Chapter 1

Basic Linear Algebra

Linear equations form the basis of linear programming. If you have a good understanding of the Gauss-Jordan method for solving linear equations, then you can also understand the solution of linear programs. In addition, this chapter introduces matrix notation and concepts that will be used in Chapters 3 and 4 (optimizing functions of several variables).

1.1 Linear Equations

The Gauss-Jordan elimination procedure is a systematic method for solving systems of linear equations. It works one variable at a time, eliminating it in all rows but one, and then moves on to the next variable. We illustrate the procedure on three examples.

Example 1.1.1 (Solving linear equations)

\[
\begin{align*}
  x_1 + 2x_2 + x_3 &= 4 \\
  2x_1 - x_2 + 3x_3 &= 3 \\
  x_1 + x_2 - x_3 &= 3
\end{align*}
\]

In the first step of the procedure, we use the first equation to eliminate \( x_1 \) from the other two. Specifically, in order to eliminate \( x_1 \) from the second equation, we multiply the first equation by 2 and subtract the result from the second equation. Similarly, to eliminate \( x_1 \) from the third equation, we subtract the first equation from the third. Such steps are called elementary row operations. We keep the first equation and the modified second and third equations. The resulting equations are

\[
\begin{align*}
  x_1 + 2x_2 + x_3 &= 4 \\
  -5x_2 + x_3 &= -5 \\
  -x_2 - 2x_3 &= -1
\end{align*}
\]

Note that only one equation was used to eliminate \( x_1 \) in all the others. This guarantees that the new system of equations has exactly the same solution(s) as the original one. In the second step of the procedure, we divide the second equation by -5 to make the coefficient of \( x_2 \) equal to 1. Then, we use this equation to eliminate \( x_2 \) from equations 1 and 3. This yields the following new system of equations.

\[
\begin{align*}
  x_1 + \frac{2}{5}x_3 &= 2 \\
  x_2 - \frac{1}{5}x_3 &= 1 \\
  -\frac{1}{5}x_3 &= 0
\end{align*}
\]
Once again, only one equation was used to eliminate \( x_2 \) in all the others and that guarantees that the new system has the same solution(s) as the original one. Finally, in the last step of the procedure, we use equation 3 to eliminate \( x_3 \) in equations 1 and 2.

\[
\begin{align*}
  x_1 & = 2 \\
  x_2 & = 1 \\
  x_3 & = 0
\end{align*}
\]

So, there is a unique solution. Note that, throughout the procedure, we were careful to keep three equations that have the same solution(s) as the original three equations. Why is it useful? Because, linear systems of equations do not always have a unique solution and it is important to identify such situations.

**Another example:**

\[
\begin{align*}
  x_1 + 2x_2 + x_3 & = 4 \\
  x_1 + x_2 + 2x_3 & = 1 \\
  2x_1 + 3x_2 + 3x_3 & = 2
\end{align*}
\]

First we eliminate \( x_1 \) from equations 2 and 3.

\[
\begin{align*}
  x_1 + 2x_2 + x_3 & = 4 \\
  -x_2 + x_3 & = -3 \\
  -x_2 + x_3 & = -6
\end{align*}
\]

Then we eliminate \( x_2 \) from equations 1 and 3.

\[
\begin{align*}
  x_1 + 3x_3 & = -2 \\
  x_2 - x_3 & = 3 \\
  0 & = -3
\end{align*}
\]

Equation 3 shows that the linear system has **no solution**.

**A third example:**

\[
\begin{align*}
  x_1 + 2x_2 + x_3 & = 4 \\
  x_1 + x_2 + 2x_3 & = 1 \\
  2x_1 + 3x_2 + 3x_3 & = 5
\end{align*}
\]

Doing the same as above, we end up with

\[
\begin{align*}
  x_1 + 3x_3 & = -2 \\
  x_2 - x_3 & = 3 \\
  0 & = 0
\end{align*}
\]

Now equation 3 is an obvious equality. It can be discarded to obtain

\[
\begin{align*}
  x_1 & = -2 - 3x_3 \\
  x_2 & = 3 + x_3
\end{align*}
\]

The situation where we can express some of the variables (here \( x_1 \) and \( x_2 \)) in terms of the remaining variables (here \( x_3 \)) is important. These variables are said to be **basic** and **nonbasic** respectively. Any choice of the nonbasic variable \( x_3 \) yields a solution of the linear system. Therefore the system has infinitely many solutions.
It is generally true that a system of $m$ linear equations in $n$ variables has either:

(a) no solution,

(b) a unique solution,

(c) infinitely many solutions.

