LINEAR ALGEBRA

W W L CHEN

Chapter 2

MATRICES

2.1. Introduction

A rectangular array of numbers of the form

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \tag{1}$$

is called an $m \times n$ matrix, with m rows and n columns. We count rows from the top and columns from the left. Hence

$$\begin{pmatrix} a_{i1} & \ldots & a_{in} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

represent respectively the i-th row and the j-th column of the matrix (1), and a_{ij} represents the entry in the matrix (1) on the i-th row and j-th column.

EXAMPLE 2.1.1. Consider the 3×4 matrix

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix}.$$

Here

$$\begin{pmatrix} 3 & 1 & 5 & 2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}$$

represent respectively the 2-nd row and the 3-rd column of the matrix, and 5 represents the entry in the matrix on the 2-nd row and 3-rd column.

We now consider the question of arithmetic involving matrices. First of all, let us study the problem of addition. A reasonable theory can be derived from the following definition.

DEFINITION. Suppose that the two matrices

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \ldots & b_{mn} \end{pmatrix}$$

both have m rows and n columns. Then we write

$$A + B = \begin{pmatrix} a_{11}+b_{11} & \ldots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \ldots & a_{mn}+b_{mn} \end{pmatrix}$$

and call this the sum of the two matrices A and B.

EXAMPLE 2.1.2. Suppose that

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 2 & -2 & 7 \\ 0 & 2 & 4 & -1 \\ -2 & 1 & 3 & 3 \end{pmatrix}.$$

Then

$$A + B = \begin{pmatrix} 2+1 & 4+2 & 3-2 & -1+7 \\ 3+0 & 1+2 & 5+4 & 2-1 \\ -1-2 & 0+1 & 7+3 & 6+3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 1 & 6 \\ 3 & 3 & 9 & 1 \\ -3 & 1 & 10 & 9 \end{pmatrix}.$$

EXAMPLE 2.1.3. We do not have a definition for "adding" the matrices

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ -1 & 0 & 7 & 6 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 2 & 4 & 3 \\ 3 & 1 & 5 \\ -1 & 0 & 7 \end{pmatrix}.$$

PROPOSITION 2A. (MATRIX ADDITION) *Suppose that A, B, C are $m \times n$ matrices. Suppose further that O represents the $m \times n$ matrix with all entries zero. Then*
(a) $A + B = B + A$;
(b) $A + (B + C) = (A + B) + C$;
(c) $A + O = A$; and
(d) there is an $m \times n$ matrix A' such that $A + A' = O$.

PROOF. Parts (a)–(c) are easy consequences of ordinary addition, as matrix addition is simply entry-wise addition. For part (d), we can consider the matrix A' obtained from A by multiplying each entry of A by -1. ◯

The theory of multiplication is rather more complicated, and includes multiplication of a matrix by a scalar as well as multiplication of two matrices.

We first study the simpler case of multiplication by scalars.

DEFINITION. Suppose that the matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

has m rows and n columns, and that $c \in \mathbb{R}$. Then we write

$$cA = \begin{pmatrix} ca_{11} & \ldots & ca_{1n} \\ \vdots & & \vdots \\ ca_{m1} & \ldots & ca_{mn} \end{pmatrix}$$

and call this the product of the matrix A by the scalar c.

EXAMPLE 2.1.4. Suppose that

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix}.$$

Then

$$2A = \begin{pmatrix} 4 & 8 & 6 & -2 \\ 6 & 2 & 10 & 4 \\ -2 & 0 & 14 & 12 \end{pmatrix}.$$

PROPOSITION 2B. (MULTIPLICATION BY SCALAR) *Suppose that A, B are $m \times n$ matrices, and that $c, d \in \mathbb{R}$. Suppose further that O represents the $m \times n$ matrix with all entries zero. Then*
(a) $c(A + B) = cA + cB$;
(b) $(c + d)A = cA + dA$;
(c) $0A = O$; and
(d) $c(dA) = (cd)A$.

PROOF. These are all easy consequences of ordinary multiplication, as multiplication by scalar c is simply entry-wise multiplication by the number c. ○

The question of multiplication of two matrices is rather more complicated. To motivate this, let us consider the representation of a system of linear equations

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1, \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{2}$$

in the form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{3}$$

represent the coefficients and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{4}$$

represents the variables. This can be written in full matrix notation by

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Can you work out the meaning of this representation?

