1. Systems of Linear Equations

1.1. Systems of equations and linear transformations.

Consider the problem of finding the solutions (if any) of the system of \( m \) equations in \( n \) unknowns:

\[
\begin{align*}
 a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n &= b_1 \\
 a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n &= b_2 \\
 & \vdots \\
 a_{m,1} x_1 + a_{m,2} x_2 + \cdots + a_{m,n} x_n &= b_m 
\end{align*}
\]  
(1.1.1)

Here the coefficients \( \{a_{j,k}\} \) and the data \( b_1, \ldots, b_m \) on the right hand side are regarded as known, and we want to solve for the unknowns \( x_1, \ldots, x_n \). We begin by reinterpreting this system as a question about linear mappings. This will allow us to say some things about the nature and number of solutions, and also say something about when solutions exist. Later we will discuss an algorithm (Gauss-Jordan elimination) for actually finding solutions. We begin with a definition.

**Definition 1.1.** The system of equations (1.1.1) is **homogeneous** if the coefficients on the right hand side of the equations are all zero. That is, the system is homogeneous if \( b_1 = b_2 = \cdots = b_m = 0 \). If the right hand side data is not all zero, the system is said to be **inhomogeneous**.

It will be convenient to think of vectors in \( \mathbb{R}^m \) or \( \mathbb{R}^n \) as column vectors. Then we can define a mapping \( A \) from \( \mathbb{R}^n \) to \( \mathbb{R}^m \) by multiplying an \( n \times 1 \) column matrix on the left by the \( m \times n \) matrix of coefficients \( \{a_{i,j}\} \). Thus

\[
A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n \\ a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n \\ \vdots \\ a_{m,1} x_1 + a_{m,2} x_2 + \cdots + a_{m,n} x_n \end{pmatrix}.
\]

It is easy to check that \( A \) is indeed a linear transformation, and the system of equations (1.1.1) is then equivalent to the single vector equation

\[
A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.
\]  
(1.1.2)

(Since column vectors take up a lot of space to print, it is convenient to introduce the following notation:

\[
[x_1, x_2, \ldots, x_n]^t = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad [b_1, b_2, \ldots, b_m]^t = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.
\]
These are special cases of the ‘transpose’ $A^t$ of an $m \times n$ matrix $A$, where $A^t$ is obtained from $A$ by interchanging rows and columns.

The following Proposition translates concepts related to solving equations (1.1.1) or (1.1.2) into concepts that we have used in discussing linear transformations.

**Proposition 1.2.**

(a) The numbers $x_1, \ldots, x_n$ are a solution of the homogeneous system

\[
\begin{align*}
    a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= 0 \\
    a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= 0 \\
    \vdots \\
    a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= 0
\end{align*}
\]

if and only if the vector $[x_1, \ldots, x_n]^t$ belongs to the null space $N(A)$ of the linear transformation $A$.

(b) The system (1.1.1) has a solution if and only if the vector $[b_1, b_2, \ldots, b_m]^t$ belongs to the range $R(A)$ of the linear transformation $A$.

We also have the following important general result:

**Lemma 1.3.** Let $b = [b_1, \ldots, b_m]^t \in R(A)$ and suppose that $x_p$ is one particular solution of the system (1.1.1). That is, suppose $A(x_p) = b$. Then the most general solution of the system (1.1.1) is $x = x_p + v$ where $v$ is any element of the null space $N(A)$.

**Proof.** First suppose that $x_p$ satisfies $A(x_p) = b$ and that $v$ is in the null space $N(A)$ so that $A(v) = 0$. Then since $A$ is linear, we have

\[
A(x_p + v) = A(x_p) + A(v) = b + 0 = b,
\]

and this shows that $x_p + v$ is also a solution of the equation $A(x) = b$.

Conversely, suppose $x$ is any solution of the equation $A(x) = b$. Then since $x_p$ is also a solution, and since $A$ is linear, we have

\[
A(x - x_p) = A(x) - A(x_p) = 0 - 0 = 0.
\]

This shows that the vector $v = x - x_p$ belongs to the null space $N(A)$, and thus $x = x_p + v$. Thus an arbitrary solution to the equation $A(x) = b$ is equal to the particular solution plus an element of the null space. This completes the proof. \qed

### 1.2. Gauss-Jordan Elimination.

To find a systematic method for solving the system (1.1.1), we begin by introducing the following standard notation. We do not really need to write down the unknown $x_1, \ldots, x_n$, and we can abbreviate the system (1.1.1) by writing down the $m \times (n+1)$ matrix

\[
A^+ = \begin{bmatrix}
    a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\
    a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m
\end{bmatrix}.
\]

(The vertical line inside the matrix just reminds us that the last column is the data on the right hand side of the system.) We then observe that the set of solutions of the system (1.1.1) is unchanged if we do any of the following “elementary operations” on the matrix $A^+$:

(A) Interchange two rows of $A^+$;

(B) Multiply any row of $A^+$ by a non-zero scalar;

(C) Add any multiple of one row of $A^+$ to a different row of $A^+$.

Our objective is to perform a series of these elementary operations so that, at the end, the solutions are easy to find.
1.2.1. **An Example.** Consider the following system of four equations in seven unknowns:

\[
\begin{align*}
3x_3 - 3x_4 - 6x_5 - 6x_6 - 30x_7 &= a \\
3x_1 - 6x_2 + 6x_3 + 9x_4 + 9x_5 - 45x_6 + 21x_7 &= b \\
2x_1 - 4x_2 + 4x_3 + 6x_4 + 10x_5 - 26x_6 + 34x_7 &= c \\
-x_1 + 2x_2 - x_3 - 4x_4 - 2x_5 + 16x_6 - 2x_7 &= d
\end{align*}
\]

(1.2.2)

(We have used letters rather than numbers on the right hand side in order to find out when this system has a solution.) We represent this system as a \(4 \times 8\) matrix:

\[
\begin{bmatrix}
0 & 0 & 3 & -3 & -6 & -6 & -30 & a \\
3 & -6 & 6 & 9 & 9 & -45 & 21 & b \\
2 & -4 & 4 & 6 & 10 & -26 & 34 & c \\
-1 & 2 & -1 & -4 & -2 & 16 & -2 & d
\end{bmatrix}
\]

(1.2.3)

We now use operations (A), (B), and (C) over and over again to reduce the matrix.

