III. First-Order Systems of Ordinary Differential Equations
3. Supplement: Matrices and Vectors

C. David Levermore
Department of Mathematics
University of Maryland

August 2, 2022

Contents

Return to Main Webpage

3. Supplement: Matrices and Vectors

3.1. **Matrix Operations.** An $m\times n$ matrix $\mathbf{A}$ consists of a rectangular array of entries arranged in m rows and n columns

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \big(a_{jk}\big)\,. \tag{3.1}$$

We call a_{jk} the jk-entry of $\mathbf{A}$, m the row-dimension of $\mathbf{A}$, n the column-dimension of $\mathbf{A}$, and $m\times n$ the dimensions of $\mathbf{A}$. The entries of a matrix can be drawn from any set, but in this course they will often be real or complex numbers.

Two matrices are said to be *equal* if they have the same dimensions and their corresponding entries are equal. Specifically, two $m\times n$ matrices $\mathbf{A} = (a_{jk})$ and $\mathbf{B} = (b_{jk})$ are equal if $a_{jk} = b_{jk}$ for every $j = 1, 2, \cdots, m$ and $k = 1, 2, \cdots, n$. In this case we write $\mathbf{A} = \mathbf{B}$. Otherwise we write $\mathbf{A} \neq \mathbf{B}$.

We will use $\mathbf{0}$ to denote any matrix or vector that has every entry equal to zero. A matrix or vector is said to be *nonzero* if at least one of its entries is not equal to zero — i.e. a matrix $\mathbf{A}$ is nonzero if $\mathbf{A} \neq \mathbf{0}$.

Special kinds of matrices include:

- $n\times n$ matrices are called *square matrices*;
- $1\times m$ matrices are called *row vectors*;
- $n\times 1$ matrices are called *column vectors*.

In this text we will use boldface lower case letters like $\mathbf{a}$ and $\mathbf{b}$ to denote column vectors. We will use boldface upper case letters like $\mathbf{A}$ and $\mathbf{B}$ to denote matrices of any dimension.

Remark. Because boldface letters are hard to render by hand, when working by hand you can adopt other notations. For example, you can write

$$\vec{a} \text{ and } \vec{b} \text{ for column vectors and } \vec{A} \text{ and } \vec{B} \text{ for matrices.}$$

This arrow notation has long been used to denote vectors. It and a variant of it with two arrows are commonly used to denote matrices. Alternatively, you can write

$$\underline{a} \text{ and } \underline{b} \text{ for column vectors and } \underline{A} \text{ and } \underline{B} \text{ for matrices.}$$

This underline notation has long been used to tell book typesetters that a letter should be printed in boldface.

Remark. The use of boldface letters or other notational devices to denote vectors and matrices is designed to help you keep track of what the letters represent. In many advanced books such devices are not used, thereby requiring the reader to recall what each letter represents from where it is introduced.

3.1.1. *Matrix Addition of Two Matrices.* Two $m{\times}n$ matrices $\mathbf{A} = (a_{jk})$ and $\mathbf{B} = (b_{jk})$ can be added to create a new $m{\times}n$ matrix sum $\mathbf{A} + \mathbf{B}$, called the *sum* of $\mathbf{A}$ and $\mathbf{B}$, defined by

$$\mathbf{A} + \mathbf{B} = \big(a_{jk} + b_{jk}\big) \,. \tag{3.2}$$

If matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ have the same dimensions then matrix addition satisfies

$$\begin{aligned}
\mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} && \text{—commutativity}\,,\\
(\mathbf{A} + \mathbf{B}) + \mathbf{C} &= \mathbf{A} + (\mathbf{B} + \mathbf{C}) && \text{—associativity}\,,\\
\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} &= \mathbf{A} && \text{—additive identity}\,,\\
\mathbf{A} + (-\mathbf{A}) = (-\mathbf{A}) + \mathbf{A} &= \mathbf{0} && \text{—additive inverse}\,.
\end{aligned}$$

Here the matrix $-\mathbf{A}$ is defined by $-\mathbf{A} = \big(-a_{jk}\big)$ when $\mathbf{A} = \big(a_{jk}\big)$. We define *matrix subtraction* by $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$.

