Matrices and Linear Systems

- Roughly speaking, matrix is a rectangle array
- We shall discuss existence and uniqueness of solution for a system of linear equation.
- The method of Gauss elimination will be given to solve the system.
We shall denote matrices by capital boldface letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \cdots$, or by writing the general entry in brackets; thus $\mathbf{A} = [a_{jk}]$, and so on. By an $m \times n$ matrix (read $m$ by $n$ matrix) we mean a matrix with $m$ rows and $n$ columns—rows come always first! $m \times n$ is called the size of the matrix. Thus an $m \times n$ matrix is of the form

$$
\mathbf{A} = [a_{jk}] = \\
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}.
$$
Each entry in (2) has two subscripts. The first is the row number and the second is the column number. Thus $a_{21}$ is the entry in Row 2 and Column 1.

If $m = n$, we call $A$ an $n \times n$ square matrix. Then its diagonal containing the entries $a_{11}, a_{22}, \cdots, a_{nn}$ is called the main diagonal of $A$. 
A vector is a matrix with only one row or column. Its entries are called the components of the vector. We shall denote vectors by lowercase boldface letters \( \mathbf{a}, \mathbf{b}, \cdots \) or by its general component in brackets, \( \mathbf{a} = [a_j] \), and so on. Our special vectors in (1) suggest that a (general) row vector is of the form

\[
\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_n]. \quad \text{For instance,} \quad \mathbf{a} = [-2 \ 5 \ 0.8 \ 0 \ 1].
\]

A column vector is of the form

\[
\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \quad \text{For instance,} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ -7 \end{bmatrix}.
\]
**Definition**

**Equality of Matrices**

Two matrices \( A = [a_{jk}] \) and \( B = [b_{jk}] \) are **equal**, written \( A = B \), if and only if they have the same size and the corresponding entries are equal, that is, \( a_{11} = b_{11}, a_{12} = b_{12}, \) and so on. Matrices that are not equal are called **different**. Thus, matrices of different sizes are always different.
DEFINITION

Addition of Matrices

The **sum** of two matrices $A = [a_{jk}]$ and $B = [b_{jk}]$ **of the same size** is written $A + B$ and has the entries $a_{jk} + b_{jk}$ obtained by adding the corresponding entries of $A$ and $B$. Matrices of different sizes cannot be added.
**Definition**

Scalar Multiplication (Multiplication by a Number)

The **product** of any $m \times n$ matrix $A = [a_{jk}]$ and any **scalar** $c$ (number $c$) is written $cA$ and is the $m \times n$ matrix $cA = [ca_{jk}]$ obtained by multiplying each entry of $A$ by $c$. 
Rules for Matrix Addition and Scalar Multiplication. From the familiar laws for the addition of numbers we obtain similar laws for the addition of matrices of the same size $m \times n$, namely,

(a) \[ A + B = B + A \]

(b) \[ (A + B) + C = A + (B + C) \]  
    (written $A + B + C$)

(c) \[ A + 0 = A \]

(d) \[ A + (-A) = 0. \]

Here 0 denotes the **zero matrix** (of size $m \times n$), that is, the $m \times n$ matrix with all entries zero. (The last matrix in Example 5 is a zero matrix.)

Hence matrix addition is **commutative** and **associative** [by (3a) and (3b)].

Similarly, for scalar multiplication we obtain the rules

(a) \[ c(A + B) = cA + cB \]

(b) \[ (c + k)A = cA + kA \]

(c) \[ c(kA) = (ck)A \]  
    (written $ckA$)

(d) \[ 1A = A. \]
**Multiplication of a Matrix by a Matrix**

The **product** $C = AB$ (in this order) of an $m \times n$ matrix $A = [a_{jk}]$ times an $r \times p$ matrix $B = [b_{jk}]$ is defined if and only if $r = n$ and is then the $m \times p$ matrix $C = [c_{jk}]$ with entries

\[
(1) \quad c_{jk} = \sum_{l=1}^{n} a_{jl} b_{lk} = a_{j1} b_{1k} + a_{j2} b_{2k} + \cdots + a_{jn} b_{nk} \quad j = 1, \ldots, m \quad k = 1, \ldots, p.
\]

The condition $r = n$ means that the second factor, $B$, must have as many rows as the first factor has columns, namely $n$. As a diagram of sizes (denoted as shown):

\[
\begin{align*}
A & \quad B \\
[m \times n] & \quad [n \times r] = [m \times r].
\end{align*}
\]
$c_{jk}$ in (1) is obtained by multiplying each entry in the $j$th row of $A$ by the corresponding entry in the $k$th column of $B$ and then adding these $n$ products. For instance, $c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}$, and so on. One calls this briefly a "multiplication of rows into columns." See the illustration in Fig. 155, where $n = 3$.

**Fig. 155.** Notations in a product $AB = C$

### Example 1
**Matrix Multiplication**

$$AB = \begin{bmatrix}
3 & 5 & -1 \\
4 & 0 & 2 \\
-6 & -3 & 2
\end{bmatrix} \begin{bmatrix}
2 & -2 & 3 & 1 \\
5 & 0 & 7 & 8 \\
9 & -4 & 1 & 1
\end{bmatrix} = \begin{bmatrix}
22 & -2 & 43 & 42 \\
26 & -16 & 14 & 6 \\
-9 & 4 & -37 & -28
\end{bmatrix}$$

Here $c_{11} = 3 \cdot 2 + 5 \cdot 5 + (-1) \cdot 9 = 22$, and so on. The entry in the box is $c_{23} = 4 \cdot 3 + 0 \cdot 7 + 2 \cdot 1 = 14$. The product $BA$ is not defined.
EXAMPLE 4

CAUTION! Matrix Multiplication Is Not Commutative, $AB \neq BA$ in General

This is illustrated by Examples 1 and 2, where one of the two products is not even defined, and by Example 3, where the two products have different sizes. But it also holds for square matrices. For instance,

$$
\begin{bmatrix}
1 & 1 \\
100 & 100
\end{bmatrix}
\begin{bmatrix}
-1 & 1 \\
1 & -1
\end{bmatrix}
= \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix}
$$

but

$$
\begin{bmatrix}
-1 & 1 \\
1 & -1
\end{bmatrix}
\begin{bmatrix}
1 & 1 \\
100 & 100
\end{bmatrix}
= \begin{bmatrix}
99 & 99 \\
-99 & -99
\end{bmatrix}.
$$

It is interesting that this also shows that $AB = 0$ does not necessarily imply $BA = 0$ or $A = 0$ or $B = 0$. We shall discuss this further in Sec. 7.8, along with reasons when this happens.
Our examples show that the order of factors in matrix products must always be observed very carefully. Otherwise matrix multiplication satisfies rules similar to those for numbers, namely.

