and the matrix A is said to be an $m \times n$ matrix.

The *ith* column of A is the representation of the image $L(v_i)$ in the basis $\{w_1, w_2, \cdots, w_m\}$ i.e. components of $L(v_i)$ are written as the *ith* column of A. The elements of the matrix A are usually called entries of the matrix and the general entry is denoted a_{ij} and is a scalar of F.

Example 56 *Consider the linear mapping $T : \mathbb{R}^2(\mathbb{R}) \rightarrow \mathbb{R}^2(\mathbb{R})$ defined by $T(x, y) = (x, 2y)$. To find a matrix representation A of T , we need first to select a basis for $\mathbb{R}^2(\mathbb{R})$. Let us use the standard basis $\{e_1, e_2\}$ for simplicity. The first column of A is the representation of $T(e_1) = T(1, 0) = (1, 0) = e_1$ or $T(e_1) = 1.e_1 + 0.e_2$ while the second column is the representation of $T(e_2)$ i.e. $T(e_2) = T(0, 1) = (0, 2) = 0e_1 + 2e_2$ leading to $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Now if select a new basis as $\{v_1, v_2\}$ where $v_1 = (1, 0)$ and $v_2 = (1, 1)$, we will get $T(v_1) = T(1, 0) = (1, 0) = v_1$ and $T(v_2) = T(1, 1) = (1, 2) = -1.v_1 + 2.v_2$ leading to a second matrix representation as $B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$*

The matrix representation of a linear mapping is **unique** once a pair of bases is selected for the domain and the codomain. However this matrix representation changes if either or both of the bases are changed.

Example 57 *Consider $L : \mathbb{R}^2(\mathbb{R}) \rightarrow \mathbb{R}^3(\mathbb{R})$ defined by $L(x, y) = (0, x, x + y)$ and the standard basis. Notice that the standard basis for $\mathbb{R}^2(\mathbb{R})$ differs from the standard basis for $\mathbb{R}^3(\mathbb{R})$: $e_1 = (1, 0), e_2 = (0, 1)$ for $\mathbb{R}^2(\mathbb{R})$ and $e'_1 = (1, 0, 0), e'_2 = (0, 1, 0), e'_3 = (0, 0, 1)$ for $\mathbb{R}^3(\mathbb{R})$. Now $L(e_1) = L(1, 0) = (0, 1, 1) = 0.\ e'_1 + 1.\ e'_2 + 1.\ e'_3$, $L(e_2) = L(0, 1) = (0, 0, 1) = 0.\ e'_1 + 0.\ e'_2 + 1.\ e'_3$ leading to the matrix representation $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$*

4.2 Operations on Matrices

Matrices can be used to do operations on linear mappings in a simpler and more efficient way: addition, scalar multiplication, compositions,etc...

4.2.1 Addition of matrices

Definition 58 *Given the $m \times n$ matrices A, B we define their sum as the matrix $C = A + B$ where $c_{ij} = a_{ij} + b_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

Clearly the addition of matrices is defined only if they have the same dimensions (rows and columns). The operation is performed by adding entries term to term.

Example 59 *Given* $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 3 & 2 \\ -2 & 5 & 1 \end{pmatrix}$ *then* $C = A + B = \begin{pmatrix} 1+2 & 2+3 & -1+2 \\ 2+(-2) & 1+5 & 4+1 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 1 \\ 0 & 6 & 5 \end{pmatrix}$

The addition of two matrices is the representation of the addition of two linear mappings $L : V_n(F) \to W_m(F)$ and $M : V_n(F) \to W_m(F)$.

4.2.2 Scalar multiplication

Definition 60 *Given the $m \times n$ matrix A and the scalar $\alpha \in F$, the scalar multiplication of A is defined as the $m \times n$ matrix $C = \alpha A$ where $c_{ij} = \alpha a_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

Every entry of A is multiplied by the scalar α to obtain the matrix C. It allows, in particular, to define the substraction operation: To compute $A-B$, we compute the scalar multiple $(-B) = (-1)B$, followed by the addition $A+(-B)$. In practice we compute $A - B$ by substracting entries term to term.

The set of all $m \times n$ matrices over the field F along with the two operations addition and scalar multiplication satisfies all the axioms of a vector space (Check it) with the zero vector being the **zero matrix** i.e. a matrix where all entries are the zer0 scalar 0.

Example 61 $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \end{pmatrix} - \begin{pmatrix} 2 & 3 & 2 \\ -2 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 1-2 & 2-3 & -1-2 \\ 2-(-2) & 1-5 & 4-1 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -3 \\ 4 & -4 & 3 \end{pmatrix}$

4.2.3 Matrix multiplication

Definition 62 *Given the $m \times n$ matrix A and the $n \times p$ matrix B, we define the product of as the $m \times p$ matrix $C = AB$ where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ where $1 \leq i \leq m$ and $1 \leq j \leq p$*

The product AB is defined only if the number of columns of A equals the number of rows of B.Every entry of C is a sum of products of entries of A and B: entry c_{ij} is the sum of products of the ith row entrees of A by the jth column entrees of B.