3.1.2. *Scalar Multiplication of a Matrix.* A number α and an $m{\times}n$ matrix $\mathbf{A} = (a_{jk})$ can be multiplied to create a new $m{\times}n$ matrix $\alpha\mathbf{A}$, called the *multiple* of $\mathbf{A}$ by α, defined by

$$\alpha\mathbf{A} = \big(\alpha a_{jk}\big) \,. \tag{3.3}$$

If matrices $\mathbf{A}$ and $\mathbf{B}$ have the same dimensions then scalar multiplication satisfies

$$\begin{aligned}
\alpha(\beta\mathbf{A}) &= (\alpha\beta)\mathbf{A} && \text{—associativity}\,,\\
\alpha(\mathbf{A} + \mathbf{B}) &= \alpha\mathbf{A} + \alpha\mathbf{B} && \text{—distributivity over matrix addition}\,,\\
(\alpha + \beta)\mathbf{A} &= \alpha\mathbf{A} + \beta\mathbf{A} && \text{—distributivity over scalar addition}\,,\\
1\mathbf{A} = \mathbf{A}\,, &\quad -1\mathbf{A} = -\mathbf{A} && \text{—scalar identity}\,,\\
0\mathbf{A} = \mathbf{0}\,, &\quad \alpha\mathbf{0} = \mathbf{0} && \text{—scalar multiplicative nullity}\,.
\end{aligned}$$

3.1.3. *Matrix Multiplication of Two Matrices.* An $l{\times}m$ matrix $\mathbf{A}$ and an $m{\times}n$ matrix $\mathbf{B}$ can be multiplied to create a new $l{\times}n$ matrix $\mathbf{AB}$, called the *product* of $\mathbf{A}$ and $\mathbf{B}$, defined by

$$\mathbf{AB} = \big(c_{ik}\big) \,, \qquad \text{where} \qquad c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \,. \tag{3.4}$$

An important special case is when $l = n = 1$. In that case $\mathbf{A}$ is a row vector, $\mathbf{B}$ is a column vector, and their product $\mathbf{AB}$ is scalar-valued. For example, if

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_m \end{pmatrix} \,, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \,,$$

then (3.4) reduces to

$$\mathbf{ab} = \sum_{j=1}^{m} a_j b_j \,.$$

This special case can be used to understand the general case given by (3.4).

The formula for matrix mulitplication given by (3.4) can be remembered if you think of the $l{\times}m$ matrix $\mathbf{A}$ as given in terms of its l rows by

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_l \end{pmatrix} ,$$

where each $\mathbf{a}_i$ is the row vector given by

$$\mathbf{a}_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} ,$$

and you think of the $m{\times}n$ matrix $\mathbf{B}$ as given in terms of its n columns by

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{pmatrix} ,$$

where each $\mathbf{b}_k$ is the column vector given by

$$\mathbf{b}_j = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{mk} \end{pmatrix} .$$

Then formula (3.4) becomes

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_l\mathbf{b}_1 & \mathbf{a}_l\mathbf{b}_2 & \cdots & \mathbf{a}_l\mathbf{b}_n \end{pmatrix} .$$

In other words, the ik entry of $\mathbf{AB}$ is $\mathbf{a}_i\mathbf{b}_k$, which is just the scalar-valued product of the $1{\times}m$ row vector $\mathbf{a}_i$ with the $m{\times}1$ column vector $\mathbf{b}_k$.

Example. For the matrices

$$\mathbf{A} = \begin{pmatrix} 3 & 5+i6 \\ 2 & i3 \\ 2+i3 & 2-i \end{pmatrix} , \qquad \mathbf{B} = \begin{pmatrix} 2-i5 & -4 \\ -1 & 6+i \end{pmatrix} ,$$

the product $\mathbf{AB}$ exists with

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} 3\cdot(2-i5)+(5+i6)\cdot(-1) & 3\cdot(-4)+(5+i6)\cdot(6+i) \\ 2\cdot(2-i5)+(i3)\cdot(-1) & 2\cdot(-4)+(i3)\cdot(6+i) \\ (2+i3)\cdot(2-i5)+(2-i)\cdot(-1) & (2+i3)\cdot(-4)+(2-i)\cdot(6+i) \end{pmatrix} \\ &= \begin{pmatrix} 12-i20 & 12+i41 \\ 7-i10 & -11+i18 \\ 18-i6 & 5-i16 \end{pmatrix} , \end{aligned}$$

while the product $\mathbf{BA}$ does not exist.

Remark. Notice that for some matrices $\mathbf{A}$ and $\mathbf{B}$, depending only on their dimensions: neither $\mathbf{AB}$ nor $\mathbf{BA}$ exist; exactly one of $\mathbf{AB}$ or $\mathbf{BA}$ exists; or both $\mathbf{AB}$ and $\mathbf{BA}$ exist.

