Matrices

0.0.1 Matrices

A real $m \times n$ **matrix** is a rectangular array

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{0.1}$$

The **diagonal entries** are the a_{ij} with $i = j$, i.e. the a_{ii}; the i**th row** is composed of entries of the form $a_{i1}, a_{i2}, \ldots, a_{in}$, and the j**th column** is composed of entries of the form $a_{1j}, a_{2j}, \ldots, a_{mj}$. (More generally, we may take the entries a_{ij} from an abstract field F or a ring R, but for now we suppose $F = \mathbb{R}$, the real numbers.)

The set of real $m \times n$ matrices is denoted

$$M_{m,n}(\mathbb{R}) \quad \text{or} \quad M_{m\times n}(\mathbb{R}) \quad \text{or} \quad \mathbb{R}^{m\times n} \tag{0.2}$$

Similarly, the set of $m \times n$ matrices over a ring R is denoted $M_{m,n}(R)$ and $M_{m\times n}(R)$, etc., and likewise with matrices over a field F. In the case that $m = n$ we also write $M_n(\mathbb{R})$, and we call a matrix in $M_n(\mathbb{R})$ a **square matrix**.

Example 0.1 *For example, $M_3(\mathbb{R})$ is the set of all real 3×3 matrices, and $M_{5,7}(2\mathbb{Z})$ is the set of all 5×7 matrices with even integer entries.* ∎

We can view these arrays as mathematical objects and impose on the set $\mathbb{R}^{m\times n}$ of them some algebraic structure. First, we define **addition** in $\mathbb{R}^{m\times n}$,

$$+ : \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n} \tag{0.3}$$

as follows: if $A, B \in \mathbb{R}^{m\times n}$, we define $+$ componentwise to be the matrix

$$A + B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \tag{0.4}$$

Remark 0.2 *Note that this will make $\left(\mathbb{R}^{m\times n}, +\right)$ into an abelian group, abelianness coming from the abelianness of the additive group $(\mathbb{R}, +)$–that is, addition of real numbers is commutative, and since matrix addition is defined componentwise, it, too, is commutative.* ∎

The operation of **scalar multiplication** on $M_{m,n}(\mathbb{R})$,

$$\cdot_s : \mathbb{R} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n} \tag{0.5}$$

is also defined componentwise: if $c \in \mathbb{R}$ and $A \in \mathbb{R}^{m\times n}$, then cA is the matrix

$$cA = c\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} := \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix} \tag{0.6}$$

In addition, we can endow the cartesian product of two (possibly different) sets of matrices with a binary **matrix multiplication** function,

$$\cdot : \mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p} \to \mathbb{R}^{m\times p} \tag{0.7}$$

taking an $m \times n$ matrix A and an $n \times p$ matrix B and giving an $m \times p$ matrix C – as long as n is a common integer for $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{n\times p}$, otherwise the operation is undefined. Formally, if $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$, then $\cdot$ is denoted by juxtaposition and is given by

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}b_{k1} & \cdots & \sum_{k=1}^n a_{1k}b_{kp} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{mk}b_{k1} & \cdots & \sum_{k=1}^n a_{mk}b_{kp} \end{pmatrix} \tag{0.8}$$

Remark 0.3 *In general* $\left(\mathbb{R}^{m\times n}, \cdot\right)$ *will not be a group, since not every matrix has a multiplicative inverse. This is clear if* $m \neq n$*. However,* $\left(\mathbb{R}^{m\times n}, +, \cdot\right)$*, is a ring.* ■

For any matrix $A \in \mathbb{R}^{m\times n}$, we may define another matrix, the **transpose** of A, which is the matrix $A^T \in \mathbb{R}^{n\times m}$ whose entries are given by $A^T_{ji} = A_{ij}$, for $1 \leq i \leq m$, $1 \leq j \leq n$. That is,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \longrightarrow \qquad A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ \vdots & \vdots & \ldots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \tag{0.9}$$

Indeed, we may view the transpose of A as a function, $T : \mathbb{R}^{m\times n} \to \mathbb{R}^{n\times m}$, which we will see is a linear transformation.

The **trace** of a *square* matrix $A \in \mathbb{R}^{n\times n}$ is a function $\mathrm{tr} : \mathbb{R}^{n\times n} \to F$ given by

$$\mathrm{tr}(A) = a_{11} + \cdots + a_{nn} = \sum_{i=1}^n a_{ii} \tag{0.10}$$

i.e. by summing the diagonal entries of A.