The Gauss-Jordan elimination procedure solves the system of linear equations using two elementary row operations:

- modify some equation by multiplying it by a nonzero scalar (a *scalar* is an actual real number, such as $\frac{1}{2}$ or $-2$; it cannot be one of the variables in the problem),

- modify some equation by adding to it a scalar multiple of another equation.

The resulting system of $m$ linear equations has the same solution(s) as the original system. If an equation $0 = 0$ is produced, it is discarded and the procedure is continued. If an equation $0 = a$ is produced where $a$ is a nonzero scalar, the procedure is stopped: in this case, the system has no solution. At each step of the procedure, a new variable is made basic: it has coefficient 1 in one of the equations and 0 in all the others. The procedure stops when each equation has a basic variable associated with it. Say $p$ equations remain (remember that some of the original $m$ equations may have been discarded). When $n = p$, the system has a unique solution. When $n > p$, then $p$ variables are basic and the remaining $n - p$ are nonbasic. In this case, the system has infinitely many solutions.

**Exercise 1** Solve the following systems of linear equations using the Gauss-Jordan elimination procedure and state whether case (a), (b) or (c) holds.

(1) \[
\begin{align*}
3x_1 - 4x_3 &= 2 \\
x_1 + x_2 + x_3 &= 4 \\
2x_2 + x_3 &= 3
\end{align*}
\]

(2) \[
\begin{align*}
2x_1 + 2x_2 - x_3 &= 1 \\
4x_1 + x_3 &= 2 \\
x_1 - x_2 + x_3 &= 2
\end{align*}
\]

(3) \[
\begin{align*}
x_1 - x_2 + x_3 &= 1 \\
-2x_1 + 2x_2 - 2x_3 &= -2 \\
-x_1 + x_2 - x_3 &= -1
\end{align*}
\]
Exercise 2 Indicate whether the following linear system of equations has 0, 1 or infinitely many solutions.

\[
\begin{align*}
\begin{align*}
x_1 + 2x_2 + 4x_3 + x_4 + 3x_5 &= 2 \\
2x_1 + x_2 + x_3 + 3x_4 + x_5 &= 1 \\
3x_2 + 7x_3 - x_4 + 5x_5 &= 6
\end{align*}
\end{align*}
\]

1.2 Operations on Vectors and Matrices

It is useful to formalize the operations on vectors and matrices that form the basis of linear algebra. For our purpose, the most useful definitions are the following.

A matrix is a rectangular array of numbers written in the form

\[
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix}
\]

The matrix \(A\) has dimensions \(m \times n\) if it has \(m\) rows and \(n\) columns. When \(m = 1\), the matrix is called a row vector; when \(n = 1\), the matrix is called a column vector. A vector can be represented either by a row vector or a column vector.

Equality of two matrices of same dimensions:

Let \(A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}\end{pmatrix}\) and \(B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\
\vdots & \ddots & \vdots \\
b_{m1} & \cdots & b_{mn}\end{pmatrix}\).

Then \(A = B\) means that \(a_{ij} = b_{ij}\) for all \(i, j\).

Multiplication of a matrix \(A\) by a scalar \(k\):

\[
kA = \begin{pmatrix} k a_{11} & k a_{12} & \cdots & k a_{1n} \\
\vdots & \vdots & & \vdots \\
k a_{m1} & k a_{m2} & \cdots & k a_{mn}\end{pmatrix}.
\]

Addition of two matrices of same dimensions:

Let \(A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn}\end{pmatrix}\) and \(B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\
\vdots & \ddots & \vdots \\
b_{m1} & \cdots & b_{mn}\end{pmatrix}\).

Then \(A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\
\vdots & & \vdots \\
a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn}\end{pmatrix}\).

Note that \(A + B\) is not defined when \(A\) and \(B\) have different dimensions.

Exercise 3 Compute \(3 \begin{pmatrix} 2 & 3 & 1 \\
0 & 5 & 4\end{pmatrix} - 2 \begin{pmatrix} 0 & 4 & 1 \\
1 & 6 & 6\end{pmatrix}\).
Multiplication of a matrix of dimensions \( m \times n \) by a matrix of dimensions \( n \times p \):

Let 
\[
A = \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & & \vdots \\
a_i & \cdots & a_{in} \\
\vdots & & \vdots \\
a_{m1} & \cdots & a_{mn}
\end{pmatrix}
\]
and 
\[
B = \begin{pmatrix}
b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\
\vdots & & \vdots & & \vdots \\
b_i & \cdots & b_{nj} & \cdots & b_{jn} \\
\vdots & & \vdots & & \vdots \\
b_{m1} & \cdots & b_{mj} & \cdots & b_{mn}
\end{pmatrix}.
\]