Remark. Notice that if $\mathbf{A}$ and $\mathbf{B}$ are $n\times n$ then $\mathbf{AB}$ and $\mathbf{BA}$ both exist and are $n\times n$, but in general

$$\mathbf{AB} \neq \mathbf{BA}\,!$$

Example. For the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -3 & 9 \\ 1 & -3 \end{pmatrix},$$

we see that

$$\mathbf{AB} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} -3 & 9 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$
$$\mathbf{BA} = \begin{pmatrix} -3 & 9 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} = \begin{pmatrix} 15 & 45 \\ -5 & -15 \end{pmatrix},$$

whereby $\mathbf{AB} \neq \mathbf{BA}$.

Remark. The above example also shows that

$$\mathbf{AB} = \mathbf{0} \quad \text{does not imply that either } \mathbf{A} = \mathbf{0} \text{ or } \mathbf{B} = \mathbf{0}\,!$$

If α is a number and $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are matrices that have the dimensions indicated on the left then matrix multiplication satisfies

$\mathbf{A}$	$\mathbf{B}$	$\mathbf{C}$		
$l \times m$	$m \times n$		$(\alpha\mathbf{A})\mathbf{B} = \alpha(\mathbf{AB})$	—associativity,
$l \times m$	$m \times n$	$n \times k$	$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$	—associativity,
$l \times m$	$m \times n$	$m \times n$	$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$	—left distributivity,
$l \times m$	$l \times m$	$m \times n$	$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$	—right distributivity.

An *identity matrix* is a square matrix $\mathbf{I}$ in the form

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

We will use $\mathbf{I}$ to denote any identity matrix; the dimensions of $\mathbf{I}$ will always be clear from the context. Identity matrices have the property for any $m\times n$ matrix $\mathbf{A}$

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A} \qquad \text{—multiplicative identity}.$$

Notice that the first $\mathbf{I}$ above is $m\times m$ while the second $\mathbf{I}$ is $n\times n$.

Similarly, zero matrices have the property for any $m\times n$ matrix $\mathbf{A}$

$$\mathbf{0A} = \mathbf{0}, \qquad \mathbf{A0} = \mathbf{0} \qquad \text{—multiplicative nullity}.$$

If the first $\mathbf{0}$ is $l\times m$ then the second $\mathbf{0}$ is $l\times n$. If the third $\mathbf{0}$ is $n\times k$ then the fourth $\mathbf{0}$ is $m\times k$.

3.1.4. *Matrix Conjugate.* The conjugate of the $m{\times}n$ matrix $\mathbf{A}$ given by (3.1) is the $m{\times}n$ matrix $\overline{A}$ given by

$$\overline{A} = \big(\overline{a}_{jk}\big)\,. \tag{3.5}$$

If α is a number and $\mathbf{A}$ and $\mathbf{B}$ are matrices that have the dimensions indicated on the left then matrix conjugate satisfies

$$\begin{array}{ccrcl}
\mathbf{A} & \mathbf{B} & & & \\
m\times n & m\times n & \overline{(\mathbf{A}+\mathbf{B})} & = & \overline{\mathbf{A}}+\overline{\mathbf{B}}\,, \\
m\times n & & \overline{(\alpha\mathbf{A})} & = & \overline{\alpha}\,\overline{\mathbf{A}}\,, \\
l\times m & m\times n & \overline{\mathbf{A}\mathbf{B}} & = & \overline{\mathbf{A}}\,\overline{\mathbf{B}} \quad \text{(note no flip)}\,, \\
m\times n & & \overline{(\overline{\mathbf{A}})} & = & \mathbf{A}\,.
\end{array}$$

3.1.5. *Matrix Transpose.* The transpose of the $m{\times}n$ matrix $\mathbf{A}$ given by (3.1) is the $n{\times}m$ matrix A^{T} given by

$$\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}. \tag{3.6}$$

If α is a number and $\mathbf{A}$ and $\mathbf{B}$ are matrices that have the dimensions indicated on the left then matrix transpose satisfies

$$\begin{array}{ccrcl}
\mathbf{A} & \mathbf{B} & & & \\
m\times n & m\times n & (\mathbf{A}+\mathbf{B})^{\mathrm{T}} & = & \mathbf{A}^{\mathrm{T}}+\mathbf{B}^{\mathrm{T}}\,, \\
m\times n & & (\alpha\mathbf{A})^{\mathrm{T}} & = & \alpha\mathbf{A}^{\mathrm{T}}\,, \\
l\times m & m\times n & (\mathbf{A}\mathbf{B})^{\mathrm{T}} & = & \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \quad \text{(note flip)}\,, \\
m\times n & & \big(\mathbf{A}^{\mathrm{T}}\big)^{\mathrm{T}} & = & \mathbf{A}\,.
\end{array}$$