**Step 1.** We want a non-zero entry in the upper left-hand corner. We can do this by interchanging equations 1 and 2. We get:

\[
\begin{bmatrix}
3 & -6 & 6 & 9 & 9 & -45 & 21 & b \\
0 & 0 & 3 & -3 & -6 & -6 & -30 & a \\
2 & -4 & 4 & 6 & 10 & -26 & 34 & c \\
-1 & 2 & -1 & -4 & -2 & 16 & -2 & d
\end{bmatrix}
\]

**Step 2.** We want to make the entry in the upper right hand corner equal to 1. We can do this by dividing equation 1 by 3. We get:

\[
\begin{bmatrix}
1 & -2 & +2 & +3 & +3 & -15 & +7 & \frac{1}{3}b \\
0 & 0 & 3 & -3 & -6 & -6 & -30 & a \\
2 & -4 & 4 & 6 & 10 & -26 & +34 & c \\
-1 & 2 & -1 & -4 & -2 & 16 & -2 & d
\end{bmatrix}
\]

**Step 3.** Now we want to make all the entries under this 1 equal to zero. We can do this by subtracting a multiple of equation 1 from the other equations. Thus replacing equation 3 by equation 3 minus 2 times equation 1 gives:

\[
\begin{bmatrix}
1 & -2 & 2 & 3 & 3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & 3 & -3 & -6 & -6 & -30 & a \\
0 & 0 & 0 & +4 & +4 & 20 & c - \frac{2}{3}b \\
-1 & 2 & -1 & -4 & -2 & 16 & -2 & d
\end{bmatrix}
\]

**Step 4.** We can also make the first entry of the last row equal to zero by replacing equation 4 by equation 4 plus equation 1. We get:

\[
\begin{bmatrix}
1 & -2 & 2 & 3 & 3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & +3 & -3 & -6 & -6 & -30 & a \\
0 & 0 & 0 & +4 & +4 & 20 & c - \frac{2}{3}b \\
0 & 0 & +1 & -1 & +1 & +5 & d + \frac{1}{3}b
\end{bmatrix}
\]

**Step 5.** We are now finished with the first column (and at least for the moment, with the first row as well). We look for the first non-zero entry in the second row or below which is as far to the left as possible. In this case, such a term appears in the third column. (If necessary, we could switch rows to bring the element up to the second row.) We want to make this entry equal to 1. We do this by dividing equation 2 by 3. We get:

\[
\begin{bmatrix}
1 & -2 & 2 & +3 & 3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & +1 & -1 & -2 & -2 & -10 \frac{1}{3}a \\
0 & 0 & 0 & +4 & +4 & 20 & c - \frac{2}{3}b \\
0 & 0 & +1 & -1 & +1 & +5 & d + \frac{1}{3}b
\end{bmatrix}
\]
Step 6. We now want to make all the entries under this 1 equal to zero. We can do this by adding an appropriate multiple of this row to rows below. We do not have to do anything to the third row, but we deal with the 4th row by replacing equation 4 by equation 4 minus equation 2. We get:

\[
\begin{pmatrix}
1 & -2 & +2 & +3 & +3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & 1 & -1 & -2 & -2 & -10 & \frac{1}{3}a \\
0 & 0 & 0 & 0 & +4 & +4 & 20 & c - \frac{2}{3}b \\
0 & 0 & 0 & 0 & +3 & +3 & 15 & d + \frac{1}{3}b - \frac{1}{3}a
\end{pmatrix}
\]

Step 7. We are now done with the first three columns (and at least for the moment, with the first two rows). Again we look for the first non-zero entry in the third row and below which is as far to the left as possible. In this case it is the 4 in the third row. We want to make this entry equal to 1, which we can do by dividing equation 3 by 4. We get:

\[
\begin{pmatrix}
1 & -2 & 2 & 3 & 3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & 1 & -1 & -2 & -2 & -10 & \frac{1}{3}a \\
0 & 0 & 0 & 0 & 1 & 1 & 5 & \frac{1}{4}c - \frac{1}{6}b \\
0 & 0 & 0 & 0 & +3 & +3 & 15 & d + \frac{1}{6}b - \frac{1}{3}a
\end{pmatrix}
\]

Step 8. Again we want to make all the entries under this 1 equal to zero. We do this by replacing equation 4 by equation 4 minus 3 times equation 3. We get:

\[
\begin{pmatrix}
1 & -2 & 2 & 3 & 3 & -15 & 7 & \frac{1}{3}b \\
0 & 0 & 1 & -1 & -2 & -2 & -10 & \frac{1}{3}a \\
0 & 0 & 0 & 0 & 1 & 1 & 5 & \frac{1}{4}c - \frac{1}{6}b \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & d - \frac{1}{6}b - \frac{1}{3}a - \frac{3}{4}c
\end{pmatrix}
\]

At this point, what information can we obtain from this matrix?

(a) Since the 4th row consists entirely of zeros, it is clear that in order for the vector \([a, b, c, d]^t\) to be in the range, we must have

\[-\frac{1}{3}a - \frac{1}{6}b - \frac{3}{4}c + d = 0.\]

More generally, each row consisting entirely of zeros leads to a conditions on the right-hand side data to insure that the data is in the range of the linear transformation.

(b) If this condition is satisfied, then we can always find a particular solution \(x_p\). We consider a tuple with zeros except in the columns corresponding to leading 1’s in various rows. Thus in this case, since the leading 1’s occur in columns 1, 3, and 5, we consider \(x_p\) of the form

\[x_p = (x_1, 0, x_3, 0, x_5, 0, 0).\]

The equations are then

\[x_1 + 2x_2 + 3x_5 = \frac{1}{3}b\]
\[x_3 - 2x_5 = \frac{1}{3}a\]
\[x_5 = \frac{1}{4}c - \frac{1}{6}b.\]

We can easily solve for \(x_1, x_3, x_5\) by back substitution:

\[x_5 = \frac{1}{4}c - \frac{1}{6}b,\]
\[x_3 = \frac{1}{3}a + \frac{1}{2}c - \frac{1}{3}b,\]
\[x_1 = \frac{1}{3}b - 3x_5 - 2x_3, \text{ etc.}\]
The point is that, with the reductions we have done, we can easily find necessary and sufficient conditions on the right-hand-side date to guarantee that it is in the range, and hence that the system is solvable. We can also find a particular solution by back substitution.

We can also further reduce our matrix by making all the terms in a column containing a leading 1 equal to zero except for the 1 itself. Thus replacing equation 1 by equation 1 minus 2 times equation 2, we get:

**Step 9.** Replacing equation 1 by equation 1 minus 2 times equation 2, we get:

\[
\begin{bmatrix}
1 & -2 & 3 & 3 & -15 & 7 \\
0 & 0 & 1 & -1 & -2 & -2 & -10 \\
0 & 0 & 0 & 0 & 1 & 1 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\frac{1}{7}b \\
\frac{1}{7}a \\
\frac{1}{7}c - \frac{1}{4}b \\
d - \frac{1}{6}b - \frac{1}{3}a - \frac{3}{4}c
\end{bmatrix}
\]

\[
\rightarrow
\begin{bmatrix}
1 & -2 & 0 & 3 & 7 & -11 & -13 \\
0 & 0 & 1 & -1 & -2 & -2 & -10 \\
0 & 0 & 0 & 0 & 1 & 1 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\frac{1}{7}b - \frac{2}{3}a \\
\frac{1}{7}a \\
\frac{1}{7}c - \frac{1}{4}b \\
d - \frac{1}{6}b - \frac{1}{3}a - \frac{3}{4}c
\end{bmatrix}
\]

**Step 10.** We next subtract 7 times equation 3 from equation 1 to get:

\[
\rightarrow
\begin{bmatrix}
1 & -2 & 0 & 3 & 0 & -18 & -48 \\
0 & 0 & 1 & -1 & -2 & -2 & -10 \\
0 & 0 & 0 & 0 & 1 & 1 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\frac{1}{7}b - \frac{2}{3}a - \frac{7}{4}c - \frac{7}{2}b \\
\frac{1}{7}a \\
\frac{1}{7}c - \frac{1}{4}b \\
d - \frac{1}{6}b - \frac{1}{3}a - \frac{3}{4}c
\end{bmatrix}
\]