If the matrix A represents the linear mapping $L : V_n(F) \to W_m(F)$ and the matrix B the linear mapping $M : U_p(F) \to Z_m(F)$, then the matrix $C = AB$ represents the linear mapping $(L \circ M)$. A simple illustration of the product is the following:

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1\times 1 + 2 \times (-1) + (-1) \times 1 & 1 \times 1 + 2 \times 2 + (-1) \times 3 \\ 2 \times 1 + 1 \times (-1) + 0 \times 1 & 2 \times 1 + 1 \times 2 + 0 \times 3 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 1 & 4 \end{pmatrix}$$

4.2.4 Matrix transposition

Definition 63 *Given the $m \times n$ matrix A , we define the matrix transpose of A as the matrix $C = A^T$ where $c_{ij} = a_{ji}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

The matrix transpose of A is obtained by interverting rows and columns i.e. row i becomes column i for $1 \leq i \leq m$.

Given $A = \begin{pmatrix} 1 & 1 & -2 \\ 3 & 1 & 4 \end{pmatrix}$, its matrix transpose is $A^T = \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ -2 & 4 \end{pmatrix}$

4.3 Rank/nullity of a matrix

Because a matrix A is a representation of a linear mapping $L : V_n(F) \to W_m(F)$, we can define as well fundamental subspaces of the matrix A. Recall that for every linear mapping L we defined two fundamental subspaces: the range space $\mathcal{R}(L)$ and null space or Kernel $\mathcal{N}(L)$.

Definition 64 *The range space of the $m \times n$ matrix A is defined as $\mathcal{R}(A) = \{Y/Y = AX\}$ and the dimension of $\mathcal{R}(A)$ is called the rank of A denoted $r(A)$.*

The range space $\mathcal{R}(A)$ is, in fact, nomore than the representation of the range space $\mathcal{R}(L)$.

The equation $Y = AX$ is the representation of $w = L(v)$ where Y is the representation of the image w in some selected basis of $W_m(F)$, X is the representation of some vector $v \in V_n(F)$ in some selected basis for $V_n(F)$ and A the representation of L relative to these two selected bases. The advantage of using $Y = AX$ instead of $w = L(v)$ stems from the fact that the image Y is obtained as a simple product AX where A is an $m \times n$ matrix and X, which a column vector with n components, can be considered as an $n \times 1$ matrix.

Consider again the general form for the matrix representation of $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & a_{11} \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{pmatrix}$ of the mapping $L : V_n(F) \to W_m(F)$. This matrix can be rewritten conveniently as $A = \begin{pmatrix} A_1 & A_2 & \cdots & A_n \end{pmatrix}$ where A_i is the ith column of A. Clearly , as seen previously, A_i is the representation of $L(v_i)$ i.e. the representation of the image of the *ith* vector of the selected basis for $V_n(F)$ in the selected basis for $W_m(F)$. We have established earlier that $\mathcal{R}(L) = [L(v_1), L(v_2), \cdots, L(v_n)]$ where $\{v_1, v_2, \cdots, v_n\}$ is some selected basis for $V_n(F)$. Since A_i is the representation of $L(v_i)$ for $1 \leq i \leq n$ and $\mathcal{R}(A)$ is the representation of $\mathcal{R}(L)$, we can deduce the following important result: $\mathcal{R}(A) = [A_1, A_2, \cdots, A_n]$ i.e. **the range space of a matrix is spanned by the columns of the matrix**. Furthermore, we have established that a span remains unchanged if we remove linearly dependent vectors. It follows that $\mathcal{R}(A) = [A_1, A_2, \cdots, A_n]$ remains unchanged if remove linearly dependent columns. Clearly if we remove linearly dependent columns we are left with a set of columns that are **linearly independent and span** $\mathcal{R}(A)$ which is precisely the definition of a basis. Hence the following result:

Theorem 65 *Given the $m \times n$ matrix A, then the range space $\mathcal{R}(A)$ is spanned by the linearly independent columns of A and its rank $r(A)$ is given by the number of linearly independent columns.*
Proof. *see development above* ■

Checking for linearly independent columns is done in the same manner : set a linear combination of columns to a zero column and solve for the coefficients of the linear combination. The set of columns is linearly independent if and only if all coefficients are zero.