3.1.6. *Hermitian Transpose.* The *Hermitian transpose* of the $m{\times}n$ matrix $\mathbf{A}$ given by (3.1) is the $n{\times}m$ matrix

$$\mathbf{A}^{\mathrm{H}} = \overline{\mathbf{A}}^{\mathrm{T}} = \overline{\mathbf{A}^{\mathrm{T}}}\,. \tag{3.7}$$

If α is a number and $\mathbf{A}$ and $\mathbf{B}$ are matrices that have the dimensions indicated on the left then Hermitian transpose satisfies

$$\begin{array}{ccrcl}
\mathbf{A} & \mathbf{B} & & & \\
m\times n & m\times n & (\mathbf{A}+\mathbf{B})^{\mathrm{H}} & = & \mathbf{A}^{\mathrm{H}}+\mathbf{B}^{\mathrm{H}}\,, \\
m\times n & & (\alpha\mathbf{A})^{\mathrm{H}} & = & \overline{\alpha}\mathbf{A}^{\mathrm{H}}\,, \\
l\times m & m\times n & (\mathbf{A}\mathbf{B})^{\mathrm{H}} & = & \mathbf{B}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}} \quad \text{(note flip)}\,, \\
m\times n & & \big(\mathbf{A}^{\mathrm{H}}\big)^{\mathrm{H}} & = & \mathbf{A}\,.
\end{array}$$

Examples. For the matrices

$$\mathbf{A} = \begin{pmatrix} 3 & 5+i6 \\ 2 & i3 \\ 2+i3 & 2-i \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 2-i5 & -4 \\ -1 & 6+i \end{pmatrix},$$

we have

$$\overline{\mathbf{A}} = \begin{pmatrix} 3 & 5-i6 \\ 2 & -i3 \\ 2-i3 & 2+i \end{pmatrix}, \qquad \overline{\mathbf{B}} = \begin{pmatrix} 2+i5 & -4 \\ -1 & 6-i \end{pmatrix},$$

$$\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 3 & 2 & 2+i3 \\ 5+i6 & i3 & 2-i \end{pmatrix}, \qquad \mathbf{B}^{\mathrm{T}} = \begin{pmatrix} 2-i5 & -1 \\ -4 & 6+i \end{pmatrix},$$

$$\mathbf{A}^{\mathrm{H}} = \begin{pmatrix} 3 & 2 & 2-i3 \\ 5-i6 & -i3 & 2+i \end{pmatrix}, \qquad \mathbf{B}^{\mathrm{H}} = \begin{pmatrix} 2+i5 & -1 \\ -4 & 6-i \end{pmatrix},$$

3.1.7. *Matrix Symmetries.* There are symmetries sssociated with each of the matrix operations of conjugation, transposition, and Hermitian transposition.

- If $\overline{\mathbf{A}} = \mathbf{A}$ then each entry of $\mathbf{A}$ is real and $\mathbf{A}$ is called a *real matrix*.
- If $\overline{\mathbf{A}} = -\mathbf{A}$ then each entry of $\mathbf{A}$ is imaginary and $\mathbf{A}$ is called an *imaginary matrix*.
- If $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$ then $\mathbf{A}$ is called a *symmetric matrix*. Symmetric matrices must be square.
- If $\mathbf{A}^{\mathrm{T}} = -\mathbf{A}$ then $\mathbf{A}$ is called a *skew-symmetric matrix* or an *anti-symmetric matirx*. Skew-symmetric matrices must be square and have diagonal entries that are zero.
- If $\mathbf{A}^{\mathrm{H}} = \mathbf{A}$ then $\mathbf{A}$ is called a *Hermitian matrix*. Hermitian matrices must be square and have diagonal entries that are real.
- If $\mathbf{A}^{\mathrm{H}} = -\mathbf{A}$ then $\mathbf{A}$ is called a *skew-Hermitian matrix* or an *anti-Hermitian matirx*. Skew-Hermitian matrices must be square and have diagonal entries that are imaginary.

Examples of real symmetric matrices are

$$\begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 4 & 2 \\ 2 & -1 \end{pmatrix}.$$

Examples of real skew-symmetric matrices are

$$\begin{pmatrix} 0 & 5 \\ -5 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -2 & -1 \\ 2 & 0 & 5 \\ 1 & -5 & 0 \end{pmatrix}.$$

3.2. **Invertibility and Inverses.** An $n\times n$ matrix $\mathbf{A}$ is said to be *invertible* if there exists another $n\times n$ matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}\,,$$

in which case $\mathbf{B}$ is said to be an *inverse* of $\mathbf{A}$.

Fact. A matrix can have at most one inverse.

Reason. Suppose that $\mathbf{B}$ and $\mathbf{C}$ are inverses of $\mathbf{A}$. Then

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}\,.$$

If $\mathbf{A}$ is invertible then its unique inverse is denoted $\mathbf{A}^{-1}$.