(a) \((kA)B = k(AB) = A(kB)\) written \(kAB\) or \(AkB\)

(b) \(A(BC) = (AB)C\) written \(ABC\)

(c) \((A + B)C = AC + BC\)

(d) \(C(A + B) = CA + CB\)

provided \(A, B,\) and \(C\) are such that the expressions on the left are defined; here, \(k\) is any scalar. (2b) is called the associative law. (2c) and (2d) are called the distributive laws.
**Definition**

**Transposition of Matrices and Vectors**

The transpose of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^T$ (read $A$ transpose) that has the first row of $\mathbf{A}$ as its first column, the second row of $\mathbf{A}$ as its second column, and so on. Thus the transpose of $\mathbf{A}$ in (2) is $\mathbf{A}^T = [a_{kj}]$, written out as

\[
\mathbf{A}^T = [a_{kj}] = \begin{bmatrix}
a_{11} & a_{21} & \cdots & a_{m1} \\
a_{12} & a_{22} & \cdots & a_{m2} \\
\cdots & \cdots & \cdots & \cdots \\
a_{1n} & a_{2n} & \cdots & a_{mn}
\end{bmatrix}.
\]

As a special case, transposition converts row vectors to column vectors and conversely.
Rules for transposition are

(a) \((A^T)^T = A\)

(b) \((A + B)^T = A^T + B^T\)

(c) \((cA)^T = cA^T\)

(d) \((AB)^T = B^T A^T\).
Symmetric and Skew-Symmetric Matrices. Transposition gives rise to two useful classes of matrices, as follows. *Symmetric matrices* and *skew-symmetric matrices* are *square* matrices whose transpose equals the matrix itself or minus the matrix, respectively:

\[
\begin{align*}
A^T &= A \quad \text{(thus } a_{kj} = a_{jk}), \\
A^T &= -A \quad \text{(thus } a_{kj} = -a_{jk}, \text{ hence } a_{jj} = 0). 
\end{align*}
\]

Symmetric Matrix

Skew-Symmetric Matrix
**Example 9**

Upper and Lower Triangular Matrices

\[
\begin{bmatrix}
1 & 3 \\
0 & 2 \\
\end{bmatrix}, \quad \begin{bmatrix}
1 & 4 & 2 \\
0 & 3 & 2 \\
\end{bmatrix}, \quad \begin{bmatrix}
2 & 0 & 0 \\
8 & -1 & 0 \\
7 & 6 & 8 \\
\end{bmatrix}, \quad \begin{bmatrix}
3 & 0 & 0 & 0 \\
9 & -3 & 0 & 0 \\
1 & 0 & 2 & 0 \\
1 & 9 & 3 & 6 \\
\end{bmatrix}.
\]

Upper triangular \quad Lower triangular
Diagonal Matrices. These are square matrices that can have nonzero entries only on the main diagonal. Any entry above or below the main diagonal must be zero.
In particular, a scalar matrix whose entries on the main diagonal are all 1 is called a **unit matrix** (or **identity matrix**) and is denoted by $I_n$ or simply by $I$. For $I$, formula (12) becomes

$$AI = IA = A.$$
A linear system of $m$ equations in $n$ unknowns $x_1, \cdots, x_n$ is a set of equations of the form

\begin{align*}
a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\
& \quad \vdots \quad \ddots \\
a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m
\end{align*}

(1)

The system is called linear because each variable $x_j$ appears in the first power only, just as in the equation of a straight line. $a_{11}, \cdots, a_{mn}$ are given numbers, called the coefficients of the system. $b_1, \cdots, b_m$ on the right are also given numbers. If all the $b_j$ are zero, then (1) is called a homogeneous system. If at least one $b_j$ is not zero, then (1) is called a nonhomogeneous system.
Matrix Form of the Linear System (1). From the definition of matrix multiplication we see that the \( m \) equations of (1) may be written as a single vector equation

\[
\begin{align*}
Ax = b
\end{align*}
\]

(2)

where the coefficient matrix \( A = [a_{jk}] \) is the \( m \times n \) matrix

\[
A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}, \text{ and } \begin{bmatrix}
    x_1 \\
    \vdots \\
    x_n
\end{bmatrix} \text{ and } b = \begin{bmatrix}
    b_1 \\
    \vdots \\
    b_m
\end{bmatrix}
\]

are column vectors. We assume that the coefficients \( a_{jk} \) are not all zero, so that \( A \) is not a zero matrix. Note that \( x \) has \( n \) components, whereas \( b \) has \( m \) components. The matrix

\textit{Continued}
is called the \textbf{augmented matrix} of the system (1). The dashed vertical line could be omitted (as we shall do later); it is merely a reminder that the last column of $\tilde{A}$ does not belong to $A$. 
Solution by Gauss Elimination. This system could be solved rather quickly by noticing its particular form. But this is not the point. The point is that the Gauss elimination is systematic and will work in general, also for large systems. We apply it to our system and then do back substitution. As indicated let us write the augmented matrix of the system first and then the system itself:

Augmented Matrix $\tilde{A}$

$\begin{bmatrix}
1 & -1 & 1 & | & 0 \\
-1 & 1 & -1 & | & 0 \\
 0 & 10 & 25 & | & 90 \\
 20 & 10 & 0 & | & 80
\end{bmatrix}$

Equations

$\begin{bmatrix}
\begin{array}{c}
 x_1 \\
-x_1 \\
 20x_1
\end{array}
\end{bmatrix}$

$\begin{array}{c}
x_2 + x_3 = 0 \\
-x_1 + x_2 - x_3 = 0 \\
10x_2 + 25x_3 = 90 \\
20x_1 + 10x_2 = 80.
\end{array}$

Step 1. Elimination of $x_1$

Call the first row of $A$ the pivot row and the first equation the pivot equation. Call the coefficient 1 of its $x_1$-term the pivot in this step. Use this equation to eliminate $x_1$ (get rid of $x_1$) in the other equations. For this, do:

- Add 1 times the pivot equation to the second equation.
- Add $-20$ times the pivot equation to the fourth equation.
This corresponds to row operations on the augmented matrix as indicated in BLUE behind the new matrix in (3). So the operations are performed on the preceding matrix. The result is

\[
\begin{bmatrix}
1 & -1 & 1 & | & 0 \\
0 & 0 & 0 & | & 0 \\
0 & 10 & 25 & | & 90 \\
0 & 30 & -20 & | & 80 \\
\end{bmatrix}
\]

\( \text{Row 2 + Row 1} \quad x_1 - x_2 + x_3 = 0 \)

\( \text{Row 4 - 20 Row 1} \quad 10x_2 + 25x_3 = 90 \)

\( 30x_2 - 20x_3 = 80 \)

**Step 2. Elimination of \( x_2 \)**

The first equation remains as it is. We want the new second equation to serve as the next pivot equation. But since it has no \( x_2 \)-term (in fact, it is \( 0 = 0 \)), we must first change the order of the equations and the corresponding rows of the new matrix. We put \( 0 = 0 \) at the end and move the third equation and the fourth equation one place up. This is called **partial pivoting** (as opposed to the rarely used total pivoting, in which also the order of the unknowns is changed). It gives

\[
\begin{bmatrix}
1 & -1 & 1 & | & 0 \\
0 & 10 & 25 & | & 90 \\
0 & 30 & -20 & | & 80 \\
0 & 0 & 0 & | & 0 \\
\end{bmatrix}
\]

\( \text{Pivot 10} \quad 10x_2 + 25x_3 = 90 \)

\( \text{Eliminate 30} \quad 30x_2 - 20x_3 = 80 \)

\( 0 = 0 \)

Continued
To eliminate \( x_2 \), do:

Add \(-3\) times the pivot equation to the third equation.

The result is

\[
\begin{bmatrix}
1 & -1 & 1 & | & 0 \\
0 & 10 & 25 & | & 90 \\
0 & 0 & -95 & | & -190 \\
0 & 0 & 0 & | & 0
\end{bmatrix}
\]

\[\begin{align*}
x_1 - x_2 + x_3 &= 0 \\
10x_2 + 25x_3 &= 90 \\
-95x_3 &= -190
\end{align*}\]

(4)

\(\text{Row 3} - 3 \text{ Row 2}\)

\[\begin{align*}
0 &= 0
\end{align*}\]

**Back Substitution. Determination of \( x_3, x_2, x_1 \) (in this order)**

Working backward from the last to the first equation of this “triangular” system (4), we can now readily find \( x_3 \), then \( x_2 \), and then \( x_1 \):

\[-95x_3 = -190 \quad \text{Rightarrow} \quad x_3 = i_3 = 2 \ [\text{A}]\]

\[10x_2 + 25x_3 = 90 \quad \text{Rightarrow} \quad x_2 = \frac{1}{10}(90 - 25x_3) = i_2 = 4 \ [\text{A}]\]

\[x_1 - x_2 + x_3 = 0 \quad \text{Rightarrow} \quad x_1 = x_2 - x_3 = i_1 = 2 \ [\text{A}]\]

where A stands for “amperes.” This is the answer to our problem. The solution is unique.
Elementary Row Operations for Matrices:

Interchange of two rows
Addition of a constant multiple of one row to another row
Multiplication of a row by a \textit{nonzero} constant \( c \).

Elementary Operations for Equations:

Interchange of two equations
Addition of a constant multiple of one equation to another equation
Multiplication of an equation by a \textit{nonzero} constant \( c \).
We now call a linear system $S_1$ **row-equivalent** to a linear system $S_2$ if $S_1$ can be obtained from $S_2$ by (finitely many!) row operations. Thus we have proved the following result, which also justifies the Gauss elimination.

**Theorem 1**

**Row-Equivalent Systems**

*Row-equivalent linear systems have the same set of solutions.*
At the end of the Gauss elimination the form of the coefficient matrix, the augmented matrix, and the system itself are called the **row echelon form**. In it, rows of zeros, if present, are the last rows, and in each nonzero row the leftmost nonzero entry is farther to the right than in the previous row. For instance, in Example 4 the coefficient matrix and its augmented in row echelon form are

\[
\begin{bmatrix}
3 & 2 & 1 \\
0 & -\frac{1}{3} & \frac{1}{3} \\
0 & 0 & 0
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
3 & 2 & 1 & 3 \\
0 & -\frac{1}{3} & \frac{1}{3} & -2 \\
0 & 0 & 0 & 12
\end{bmatrix}.
\]

Note that we do not require that the leftmost nonzero entries be 1 since this would have no theoretic or numeric advantage. (The so-called **reduced echelon form**, in which those entries are 1, will be discussed in Sec. 7.8.)

**the row reduced echelon form**

\[
\begin{bmatrix}
1 & 2/3 & 1/3 \\
0 & 1 & -1 \\
0 & 0 & 0
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & -1 \\
0 & 0 & 0
\end{bmatrix}
\]

---

*Advanced Engineering Mathematics* by Erwin Kreyszig

Copyright © 2007 John Wiley & Sons, Inc. All rights reserved.
Example: \[
\begin{bmatrix}
1 & 3 & 0 & 5 & -1 \\
0 & 0 & 1 & 0 & 2 \\
\end{bmatrix}
\]
is a row reduced echelon matrix.

1. Find row reduced echelon forms of
\[
\begin{bmatrix}
-1 & 4 & 1 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}, \quad
\begin{bmatrix}
-2 & 1 & 3 \\
0 & 1 & 1 \\
2 & 0 & 1 \\
\end{bmatrix}
\]

2. Show that they are inconsistent systems.

\[
\begin{align*}
2x_1 + -3x_2 &= 1 & 2x_1 - 3x_2 &= 6 \\
x_1 + 3x_2 &= 0 & 4x_1 - 6x_2 &= 18 \\
x_1 - 4x_2 &= 3
\end{align*}
\]
At the end of the Gauss elimination (before the back substitution) the row echelon form of the augmented matrix will be

\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} & | & b_1 \\
  c_{22} & \cdots & \cdots & c_{2n} & | & \tilde{b}_2 \\
  \vdots & \cdots & \cdots & \vdots & | & \vdots \\
  k_{rr} & \cdots & k_{rn} & | & \tilde{b}_r \\
  \vdots & \cdots & \vdots & \vdots & | & \vdots \\
  \vdots & \cdots & \vdots & \vdots & \vdots & \tilde{b}_{r+1} \\
  \vdots & \cdots & \vdots & \vdots & \vdots & \tilde{b}_m \\
\end{bmatrix}
\]

(8)

Here, \( r \leq m \) and \( a_{11} \neq 0, c_{22} \neq 0, \cdots, k_{rr} \neq 0 \), and all the entries in the blue triangle as well as in the blue rectangle are zero. From this we see that with respect to solutions of the system with augmented matrix (8) (and thus with respect to the originally given system) there are three possible cases:

Continued
(a) **Exactly one solution** if \( r = n \) and \( \tilde{b}_{r+1}, \ldots, \tilde{b}_m \), if present, are zero. To get the solution, solve the \( n \)th equation corresponding to (8) (which is \( k_{nn}x_n = \tilde{b}_n \)) for \( x_n \), then the \((n - 1)\)st equation for \( x_{n-1} \), and so on up the line. See Example 2, where \( r = n = 3 \) and \( m = 4 \).