**Step 11.** Finally we add 2 times equation 3 to equation 2 to get:

\[
\rightarrow
\begin{bmatrix}
1 & -2 & 0 & 3 & 0 & -18 & -48 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\frac{1}{7}b - \frac{3}{4}a - \frac{7}{4}c - \frac{7}{2}b \\
\frac{1}{7}a + \frac{1}{3}c - \frac{7}{3}b \\
\frac{1}{7}c - \frac{1}{4}b \\
d - \frac{1}{6}b - \frac{1}{3}a - \frac{3}{4}c
\end{bmatrix}
\]

Now suppose that \([x_1, x_2, x_3, x_4, x_5, x_6, x_7]^t\) belongs to the null space. Then we have

\[
x_5 = -x_6 - 5x_7, \\
x_3 = x_4, \\
x_1 = 2x_2 - 3x_4 + 18x_6 + 48x_7
\]

Thus an element of the null space has the form

\[
(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (2x_2 - 3x_4 + 18x_6 + 48x_7, x_2, x_4, -x_6 - 5x_7, x_6, x_7)
\]

\[
= x_2(2, 1, 0, 0, 0, 0, 0) + x_4(-3, 0, 1, 0, 0, 0, 0) + x_6(18, 0, 0, 0, -1, 1, 0) + x_7(48, 0, 0, 0, 0, 0, 1).
\]

This gives us a basis for the null space.

### 2. Determinants

#### 2.1. Definitions and basic properties.

We want to associate to each \(n \times n\) matrix of numbers (real or complex) another number, called the determinant of \(A\). You have probably seen the concept of the determinant of a \(2 \times 2\) matrix or a \(3 \times 3\)
matrix, and we begin by reviewing these notions. We have

\[
\det \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1;
\]

\[
\det \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \det \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \det \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \det \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}
= a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1.
\]

In order to understand the meaning of these expressions, it helps to think of the determinant not as a function of 4 or 9 entries, but rather as functions of vectors, for example the 2 row vectors in a $2 \times 2$ matrix or the 3 row vectors of a $3 \times 3$ matrix. Thus we think of the determinant as follows:

\[
\det \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \det \begin{vmatrix} (a_1, a_2) \\ (b_1, b_2) \end{vmatrix} = a_1 b_2 - a_2 b_1
\]

and similarly

\[
\det \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \det \begin{vmatrix} (a_1, a_2, a_3) \\ (b_1, b_2, b_3) \\ (c_1, c_2, c_3) \end{vmatrix}
= a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1.
\]

It is not too hard the check that $2 \times 2$ and $3 \times 3$ determinants have the following properties:

(a) If we multiply any row vector by a scalar, the determinant is then multiplies by the same scalar. The point here is that in the sum of products on the right hand side, each term contains exactly one entry from each vector.

(b) If we replace any row vector by a sum of two vectors, then the resulting determinant is the sum of the corresponding determinants.

(c) If any two of the row vectors are the same, then the determinant is equal to zero.

(d) The determinant of the identity matrix is equal to 1.

Thus if we were to write (say) the $3 \times 3$ matrix as $\det \begin{vmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{vmatrix}$ where $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are vectors in $\mathbb{R}^3$ then these four properties say that:

(a) If $\lambda$ is a scalar, then

\[
\det \begin{vmatrix} \lambda \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{vmatrix} = \det \begin{vmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{vmatrix} = \lambda \det \begin{vmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{vmatrix}.
\]
(b)\[
\det \begin{vmatrix}
v_1' + v_2' \\
v_2 \\
v_3 
\end{vmatrix} = \det \begin{vmatrix}
v_1' \\
v_2 \\
v_3 
\end{vmatrix} + \det \begin{vmatrix}
v_1' \\
v_2' \\
v_3 
\end{vmatrix},
\]
\[
\det \begin{vmatrix}
v_1 \\
v_2' + v_3'' \\
v_3 
\end{vmatrix} = \det \begin{vmatrix}
v_1 \\
v_2' \\
v_3 
\end{vmatrix} + \det \begin{vmatrix}
v_1 \\
v_2'' \\
v_3 
\end{vmatrix},
\]
\[
\det \begin{vmatrix}
v_1 \\
v_2' \\
v_3' + v_3'' 
\end{vmatrix} = \det \begin{vmatrix}
v_1 \\
v_2' \\
v_3' 
\end{vmatrix} + \det \begin{vmatrix}
v_1 \\
v_2'' \\
v_3' 
\end{vmatrix},
\]
\[
\det \begin{vmatrix}
v_1' \\
v_2 \\
v_3 
\end{vmatrix} = \det \begin{vmatrix}
v_1 \\
v_2 \\
v_3 
\end{vmatrix} + \det \begin{vmatrix}
v_1 \\
v_2' \\
v_3 
\end{vmatrix} + \det \begin{vmatrix}
v_1 \\
v_2 \\
v_3' 
\end{vmatrix} + \det \begin{vmatrix}
v_1 \\
v_2' \\
v_3' 
\end{vmatrix}.
\]

(c)\[
\det \begin{vmatrix}
v_1' \\
v_2 \\
v_3 
\end{vmatrix} = \det \begin{vmatrix}
v_1 \\
v_2 \\
v_3 
\end{vmatrix} = \det \begin{vmatrix}
v_1 \\
v_2 \\
v_3 
\end{vmatrix} = 0.
\]

(d)\[
\begin{vmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 
\end{vmatrix} = 1.
\]

We can use the analogues of these four properties to generalize the definition of determinant to arbitrary square matrices. Thus let

\[
A = \begin{bmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n,1} & a_{n,2} & \cdots & a_{n,n} 
\end{bmatrix} = \begin{bmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_n 
\end{bmatrix}
\]

be an \(n \times n\) matrix, where \(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n\) are the rows of \(A\). Thus

\[
\bar{A}_1 = (a_{1,1}, a_{1,2}, \ldots, a_{1,n-1}, a_{1,n}),
\]
\[
\bar{A}_2 = (a_{2,1}, a_{2,2}, \ldots, a_{2,n-1}, a_{2,n}),
\]
\[
\vdots
\]
\[
\bar{A}_j = (a_{j,1}, a_{j,2}, \ldots, a_{j,n-1}, a_{j,n}),
\]
\[
\vdots
\]
\[
\bar{A}_n = (a_{n,1}, a_{n,2}, \ldots, a_{n,n-1}, a_{n,n}).
\]

**Theorem 2.1.** There exists one and only one function, called the **determinant**, which is defined for all \(n \times n\) matrices \(A\) and which has the following properties:

(A) For any \(1 \leq j \leq n\), \(\det \begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\lambda \bar{A}_j \\
\vdots \\
\bar{A}_n 
\end{vmatrix} = \lambda \det \begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_j \\
\vdots \\
\bar{A}_n 
\end{vmatrix}.
\)
(B) For any \(1 \leq j \leq n\),

\[
\begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_j' + \bar{A}_j'' \\
\vdots \\
\bar{A}_n
\end{vmatrix} = \begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_j' \\
\vdots \\
\bar{A}_n
\end{vmatrix} + \begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_j'' \\
\vdots \\
\bar{A}_n
\end{vmatrix}.
\]

(C) For any \(1 \leq j \leq n - 1\), if \(\bar{A}_j = \bar{A}_{j+1}\), then \(\begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_n
\end{vmatrix} = 0\); i.e. if any two adjacent rows are equal, the determinant is equal to zero.

(D)

\[
\det \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix} = 1.
\]

Note that properties (A) and (B) say that, if we keep all rows except the \(j^{th}\) fixed, then the determinant is a real-valued linear mapping of the \(j^{th}\) row.