3.2.1. *Properties of Invertibility.* **Fact.** If $\mathbf{A}$ and $\mathbf{B}$ are invertible $n\times n$ matrices and $\alpha \neq 0$ then

- $\alpha\mathbf{A}$ is invertible with $(\alpha\mathbf{A})^{-1} = \dfrac{1}{\alpha}\,\mathbf{A}^{-1}$;
- $\mathbf{AB}$ is invertible with $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (notice the flip);
- $\mathbf{A}^{-1}$ is invertible with $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$;
- $\overline{\mathbf{A}}$ is invertible with $\left(\overline{\mathbf{A}}\right)^{-1} = \overline{\mathbf{A}^{-1}}$;
- $\mathbf{A}^{\mathrm{T}}$ is invertible with $\left(\mathbf{A}^{\mathrm{T}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{T}}$;
- $\mathbf{A}^{\mathrm{H}}$ is invertible with $\left(\mathbf{A}^{\mathrm{H}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{H}}$.

Reason. Each of these facts can be checked by direct calculation. For example, the second fact is checked by

$$(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \left((\mathbf{AB})\mathbf{B}^{-1}\right)\mathbf{A}^{-1} = \left(\mathbf{A}\left(\mathbf{BB}^{-1}\right)\right)\mathbf{A}^{-1} = (\mathbf{AI})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}\,.$$

Fact. If $\mathbf{A}$ and $\mathbf{B}$ are $n\times n$ matrices and $\mathbf{AB}$ is invertible then both $\mathbf{A}$ and $\mathbf{B}$ are invertible with

$$\mathbf{A}^{-1} = \mathbf{B}(\mathbf{AB})^{-1}\,, \qquad \mathbf{B}^{-1} = (\mathbf{AB})^{-1}\mathbf{A}\,.$$

Reason. Each of these facts can be checked by direct calculation.

Fact. If $\mathbf{A}$ is invertible and $\mathbf{AB} = \mathbf{0}$ then $\mathbf{B} = \mathbf{0}$.

Reason.

$$\mathbf{B} = \mathbf{IB} = (\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}\,.$$

Fact. Not all nonzero square matrices are invertible.

Reason. Earlier we gave an example of two nonzero matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{0}$. The previous fact then implies that $\mathbf{A}$ is not invertible.

3.2.2. *Determinants and Invertibility.* The invertability of a square matrix is characterized by its determinant.

Fact. A matrix $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

Reason. This fact is proved in Linear Algebra courses. We will prove it for the special case of $2{\times}2$ matrices. In that case the inverse is easy to compute when it exists. If

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then $\det(\mathbf{A}) = ad - bc$. If $\det(\mathbf{A}) = ad - bc \neq 0$ then $\mathbf{A}$ is invertible with

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

This result follows from the calculation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ab - cd & -ab + ba \\ -cd + dc & -cb + da \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det(\mathbf{A})\mathbf{I}.$$

This calculation also shows that if $\det(\mathbf{A}) = ad - bc = 0$ then $\mathbf{A}$ is not invertible.

The following is a very important fact about determinants.

Fact. If $\mathbf{A}$ and $\mathbf{B}$ are $n{\times}n$ matrices then $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.

Reason. We can prove this fact for $2{\times}2$ matrices by direct calculation. If

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

then

$$\begin{aligned}
\det(\mathbf{AB}) &= \det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right) = \det\left(\begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix} \right) \\
&= (ae + bg)(cf + dh) - (ce + dg)(af + bh) \\
&= aecf + aedh + bgcf + bgdh - ceaf - cebh - dgaf - dgbh \\
&= adeh + bcfg - bceh - adfg = (ad - bc)(eh - fg) = \det(\mathbf{A})\det(\mathbf{B}).
\end{aligned}$$

This remarkable fact is proved for $n{\times}n$ matrices in linear algebra courses.

Remark. Notice that if $\det(\mathbf{AB}) \neq 0$ then the above fact implies that both $\det(\mathbf{A}) \neq 0$ and $\det(\mathbf{B}) \neq 0$. This is consistent with the fact given earlier that if $\mathbf{AB}$ is invertible then both $\mathbf{A}$ and $\mathbf{B}$ are invertible

Other important facts about determinants include the following.

Fact. If $\mathbf{A}$ is an $n{\times}n$ matrix then

$$\det\big(\overline{\mathbf{A}}\big) = \overline{\det(\mathbf{A})}, \qquad \det\big(\mathbf{A}^{\mathrm{T}}\big) = \det(\mathbf{A}), \qquad \det\big(\mathbf{A}^{\mathrm{H}}\big) = \overline{\det(\mathbf{A})}.$$

Reason. It is easy to prove these facts for $2{\times}2$ matrices. These facts are proved for $n{\times}n$ matrices in linear algebra courses.