(b) **Infinitely many solutions** if \( r < n \) and \( \tilde{b}_{r+1}, \ldots, \tilde{b}_m \), if present, are zero. To obtain any of these solutions, choose values of \( x_{r+1}, \ldots, x_n \) arbitrarily. Then solve the \( r \)th equation for \( x_r \), then the \((r - 1)\)st equation for \( x_{r-1} \), and so on up the line. See Example 3.

(c) **No solution** if \( r < m \) and one of the entries \( \tilde{b}_{r+1}, \ldots, \tilde{b}_m \) is not zero. See Example 4, where \( r = 2 < m = 3 \) and \( \tilde{b}_{r+1} = \tilde{b}_3 = 12 \).
Vector Spaces

- A quantity such as work, area or energy which is described in terms of magnitude alone is called a scalar.

- A quantity which has both magnitude and direction for its description is called a vector.

A vector is an element of vector space.
Definition: A vector space $V$ in $\mathbb{R}$ is the set satisfying

1. If $u, v \in V$ and $a \in \mathbb{R}$, then $u + v, au \in V$.
2. $u + v = v + u$
3. $u + (v + w) = (u + v) + w$
4. $u + 0 = u$ (a unique zero element)
5. $u + (-u) = 0$ (a unique additive inverse)
6. $(a + b)u = au + bu$
7. $(ab)u = a(bu)$
8. $1u = u$ (1 is identity scalar)
9. $a(u + v) = au + av$
Examples for vector spaces

1. $V = \{0\}$
2. $\mathbb{R}^n$
3. $\mathbb{R}^{m \times n}$ (space of matrices)
4. $P_n[x]$ (space of polynomials)
5. $F[a, b]$ (function - space on $[a, b]$)
6. $C^n[a, b]$ (continuously differentiable func.)
Define a linear combination of nonzero vectors 
\[ c_1 u_1 + c_2 u_2 + \ldots + c_n u_n, \]
where \( u_i \in V, c_i \in R, i = 1,2,3,\ldots \)

If the equation 
\[ c_1 u_1 + c_2 u_2 + \ldots c_n u_n = 0 \]
has only trivial solution for all \( c_i \), then 
\( u_1, u_2, \ldots u_n \) are linearly independent. Otherwise, 
if any \( c_i \) is zero then they are linearly dependent.
DEFINITION

The **rank** of a matrix $A$ is the maximum number of linearly independent row vectors of $A$. It is denoted by $\text{rank } A$. 
EXAMPLE 3

Determination of Rank

For the matrix in Example 2 we obtain successively

\[ A = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix} \text{ (given)} \]

Row 2 + 2 Row 1

\[ \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & -21 & -14 & -29 \end{bmatrix} \]

Row 3 - 7 Row 1

\[ \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ Row 3 + \frac{1}{2} Row 2} \]

Rank of A is 2 because the first two rows are linearly independent.
**Row-Equivalent Matrices**

Row-equivalent matrices have the same rank.

**Example**

\[
A = \begin{bmatrix}
-2 & 1 & 3 \\
0 & 1 & 1 \\
2 & 0 & 1
\end{bmatrix}
\text{ is row equivalent to } 
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

\[
\text{rank} A = 3
\]
THEOREM 2

Linear Independence and Dependence of Vectors

$p$ vectors with $n$ components each are linearly independent if the matrix with these vectors as row vectors has rank $p$, but they are linearly dependent if that rank is less than $p$. 
THEOREM 3

Rank in Terms of Column Vectors

The rank $r$ of a matrix $A$ equals the maximum number of linearly independent column vectors of $A$.

Hence $A$ and its transpose $A^T$ have the same rank.
Dimension of a vector space $V$

- $\text{Span}_S = \text{All linear combinations of vectors of the subset } S \text{ of } V$.
- A **basis** for $V$ is a linearly independent subset $S$ of $V$ which spans the space $V$.
- That is, $\text{Span}_S = V$ where $S$ is lin. Indep.

$\text{dim}V = \text{The number of vectors in any basis for } V$.

$V$ is **finite-dimensional** if $V$ has a basis consisting of a finite number of vectors.
**Theorem 6**

Row Space and Column Space

The row space and the column space of a matrix $A$ have the same dimension, equal to rank $A$.

Finally, for a given matrix $A$ the solution set of the homogeneous system $Ax = 0$ is a vector space, called the **null space** of $A$, and its dimension is called the **nullity** of $A$. In the next section we motivate and prove the basic relation

$$\text{rank } A + \text{nullity } A = \text{Number of columns of } A.$$  

Note: (6) is known as dimension theorem
THEOREM 1

Fundamental Theorem for Linear Systems

(a) Existence. A linear system of $m$ equations in $n$ unknowns $x_1, \cdots, x_n$

\[ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \]

\[ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \]

\[ \cdots \cdots \cdots \cdots \cdots \cdots \cdots \]

\[ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \]

is consistent, that is, has solutions, if and only if the coefficient matrix $A$ and the augmented matrix $\tilde{A}$ have the same rank. Here,
\[ A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \tilde{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & | & b_1 \\ \cdot & \cdot & \cdot & | & \cdot \\ \cdot & \cdot & \cdot & | & \cdot \\ a_{m1} & \cdots & a_{mn} & | & b_m \end{bmatrix} \]

(b) **Uniqueness.** The system (1) has precisely one solution if and only if this common rank \( r \) of \( A \) and \( \tilde{A} \) equals \( n \).