We shall defer the proof of the existence of the determinant function for a while, and explore some of the consequences of properties (A) through (D).

**Theorem 2.2.** The determinant function has the following properties:

(a) If any row is the zero vector, then the determinant equals zero.

(b) The determinant change sign if any two adjacent rows are interchanged.

(c) The determinant changes sign if any two distinct rows are interchanged.

(d) If any two rows of the matrix are equal, then the determinant equals zero.

(e) If the rows of the matrix are linearly dependent, then the determinant is zero.

**Proof.** Assertion (a) follows from property (A) of Theorem 2.1 since the zero row \(0\) is equal to the scalar 0 times the zero row \(0\). To establish assertion (b), first note that it follows from property (C) of Theorem 2.1 that

\[
\det \begin{vmatrix}
\bar{A}_1 \\
\bar{A}_2 \\
\vdots \\
\bar{A}_j + \bar{A}_{j+1} \\
\vdots \\
\bar{A}_n
\end{vmatrix} = 0.
\]
since we have the same vector in the $j^{th}$ and $(j+1)^{st}$ rows. Then using properties (A) and (B) of Theorem 2.1, we can expand this determinant and get

$$0 = \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_n \end{vmatrix} + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_{j+1} \\ \vdots \\ \bar{A}_n \end{vmatrix} + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_{j+1} \end{vmatrix} + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_n \end{vmatrix} = 0 + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_{j+1} \\ \vdots \\ \bar{A}_n \end{vmatrix} + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_{j+1} \end{vmatrix} + \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_n \end{vmatrix}$$

This shows that interchanging adjacent rows changes the sign of the determinant.

Assertion (c) follows since one can interchange any two rows by making an odd number of interchanges of adjacent rows. Assertion (d) follows from assertion (c) since the only number which is equal to its negative is zero.

Finally, suppose that the rows of the matrix $A$ are linearly dependent. Then there is a relation

$$\alpha_1 \bar{A}_1 + \cdots + \alpha_n \bar{A}_n = 0$$

with at least one of the coefficients $\alpha_j \neq 0$. Then we can solve for $\bar{A}_j$ in terms of the other rows:

$$\bar{A}_j = -\frac{\alpha_1}{\alpha_j} \bar{A}_1 - \cdots - \frac{\alpha_{j-1}}{\alpha_j} \bar{A}_{j-1} - \frac{\alpha_{j+1}}{\alpha_j} \bar{A}_{j+1} - \cdots - \frac{\alpha_n}{\alpha_j} \bar{A}_n.$$

But then using the fact that the determinant is a linear function of the $j^{th}$ row, we have

$$\det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_n \end{vmatrix} = -\frac{\alpha_1}{\alpha_j} \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_j \\ \vdots \\ \bar{A}_n \end{vmatrix} - \frac{\alpha_{j-1}}{\alpha_j} \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_{j-1} \\ \vdots \\ \bar{A}_n \end{vmatrix} - \frac{\alpha_{j+1}}{\alpha_j} \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_{j+1} \\ \vdots \\ \bar{A}_n \end{vmatrix} - \cdots - \frac{\alpha_n}{\alpha_j} \det \begin{vmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_n \\ \vdots \\ \bar{A}_n \end{vmatrix}$$

$$= 0$$

since all of the determinants on the right hand side have a repeated row. This establishes property (e) and completes the proof. \[\square\]
2.2. Computing a $3 \times 3$ determinant.

We next want to show that the properties listed in Theorem 2.1 do in fact define a unique function. (We still assume the existence of at least one such function.) Rather than doing the argument for general $n$, which would be very messy, we will consider the case of a $3 \times 3$ determinant. However, one should be able to see what will happen in general.

Thus we now use the properties of determinants to compute $\det |M| = \det \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$. We write the rows of the matrix $M$ as

$$(a_1, a_2, a_3) = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3,$$

$$(b_1, b_2, b_3) = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3,$$

$$(c_1, c_2, c_3) = c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3.$$

Then using the linearity in each row, we have

$$\det |M| = \det \begin{vmatrix} a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3 \\ b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix}$$

$$= a_1 \det \begin{vmatrix} \mathbf{e}_1 \\ b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix} + a_2 \det \begin{vmatrix} \mathbf{e}_1 \\ b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix} + a_3 \det \begin{vmatrix} \mathbf{e}_1 \\ b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix}$$

$$= a_1 b_1 \det \begin{vmatrix} \mathbf{e}_1 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \\ \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix} + a_2 b_2 \det \begin{vmatrix} \mathbf{e}_1 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \\ \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix} + a_3 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \\ \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 \end{vmatrix}$$

$$= a_1 b_1 c_1 \det \begin{vmatrix} \mathbf{e}_1 \\ e_1 \\ e_2 \end{vmatrix} + a_1 b_1 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_1 \\ e_3 \end{vmatrix} + a_1 b_1 c_1 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_1 b_2 c_2 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_1 b_2 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_1 b_2 c_1 \det \begin{vmatrix} \mathbf{e}_1 \\ e_3 \\ e_3 \end{vmatrix} + a_2 b_2 c_2 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_2 b_2 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_2 b_2 c_1 \det \begin{vmatrix} \mathbf{e}_1 \\ e_3 \\ e_3 \end{vmatrix} + a_2 b_3 c_2 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_2 b_3 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_2 \\ e_3 \end{vmatrix} + a_2 b_3 c_1 \det \begin{vmatrix} \mathbf{e}_1 \\ e_3 \\ e_3 \end{vmatrix} + a_3 c_2 \det \begin{vmatrix} \mathbf{e}_1 \\ e_1 \\ e_3 \end{vmatrix} + a_3 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_1 \\ e_3 \end{vmatrix} + a_3 c_3 \det \begin{vmatrix} \mathbf{e}_1 \\ e_1 \\ e_2 \end{vmatrix}.$$
This leaves only six terms:

\[ + a_3 b_1 c_1 \det \begin{vmatrix} e_3 \\ e_1 \end{vmatrix} + a_3 b_1 c_2 \det \begin{vmatrix} e_3 \\ e_2 \end{vmatrix} + a_3 b_1 c_3 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} \\
+ a_3 b_2 c_1 \det \begin{vmatrix} e_3 \\ e_2 \end{vmatrix} + a_3 b_2 c_2 \det \begin{vmatrix} e_3 \\ e_2 \end{vmatrix} + a_3 b_2 c_3 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} \\
+ a_3 b_3 c_1 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} + a_3 b_3 c_2 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} + a_3 b_3 c_3 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} \]

 Altogether there are 27 terms on the right hand side. However, any determinant with repeated rows is zero. This leaves only six terms:

\[ \det |M| = a_1 b_2 c_3 \det \begin{vmatrix} e_1 \\ e_3 \end{vmatrix} + a_1 b_3 c_2 \det \begin{vmatrix} e_1 \\ e_2 \end{vmatrix} + a_2 b_1 c_3 \det \begin{vmatrix} e_1 \\ e_3 \end{vmatrix} \\
+ a_2 b_3 c_1 \det \begin{vmatrix} e_2 \\ e_3 \end{vmatrix} + a_3 b_1 c_2 \det \begin{vmatrix} e_3 \\ e_2 \end{vmatrix} + a_3 b_2 c_3 \det \begin{vmatrix} e_3 \\ e_3 \end{vmatrix} \]

