E. Kreyszig, Advanced Engineering Mathematics, 9^{th} **ed.** Section N/A, pgs. N/A

Lecture: Basics of Matrices and Their Algebra Module: 01

Suggested Problem Set: {NULL} January 12, 2010

Quote of Lecture Notes One
Bertrand: Everything is vague to a degree you do not realize till you have tried to make it precise.
Bertrand Russell : The Philosophy of Logical Atomism (1918)

Basic Definitions

Definition: Matrix - A *matrix* is a set of *elements* 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 $n, m \in \mathbb{N}$.[1] At the element level we have that:

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

- 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, $n \times 1$ where $n \in \mathbb{N}$. A *row vector* is matrix of size, $1 \times n$ where $n \in \mathbb{N}$. At the element level we have that:

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

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

Definition: Scalar - A *scalar* is a matrix whose size is 1×1. In this case that this scalar is an object from the real numbers we write $a \in \mathbb{R}$.

Definition: 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 $j = 1,2,3,\dots,n$.

[1] Often it is useful to consider elements, which are functions. However, it is traditional and straightforward to first consider matrices of numbers.

Unitary Operations

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

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix} \tag{4}$$

- If $\mathbf{A}$ is such that $\mathbf{A}= \mathbf{A}^{\mathrm{T}}$ then the matrix $\mathbf{A}$ is called symmetric. [2]
- If $\mathbf{A}$ is such that $\mathbf{A}^{\mathrm{T}} = -\mathbf{A}$ then the matrix $\mathbf{A}$ is called skew-symmetric. [3]
- Using the previous definitions one can quickly show that $(\mathbf{A}+\mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}$ assuming that the matrices are such that their addition is well-defined. [4]

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

$$\bar{\mathbf{A}} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} & \cdots & \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} & \cdots & \bar{a}_{2n} \\ \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} & \cdots & \bar{a}_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{m1} & \bar{a}_{m2} & \bar{a}_{m3} & \cdots & \bar{a}_{mn} \end{bmatrix}. \tag{5}$$

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

Definition: Adjoint - Given $\mathbf{A} \in \mathbb{C}^{m \times n}$, define the adjoint or Hermitian of $\mathbf{A}$ to be the matrix $\mathbf{A}^{\mathrm{H}} \in \mathbb{C}^{m \times n}$ such that $\mathbf{A}^{\mathrm{H}} = (\bar{\mathbf{A}})^{\mathrm{T}} = \overline{(\mathbf{A}^{\mathrm{T}})}$. [5]

- 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 if $\mathbf{A}^{\mathrm{H}} = \mathbf{A}$. [6]
- A matrix is called skew-adjoint if $\mathbf{A}^{\mathrm{H}} = -\mathbf{A}$. [7]

Binary Operations

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

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$

[2] It can be shown that the eigenvalues of symmetric matrices are always real numbers.

[3] It can be shown that the eigenvalues of skew-symmetric matrices are always imaginary numbers or the number zero.

[4] From this it follows that a matrix can always be written as the sum of a symmetric and skew-symmetric matrix. To show this note that $\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathrm{T}})$.

[5] It is often the case that the Hermitian is denoted $\mathbf{A}^{\dagger}$.

[6] It can be shown that the eigenvalues of self-adjoint matrices are always real numbers.

[7] It can be shown that the eigenvalues of skew-adjoint matrices are always imaginary numbers or the number zero.

4. $\mathbf{A} + -1 \cdot \mathbf{A} = \mathbf{0}$ where $\mathbf{0}$ denotes an $m \times n$ matrix whose elements are the scalar zero.
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 $r, s \in \mathbb{C}$

Definition: Matrix Product - 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. $\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$

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

- 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 with the scalar identity, i.e. the number one, on the main diagonal. That is $[\mathbf{I}_{k\times k}]_{ij} = 1$ 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 $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

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

$$\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y} = \sum_{i=1}^{n} x_i y_i \tag{6}$$

- Using the inner product it is possible to define matrix multiplication as $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \mathbf{a}_i \cdot \mathbf{b}_j$ where $\mathbf{a}_i$ is the ith row of $\mathbf{A}$ and $\mathbf{b}_j$ is the jth column of $\mathbf{B}$.
- When working with complex <u>vectors</u> then it is typical to define the inner product to be $\mathbf{x}^{\mathrm{H}}\mathbf{y}$. It is rare to multiply matrices with this definition.

Definition: Outer Product - Given $\mathbf{x} \in \mathbb{R}^{n\times 1}$ and $\mathbf{y} \in \mathbb{R}^{n\times 1}$ define the outer product of $\mathbf{x}$ and $\mathbf{y}$ to be $\mathbf{xy}^{\mathrm{T}}$. It is easily verified that this product results in an $n \times n$ matrix.

- If we take on faith that $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ then we can also see that the outer product produces a symmetric matrix. [8]
- When working with complex <u>vectors</u> then it is typical to define the outer product to be $\mathbf{xy}^{\mathrm{H}}$.

[8] To prove the aforementioned equality note that $[\mathbf{AB}]_{ij} = \mathbf{a}_i \cdot \mathbf{b}_j$ thus the $i, j-$element of the transpose of $\mathbf{AB}$ is $\mathbf{a}_j \cdot \mathbf{b}_i$, which is the product of the $j^{th}-$row of $\mathbf{A}$ and $i^{th}-$column of $\mathbf{B}$. Since the $i^{th}-$column of $\mathbf{B}$ is the $i^{th}-$row of $\mathbf{B}^{\mathrm{T}}$ and the $j^{th}-$row of $\mathbf{A}$ is the $j^{th}-$column of $\mathbf{A}^{\mathrm{T}}$ the desired equality follows.