Example 66 *Given the matrix* $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ --1 & 0 & -1 \end{pmatrix}$ *, we want to determine its rank $r(A)$. To do so we search among the columns to determine the number of linearly independent ones. Starting with $\{A_1\}$: $A_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \neq \theta$, then $\{A_1\}$ is linearly independent . Consider now $\{A_1, A_2\}$: $\alpha_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ leading to $\alpha_1 + \alpha_2 = 0$, $\alpha_1 + 2\alpha_2 = 0$ and $-\alpha_1 = 0$, thus $\alpha_1 = \alpha_2 = 0$ and the set $\{A_1, A_2\}$ is linearly independent. Finally check $\{A_1, A_2, A_3\}$: $\alpha_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ leading to $\alpha_1 + \alpha_2 = 0$, $\alpha_1 + 2\alpha_2 - \alpha_3 = 0$ and $-\alpha_1 - \alpha_3 = 0$ or $\alpha_1 = -\alpha_2 = -\alpha_3$ hence $\{A_1, A_2, A_3\}$ linearly dependent so the rank of A is $r(A) = 2$. Clearly we can avoid these computations if we notice that $A_3 = A_1 - A_2$.*

Definition 67 *The null space or Kernel of the $m \times n$ matrix A is defined as $\mathcal{N}(A) = \{X/AX = \theta\}$ and the dimension of $\mathcal{N}(A)$ is called nullity of A and denoted $n(A)$.*

Clearly $\mathcal{N}(A) = \{X/AX = \theta\}$ is the representation of $\mathcal{N}(L) = \{v/L(v) = \theta\}$ if A is the matrix representation of the linear mapping $L : V_n(F) \to W_m(F)$ using a pair of bases for $V_n(F)$ and $W_m(F)$. From the established relationship $r(L) + n(L) = n$, we conclude that $r(A) + n(A) = n$ where n is the number of columns of A. It follows that we can deduce the nullity from the rank and vice-versa.

$\mathcal{N}(A) = \{X/AX = \theta\}$ indicates that $\mathcal{N}(A)$ is constituted of **all** the vectors satisfying the equation $AX = \theta$ which is a short notation for a set of m equations where the unknowns are the n components of X. Equivalently $\mathcal{N}(A)$ is then the set of **all solutions** of the homogeneous set of equations $AX = \theta$ and $n(A)$ is consequently the number of nearly independent solutions of $AX = \theta$.

Example 68 *Consider again* $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ --1 & 0 & -1 \end{pmatrix}$. *We found in the previous example that* $r(A) = 2$. *We can deduce then that* $n(A) = n - r(A) = 3 - 2 = 1$.

Now consider the equation $AX = \theta$ *leading to* $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ --1 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ *or* $\begin{array}{rcrcrcl} x_1 & + & x_2 & & & = & 0 \\ x_1 & + & 2x_2 & - & x_3 & = & 0 \\ - & x_1 & & - & x_3 & = & 0 \end{array}$ *whose solution is* $x_1 = -x_2 = -x_3$. *We have obviously an infinity of solutions but we have only one* ***degree of freedom*** *in the choice of components of* X *indicating that the dimension of* $\mathcal{N}(A)$ *is* 1 *i.e.* $n(A) = 1$

4.4 Nonsingular matrices

Definition 69 *The* $n \times n$ *matrix* A *is* ***nonsingular*** *if and only if there exists an* $n \times n$ *matrix* A^{-1} *, called inverse of* A*, such that* $AA^{-1} = A^{-1}A = I_n$ *where* I_n *is the* $n \times n$ *identity matrix, otherwise it is* ***singular****.*

The identity matrix I_n is an $n \times n$ matrix were all entries are 0 except the diagonal entries which are all $1's$ such as $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, etc...

It must noted that according to the definition, singularity/nonsingularity concerns **only square matrices**, that is matrices where the number of rows equals the number of columns. A nonsingular matrix A is the matrix representation of a nonsingular linear transformation and as such inherits the properties of a nonsingular transformation.

It follows that the following statements are equivalent:

i) A nonsingular

ii) A^{-1} exists

iii) $\mathcal{N}(A) = [\theta]$ or $n(A) = 0$

iv) $r(A) = n$

v) A preserves linear independence

As for linear mappings, we can use any of these statements to check the singularity/nonsingularity of a matrix. For instance we can use statement *iii)* by setting $AX = \theta$ and solving for X : If the **unique solution** is $X = \theta$, A is nonsingular otherwise it is singular.

Example 70 *Consider the matrix* $A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$*; Is it nonsingular?*
Using the rank: testing for linearly independent columns we find $r(A) = 3$*.*
If we use $AX = \theta$*, we have* $\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ *leading to* $x_1 = x_2 = x_3 = 0$*, hence the unique solution is* $X = \theta$ *or* $\mathcal{N}(A) = [\theta]$ *or* $n(A) = 0$ *and* A *is nonsingular.*

Consider the product AB where A is an $n \times n$ matrix and B is an $n \times m$ matrix ; Let $B = \begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix}$ where b_i is the *ith* column of B so we can write $AB = A\begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix} = \begin{pmatrix} Ab_1 & Ab_2 & \cdots & Ab_m \end{pmatrix}$. Since a nonsingular matrix preserves linear independence, we conclude that $r(AB) = r(B)$, that is nonsingular matrices **preserve the rank upon multiplication**.