Finally, we can compute each of the six remaining determinants since we can permute the rows to make them all the identity matrix.

\[ \det \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = +1, \quad \det \begin{vmatrix} e_2 \\ e_3 \\ e_1 \end{vmatrix} = -\det \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = -1, \quad \det \begin{vmatrix} e_1 \\ e_3 \\ e_2 \end{vmatrix} = -\det \begin{vmatrix} e_2 \\ e_3 \\ e_1 \end{vmatrix} = -1, \quad \det \begin{vmatrix} e_2 \\ e_1 \\ e_3 \end{vmatrix} = +\det \begin{vmatrix} e_3 \\ e_2 \\ e_1 \end{vmatrix} = +1 \]

\[ \det \begin{vmatrix} e_3 \\ e_2 \\ e_1 \end{vmatrix} = +\det \begin{vmatrix} e_1 \\ e_2 \\ e_3 \end{vmatrix} = +1, \quad \det \begin{vmatrix} e_3 \\ e_1 \\ e_2 \end{vmatrix} = +\det \begin{vmatrix} e_2 \\ e_3 \\ e_1 \end{vmatrix} = +1, \quad \det \begin{vmatrix} e_1 \\ e_3 \\ e_2 \end{vmatrix} = -\det \begin{vmatrix} e_2 \\ e_1 \\ e_3 \end{vmatrix} = -1 \]

Therefore

\[ \det |M| = a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1 \]

\[ = a_1 (b_2 c_3 - b_3 c_2) - a_2 (b_1 c_3 - b_3 c_1) + a_3 (b_1 c_2 - b_2 c_1). \]

This is the formula we started with for \(3 \times 3\) determinants.

3. Permutations

**Definition 3.1.** A permutation of the set \(\{1, \ldots, n\}\) is a mapping or function \(\sigma : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}\) which is one-to-one (and hence onto), or else onto (and hence one-to-one). (We often think of a permutation as a rearrangement of the set \(\{1, \ldots, n\}\).) If \(\sigma\) and \(\tau\) are two permutations of the set \(\{1, \ldots, n\}\) then the product \(\sigma \tau\) is the composition of the two mappings, so that \(\sigma \tau(j) = \sigma(\tau(j))\).

Note that there are \(n! = n(n - 1) \cdots 3 \cdot 2 \cdot 1\) distinct permutations of the set \(\{1, \ldots, n\}\). This is because a one-to-one mapping \(\sigma\) of this set to itself can take 1 to any of the \(n\) elements of \(\{1, \ldots, n\}\), but then can take 2 to only the \((n - 1)\) elements which are not equal to \(\sigma(1)\), can take 3 to only the \((n - 2)\) elements which are not equal to \(\sigma(1)\) or \(\sigma(2)\), etc.

3.1. Examples.

If \(n = 2\), there are exactly two permutations of the set \(\{1, 2\}\). The first is the identity mapping \(\iota\) given by \(\iota(1) = 1\) and \(\iota(2) = 2\); the second is the map \(\sigma\) which interchanges the two elements so that \(\sigma(1) = 2\) and \(\sigma(2) = 1\). The rules for multiplying these two permutations are:

\[ \iota \iota = \iota \quad \iota \sigma = \sigma \]
\[ \sigma \iota = \sigma \quad \sigma \sigma = \iota \]
If \( n = 3 \) there are \( 3! = 6 \) permutations of the set \( \{1, 2, 3\} \). To write them down, we use the following notation:

\[
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\sigma}{\rightarrow}
\begin{bmatrix}
a \\
b \\
c
\end{bmatrix}
\]

means that

\( \sigma(1) = a, \quad \sigma(2) = b, \quad \sigma(3) = c, \)

where \( a, b, c \) are the numbers 1, 2, 3 in some possibly different order. Then the six permutations of \( \{1, 2, 3\} \) are:

\[
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\iota}{\rightarrow}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\sigma}{\rightarrow}
\begin{bmatrix}
1 \\
3 \\
2
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\tau}{\rightarrow}
\begin{bmatrix}
2 \\
1 \\
3
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\lambda}{\rightarrow}
\begin{bmatrix}
2 \\
3 \\
1
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\mu}{\rightarrow}
\begin{bmatrix}
3 \\
1 \\
2
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\overset{\eta}{\rightarrow}
\begin{bmatrix}
3 \\
2 \\
1
\end{bmatrix}
\]

We can compute the 36 possible products of these permutations, and put them in a ‘multiplication table’:

<table>
<thead>
<tr>
<th></th>
<th>( \iota )</th>
<th>( \sigma )</th>
<th>( \tau )</th>
<th>( \lambda )</th>
<th>( \mu )</th>
<th>( \eta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \iota )</td>
<td>( \iota )</td>
<td>( \sigma )</td>
<td>( \tau )</td>
<td>( \lambda )</td>
<td>( \mu )</td>
<td>( \eta )</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>( \sigma )</td>
<td>( \iota )</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>( \tau )</td>
<td>( \tau )</td>
<td>?</td>
<td>( \iota )</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>( \lambda )</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>( \iota )</td>
<td>?</td>
</tr>
<tr>
<td>( \mu )</td>
<td>( \mu )</td>
<td>?</td>
<td>?</td>
<td>( \iota )</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>( \eta )</td>
<td>( \eta )</td>
<td>?</td>
<td>( \tau )</td>
<td>?</td>
<td>?</td>
<td>( \iota )</td>
</tr>
</tbody>
</table>

Here we have multiplies an element of the top row times an element of the left column, in that order. This means we do the permutation from the left column first, and then follow it with the permutation from the top row. You should fill in all the question marks.

3.2. **Groups.**

If calculated correctly, the set of permutations \( \{\iota, \sigma, \tau, \lambda, \mu, \eta\} \) and the multiplication given in the table have the following properties:

(A) The product of any two permutations is again a permutation. (That is, this set is closed under the operation of multiplication.)

(B) There is a permutation \( \iota \) so that for any permutation \( f \), we have

\[
\iota f = f \iota = f.
\]

(That is, the permutation \( \iota \) acts like an identity, or the number 1.)

(C) For any permutation \( f \) there is a permutation \( g \) so that \( fg = gf = \iota \). (That is, for each permutation, there is an inverse.)

(D) For any three permutations \( f, g, h \), we have \( (fg)h = f(gh) \). (That is, the multiplication is associative.)

This is an example of what is called a **group**.
Definition 3.2. Let G be a non-empty set. Then G is a group if the following four properties hold:

(A) If x, y are any two elements of G, there is a unique product xy which is again an element of G.

(B) There is a unique element e ∈ G such that for any x ∈ G, ex = xe = x.

(C) For any x ∈ G there is a unique y ∈ G so that xy = yx = e.

(D) For any three elements x, y, z ∈ G, we have (xy)z = x(yz).

The group is said to be commutative or Abelian if it is also true that

(E) For any two elements x, y ∈ G, we have xy = yx.