Remark. These are consistent with the facts given earlier that if $\mathbf{A}$ is invertible then so are $\overline{\mathbf{A}}$, $\mathbf{A}^{\mathrm{T}}$, and $\mathbf{A}^{\mathrm{H}}$.

3.3. **Vector Operations.** Our geometric intuition about *vectors* is grounded upon pictures of arrows that are either drawn in the plane or visualized in three-dimensional space. Most vectors can be thought of as being specified by a positive length and a direction. The exception is the so-called *trivial vector* that has zero length and therefore does not have a direction associated with it.

3.3.1. *Cartesian Coordinates.* Every vector can be associated with a unique point in space. Associate the trivial vector with the origin. Every nontrivial vector can be represented by an arrow that points in the direction of the vector and whose length is the length of the vector. If the "tail" of this arrow is placed at the origin then the "tip" of the arrow will be located at some point in space that is not the origin. Associate the vector with this point. In this way every vector is associated with one point and every point is associated with one vector.

This means that every vector is uniquely specified by the *Cartesian coordinates* of the point associated with it. In the plane those coordinates are given by an ordered pair of real numbers (x, y). In three-dimensional space those coordinates are given by an ordered triple of real numbers (x, y, z). More generally, in n-dimensional space those coordinates are given by an ordered n-tuple of real numbers $(x_1, x_2, \cdots, x_n)$. By identifying vectors with ordered n-tuples, geometric notions about vectors that are grounded upon our intuition in two and three dimensions can be easily extended to higher dimensions.

Because the Cartesian coordinates of any vector are given by an n-tuple of real numbers $(x_1, x_2, \cdots, x_n)$, and because every such n-tuple is associated with the column vector $\mathbf{x}$ and the row vector $\mathbf{x}^{\mathrm{T}}$ given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{x}^{\mathrm{T}} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} .$$

we see that vectors are thereby identified with the *matrices* that we have dubbed column vectors and row vectors. We will use this fact in two ways. First, we will use our geometric intuition about vectors to build a better understanding of matrices. Second, we will use our algebraic understanding of matrices to gain new insights about vectors. Here we will limit our presentation to ideas relevant to our study of first-order systems of ordinary differential equations. These ideas are developed further in *linear algebra* courses. There is no part of mathematics, statistics, science, or engineering in which ideas from linear algebra do not play a major role.

3.3.2. *Vector Addition of Two Vectors.* The most important notion about vectors is that *any two of them can be added.* The geometric intuition is to fix one of the vectors with its tail at the origin while moving the second vector without changing either its length or its direction so that its tail is located at the tip of the first vector. The sum of the two vectors is then the vector associated with the location of the tip of the translated second vector. A picture should convince you that the result does not depend on which vector is held fixed and which vector is moved.

The addition of two vectors can be expressed algebraically in terms of their Cartesian coordinates. If the Cartesian coordinates of the first vector are $(x_1, x_2, \cdots, x_n)$ and the Cartesian coordinates of the second vector are $(y_1, y_2, \cdots, y_n)$ then the Cartesian coordinates of their sum are $(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$. Therefore the algebraic operation of *vector addition* is defined by the formula

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) .$$

This formula confirms that the order in which the vectors are added does not change the result.

The operation of vector addition for n-tuples can be thought of as special case of matrix addition. For example, the n-tuples $(a_1, a_2, \cdots, a_n)$ and $(b_1, b_2, \cdots, b_n)$ are associated with the column vectors

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^{\mathrm{T}} , \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}^{\mathrm{T}} .$$

The vector addition of these n-tuples is defined as

$$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n) .$$

This result is the n-tuple associated with the column vector $\mathbf{a} + \mathbf{b}$.

3.3.3. *Scalar Multiplication of a Vector.* The second important notion about vectors is that *any vector can be rescaled by any real number* α. If $\alpha > 0$ then the geometric intuition is that the direction of the rescaled vector will be the same as for the original vector while its length is changed by a factor of α. If $\alpha < 0$ then the geometric intuition is that the direction of the rescaled vector will be the opposite to that for the original vector while its length is changed by a factor of $|\alpha|$. If $\alpha = 0$ then the rescaled vector will be the trivial vector, which is also called the *zero vector* because its Cartesian coordinates are all zero. Real numbers are commonly called *scalars* in the context of vectors because that is role they play in this rescaling.