(c) **Infinitely many solutions.** If this common rank \( r \) is less than \( n \), the system (1) has infinitely many solutions. All of these solutions are obtained by determining \( r \) suitable unknowns (whose submatrix of coefficients must have rank \( r \)) in terms of the remaining \( n - r \) unknowns, to which arbitrary values can be assigned. (See Example 3 in Sec. 7.3.)

(d) **Gauss elimination (Sec. 7.3).** If solutions exist, they can all be obtained by the Gauss elimination. (This method will automatically reveal whether or not solutions exist; see Sec. 7.3.)
**Homogeneous Linear System**

A homogeneous linear system

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
    \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{align*}
\]

always has the trivial solution \(x_1 = 0, \cdots, x_n = 0\). Nontrivial solutions exist if and only if rank \(A < n\). If rank \(A = r < n\), these solutions, together with \(x = 0\), form a vector space (see Sec. 7.4) of dimension \(n - r\), called the solution space of (4).

In particular, if \(x_{(1)}\) and \(x_{(2)}\) are solution vectors of (4), then \(x = c_1x_{(1)} + c_2x_{(2)}\) with any scalars \(c_1\) and \(c_2\) is a solution vector of (4). (This does not hold for nonhomogeneous systems. Also, the term solution space is used for homogeneous systems only.)
The solution space of (4) is also called the null space of \( A \) because \( Ax = 0 \) for every \( x \) in the solution space of (4). Its dimension is called the nullity of \( A \). Hence Theorem 2 states that

\[(5) \quad \text{rank } A + \text{nullity } A = n\]

where \( n \) is the number of unknowns (number of columns of \( A \)).
Homogeneous Linear System with Fewer Equations Than Unknowns

A homogeneous linear system with fewer equations than unknowns has always nontrivial solutions.
**Theorem 4**

**Nonhomogeneous Linear System**

*If a nonhomogeneous linear system (1) is consistent, then all of its solutions are obtained as*

\[(6) \quad x = x_0 + x_h\]

*where* \(x_0\) *is any (fixed) solution of (1) and* \(x_h\) *runs through all the solutions of the corresponding homogeneous system (4).*
Example

\[3x_1 + 2x_2 + 3x_3 - 2x_4 = 1\]
\[x_1 + x_2 + x_3 + 0x_4 = 3\]
\[x_1 + 2x_2 + x_3 - x_4 = 2\]

\[x_0 = \begin{bmatrix} 1 & 2 & 0 & 3 \end{bmatrix}^T, \quad x_h = c \begin{bmatrix} -1 & 0 & 1 & 0 \end{bmatrix}^T\]
Determinant

Determinant is a function form square matrices to scalars.

Our efficient computational procedure will be cofactor expansion.
A determinant of second order is denoted and defined by

$$D = \det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (1)$$

So here we have *bars* (whereas a matrix has *brackets*).

**Cramer’s rule** for solving linear systems of two equations in two unknowns

(a) \[ a_{11}x_1 + a_{12}x_2 = b_1 \]

(b) \[ a_{21}x_1 + a_{22}x_2 = b_2 \]

is

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{D} = \frac{b_1a_{22} - a_{12}b_2}{D}.$$
(3) \[ x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{D} = \frac{a_{11}b_2 - b_1a_{21}}{D}. \]

with \( D \) as in (1), provided \( D \neq 0 \).
A **determinant of third order** can be defined by

\[ D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \]

Note the following. The signs on the right are $+ - +$. Each of the three terms on the right is an entry in the first column of $D$ times its **minor**, that is, the second-order determinant obtained from $D$ by deleting the row and column of that entry; thus, for $a_{11}$ delete the first row and first column, and so on.

If we write out the minors in (4), we obtain

\[ D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{21}a_{13}a_{32} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}. \]
Cramer’s Rule for Linear Systems of Three Equations

\[ a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \]

\[ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \]

\[ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \]

is

\[ x_1 = \frac{D_1}{D} , \quad x_2 = \frac{D_2}{D} , \quad x_3 = \frac{D_3}{D} \quad (D \neq 0) \]

with the determinant \( D \) of the system given by (4) and

\[ D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} , \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} , \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} . \]

Note that \( D_1, D_2, D_3 \) are obtained by replacing Columns 1, 2, 3, respectively, by the column of the right sides of (5).
A **determinant of order** $n$ is a scalar associated with an $n \times n$ (hence *square*) matrix $A = [a_{jk}]$, which is written

\[
D = \det A = \begin{vmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \cdot & \cdot & \cdots & \cdot \\
  \cdot & \cdot & \cdots & \cdot \\
  a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
\]

and is defined for $n = 1$ by

\[
(1) \quad D = a_{11}
\]

and for $n \geq 2$ by

\[
(3a) \quad D = a_{j1}C_{j1} + a_{j2}C_{j2} + \cdots + a_{jn}C_{jn} \quad (j = 1, 2, \cdots, \text{or } n)
\]
or

\[(3b) \quad D = a_{1k} C_{1k} + a_{2k} C_{2k} + \cdots + a_{nk} C_{nk} \quad (k = 1, 2, \ldots, \text{or } n)\]

Here,

\[C_{jk} = (-1)^{j+k} M_{jk}\]

and \(M_{jk}\) is a determinant of order \(n - 1\), namely, the determinant of the submatrix of \(A\) obtained from \(A\) by omitting the row and column of the entry \(a_{jk}\), that is, the \(j\)th row and the \(k\)th column.

In this way, \(D\) is defined in terms of \(n\) determinants of order \(n - 1\), each of which is, in turn, defined in terms of \(n - 1\) determinants of order \(n - 2\), and so on; we finally arrive at second-order determinants, in which those submatrices consist of single entries whose determinant is defined to be the entry itself.

From the definition it follows that we may expand \(D\) by any row or column, that is, choose in (3) the entries in any row or column, similarly when expanding the \(C_{jk}\)'s in (3), and so on.