10/29/12

3.3. The sign or parity of a permutation.

It is not hard to see that any permutation of \{1, \ldots, n\} can be written as a product of simpler permutations which interchange only two elements of the set. What is true (and remarkable!) but not obvious, is that while this can be done in many ways, the number of such interchanges is either always odd, or is always even. To see this, we argue as follows:

Consider the polynomial of \(n\) variables \(x_1, x_2, \ldots, x_n\) given by

\[
P(x_1, \ldots, x_n) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4) \cdots (x_1 - x_{n-1})(x_1 - x_n) \\
\times (x_2 - x_3)(x_2 - x_4) \cdots (x_2 - x_{n-1})(x_2 - x_n) \\
\times (x_3 - x_4) \cdots (x_3 - x_{n-1})(x_3 - x_n) \\
\times \cdots \\
\times (x_{n-2} - x_{n-1})(x_{n-2} - x_n) \\
\times (x_{n-1} - x_n)
\]

\[
= \prod_{j=1}^{n-1} \prod_{k=j+1}^{n} (x_j - x_k).
\]

Note that for any pair of distinct indices \(j\) and \(k\), either the factor \((x_j - x_k)\) or the factor \((x_k - x_j)\) occurs in \(P\), and only one of them occurs. Now let \(\sigma\) be a permutation of the set \(\{1, \ldots, n\}\), and consider the polynomial

\[
P_\sigma(x_1, \ldots, x_n) = P(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).
\]

Since the mapping \(\sigma\) is one-to-one, it is again true that for any pair of distinct indices \(j\) and \(k\), either the factor \((x_j - x_k)\) or the factor \((x_k - x_j)\) occurs in \(P_\sigma\), and only one of them appears. Thus the factors of \(P_\sigma\) are the same as the factors of \(P\), except for possible changes of sign. In particular, the ratio of the two polynomials is \(\pm 1\).

Definition 3.3. The sign or parity of the permutation \(\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}\) is

\[
\text{sgn}(\sigma) = \frac{P(x_1, \ldots, x_n)}{P_\sigma(x_1, \ldots, x_n)}.
\]

Example: Consider the permutation \(\sigma\) of \(\{1, 2, 3, 4\}\) given by

\[
\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \\ 3 & 1 & 4 & 2 \\ 4 & 2 & 1 & 3 \end{bmatrix}.
\]

Then

\[
P(x_1, x_2, x_3, x_4) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4),
\]

\[
P_\sigma(x_1, x_2, x_3, x_4) = (x_3 - x_1)(x_3 - x_4)(x_3 - x_2)(x_1 - x_4)(x_1 - x_2)(x_4 - x_2),
\]

and so

\[
\frac{P(x_1, x_2, x_3, x_4)}{P_\sigma(x_1, x_2, x_3, x_4)} = -1 \text{ since there are three factors which have changed sign.}
\]
Proposition 3.4. Suppose that $\sigma$ is permutation of $\{1, \ldots, n\}$, and that $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ are two sets of variables. Then

$$\frac{P(x_1, \ldots, x_n)}{P_{\sigma}(x_1, \ldots, x_n)} = \frac{P(y_1, \ldots, y_n)}{P_{\sigma}(y_1, \ldots, y_n)}.$$  

Proof. This is clear, since the value of the quotient does not depend on the name of the variables. \hfill $\Box$

Proposition 3.5. Suppose that $\sigma$ and $\tau$ are permutations of $\{1, \ldots, n\}$. Then

$$\frac{P(x_1, \ldots, x_n)}{P_{\sigma}(x_1, \ldots, x_n)} = \frac{P_{\tau}(x_1, \ldots, x_n)}{P_{\sigma \tau}(x_1, \ldots, x_n)}.$$  

Proof. for $1 \leq j \leq n$ let $y_j = x_{\tau(j)}$. Then

$$\frac{P_{\tau}(x_1, \ldots, x_n)}{P_{\sigma \tau}(x_1, \ldots, x_n)} = \frac{P(x_{\tau(1)}, \ldots, x_{\tau(n)})}{P_{\sigma \tau}(x_1, \ldots, x_n)} = \frac{P(x_{\tau(1)}, \ldots, x_{\tau(n)})}{P_{\sigma}(x_{\tau(1)}, \ldots, x_{\tau(n)})} = \frac{P(y_1, \ldots, y_n)}{P_{\sigma}(y_1, \ldots, y_n)} = \frac{P(x_1, \ldots, x_n)}{P_{\sigma}(x_1, \ldots, x_n)}$$

as asserted. \hfill $\Box$

Lemma 3.6. Suppose that $\sigma$ and $\tau$ are permutations of $\{1, \ldots, n\}$. Then $\text{sgn}(\sigma \tau) = \text{sgn}(\sigma) \text{sgn}(\tau)$.

Proof. Using Proposition 3.5 we have

$$\text{sgn}(\sigma \tau) = \frac{P(x_1, \ldots, x_n)}{P_{\sigma \tau}(x_1, \ldots, x_n)} = \frac{P(x_1, \ldots, x_n)}{P_{\tau}(x_1, \ldots, x_n) P_{\sigma}(x_1, \ldots, x_n)} = \frac{P_{\tau}(x_1, \ldots, x_n)}{P_{\sigma}(x_1, \ldots, x_n) P_{\tau}(x_1, \ldots, x_n)} = \frac{\text{sgn}(\tau) \text{sgn}(\sigma)}{\text{sgn}(\tau)}$$

as asserted. \hfill $\Box$

Corollary 3.7. If $\sigma, \tau$ are permutations of $\{1, \ldots, n\}$ such that $\sigma \tau$ is the identity, then $\text{sgn}(\sigma) = |\text{sgn}(\tau)|^{-1}$.

Let $\sigma$ be the permutation of $\{1, \ldots, n\}$ which interchanges $n-1$ and $n$, and leaves everything else alone. The only factors in $P(x_1, \ldots, x_n)$ which are changed in $P_{\sigma}(x_1, \ldots, x_n)$ are

$$(x_1 - x_{n-1})(x_1 - x_n) \cdots (x_{n-2} - x_{n-1})(x_{n-2} - x_n)(x_{n-1} - x_n).$$

The first $2n-2$ of these factors are reordered but not changed in $P_{\sigma}$, while the last factor changes sign. Thus we see that $\text{sgn}(\sigma) = -1$.

Now let $j, k \in \{1, \ldots, n\}$ with $j \neq k$.

(I) Suppose first that $\{j, k\} \cap \{n-1, n\} = \emptyset$. Let $\tau$ be the permutation which interchanges $j$ with $n-1$ and $k$ with $n$. Then

$$\tau \sigma \tau(j) = \tau(\sigma(\tau(j))) = \tau(\sigma(n-1)) = \tau(n) = k,$$

$$\tau \sigma \tau(k) = \tau(\sigma(\tau(k))) = \tau(\sigma(n)) = \tau(n-1) = j,$$

$$\tau \sigma \tau(n-1) = \tau(\sigma(\tau(n-1))) = \tau(\sigma(j)) = \tau(j) = n-1,$$

$$\tau \sigma \tau(n) = \tau(\sigma(\tau(n))) = \tau(\sigma(k)) = \tau(j) = n.$$

Thus $\tau \sigma \tau$ is the permutation which interchanges $j$ and $k$ and leaves everything else fixed.
(II) Suppose that \( j \notin \{n - 1, n\} \). Let \( \tau_1 \) be the permutation which interchanges \( j \) with \( n - 1 \), and leaves everything else fixed. Then
\[
\tau_1 \sigma \tau_1(j) = \tau_1(\sigma(\tau_1(j))) = \tau_1(\sigma(n - 1)) = \tau_1(n) = n,
\]
\[
\tau_1 \sigma \tau_1(n - 1) = \tau_1(\sigma(\tau_1(n - 1))) = \tau_1(\sigma(j)) = \tau_1(j) = n - 1,
\]
\[
\tau_1 \sigma \tau_1(n) = \tau_1(\sigma(\tau_1(n))) = \tau_1(\sigma(n)) = \tau_1(n - 1) = j.
\]
Thus \( \tau_1 \sigma \tau_1 \) is the permutation which interchanges \( j \) and \( k \) and leaves everything else fixed.