Then \( AB \) is a matrix of dimensions \( m \times p \) computed as follows:

\[
AB = \begin{pmatrix}
a_{11}b_{11} + \cdots + a_{1n}b_{m1} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{mp} \\
\vdots & & \vdots \\
a_i b_{11} + \cdots + a_{in}b_{m1} & \cdots & a_i b_{1p} + \cdots + a_{in}b_{mp} \\
\vdots & & \vdots \\
a_{m1}b_{11} + \cdots + a_{mn}b_{m1} & \cdots & a_{m1}b_{1p} + \cdots + a_{mn}b_{mp}
\end{pmatrix}.
\]

As an example, let us multiply the matrices

\[
A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 1 & 0 \end{pmatrix}
\text{ and } B = \begin{pmatrix} 11 & 12 \\ 14 & 15 \\ 17 & 18 \end{pmatrix}.
\]

The result is

\[
AB = \begin{pmatrix} 2 \times 11 + 3 \times 14 + 4 \times 17 & 2 \times 12 + 3 \times 15 + 4 \times 18 \\ 5 \times 11 + 1 \times 14 + 0 \times 17 & 5 \times 12 + 1 \times 15 + 0 \times 18 \end{pmatrix} = \begin{pmatrix} 132 & 141 \\ 69 & 75 \end{pmatrix}.
\]

Note that \( AB \) is defined only when the number of columns of \( A \) equals the number of rows of \( B \). An important remark: even when both \( AB \) and \( BA \) are defined, the results are usually different. A property of matrix multiplication is the following:

\[
(AB)C = A(BC).
\]

That is, if you have three matrices \( A, B, C \) to multiply and the product is legal (the number of columns of \( A \) equals the number of rows of \( B \) and the number of columns of \( B \) equals the number of rows of \( C \)), then you have two possibilities: you can first compute \( AB \) and multiply the result by \( C \), or you can first compute \( BC \) and multiply \( A \) by the result.

**Exercise 4** Consider the following matrices

\[
x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},
\]

\[
y = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix},
\]

\[
z = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.
\]

When possible, compute the following quantities:

(a) \( xz \)
(b) \( zx \)
(c) \( yzx \)
(d) \( xzy \)
(e) \( (x + y)z \)
(f) \((xz) + (yz)\).

**Remark:** A system of linear equations can be written conveniently using matrix notation. Namely,

\[
\begin{align*}
    a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
    \vdots & \quad \vdots \\
    a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]

can be written as

\[
\begin{pmatrix}
    a_{11} & \cdots & a_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    \vdots \\
    x_n
\end{pmatrix}
= 
\begin{pmatrix}
    b_1 \\
    \vdots \\
    b_m
\end{pmatrix}
\]

or as

\[
\begin{pmatrix}
    a_{11} \\
    \vdots \\
    a_{m1}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    \vdots \\
    x_n
\end{pmatrix}
+ 
\begin{pmatrix}
    a_{1n} \\
    \vdots \\
    a_{mn}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    \vdots \\
    x_n
\end{pmatrix}
= 
\begin{pmatrix}
    b_1 \\
    \vdots \\
    b_m
\end{pmatrix}
\]

So a matrix equation \(Ax = b\) where \(A\) is a given \(m \times n\) matrix, \(b\) is a given \(m\)-column vector and \(x\) is an unknown \(n\)-column vector, is a linear system of \(m\) equations in \(n\) variables. Similarly, a vector equation \(a_1x_1 + \cdots + a_nx_n = b\) where \(a_1, \ldots, a_n, b\) are given \(m\)-column vectors and \(x_1, \ldots, x_n\) are \(n\) unknown real numbers, is also a system of \(m\) equations in \(n\) variables.

**Exercise 5** (a) Solve the matrix equation

\[
\begin{pmatrix}
    1 & 2 & 1 \\
    2 & -1 & 3 \\
    1 & 1 & -1
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{pmatrix}
= 
\begin{pmatrix}
    4 \\
    3 \\
    3
\end{pmatrix}
\]

(b) Solve the vector equation

\[
\begin{pmatrix}
    1 \\
    2 \\
    1
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{pmatrix}
+ 
\begin{pmatrix}
    2 \\
    -1 \\
    1
\end{pmatrix}
\begin{pmatrix}
    x_2 \\
    x_3 \\
    x_3
\end{pmatrix}
= 
\begin{pmatrix}
    4 \\
    3 \\
    3
\end{pmatrix}
\]

[Hint: Use Example 1.1.1]

The following standard definitions will be useful:

A **square** matrix is a matrix with same number of rows as columns.