**This definition is unambiguous**, that is, yields the same value for \(D\) no matter which columns or rows we choose in expanding.
$M_{jk}$ is called the **minor** of $a_{jk}$ in $D$, and $C_{jk}$ the **cofactor** of $a_{jk}$ in $D$. For later use we note that (3) may also be written in terms of minors

\[(4a)\]

\[
D = \sum_{k=1}^{n} (-1)^{j+k} a_{jk} M_{jk} \quad (j = 1, 2, \cdots, \text{or } n)
\]

\[(4b)\]

\[
D = \sum_{j=1}^{n} (-1)^{j+k} a_{jk} M_{jk} \quad (k = 1, 2, \cdots, \text{or } n).
\]
EXAMPLE 1

Minors and Cofactors of a Third-Order Determinant

In (4) of the previous section the minors and cofactors of the entries in the first column can be seen directly. For the entries in the second row the minors are

\[
M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}
\]

and the cofactors are \( C_{21} = -M_{21}, \ C_{22} = +M_{22}, \) and \( C_{23} = -M_{23} \). Similarly for the third row—write these down yourself. And verify that the signs in \( C_{jk} \) form a checkerboard pattern

\[
+ \ - \ + \\
- \ + \ - \\
+ \ - \ +
\]
EXAMPLE 2  

Expansions of a Third-Order Determinant

\[
D = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 1 \begin{vmatrix} 6 & 4 \\ 0 & 2 \end{vmatrix} - 3 \begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} + 0 \begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} \\
= 1(12 - 0) - 3(4 + 4) + 0(0 + 6) = -12.
\]

This is the expansion by the first row. The expansion by the third column is

\[
D = 0 \begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} - 4 \begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} + 2 \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0 - 12 + 0 = -12,
\]

Verify that the other four expansions also give the value \(-12\).
Theorem 1

Behavior of an *n*-th Order Determinant under Elementary Row Operations

(a) *Interchange of two rows multiplies the value of the determinant by* $-1$.

(b) *Addition of a multiple of a row to another row does not alter the value of the determinant.*

(c) *Multiplication of a row by a nonzero constant* $c$ *multiplies the value of the determinant by* $c$. (This holds also when* $c = 0$, *but gives no longer an elementary row operation.*)
Further Properties of $n$th-Order Determinants

(a)–(c) in Theorem 1 hold also for columns.

(d) Transposition leaves the value of a determinant unaltered.

(e) A zero row or column renders the value of a determinant zero.

(f) Proportional rows or columns render the value of a determinant zero. In particular, a determinant with two identical rows or columns has the value zero.
THEOREM 3

Rank in Terms of Determinants

An $m \times n$ matrix $A = [a_{jk}]$ has rank $r \geq 1$ if and only if $A$ has an $r \times r$ submatrix with nonzero determinant, whereas every square submatrix with more than $r$ rows that $A$ has (or does not have!) has determinant equal to zero.

In particular, if $A$ is square, $n \times n$, it has rank $n$ if and only if

$$\det A \neq 0.$$
Cramer’s Theorem (Solution of Linear Systems by Determinants)

(a) If a linear system of \( n \) equations in the same number of unknowns \( x_1, \ldots, x_n \)

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
    &\vdots \\
    a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{align*}
\]

has a nonzero coefficient determinant \( D = \det A \), the system has precisely one solution. This solution is given by the formulas

(7) \[
    x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \cdots, \quad x_n = \frac{D_n}{D} \quad \text{(Cramer’s rule)}
\]

where \( D_k \) is the determinant obtained from \( D \) by replacing in \( D \) the \( k \)th column by the column with the entries \( b_1, \ldots, b_n \).
**Theorem 1**

**Existence of the Inverse**

The inverse $A^{-1}$ of an $n \times n$ matrix $A$ exists if and only if rank $A = n$, thus (by Theorem 3, Sec. 7.7) if and only if $\det A \neq 0$. Hence $A$ is nonsingular if rank $A = n$, and is singular if rank $A < n$. 
EXAMPLE 1

Inverse of a Matrix. Gauss–Jordan Elimination

Determine the inverse $A^{-1}$ of

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}. $$
Solution. We apply the Gauss elimination (Sec. 7.3) to the following $n \times 2n = 3 \times 6$ matrix, where BLUE always refers to the previous matrix.

$$[A \; I] = \begin{bmatrix}
-1 & 1 & 2 & | & 1 & 0 & 0 \\
3 & -1 & 1 & | & 0 & 1 & 0 \\
-1 & 3 & 4 & | & 0 & 0 & 1 \\
\end{bmatrix}$$

$$\begin{bmatrix}
-1 & 1 & 2 & | & 1 & 0 & 0 \\
0 & 2 & 7 & | & 3 & 1 & 0 \\
0 & 2 & 2 & | & -1 & 0 & 1 \\
\end{bmatrix} \quad \text{Row 2 + 3 Row 1}$$

$$\begin{bmatrix}
-1 & 1 & 2 & | & 1 & 0 & 0 \\
0 & 2 & 7 & | & 3 & 1 & 0 \\
0 & 0 & -5 & | & -4 & -1 & 1 \\
\end{bmatrix} \quad \text{Row 3 – Row 2}$$
This is $[U \quad H]$ as produced by the Gauss elimination. Now follow the additional Gauss–Jordan steps, reducing $U$ to $I$, that is, to diagonal form with entries 1 on the main diagonal.

\[
\begin{bmatrix}
1 & -1 & -2 & | & -1 & 0 & 0 \\
0 & 1 & 3.5 & | & 1.5 & 0.5 & 0 \\
0 & 0 & 1 & | & 0.8 & 0.2 & -0.2
\end{bmatrix} - \text{ Row 1}
\]

\[
\begin{bmatrix}
1 & -1 & 0 & | & 0.6 & 0.4 & -0.4 \\
0 & 1 & 0 & | & -1.3 & -0.2 & 0.7 \\
0 & 0 & 1 & | & 0.8 & 0.2 & -0.2
\end{bmatrix} - 0.5 \text{ Row 2}
\]

\[
\begin{bmatrix}
1 & -1 & 0 & | & -0.7 & 0.2 & 0.3 \\
0 & 1 & 0 & | & -1.3 & -0.2 & 0.7 \\
0 & 0 & 1 & | & 0.8 & 0.2 & -0.2
\end{bmatrix} - \text{ Row 1 + Row 2}
\]

The last three columns constitute $A^{-1}$. Check:

\[
\begin{bmatrix}
-1 & 1 & 2 \\
3 & -1 & 1 \\
-1 & 3 & 4
\end{bmatrix}
\begin{bmatrix}
-0.7 & 0.2 & 0.3 \\
1.3 & -0.2 & 0.7 \\
0.8 & 0.2 & -0.2
\end{bmatrix}
= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

Hence $AA^{-1} = I$. Similarly, $A^{-1}A = I$. 