Under the same assumption that \( j \notin \{n - 1, n\} \), let \( \tau_2 \) be the permutation which interchanges \( j \) with \( n - 1 \), and leaves everything else fixed. Then
\[
\tau_2 \sigma \tau_2(j) = \tau_2(\sigma(\tau_2(j))) = \tau_2(\sigma(n)) = \tau_2(n - 1) = n - 1,
\]
\[
\tau_2 \sigma \tau_2(n) = \tau_2(\sigma(\tau_2(n))) = \tau_2(\sigma(j)) = \tau_2(j) = n,
\]
\[
\tau_2 \sigma \tau_2(n - 1) = \tau_2(\sigma(\tau_2(n - 1))) = \tau_2(\sigma(n - 1)) = \tau_2(n) = j.
\]
Thus \( \tau_2 \sigma \tau_2 \) is the permutation which interchanges \( j \) and \( n - 1 \) and leaves everything else fixed.

In either case (I) or case (II), we can apply the Corollary and get
\[
\text{sgn}(\tau \sigma \tau) = \text{sgn}(\tau)\text{sgn}(\sigma)\text{sgn}(\tau) = \text{sgn}(\sigma)[\text{sgn}(\tau)]^2 = \text{sgn}(\sigma) = -1.
\]
Thus we have proved

**Lemma 3.8.** If \( \mu \) is a permutation which interchanges two distinct elements of \( \{1, \ldots, n\} \) then \( \text{sgn}(\mu) = -1 \).

**Corollary 3.9.** If \( \lambda \) is any permutation of \( \{1, \ldots, n\} \) and if \( \lambda \) can be achieved by making \( k \) interchanges of two elements, then
\[
\text{sgn}(\lambda) = (-1)^k.
\]
In particular, the number of interchanges is either always odd or always even.

### 3.4. The definition of an \( n \times n \) determinant.

Let \( \mathfrak{S}_n \) be the group of all permutations of the set \( \{1, \ldots, n\} \).

**Definition 3.10.** Let \( A = \{a_{j,k}\} \) be an \( n \times n \) matrix of scalars. Then
\[
\det(A) = \det \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} = \sum_{\sigma \in \mathfrak{S}_n} a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)} \text{sgn}(\sigma) = \sum_{\sigma \in \mathfrak{S}_n} \left( \prod_{j=1}^{n} a_{j,\sigma(j)} \right) \text{sgn}(\sigma).
\]

### 3.5. The basic properties of determinants.

**Theorem 3.11.** The determinant satisfies the following properties:

(a) \[
\det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = 1.
\]

(b) If any \( n - 1 \) rows of an \( n \times n \) matrix are held fixed, the determinant is a scalar-valued linear function of the remaining row.

(c) If any two rows of an \( n \times n \) matrix are equal, the determinant equals zero.

(i) If any two rows of an \( n \times n \) matrix are interchanged, the determinant is multiplied by \((-1)\).

(ii) If the rows of a matrix are linearly dependent, the determinant equals zero.
(d) If $A'$ is the transpose of the matrix $A$ (so that rows and columns are interchanged, then $\det |A| = \det |A'|$.

(i) If any two columns of an $n \times n$ matrix are equal, the determinant equals zero.

(ii) If any two columns of an $n \times n$ matrix are interchanged, the determinant is multiplied by $(-1)$.

(iii) If the columns of a matrix are linearly dependent, the determinant equals zero.

(c) If $A$ and $B$ are $n \times n$ matrices, then

$$\det |AB| = \det |A| \det |B|.$$  

(f) Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for $\mathbb{R}^n$. If $\mathbf{v} \in \mathbb{R}^n$ then $\mathbf{v} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_n \mathbf{e}_n$ where

$$\alpha_j = \frac{\det(\mathbf{e}_1, \ldots, \mathbf{e}_{j-1}, \mathbf{v}, \mathbf{e}_{j+1}, \ldots, \mathbf{e}_n)}{\det(\mathbf{e}_1, \ldots, \mathbf{e}_{j-1}, \mathbf{e}_j, \mathbf{e}_{j+1}, \ldots, \mathbf{e}_n)}$$

where we regard the vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathbf{v}$ either as rows of a matrix or as columns of a matrix.

Proof of (a) and (b). Property (a) follows from the formula in Definition 3.10 since the only product which does not have a zero in it is the one with the identity permutation. To establish property (b), suppose that the $j^{th}$ row $\mathbf{A}_j$ is given by by $\mathbf{A}_j = \alpha \mathbf{A}'_j + \beta \mathbf{A}_j''$. Then for $1 \leq k \leq n$ we have

$$a_{j,k} = a_{j,k}' + \beta a_{j,k}''.$$  

Then

$$a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{j,\sigma(j)} \cdots a_{n,\sigma(n)} \text{sgn}(\sigma) = a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots (a_{j,k}' + \beta a_{j,k}'') \cdots a_{n,\sigma(n)} \text{sgn}(\sigma)$$

$$= \alpha a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{j,k}' \cdots a_{n,\sigma(n)} \text{sgn}(\sigma)$$

$$+ \beta a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{j,k}'' \cdots a_{n,\sigma(n)} \text{sgn}(\sigma)$$

The linearity is now clear. \hfill $\Box$

Proof of (c) and its consequences. To prove property (c), note that if two rows are equal, (say the $j^{th}$ and $k^{th}$ rows), then each product $a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{j,\sigma(j)} \cdots a_{k,\sigma(j)} \cdots a_{n,\sigma(n)}$ appears twice, but the signs of the corresponding permutations will be opposite. Thus the sum will be zero. We have already seen how properties (c[i]) and (c[ii]) follow from (c). \hfill $\Box$

Proof of (d) and its consequences. To show that $\det |A| = \det |A'|$ we must show that

$$\sum_{\sigma \in \mathfrak{S}_n} \left( \prod_{j=1}^{n} a_{j,\sigma(j)} \right) \text{sgn}(\sigma) = \sum_{\tau \in \mathfrak{S}_n} \left( \prod_{j=1}^{n} a_{\tau(j),j} \right) \text{sgn}(\tau).$$

But commuting the terms in any one of the products on the left hand side we see that

$$\prod_{j=1}^{n} a_{\tau(j),j} \text{sgn}(\tau) = \prod_{j=1}^{n} a_{j,\tau^{-1}(j)} \text{sgn}(\tau) = \prod_{j=1}^{n} a_{j,\tau^{-1}(j)} \text{sgn}(\tau^{-1})$$

since $\text{sgn}(\tau) = \text{sgn}(\tau^{-1})$. Now as $\tau$ runs over the elements of $\mathfrak{S}_n$, the inverses $\tau^{-1}$ also run over the elements of $\mathfrak{S}_n$. This establishes the desired equality. The consequences (i), (ii), and (iii) now follow from the corresponding facts for rows. \hfill $\Box$

If $A$ is an $n \times n$ matrix, let $\mathbf{A}_j$ denote the $j^{th}$ row of $A$. Before proving (e), we first establish the following result.

