

E. Kreyszig, Advanced Engineering Mathematics, 8th ed.

Section 7.1, pgs. 272-278

Lecture: Matrices, Vectors: Algebra of '+'Module: 01

Suggested Problem Set: Suggested Problems : {5, 7 }

August 22, 2008

E. Kreyszig, Advanced Engineering Mathematics, 8th ed.

Section 7.2, pgs. 278-287

Lecture: Matrix Multiplication : Algebra of '·'Module: 01

Suggested Problem Set: Suggested Problems : {3, 5, 8, 13, 19, 20, 22 }

August 22, 2008

Quote of Lecture 1	
Remember where the thought is. I brought all this so you can survive when law is lawless.	
	Gorillaz: Clint Eastwood (2001)

We begin the course with the algebra of matrices. To do this we must do the following:

- Define a mathematical object called a **matrix**.
- Define how groups of **matrices** behave with respect to the binary operations + and ·.

once this is completed we can begin to study how this algebra relates to linear systems and how we should think about solutions to linear problems. You have studied linear equations before in the sense that you have asked the question, for what values of x does the equation,

$$ax = b, \quad a, b \in \mathbb{R}, \quad (1)$$

have a solution? It should be straightforward to see that for $a \neq 0$ the problem has the unique solution,

$$x = \frac{b}{a}, \quad a \neq 0. \quad (2)$$

What about the case where $a = 0$? Well, that is trickier. If $b = 0$ then the value of x doesn't matter. We always have equality, there are infinitely many choice for x ! However, in the case where $b \neq 0$ we have the inconsistent equation, $0 \cdot x = b \neq 0$. There are no values of x , which will satisfy the equation.

However, we are getting ahead of ourselves. The question we need to ask now is given a lot of data how can we systematically organize it? Also, once this is done how can we manipulate groups of them? Without answers to these questions we have no hope of setting up equations (1)-(2) for large data. In what follows will record the definitions of **matrices** and their **algebraic structure**.¹

¹From the perspective of *abstract algebra*, a field that studies collections of objects and their properties with respect to *binary operators*, we would call this sort of structure a *non-commutative ring*.

Definition: Matrix - A *matrix* is a set of objects organized by two indices into a rectangular array. In the case that these objects exist in the set of complex numbers we write $\mathbf{A} \in \mathbb{C}^{m \times n}$, where $n, m \in \mathbb{N}$. At the element level we have that:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}, \text{ where } [\mathbf{A}]_{ij} = a_{ij}, a_{ij} \in \mathbb{C}, \text{ for } i = 1, 2, 3, \dots, m \text{ and } j = 1, 2, 3, \dots, n. \quad (3)$$

- In the case that $n = m$ we call the matrix **square**. Otherwise it is called **rectangular**.
- For a square matrix the entries running from the upper left to the lower right are called the **main diagonal** entries.

Definition: Vector - A *column vector*, or just vector, is matrix of size, $m \times 1$ where $m \in \mathbb{N}$. A *row vector* is matrix of size, $1 \times n$ where $n \in \mathbb{N}$ and if $\mathbf{v} \in \mathbb{C}^{m \times 1}$ or $\mathbf{r} \in \mathbb{C}^{1 \times n}$ and at the element level we have that:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_m \end{bmatrix}, \text{ where } v_i \in \mathbb{C} \text{ for } i = 1, 2, 3, \dots, m. \quad (4)$$

$$\mathbf{r} = \begin{bmatrix} r_1 & r_2 & r_3 & \dots & r_n \end{bmatrix}, \text{ where } r_j \in \mathbb{C}, \text{ for } j = 1, 2, 3, \dots, n. \quad (5)$$

Definition: Scalar - A *scalar* is a matrix whose size is 1×1 . That is, a *scalar* is an element of the complex number system, $s \in \mathbb{C}$.

Defintion: Equality of Matrices - Two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$ are said to be equal if and only if $a_{ij} = b_{ij}$ for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$.