---

*Advanced Engineering Mathematics* by Erwin Kreyszig
**Theorem 2**

**Inverse of a Matrix**

The inverse of a nonsingular $n \times n$ matrix $A = [a_{jk}]$ is given by

$$
A^{-1} = \frac{1}{\det A} [C_{jk}]^T = \frac{1}{\det A} \begin{bmatrix}
C_{11} & C_{21} & \cdots & C_{n1} \\
C_{12} & C_{22} & \cdots & C_{n2} \\
\vdots & \vdots & \ddots & \vdots \\
C_{1n} & C_{2n} & \cdots & C_{nn}
\end{bmatrix},
$$

where $C_{jk}$ is the cofactor of $a_{jk}$ in $\det A$ (see Sec. 7.7). (CAUTION! Note well that in $A^{-1}$, the cofactor $C_{jk}$ occupies the same place as $a_{kj}$ (not $a_{jk}$) does in $A$.)

In particular, the inverse of

$$(4^*) \quad A = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} \quad \text{is} \quad A^{-1} = \frac{1}{\det A} \begin{bmatrix}
a_{22} & -a_{12} \\
-a_{21} & a_{11}
\end{bmatrix}. $$
EXAMPLE 2
Inverse of a $2 \times 2$ Matrix

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, \quad A^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}$$

EXAMPLE 3
Further Illustration of Theorem 2

Using (4), find the inverse of

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$
Solution. We obtain \( \det A = -1(-7) - 1 \cdot 13 + 2 \cdot 8 = 10 \), and in (4),

\[
C_{11} = \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -7, \quad C_{21} = -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2, \quad C_{31} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3,
\]

\[
C_{12} = -\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} = -13, \quad C_{22} = \begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix} = -2, \quad C_{32} = -\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = 7,
\]

\[
C_{13} = \begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} = 8, \quad C_{23} = -\begin{vmatrix} -1 & 1 \\ -1 & 3 \end{vmatrix} = 2, \quad C_{33} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = -2,
\]

so that by (4), in agreement with Example 1,

\[
A^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.
\]
Diagonal matrices $A = [a_{jk}]$, $a_{jk} = 0$ when $j \neq k$, have an inverse if and only if all $a_{jj} \neq 0$. Then $A^{-1}$ is diagonal, too, with entries $1/a_{11}, \cdots, 1/a_{nn}$. 
Products can be inverted by taking the inverse of each factor and multiplying these inverses in reverse order,

(7) \[(AC)^{-1} = C^{-1}A^{-1}\].
[1.] Matrix multiplication is not commutative, that is, in general we have
\[ AB \neq BA. \]

[2.] \( AB = 0 \) does not generally imply \( A = 0 \) or \( B = 0 \) (or \( BA = 0 \)); for example,
\[
\begin{bmatrix}
1 & 1 \\
2 & 2
\end{bmatrix}
\begin{bmatrix}
-1 & 1 \\
1 & -1
\end{bmatrix}
= \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix}.
\]

[3.] \( AC = AD \) does not generally imply \( C = D \) (even when \( A \neq 0 \)).

Complete answers to [2.] and [3.] are contained in the following theorem.
THEOREM 3

Cancellation Laws

Let \( A, B, C \) be \( n \times n \) matrices. Then:

(a) If \( \text{rank } A = n \) and \( AB = AC \), then \( B = C \).

(b) If \( \text{rank } A = n \), then \( AB = 0 \) implies \( B = 0 \). Hence if \( AB = 0 \), but \( A \neq 0 \) as well as \( B \neq 0 \), then \( \text{rank } A < n \) and \( \text{rank } B < n \).

(c) If \( A \) is singular, so are \( BA \) and \( AB \).
**Theorem 4**

**Determinant of a Product of Matrices**

For any $n \times n$ matrices $A$ and $B$,

(10) \[ \det(AB) = \det(BA) = \det A \det B. \]
Real Inner Product Space

A real vector space \( V \) is called a \textbf{real inner product space} (or \textit{real pre-Hilbert} space) if it has the following property. With every pair of vectors \( a \) and \( b \) in \( V \) there is associated a real number, which is denoted by \((a, b)\) and is called the \textbf{inner product} of \( a \) and \( b \), such that the following axioms are satisfied.

\begin{enumerate}
  \item For all scalars \( q_1 \) and \( q_2 \) and all vectors \( a, b, c \) in \( V \),

    \[(q_1 \mathbf{a} + q_2 \mathbf{b}, \mathbf{c}) = q_1(a, c) + q_2(b, c)\]  
    \textit{(Linearity)}.

  \item For all vectors \( a \) and \( b \) in \( V \),

    \[(a, b) = (b, a)\]  
    \textit{(Symmetry)}.

  \item For every \( a \) in \( V \),

    \[
    \begin{cases}
    (a, a) \geq 0, \\
    (a, a) = 0 \text{ if and only if } a = 0
    \end{cases}
    \]  
    \textit{(Positive-definiteness)}.
\end{enumerate}
Vectors whose inner product is zero are called **orthogonal**. The *length* or **norm** of a vector in $V$ is defined by

\begin{equation}
\|a\| = \sqrt{a, a} \quad (\geq 0).
\end{equation}

A vector of norm 1 is called a **unit vector**.

From these axioms and from (2) one can derive the basic inequality

\begin{equation}
|(a, b)| \leq \|a\| \|b\| \quad (Cauchy–Schwarz$^5$ inequality).
\end{equation}

From this follows

\begin{equation}
\|a + b\| \leq \|a\| + \|b\| \quad (Triangle inequality).
\end{equation}

A simple direct calculation gives

\begin{equation}
\|a + b\|^2 + \|a - b\|^2 = 2(\|a\|^2 + \|b\|^2) \quad (Parallelogram equality).
\end{equation}
Examples

\[(a, b) = a^\top b = a_1 b_1 + \cdots + a_n b_n\]

\[(f, g) = \int_\alpha^\beta f(x) g(x) \, dx.\]

\[\|a\| = \sqrt{(a, a)} = \sqrt{a^\top a} = \sqrt{a_1^2 + \cdots + a_n^2}.\]

\[\|f\| = \sqrt{(f, f)} = \sqrt{\int_\alpha^\beta f(x)^2 \, dx}.\]
Linear Transformations

$F$ is called a linear mapping or linear transformation if for all vectors $v$ and $x$ in $X$ and scalars $c$,

$F(v + x) = F(v) + F(x)$

(10)

$F(cx) = cF(x)$.