**Lemma 3.12.** If $A$ and $B$ are two $n \times n$ matrices, then $(AB)_j = [\mathbf{A}_j] \cdot \mathbf{B}$, where $[\mathbf{A}_j] \cdot \mathbf{B}$ is the row vector obtained by matrix multiplication.
Proof. The \((j,k)^{th}\) entry of \(AB\) is \(\sum_{l=1}^{n} a_{j,l}b_{l,k}\), and so the \(j^{th}\) row of \((AB)\) is
\[
\left(\sum_{l=1}^{n} a_{j,l}b_{l,1}, \sum_{l=1}^{n} a_{j,l}b_{l,2}, \ldots, \sum_{l=1}^{n} a_{j,l}b_{l,k}, \ldots, \sum_{l=1}^{n} a_{j,l}b_{l,n}\right).
\]

On the other hand,
\[
[\overline{A_j}] \cdot B = [a_{j,1}, a_{j,2}, \ldots, a_{j,n}] \cdot \begin{bmatrix}
 b_{1,1} & b_{1,2} & \cdots & b_{1,k} & \cdots & b_{1,n} \\
 b_{2,1} & b_{2,2} & \cdots & b_{2,k} & \cdots & b_{2,n} \\
 \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
 b_{n,1} & b_{n,2} & \cdots & b_{n,k} & \cdots & b_{n,n}
\end{bmatrix}
\]
\[
= \left(\sum_{l=1}^{n} a_{j,l}b_{l,1}, \sum_{l=1}^{n} a_{j,l}b_{l,2}, \ldots, \sum_{l=1}^{n} a_{j,l}b_{l,k}, \ldots, \sum_{l=1}^{n} a_{j,l}b_{l,n}\right),
\]
which is the same. \(\square\)

Proof of \((\varepsilon)\). Fix the \(n \times n\) matrix \(B\), and define a new scalar-valued function \(\delta\) of \(n \times n\) matrices by setting \(\delta(A) = \det(AB)\). It follows from Lemma 3.12 that \(\delta\) is linear in each row separately. It is also clear that if two rows of \(A\) are the same then \(\delta(A) = 0\). From our discussion of uniqueness, it follows that \(\delta\) must be a constant multiple of the determinant \(\det(A)\), and we evaluate this constant by taking \(A = I\), the identity matrix. \(\square\)

Proof of \((f)\). Since \(\{e_1, \ldots, e_n\}\) is a basis for \(V\), there do exist unique constants \(\{\alpha_1, \ldots, \alpha_n\}\) so that \(v = \alpha_1 e_1 + \cdots + \alpha_n e_n\). We then have
\[
\det[e_1, \ldots, e_{j-1}, v, e_{j+1}, \ldots, e_n] = \det[e_1, \ldots, e_{j-1}, \sum_{k=1}^{n} \alpha_k e_k, e_{j+1}, \ldots, e_n]
\]
\[
= \sum_{k=1}^{n} \alpha_k \det[e_1, \ldots, e_{j-1}, e_k, e_{j+1}, \ldots, e_n]
\]
\[
= \alpha_j \det[e_1, \ldots, e_{j-1}, e_j, e_{j+1}, \ldots, e_n]
\]
since all the other determinants have a repeated row or column. This gives the desired result. \(\square\)

3.6. Inverting matrices.

Consider the system of \(n\) equations in \(n\) unknowns
\[
a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,k}x_k + \cdots + a_{1,n}x_n = c_1,
a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,k}x_k + \cdots + a_{2,n}x_n = c_2,
\]
\[
\vdots
\]
\[
a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,k}x_k + \cdots + a_{j,n}x_n = c_j,
\]
\[
\vdots
\]
\[
a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,k}x_k + \cdots + a_{n,n}x_n = c_n.
\]

We can rewrite this system of equations as a single vector equation
\[
x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{j,1} \\ \vdots \\ a_{n,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{j,2} \\ \vdots \\ a_{n,2} \end{bmatrix} + \cdots + x_k \begin{bmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{j,k} \\ \vdots \\ a_{n,k} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{j,n} \\ \vdots \\ a_{n,n} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \\ c_n \end{bmatrix}
\]
or as a matrix equation

$$\mathbf{A} \mathbf{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & a_{j,2} & \cdots & a_{j,k} & \cdots & a_{j,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \\ c_n \end{bmatrix}, \quad (3.6.3)$$

Now suppose that the $n$ columns of the matrix $\mathbf{A}$ are linearly independent in $\mathbb{R}^n$. Then these columns form a basis for $\mathbb{R}^n$ since there are $n$ of them. Hence given any right hand side columns vector $\mathbf{c}$, there exists a column vector $\mathbf{x}$ which solves the equation (3.6.3). In particular, we can find a solution when the right hand side is one of the standard basis elements. Thus we can find scalars $\{b_{j,k}\}$ so that

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & a_{j,2} & \cdots & a_{j,k} & \cdots & a_{j,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{k,1} \\ \vdots \\ b_{n,1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.6.4)$$

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & a_{j,2} & \cdots & a_{j,k} & \cdots & a_{j,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{k,2} \\ \vdots \\ b_{n,2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.6.4)$$

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & a_{j,2} & \cdots & a_{j,k} & \cdots & a_{j,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,k} \\ b_{2,k} \\ \vdots \\ b_{j,k} \\ \vdots \\ b_{n,k} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad (3.6.4)$$

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & a_{j,2} & \cdots & a_{j,k} & \cdots & a_{j,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{k,n} \\ \vdots \\ b_{n,n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.6.4)$$
If we put all these equations together in one matrix equation, we see that

\[
\begin{pmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,n}
\end{pmatrix}
\begin{pmatrix}
b_{1,1} & b_{1,2} & \cdots & b_{1,k} & \cdots & b_{1,n}
\end{pmatrix}
= I_n
\]

Thus we have proved

**Lemma 3.13.** If the columns of an \( n \times n \) matrix \( A \) are linearly independent, there is an \( n \times n \) matrix \( B \) so that \( A B = I_n \), the \( n \times n \) identity matrix.

We now have the following

**Theorem 3.14.** Let \( A \) be an \( n \times n \) matrix. Then the following statements are equivalent.

(a) The matrix \( A \) has a right inverse.
(b) The matrix \( A \) has a left inverse.
(c) The matrix \( A \) is invertible.
(d) \( \det |A| \neq 0 \).
(e) \( \det |A^t| \neq 0 \).
(f) The rows and columns of \( A^t \) are linearly independent.
(g) The rows and columns of \( A \) are linearly independent.

**Proof.** The equivalence of (a), (b), and (c) was a homework problem. If \( A B = I_n \), then \( 1 = \det |I_n| = \det |A B| = \det |A| \det |B| \), and it follows that \( \det |A| \neq 0 \). Thus (c) implies (d). Since \( \det |A| = \det |A^t| \), it is clear that (d) and (e) are equivalent. If the rows of \( A \) or of \( A^t \) were linearly dependent, we know that \( \det |A| = 0 \) or that \( \det |A^t| = 0 \). Thus (d) and (e) imply that the rows of \( A \) and \( A^t \) are linearly independent, and since the rows of one are the columns of the other, this shows that (d) and (e) imply (f) and (g). Finally, Lemma 3.13 shows that if the columns of \( A \) are linearly independent, then \( A \) has a right inverse. This completes the proof. \( \square \)