This rescaling of a vector can be expressed algebraically in terms of the scalar α and the Cartesian coordinates of the vector. If the Cartesian coordinates of the vector are $(x_1, x_2, \cdots, x_n)$ then the Cartesian coordinates of the rescaled vector are $(\alpha x_1, \alpha x_2, \cdots, \alpha x_n)$. Therefore the algebraic operation of *scalar multiplication* is defined by the formula

$$\alpha \, (x_1, x_2, \cdots, x_n) = (\alpha x_1, \alpha x_2, \cdots, \alpha x_n) .$$

The operation of scalar multiplication for n-tuples can be thought of as special case of scalar multiplication of matrices. For example, the n-tuple $(a_1, a_2, \cdots, a_n)$ is associated with the column vector

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^{\mathrm{T}} .$$

The scalar multiplication of this n-tuple by a number α is defined as

$$\alpha(a_1, a_2, \cdots, a_n) = (\alpha a_1, \alpha b_2, \cdots, \alpha a_n) .$$

This result is the n-tuple associated with the column vector $\alpha \mathbf{a}$.

3.3.4. *Scalar Product of Two Vectors.* The operations described in the preceeding subsections produced vectors. Vector addition took two vectors and produced a vector. Scalar multiplication took a scalar and a vector and produced a vector. Here we describe a so-called *scalar product* that takes two vectors and produces a scalar. If the Cartesian coordinates of two vectors are $(x_1, x_2, \cdots, x_n)$ and $(y_1, y_2, \cdots, y_n)$ then their scalar product is defined by

$$\begin{aligned}(x_1, x_2, \cdots, x_n) \cdot (y_1, y_2, \cdots, y_n) &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= \sum_{k=1}^{n} x_k y_k \,.\end{aligned}$$

This product also goes by the names *dot product* and *inner product*.

The operation of scalar product for n-tuples can be thought of as special case of matrix multiplication. For example, the n-tuples $(a_1, a_2, \cdots, a_n)$ and $(b_1, b_2, \cdots, b_n)$ are associated with the column vectors

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^{\mathrm{T}} , \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}^{\mathrm{T}} .$$

The scalar product of these n-tuples is defined as

$$(a_1, a_2, \cdots, a_n) \cdot (b_1, b_2, \cdots, b_n) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \,.$$

This is also the result of the matrix multiplication $\mathbf{a}^{\mathrm{T}}\mathbf{b}$.

Remark. This shows that the ij^{th} entry of the matrix multiplication $\mathbf{AB}$ is just the scalar product of the vector associated with the i^{th} row of $\mathbf{A}$ with the vector associated with the j^{th} column of $\mathbf{B}$.

Exercises on Matrices and Vectors

(1) Answer each of the following questions.
  (a) If $\mathbf{A}$ is an $m \times n$ matrix, how many columns does it have? How many rows?
  (b) When can you not add two matrices together?
  (c) When can you not multiply two matrices together?
  (d) Is it always the case that $\mathbf{AB} = \mathbf{BA}$?
  (e) What is $\mathbf{I}$? Given an $m \times n$ matrix $\mathbf{A}$ and an $n \times n$ matrix $\mathbf{I}$, what is $\mathbf{AI}$ and what are its dimensions? Does $\mathbf{AI} = \mathbf{IA}$?

Solution

For problems 2- 4, let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 4 & 1 & 5 \\ 1 & 2 & 1 \\ 3 & 1 & 4 \end{pmatrix}.$$

Calculate the following.

(2) $\mathbf{A} + \mathbf{B}$

Solution

(3) $2\mathbf{A} + \mathbf{B}$

Solution

(4) $\mathbf{A} - \mathbf{B}$

Solution

For problems 5- 7, let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 8 & 4 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

Calculate the following.

(5) $\mathbf{AB}$

Solution

(6) $\mathbf{BA}$

Solution

(7) $2\mathbf{A} - \mathbf{AB}$

Solution

For problems 8- 13 let,

$$\mathbf{A} = \begin{pmatrix} 1+i & 2 \\ i & 1-3i \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} i & 0 \\ 0 & 2i \end{pmatrix}.$$

Calculate the following.