The **identity matrix** \(I\) is a square matrix with 1’s on the main diagonal and 0’s elsewhere, i.e.

\[
I = \begin{pmatrix}
    1 & 0 & \cdots & 0 \\
    0 & 1 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 1
\end{pmatrix}
\]

The **zero matrix** contains only 0’s. In particular, the **zero vector** has all components equal to 0. A **nonzero vector** is one that contains at least one nonzero component.
1.3 Linear Combinations

Suppose that the vector \((1, 3)\) represents the contents of a “Regular” can of mixed nuts (1 lb cashews and 3 lb peanuts) while \((1, 1)\) represents a “Deluxe” can (1 lb cashews and 1 lb peanuts). Can you obtain a mixture of 2 lb cashews and 3 lb peanuts by combining the two mixtures in appropriate amounts? The answer is to use \(x_1\) cans of Regular and \(x_2\) cans of Deluxe in order to satisfy the equality

\[
x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.
\]

This is none other than a system of two linear equations:

\[
\begin{align*}
x_1 + x_2 &= 2 \\
3x_1 + x_2 &= 3
\end{align*}
\]

The solution of these equations (obtained by the Gauss-Jordan procedure) is \(x_1 = 1/2,\ x_2 = 3/2\). So the desired combination is to mix 1/2 can of Regular nuts with 3/2 cans of Deluxe nuts. Thus if some recipe calls for the mixture \((2, 3)\), you can substitute 1/2 can of Regular mix and 3/2 can of Deluxe mix.

Suppose now that you want to obtain 1 lb cashews and no peanuts. This poses the equations,

\[
x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.
\]

The solution is \((x_1, x_2) = (-1/2, 3/2)\). Thus you can obtain a pound of pure cashews by buying 3/2 cans of Deluxe mix and removing enough nuts to form 1/2 can Regular mix, which can be sold. In this case it is physically possible to use a negative amount of some ingredient, but in other cases it may be impossible, as when one is mixing paint.

A vector of the form \(x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}\) is called a linear combination of the vectors \(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\) and \(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\). In particular we just saw that the vector \(\begin{pmatrix} 2 \\ 3 \end{pmatrix}\) is a linear combination of the vectors \(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\) and \(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\). And so is \(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\).

The question arises: can one obtain any mixture whatever by taking the appropriate combination of Regular and Deluxe cans? Is any vector \(\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\) a linear combination of \(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\) and \(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\)?

To answer the question, solve the equations in general. You want a mixture of \(b_1\) lb cashews and \(b_2\) lb peanuts and set

\[
x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},
\]
or equivalently,

\[ x_1 + x_2 = b_1 \]
\[ 3x_1 + x_2 = b_2. \]

These equations have a unique solution, no matter what are the values of \( b_1 \) and \( b_2 \), namely,

\[ x_1 = \frac{b_2 - b_1}{2}, \quad x_2 = \frac{3b_1 - b_2}{2}. \]

No matter what vector \((b_1, b_2)\) you want, you can get it as a linear combination of \((1, 3)\) and \((1, 1)\). The vectors \((1, 3)\) and \((1, 1)\) are said to be *linearly independent*.

Not all pairs of vectors can yield an arbitrary mixture \((b_1, b_2)\). For instance, no linear combination of \((1, 1)\) and \((2, 2)\) yields \((2, 3)\). In other words, the equations

\[ x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \]

have no solution.

The reason is that \((2, 2)\) is already a multiple of \((1, 1)\), that is \((2, 2)\) is a linear combination of \((1, 1)\). The vectors \((1, 1)\) and \((2, 2)\) are said to be *linearly dependent*.

If \((2, 2)\) represents, for instance, a large Deluxe can of nuts and \((1, 1)\) a small one, it is clear that you cannot obtain any mixture you want by combining large and small Deluxe cans. In fact, once you have small cans, the large cans contribute nothing at all to the mixtures you can generate, since you can always substitute two small cans for a large one.