Let $e_{(1)}, \cdots, e_{(n)}$ be any basis for $\mathbb{R}^n$. Then every $x$ in $\mathbb{R}^n$ has a unique representation

$$x = x_1e_{(1)} + \cdots + x_ne_{(n)}.$$ 

Since $F$ is linear, this representation implies for the image $F(x)$:

$$F(x) = F(x_1e_{(1)} + \cdots + x_ne_{(n)}) = x_1F(e_{(1)}) + \cdots + x_nF(e_{(n)}).$$

Hence $F$ is uniquely determined by the images of the vectors of a basis for $\mathbb{R}^n$. We now determine an $m \times n$ matrix $A = [a_{jk}]$ such that for every $x$ in $\mathbb{R}^n$ and image $y = F(x)$ in $\mathbb{R}^m$,

$$y = F(x) = Ax.$$
Example

Interpreted as transformations of Cartesian coordinates in the plane, the matrices

\[
\begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix},
\begin{bmatrix}
1 & 0 \\
0 & -1
\end{bmatrix},
\begin{bmatrix}
-1 & 0 \\
0 & 1
\end{bmatrix},
\begin{bmatrix}
a & 0 \\
0 & 1
\end{bmatrix}
\]

represent a reflection in the line \(x_2 = x_1\), a reflection in the \(x_1\)-axis, a reflection in the origin, and a stretch (when \(a > 1\), or a contraction when \(0 < a < 1\)) in the \(x_1\)-direction, respectively.
Example:

Find the representation matrix of
the linear transformation that maps \((x_1, x_2)\) onto \((2x_1 - 5x_2, 3x_1 + 4x_2)\).

**Solution.** Obviously, the transformation is

\[
\begin{align*}
y_1 &= 2x_1 - 5x_2 \\
y_2 &= 3x_1 + 4x_2.
\end{align*}
\]

From this we can directly see that the matrix is

\[
A = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix}.
\]

Check: \[
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix}.
\]
Range and Null (Kernel) spaces

Let $F : V \rightarrow W$ be a linear transform.

$\text{Null}F = \{u : F(u) = 0, u \in V\}$.

$\text{Range}F = \{v : v = F(u), u \in V\} \subseteq W$
includes all images vectors.

$\dim(\text{Null}F) = \text{nullity}F$
$\dim(\text{Range}F) = \text{rank}F$

Theorem: $\text{rank}F + \text{nullity}F = \dim V$. 
SUMMARY OF CHAPTER 7

Linear Algebra: Matrices, Vectors, Determinants
Linear Systems of Equations

An \( m \times n \) matrix \( A = [a_{jk}] \) is a rectangular array of numbers or functions ("entries", "elements") arranged in \( m \) horizontal rows and \( n \) vertical columns. If \( m = n \), the matrix is called square. A \( 1 \times n \) matrix is called a row vector and an \( m \times 1 \) matrix a column vector (Sec. 7.1).

The sum \( A + B \) of matrices of the same size (i.e., both \( m \times n \)) is obtained by adding corresponding entries. The product of \( A \) by a scalar \( c \) is obtained by multiplying each \( a_{jk} \) by \( c \) (Sec. 7.1).

The product \( C = AB \) of an \( m \times n \) matrix \( A \) by an \( r \times p \) matrix \( B = [b_{jk}] \) is defined only when \( r = n \), and is the \( m \times p \) matrix \( C = [c_{jk}] \) with entries

\[
(1) \quad c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \quad \text{(row } j \text{ of } A \text{ times column } k \text{ of } B). 
\]
This multiplication is motivated by the composition of **linear transformations** (Sects. 7.2, 7.9). It is associative, but is **not commutative**: if \( AB \) is defined, \( BA \) may not be defined, but even if \( BA \) is defined, \( AB \neq BA \) in general. Also \( AB = 0 \) may not imply \( A = 0 \) or \( B = 0 \) or \( BA = 0 \) (Sects. 7.2, 7.8). Illustrations:

\[
\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\]

\[
\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}
\]

\[
\begin{bmatrix} 1 & 2 \\ 4 \end{bmatrix} = [11], \quad \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}.
\]

The **transpose** \( A^T \) of a matrix \( A = [a_{jk}] \) is \( A^T = [a_{kj}] \); rows become columns and conversely (Sec. 7.2). Here, \( A \) need not be square. If it is and \( A = A^T \), then \( A \) is called **symmetric**; if \( A = -A^T \), it is called **skew-symmetric**. For a product, \( (AB)^T = B^TA^T \) (Sec. 7.2).
A main application of matrices concerns **linear systems of equations**

\[ (2) \quad Ax = b \quad \text{(Sec. 7.3)} \]

\((m\) equations in \(n\) unknowns \(x_1, \ldots, x_n\); \(A\) and \(b\) given). The most important method of solution is the **Gauss elimination** (Sec. 7.3), which reduces the system to “triangular” form by **elementary row operations**, which leave the set of solutions unchanged. (Numeric aspects and variants, such as **Doolittle’s** and **Cholesky’s methods**, are discussed in Secs. 20.1 and 20.2)

**Cramer’s rule** (Secs. 7.6, 7.7) represents the unknowns in a system (2) of \(n\) equations in \(n\) unknowns as quotients of determinants; for numeric work it is impractical. **Determinants** (Sec. 7.7) have decreased in importance, but will retain their place in eigenvalue problems, elementary geometry, etc.

The **inverse** \(A^{-1}\) of a square matrix satisfies \(AA^{-1} = A^{-1}A = I\). It exists if and only if \(\det A \neq 0\). It can be computed by the **Gauss–Jordan elimination** (Sec. 7.8).

The **rank** \(r\) of a matrix \(A\) is the maximum number of linearly independent rows or columns of \(A\) or, equivalently, the number of rows of the largest square submatrix of \(A\) with nonzero determinant (Secs. 7.4, 7.7).

*Continued*
The system (2) has solutions if and only if rank $A = \text{rank } [A \ b]$, where $[A \ b]$ is the augmented matrix (Fundamental Theorem, Sec. 7.5).

The homogeneous system

\begin{equation}
Ax = 0
\end{equation}

has solutions $x \neq 0$ (“nontrivial solutions”) if and only if rank $A < n$, in the case $m = n$ equivalently if and only if $\det A = 0$ (Secs. 7.6, 7.7).

Vector spaces, inner product spaces, and linear transformations are discussed in Sec. 7.9. See also Sec. 7.4.