(8) $\mathbf{A}^T$

Solution

(9) $\overline{\mathbf{A}}$

Solution

(10) $\mathbf{A}^*$

Solution

(11) $(\mathbf{AB})^T$

Solution

(12) $\mathbf{B}^T\mathbf{A}^T$

Solution

(13) $(\mathbf{AB})^*$

Solution

(14) Show the following properties of invertibility as mentioned in the text. If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ invertible matrices and $\alpha \neq 0$, then
 (a) $(\alpha\mathbf{A})^{-1} = \dfrac{1}{\alpha}\mathbf{A}^{-1}$
 (b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
 (c) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
 (d) $(\overline{\mathbf{A}})^{-1} = \overline{\mathbf{A}^{-1}}$
 (e) $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$
 (f) $(\mathbf{A}^*)^{-1} = (\mathbf{A}^{-1})^*$

Solution

(15) Answer the following questions,
 (a) What is the definition of an inverse of a matrix $\mathbf{A}$?
 (b) If $\det(\mathbf{A}) \neq 0$, is the matrix $\mathbf{A}$ invertible? If so, find $\mathbf{A}^{-1}$ if $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Solution

(16) Show the following properties of determinants as mentioned in the text. If $\mathbf{A}$ is an 2×2 matrix then
 (a) $\det\left(\overline{\mathbf{A}}\right) = \overline{\det(\mathbf{A})}$
 (b) $\det\left(\mathbf{A}^T\right) = \det(\mathbf{A})$
 (c) $\det(\mathbf{A}^*) = \overline{\det(\mathbf{A})}$

Solution

For problems 17- 21 calculate the determinants of the given matrices. State whether they are invertible or not. If invertible, calculate the inverse.

(17) $\begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix}$

Solution

(18) $\begin{pmatrix} 3 & -1 \\ 6 & 1 \end{pmatrix}$

Solution

(19) $\begin{pmatrix} 10 & 1 \\ 9 & 2 \end{pmatrix}$

Solution

(20) $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$

Solution

(21) $\begin{pmatrix} 2 & 3 & 1 \\ -1 & 2 & 1 \\ 4 & -1 & -1 \end{pmatrix}$

Solution

For problem 22- 24 solve the system of equations

(22) $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \end{pmatrix}$

Short Answer
Solution

(23) $\begin{pmatrix} 2 & 3 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ -15 \end{pmatrix}$

Short Answer
Solution

(24) $\begin{pmatrix} 1 & 4 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$

Short Answer
Solution

For problems 25- 26 use row reduction to solve the system of equations

(25) $\begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$

Short Answer
Solution

(26) $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$

Short Answer

Solution

(27) Suppose we associate to every complex number $a+ib$ a matrix

$$\mathbf{Z}(a+ib) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

If z and w are any two complex numbers, verify the following properties:
(a) $\mathbf{Z}(z) + \mathbf{Z}(w) = \mathbf{Z}(z+w)$
(b) $\mathbf{Z}(z)\mathbf{Z}(w) = \mathbf{Z}(zw) = \mathbf{Z}(w)\mathbf{Z}(z)$
(c) $\mathbf{Z}(z)^T = \mathbf{Z}(a-ib) = \mathbf{Z}(\overline{z})$
(d) $\mathbf{Z}(z)\mathbf{Z}(z)^T = (a^2+b^2)\mathbf{I} = \mathbf{Z}(z\overline{z})$
(e) $\det \mathbf{Z}(z) = a^2+b^2 = z\overline{z}$
(f) $\mathbf{Z}(z)^{-1} = \dfrac{1}{a^2+b^2}\mathbf{Z}(z)^T = \mathbf{Z}(z^{-1}), \quad z \neq 0.$

Solution

(28) Recall that the derivative of a matrix valued function $\mathbf{A}(t)$ is defined entrywise,

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \left(\frac{\mathrm{d}a_{ij}}{\mathrm{d}t}\right).$$

If $\mathbf{A}(t)$ and $\mathbf{B}(t)$ are matrix valued functions, show that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{AB}) = \mathbf{A}\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} + \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\mathbf{B}.$$

Solution

(29) Let $\mathbf{A}$ be an *upper triangular matrix*, that is a matrix of the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

What is $\det(\mathbf{A})$? Give conditions on the entries of $\mathbf{A}$ for it to be invertible.

Solution

(30) A matrix $\mathbf{A}$ is refered to as *nilpotent* if there is a natural number $n > 0$, such that

$$\mathbf{A}^n = \underbrace{\mathbf{A}\mathbf{A}\ldots\mathbf{A}}_{n \text{ times}} = \mathbf{0}.$$

Show that a nilpotent matrix *cannot* be invertible. In particular, show that any $n \times n$ matrix $\mathbf{A}$ of the form

$$\mathbf{A} = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & a_{n-1,n} \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

is nilpotent with $\mathbf{A}^n = \mathbf{0}$ (n being the size of the matrix). Using the formula derived in the previous problem, compute $\det(\mathbf{A})$ to verify the non-invertibility claim above.

Solution

Navigation to Other Chapters

This page may not work with every browser-driven pdf viewer. When it does work then it will enable you to link directly to any chapter. When it does not work then you can link to any chapter through the main webpage.