A more interesting example is \((1, 2, 0), (1, 0, 1)\) and \((2, 2, 1)\). The third vector is clearly a linear combination of the other two (it equals their sum). Altogether, these three vectors are said to be linearly dependent. For instance, these three vectors might represent mixtures of nuts as follows:

<table>
<thead>
<tr>
<th>Brand</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>cashews</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>peanuts</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>almonds</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Then once you have brands A and B, brand C adds nothing whatever to the variety of mixtures you can concoct. This is because you can already obtain a can of brand C from brands A and B anyway. In other words, if a recipe calls for brand C, you can always substitute a mixture of 1 can brand A and 1 can brand B.

Suppose you want to obtain the mixture \((1, 2, 1)\) by combining Brands A, B, and C. The equations are

\[ x_1 + x_2 + 2x_3 = 1 \]
\[ 2x_1 + 2x_3 = 2 \]
\[ x_2 + x_3 = 1 \]

If you try to solve these equations, you will find that there is no solution. So the vector \((1, 2, 1)\) is *not* a linear combination of \((1, 2, 0), (1, 0, 1)\) and \((2, 2, 1)\).

The concepts of linear combination, linear dependence and linear independence introduced in the above examples can be defined more formally, for any number of \(n\)-component vectors. This is done as follows.
A vector $b$ having $n$ components is a linear combination of the $k$ vectors $v_1, \ldots, v_k$, each having $n$ components, if it is possible to find $k$ real numbers $x_1, \ldots, x_k$ satisfying:

$$x_1 v_1 + \ldots + x_k v_k = b \quad (1.1)$$

To find the numbers $x_1, \ldots, x_k$, view (1.1) as a system of linear equations and solve by the Gauss-Jordan method.

A set of vectors (all having $n$ components) is linearly dependent if at least one vector is a linear combination of the others. Otherwise they are linearly independent.

Given $n$ linearly independent vectors $v_1, \ldots, v_n$, each having $n$ components, any desired vector $b$ with $n$ components can be obtained as a linear combination of them:

$$x_1 v_1 + \ldots + x_n v_n = b \quad (1.2)$$

The desired weights $x_1, \ldots, x_n$ are computed by solving (1.2) with the Gauss-Jordan method: there is a unique solution whenever the vectors $v_1, \ldots, v_n$ are linearly independent.

**Exercise 6** A can of Brand A mixed nuts has 1 lb cashews, 1 lb almonds, 2 lb peanuts. Brand B has 1 lb cashews and 3 lb peanuts. Brand C has 1 lb almonds and 2 lb peanuts. Show how much of each brand to buy/sell so as to obtain a mixture containing 5 lb of each type of nut.

**Exercise 7** Determine whether the vector $\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ is a linear combination of

(a) $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$

(b) $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}$.

In linear programming, there are typically many more variables than equations. So, this case warrants looking at another example.

**Example 1.3.1 (Basic Solutions.)** A craft shop makes deluxe and regular belts. Each deluxe belt requires a strip of leather and 2 hours of labor. Each regular belt requires a leather strip and 1 hour of labor. 40 leather strips and 60 hours of labor are available. How many belts of either kind can be made?

This is really a mixing problem. Rather than peanuts and cashews, the mixture will contain leather and labor. The items to be mixed are four activities: manufacturing a deluxe belt, manufacturing a regular belt, leaving a leftover leather strip in inventory, and leaving an hour of labor
idle (or for other work). Just as each Regular can of mixed nuts contributes 1 pound of cashews and 3 pounds of peanuts to the mixture, each regular belt will consume 1 leather strip and 2 hours of labor. The aim is to combine the four activities in the right proportion so that 40 strips and 60 hours are accounted for:

\[ x_1 \left( \begin{array}{c} 1 \\ 2 \end{array} \right) + x_2 \left( \begin{array}{c} 1 \\ 1 \end{array} \right) + s_1 \left( \begin{array}{c} 1 \\ 0 \end{array} \right) + s_2 \left( \begin{array}{c} 0 \\ 1 \end{array} \right) = \left( \begin{array}{c} 40 \\ 60 \end{array} \right). \]

So,

- \( x_1 \) = number of deluxe belts made
- \( x_2 \) = number of regular belts made
- \( s_1 \) = number of leather strips left over
- \( s_2 \) = number of labor hours left over

In tableau form, the equations are:

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>60</td>
</tr>
</tbody>
</table>

(1.3)

Since there are only two equations, you can only solve for two variables. Let’s solve for \( x_1 \) and \( x_2 \), using Gauss-Jordan elimination. After the first iteration you get,

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-2</td>
<td>1</td>
<td>-20</td>
</tr>
</tbody>
</table>

(1.4)

A second iteration yields the solution,

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>-1</td>
<td>20</td>
</tr>
</tbody>
</table>

(1.5)

This tableau represents the equations

\[ x_1 - s_1 + s_2 = 20 \]
\[ x_2 + 2s_1 - s_2 = 20 \]

or,

\[ x_1 = 20 + s_1 - s_2 \]
\[ x_2 = 20 - 2s_1 + s_2 \]

(1.6)

You can’t say how many deluxe belts \( x_1 \) and regular belts \( x_2 \) the plant will make until you specify how much leather \( s_1 \) and labor \( s_2 \) will be left over. But (1.6) is a formula for computing how many belts of either type must be made, for any given \( s_1 \) and \( s_2 \). So the equations have many solutions (infinitely many).