A square matrix $A \in \mathbb{R}^{n\times n}$ is said to be **invertible** if there exists a matrix $B \in \mathbb{R}^{n\times n}$ such that $AB = I$. That is, A is invertible if it is a unit in the ring $\mathbb{R}^{n\times n}$. The **set of all** $n \times n$ **invertible matrices**, when considered as a *multiplicative group* under the matrix multiplication map, i.e. the group of units of the ring $\mathbb{R}^{n\times n}$, is called the **general linear group**, and denoted $\mathrm{GL}(n, \mathbb{R})$,

$$\begin{aligned} \mathrm{GL}(n,\mathbb{R}) \quad &:= \quad \{A \in \mathbb{R}^{n\times n} \mid A \text{ is invertible}\} && (0.11) \\ &= \quad M_n(\mathbb{R})^\times = (\mathbb{R}^{n\times n})^\times && (0.12) \end{aligned}$$

The notation $\mathrm{GL}_n(\mathbb{R})$ is also used. The identity element in $\mathrm{GL}(n, \mathbb{R})$ is the **identity matrix** $I_n \in \mathbb{R}^{n\times n}$ is given by $(I_n)_{ij} := \delta_{ij}$, that is

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

0.0.2 Special Matrices

To study matrices in further detail, we will need to perform **elementary row and column operations** on them. If $A \in \mathbb{R}^{m\times n}$, any one of the following operations on the rows or columns of A is called an elementary row (resp. column) operation

type 1: interchanging any two rows (resp. columns) of A

type 2: multiplying any row (resp. column) of A by a nonzero scalar

type 3: adding any scalar multiple of a row (resp. column) of A to another row/column

Note that $E \in \mathbb{R}^{m\times m}$ if it performs a row operation, in which case it left-multiplies A, and $E \in \mathbb{R}^{n\times n}$ if it performs a column operation, in which case it right-multiplies A. That is, EA defines a row operation, and AE defines a column operation.

We will need the **Kronecker delta** function

$$\delta : \mathbb{N} \times \mathbb{N} \to \{0, 1\} \tag{0.13}$$

in what follows, which is given by

$$\delta_{ij} = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{if } \ i \neq j \end{cases} \tag{0.14}$$

Using the Kronecker delta function we list here some other important special matrices. These matrices can be defined over any field F, not just the field of reals $\mathbb{R}$. Indeed, they may be defined over any ring R, such as the ring of polynomials $F[x]$ over a field F or indeed the ring of matrices $F^{n\times n}$.

1. A **diagonal matrix** $A \in F^{n\times n}$ is said to be diagonal if $i \neq j \implies A_{ij} = 0$.
2. An **upper triangular** $A \in F^{n\times n}$ satisfies $i > j \implies A_{ij} = 0$.
3. A **symmetric matrix** is a square matrix $A \in F^{n\times n}$ which satisfies $A = A^T$.
4. A **skew-symmetric matrix** $A \in F^{n\times n}$ satisfies $A = -A^T$.
5. An **idempotent** matrix $A \in F^{n\times n}$ satisfies $A^2 = A$. Such a matrix has only 0 and 1 as eigenvalues.
6. A **nilpotent** matrix $A \in F^{n\times n}$ is one for which there is some $k \in \mathbb{N}$ such that $A^k = O$. Such a matrix has only 0 as an eigenvalue.
7. A **scalar matrix** $A \in F^{n\times n}$ is of the form $A = \lambda I_n$ for some scalar $\lambda \in F$. All its diagonal entries are equal, and non-diagonal entries are 0.
8. An **incidence matrix** is a square matrix in which all the entries are either 0 or 1, and for convenience, all the diagonal entries are 0.
9. Two matrices $A, B \in F^{m\times n}$ are said to be **row equivalent matrices** if either one can be obtained from the other by a series of elementary row operations, that is by left-multiplication be a sequence of elementary matrices. Row equivalence is an equivalence relation on $F^{m\times n}$.
10. Two **column equivalent matrices** $A, B \in F^{m\times n}$ are such that one can be obtained from the other by a series of elementary column operations, that is by right-multiplication by a sequence of elementary matrices. Column equivalence is an equivalence relation on $F^{m\times n}$.

11. Two matrices $A, B \in F^{m\times n}$ are said to be **equivalent matrices** if there exist invertible matrices $P \in \mathrm{GL}(m, F)$ and $Q \in \mathrm{GL}(n, F)$ such that $B = PAQ^{-1}$, that is if A and B are both (or either) row and column equivalent.

12. Two square matrices $A, B \in F^{n\times n}$ are said to be **similar matrices** if $\exists P \in \mathrm{GL}(n, F)$ invertible such that $B = PAP^{-1}$.

The **direct sum of square matrices** $B_1, B_2, \ldots, B_k$ are square matrices (not necessarily of equal dimension) is defined recursively as follows: if $B_1 \in F^{n\times n}$ and $B_2 \in F^{m\times m}$, the direct sum of B_1 and B_2 is given by

$$\begin{aligned} B_1 \oplus B_2 &= A \in F^{m+n\times m+n}, \quad A_{ij} = \begin{cases} (B_1)_{ij} & \text{for } 1 \leq i, j \leq m \\ (B_2)_{i-m,j-m} & \text{for } m+1 \leq i, j \leq m+n \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix} \end{aligned}$$

The definition then extends recursively,

$$\begin{aligned} \bigoplus_{i=1}^{k} B_i &= B_1 \oplus B_2 \oplus \cdots \oplus B_k \\ &:= (B_1 \oplus B_2 \oplus \cdots \oplus B_{k-1}) \oplus B_k \\ &= \begin{pmatrix} B_1 & O & \cdots & O \\ O & B_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & B_k \end{pmatrix} \end{aligned}$$

Analogous definitions apply to matrices over a ring R. This definition will play an important role in the Jordan and rational canonical forms of matrices and linear transformations, where we will need to decompose a space into a direct sum of T-invariant subspaces, each with an associated matrix representation of T restricted to that subspace.