Defintion: Addition and Scalar Multiplication of Matrices - Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{m \times n}$ then $\mathbf{A} + \mathbf{B} = \mathbf{C}$ is defined such that $\mathbf{C} \in \mathbb{C}^{m \times n}$ where $c_{ij} = a_{ij} + b_{ij}$ for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$. Also, let $s \in \mathbb{C}$, then $s\mathbf{A} = \mathbf{C}$, where $c_{ij} = s \cdot a_{ij}$ for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$. Following from these definitions we have the rules for matrix addition and scalar multiplication as:

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$
4. $\mathbf{A} + -1 \cdot \mathbf{A} = \mathbf{0}$
5. $r(\mathbf{A} + \mathbf{B}) = r\mathbf{A} + r\mathbf{B}$
6. $(r + s)\mathbf{A} = r\mathbf{A} + s\mathbf{A}$
7. $r(s\mathbf{A}) = (rs)\mathbf{A}$
8. $1 \cdot \mathbf{A} = \mathbf{A}$

where $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{m \times n}$ and $r, s \in \mathbb{C}$

Defintion: Transposition - Given $\mathbf{A}^T \in \mathbb{R}^{n \times m}$, define the transpose of \mathbf{A} to be the matrix $\mathbf{A}^T \in \mathbb{R}^{n \times m}$, such that:

$$\mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \dots & a_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nm} \end{bmatrix} \quad (6)$$

- If \mathbf{A} is such that $\mathbf{A} = \mathbf{A}^T$ then the matrix \mathbf{A} is called symmetric.
- If \mathbf{A} is such that $-\mathbf{A} = \mathbf{A}^T$ then the matrix \mathbf{A} is called skew-symmetric.

Defintion: Conjugation - Given $\mathbf{A} \in \mathbb{C}^{n \times m}$, define the conjugate of \mathbf{A} to be the matrix $\bar{\mathbf{A}} \in \mathbb{C}^{n \times m}$ such that,

$$\bar{\mathbf{A}} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} & \dots & \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} & \dots & \bar{a}_{2n} \\ \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} & \dots & \bar{a}_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{m1} & \bar{a}_{m2} & \bar{a}_{m3} & \dots & \bar{a}_{mn} \end{bmatrix}. \quad (7)$$

- The bar implies complex conjugation. That is if $c \in \mathbb{C}$ then $c = a + bi$, $a, b \in \mathbb{R}$ and $\bar{c} = a - bi$.

Defintion: Adjoint - Given $\mathbf{A} \in \mathbb{C}^{n \times m}$, define the adjoint of \mathbf{A} to be the matrix $\mathbf{A}^H \in \mathbb{C}^{n \times m}$ such that $\mathbf{A}^H = (\bar{\mathbf{A}})^T = \overline{(\mathbf{A}^T)}$.

- The adjoint is considered as an extension of the transpose to matrices with complex numbers. Sometimes the adjoint is called the Hermitian of a matrix.
- A matrix is called self-adjoint or Hermitian if $\mathbf{A}^H = \mathbf{A}$.

Defintion: Matrix Multiplication - Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times q}$. If $n = p$ then $\mathbf{AB} = \mathbf{C}$ is defined such that $\mathbf{C} \in \mathbb{C}^{m \times q}$ where $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. The general properties for matrix products are:

1. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
2. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
3. $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$
4. $r(\mathbf{AB}) = r(\mathbf{A})\mathbf{B} = \mathbf{A}r\mathbf{B}$
5. $\mathbf{I}_n \mathbf{A} = \mathbf{A} = \mathbf{A} \mathbf{I}_m$

where \mathbf{A} , \mathbf{B} , \mathbf{C} are defined appropriately and $r \in \mathbb{C}$

- It is not necessarily the case that $\mathbf{AB} = \mathbf{BA}$. That is, matrix multiplication does not, in general, commute.
- The identity matrix \mathbf{I}_k is a square matrix such that $[\mathbf{I}_{k \times k}]_{ij} = 1$ if $i = j$ and $[\mathbf{I}_{k \times k}]_{ij} = 0$ if $i \neq j$.
- The inverse matrix of a square matrix \mathbf{A} is the square matrix \mathbf{A}^{-1} such that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Defintion: Inner Product - Given $\mathbf{x} \in \mathbb{C}^{m \times 1}$ and $\mathbf{y} \in \mathbb{C}^{m \times 1}$ define the inner product of \mathbf{x} and \mathbf{y} to be:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^H \mathbf{y} = \sum_{i=1}^m x_i y_i \quad (8)$$

- Using the inner product it is possible to define matrix multiplication as $c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} = \mathbf{a}_i \cdot \mathbf{b}_j$ where \mathbf{a}_i is the i^{th} row of \mathbf{A} and \mathbf{b}_j is the j^{th} column of \mathbf{B} .