For example, if you want to have nothing left over (\( s_1 = s_2 = 0 \)), you will make 20 of each. If you want to have 5 strips and 5 hours left over, you will make \( x_1 = 20 \) deluxe and \( x_2 = 15 \) regular belts.

The variables \( x_1, x_2 \) you solved for are called basic variables. The other variables are nonbasic. You have control of the nonbasic variables. Once you give them values, the values of the basic
variables follow. A solution in which you make all the nonbasic variables zero is called a basic solution.

Can you have basic variables other than $x_1, x_2$? Sure. Any pair of variables can be basic, provided the corresponding columns in (1.3) are linearly independent (otherwise, you can’t solve for the basic variables).

Equations (1.3), for instance, are already solved in (1.3) for basic variables $s_1, s_2$. Here the two basic activities are having leftover leather and having leftover labor. The basic solution is $(s_1, s_2) = (40, 60)$. This means that if you decide to produce no belts ($x_1 = x_2 = 0$), you must have 40 leftover leather strips and 60 leftover labor hours.

The intermediate step (1.4) solves the equations with basic variables $x_1$ and $s_2$. Here the basic solution is unrealizable. If you decide to participate only in the basic activities (making deluxe belts and having leftover labor), you must make 40 belts and have $-20$ leftover hours (i.e., use 20 more than you have), which you can’t do within your current resources.

**Exercise 8** Consider the system of equations

\[
\begin{array}{cccc|c}
  x_1 & x_2 & x_3 & x_4 & \text{RHS} \\
  1 & 1 & 2 & 4 & 100 \\
  3 & 1 & 1 & 2 & 200 \\
\end{array}
\]

where the first four columns on the left represent a Regular mixture of nuts, a Deluxe mixture, a small can of Premium mixture, and a large can of Premium.

a) Solve the system with $x_1$ and $x_2$ basic.

b) You want 100 cans of mixed nuts, each of which contains 1 lb cashews and 2 lb peanuts (i.e., you want the right-hand side of the above equation). How can you get them by blending 10 small cans of Premium with proper amounts of the Regular and Deluxe blends? *Hint.* Set $x_3 = 10$ and $x_4 = 0$ in the expression for the solution values of $x_1, x_2$ found in (a).

c) How can you obtain a small can of Premium mix by combining (and decombining) the Regular and Deluxe blends?

d) How can you obtain one can of Regular mix by combining large and small cans of Premium? If you cannot do it, why not? *Hint.* It has to do with linear dependence.

**Exercise 9** A can of paint A has 1 quart red, 1 quart yellow. A can of paint B has 1 quart red, 1 quart blue. A can of paint C has 1 quart yellow, 1 quart blue.

a) How much of each paint must be mixed to obtain a mixture of 1 quart red, 1 quart yellow and 1 quart blue?

b) How much of each paint must be mixed to obtain one quart of pure red? What do you conclude about the physical feasibility of such a mixture?

**Exercise 10** An electronics plant wants to make stereos and CB’s. Each stereo requires 1 power supply and 3 speakers. Each CB requires 1 power supply and 1 speaker. 100 power supplies and 200 speakers are in stock. How many stereos and CB’s can it make if it wants to use all the power supplies and all but 10 of the speakers? *Hint.* Use the following equations.

\[
\begin{array}{cccc|c}
  x_1 & x_2 & s_1 & s_2 & \text{RHS} \\
  1 & 1 & 1 & 0 & 100 \\
  3 & 1 & 0 & 1 & 200 \\
\end{array}
\]
Exercise 11 A construction foreman needs cement mix containing 8 cubic yards ($yd^3$) cement, 12 $yd^3$ sand and 16 $yd^3$ water. On the site are several mixer trucks containing mix $A$ (1 $yd^3$ cement, 3 $yd^3$ sand, 3 $yd^3$ water), several containing mix $B$ (2 $yd^3$ cement, 2 $yd^3$ sand, 3 $yd^3$ water) and several containing mix $C$ (2 $yd^3$ cement, 2 $yd^3$ sand, 5 $yd^3$ water).

How many truckloads of each mix should the foreman combine to obtain the desired blend?

1.4 Inverse of a Square Matrix

If $A$ and $B$ are square matrices such that $AB = I$ (the identity matrix), then $B$ is called the inverse of $A$ and is denoted by $A^{-1}$. A square matrix $A$ has either no inverse or a unique inverse $A^{-1}$. In the first case, it is said to be singular and in the second case nonsingular. Interestingly, linear independence of vectors plays a role here: a matrix is singular if its columns form a set of linearly dependent vectors; and it is nonsingular if its columns are linearly independent. Another property is the following: if $B$ is the inverse of $A$, then $A$ is the inverse of $B$.

Exercise 12 (a) Compute the matrix product
\[
\begin{pmatrix}
1 & 3 & -2 \\
0 & -5 & 4 \\
2 & -3 & 3
\end{pmatrix}
\begin{pmatrix}
-3 & -3 & 2 \\
8 & 7 & -4 \\
10 & 9 & -5
\end{pmatrix}.
\]

(b) What is the inverse of
\[
\begin{pmatrix}
-3 & -3 & 2 \\
8 & 7 & -4 \\
10 & 9 & -5
\end{pmatrix}?
\]

(c) Show that the matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ is singular.

[Hint: Assume that the inverse of $A$ is $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$ and perform the matrix product $AB$. Then show that no choice of $b_{ij}$ can make this product equal to the identity matrix $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.]

An important property of nonsingular square matrices is the following. Consider the system of linear equations
\[
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
a_{21} & \cdots & a_{2n} \\
\vdots & \vdots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix} =
\begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n
\end{pmatrix},
\]
simply written as $Ax = b$.

When $A$ is a square nonsingular matrix, this linear system has a unique solution, which can be obtained as follows. Multiply the matrix equation $Ax = b$ by $A^{-1}$ on the left:
\[
A^{-1}Ax = A^{-1}b.
\]

This yields $Ix = A^{-1}b$ and so, the unique solution to the system of linear equations is
\[
x = A^{-1}b.
\]
Exercise 13 Solve the system

\[
\begin{align*}
-3x_1 - 3x_2 + 2x_3 &= 1 \\
8x_1 + 7x_2 - 4x_3 &= -1 \\
10x_1 + 9x_2 - 5x_3 &= 1
\end{align*}
\]

using the result of Exercise 12(b).

Finding the Inverse of a Square Matrix

Given \( A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \), we must find \( B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \) such that \( AB = I \) (the identity matrix). Therefore, the first column of \( B \) must satisfy \( A \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \) (this vector is the 1st column of \( I \)). Similarly, for the other columns of \( B \). For example, the \( j \)th column of \( B \) satisfies \( A \begin{pmatrix} b_{i,j} \\ \vdots \\ b_{n,j} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \) (the \( j \)th column of \( I \)). So in order to get the inverse of an \( n \times n \) matrix, we must solve \( n \) linear systems. However, the same steps of the Gauss-Jordan elimination procedure are needed for all of these systems. So we solve them all at once, using the matrix form.

Example: Find the inverse of \( A = \begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix} \).

We need to solve the following matrix equation

\[
\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix} B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]

We divide the first row by 3 to introduce a 1 in the top left corner.

\[
\begin{pmatrix} 1 & -\frac{2}{3} \\ -4 & \frac{3}{3} \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix}
\]

Then we add four times the first row to the second row to introduce a 0 in the first column.

\[
\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ \frac{4}{3} & 1 \end{pmatrix}
\]

Multiply the second row by 3.

\[
\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ \frac{4}{3} & 3 \end{pmatrix}
\]
Add the second row to the first. (All this is classical Gauss-Jordan elimination.)

\[
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
\end{pmatrix} B = \begin{pmatrix}
3 & 2 \\
4 & 3 \\
\end{pmatrix}
\]

As \(IB = B\), we get

\[
B = \begin{pmatrix}
3 & 2 \\
4 & 3 \\
\end{pmatrix}.
\]

It is important to note that, in addition to the two elementary row operations introduced earlier in the context of the Gauss-Jordan elimination procedure, a third elementary row operation may sometimes be needed here, namely permuting two rows.

**Example:** Find the inverse of \(A = \begin{pmatrix}
0 & 1 & 0 \\
2 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}\).

\[
\begin{pmatrix}
0 & 1 & 0 \\
2 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix} B = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}
\]

Because the top left entry of \(A\) is 0, we need to permute rows 1 and 2 first.

\[
\begin{pmatrix}
2 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix} B = \begin{pmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}
\]

Now we divide the first row by 2.

\[
\begin{pmatrix}
1 & \frac{1}{2} & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix} B = \begin{pmatrix}
0 & \frac{1}{2} & 0 \\
1 & 0 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}
\]

Next we add \(-\frac{1}{2}\) the second row to the first.

\[
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix} B = \begin{pmatrix}
-\frac{1}{2} & \frac{1}{2} & 0 \\
1 & 0 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}
\]

and we are done, since the matrix in front of \(B\) is the identity.

**Exercise 14** Find the inverse of the matrix

\[
A = \begin{pmatrix}
1 & 4 & 2 \\
0 & 1 & 0 \\
-1 & -4 & 2 \\
\end{pmatrix}
\]

**Exercise 15** Find the inverse of the matrix

\[
A = \begin{pmatrix}
0 & 1 & -2 \\
4 & 0 & 5 \\
3 & 0 & 4 \\
\end{pmatrix}
\]
1.5 Determinants

To each square matrix, we associate a number, called its determinant, defined as follows:

If $A = (a_{11})$, then $\det(A) = a_{11}$.

If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, then $\det(A) = a_{11}a_{22} + a_{12}a_{21}$.

For a square matrix $A$ of dimensions $n \times n$, the determinant can be obtained as follows. First, define $A_{ij}$ as the matrix of dimensions $(n - 1) \times (n - 1)$ obtained from $A$ by deleting row $i$ and column $j$. Then

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) - \ldots a_{1n}\det(A_{1n}).$$

Note that, in this formula, the signs alternate between $+$ and $-$.

For example, if $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, then

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}).$$

Determinants have several interesting properties. For example, the following statements are equivalent for a square matrix $A$:

- $\det(A) = 0$,
- $A$ has no inverse, i.e. $A$ is singular,
- the columns of $A$ form a set of linearly dependent vectors,
- the rows of $A$ form a set of linearly dependent vectors.

For our purpose, however, determinants will be needed mainly in our discussion of classical optimization, in conjunction with the material from the following section.

Exercise 16 Compute the determinant of $A = \begin{pmatrix} 4 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

1.6 Positive Definite Matrices

When we study functions of several variables (see Chapter 3!), we will need the following matrix notions.

A square matrix $A$ is positive definite if $x^TAx > 0$ for all nonzero column vectors $x$. It is negative definite if $x^TAx < 0$ for all nonzero $x$. It is positive semidefinite if $x^TAx \geq 0$ and negative semidefinite if $x^TAx \leq 0$ for all $x$. These definitions are hard to check directly and you might as well forget them for all practical purposes.

More useful in practice are the following properties, which hold when the matrix $A$ is symmetric (that will be the case of interest to us), and which are easier to check.

The $i$th principal minor of $A$ is the matrix $A_i$ formed by the first $i$ rows and columns of $A$. So, the first principal minor of $A$ is the matrix $A_1 = (a_{11})$, the second principal minor is the matrix $A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and so on.
• The matrix $A$ is positive definite if all its principal minors $A_1, A_2, \ldots A_n$ have strictly positive determinants.

• If these determinants are nonzero and alternate in signs, starting with $\det(A_1) < 0$, then the matrix $A$ is negative definite.

• If the determinants are all nonnegative, then the matrix is positive semidefinite.

• If the determinant alternate in signs, starting with $\det(A_1) \leq 0$, then the matrix is negative semidefinite.

To fix ideas, consider a $2 \times 2$ symmetric matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$.

It is positive definite if:

(i) $\det(A_1) = a_{11} > 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} > 0$

and negative definite if:

(i) $\det(A_1) = a_{11} < 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} > 0$.

It is positive semidefinite if:

(i) $\det(A_1) = a_{11} \geq 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} \geq 0$

and negative semidefinite if:

(i) $\det(A_1) = a_{11} \leq 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} \geq 0$.

**Exercise 17** Check whether the following matrices are positive definite, negative definite, positive semidefinite, negative semidefinite or none of the above.

(a) $A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$

(b) $A = \begin{pmatrix} -2 & 4 \\ 4 & -8 \end{pmatrix}$

(c) $A = \begin{pmatrix} -2 & 2 \\ 2 & -4 \end{pmatrix}$

(d) $A = \begin{pmatrix} 2 & 4 \\ 4 & 3 \end{pmatrix}$.