Now let us define matrix multiplication more formally.

DEFINITION. Suppose that

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} b_{11} & \ldots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{np} \end{pmatrix}$$

are respectively an $m \times n$ matrix and an $n \times p$ matrix. Then the matrix product AB is given by the $m \times p$ matrix

$$AB = \begin{pmatrix} q_{11} & \ldots & q_{1p} \\ \vdots & & \vdots \\ q_{m1} & \ldots & q_{mp} \end{pmatrix},$$

where for every $i = 1, \ldots, m$ and $j = 1, \ldots, p$, we have

$$q_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + \ldots + a_{in} b_{nj}.$$

REMARK. Note first of all that the number of columns of the first matrix must be equal to the number of rows of the second matrix. On the other hand, for a simple way to work out q_{ij}, the entry in the i-th row and j-th column of AB, we observe that the i-th row of A and the j-th column of B are respectively

$$\begin{pmatrix} a_{i1} & \ldots & a_{in} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

We now multiply the corresponding entries – from a_{i1} with b_{1j}, and so on, until a_{in} with b_{nj} – and then add these products to obtain q_{ij}.

EXAMPLE 2.1.5. Consider the matrices

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & -2 \\ 3 & 1 \end{pmatrix}.$$

Note that A is a 3×4 matrix and B is a 4×2 matrix, so that the product AB is a 3×2 matrix. Let us calculate the product

$$AB = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \\ q_{31} & q_{32} \end{pmatrix}.$$

Consider first of all q_{11}. To calculate this, we need the 1-st row of A and the 1-st column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} 1 & \times \\ 2 & \times \\ 0 & \times \\ 3 & \times \end{pmatrix} = \begin{pmatrix} q_{11} & \times \\ \times & \times \\ \times & \times \end{pmatrix}.$$

From the definition, we have

$$q_{11} = 2 \cdot 1 + 4 \cdot 2 + 3 \cdot 0 + (-1) \cdot 3 = 2 + 8 + 0 - 3 = 7.$$

Consider next q_{12}. To calculate this, we need the 1-st row of A and the 2-nd column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} \times & 4 \\ \times & 3 \\ \times & -2 \\ \times & 1 \end{pmatrix} = \begin{pmatrix} \times & q_{12} \\ \times & \times \\ \times & \times \end{pmatrix}.$$

From the definition, we have

$$q_{12} = 2 \cdot 4 + 4 \cdot 3 + 3 \cdot (-2) + (-1) \cdot 1 = 8 + 12 - 6 - 1 = 13.$$

Consider next q_{21}. To calculate this, we need the 2-nd row of A and the 1-st column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} \times & \times & \times & \times \\ 3 & 1 & 5 & 2 \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} 1 & \times \\ 2 & \times \\ 0 & \times \\ 3 & \times \end{pmatrix} = \begin{pmatrix} \times & \times \\ q_{21} & \times \\ \times & \times \end{pmatrix}.$$

From the definition, we have

$$q_{21} = 3 \cdot 1 + 1 \cdot 2 + 5 \cdot 0 + 2 \cdot 3 = 3 + 2 + 0 + 6 = 11.$$

Consider next q_{22}. To calculate this, we need the 2-nd row of A and the 2-nd column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} \times & \times & \times & \times \\ 3 & 1 & 5 & 2 \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} \times & 4 \\ \times & 3 \\ \times & -2 \\ \times & 1 \end{pmatrix} = \begin{pmatrix} \times & \times \\ \times & q_{22} \\ \times & \times \end{pmatrix}.$$

From the definition, we have

$$q_{22} = 3 \cdot 4 + 1 \cdot 3 + 5 \cdot (-2) + 2 \cdot 1 = 12 + 3 - 10 + 2 = 7.$$

Consider next q_{31}. To calculate this, we need the 3-rd row of A and the 1-st column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ -1 & 0 & 7 & 6 \end{pmatrix} \begin{pmatrix} 1 & \times \\ 2 & \times \\ 0 & \times \\ 3 & \times \end{pmatrix} = \begin{pmatrix} \times & \times \\ \times & \times \\ q_{31} & \times \end{pmatrix}.$$

From the definition, we have

$$q_{31} = (-1) \cdot 1 + 0 \cdot 2 + 7 \cdot 0 + 6 \cdot 3 = -1 + 0 + 0 + 18 = 17.$$

Consider finally q_{32}. To calculate this, we need the 3-rd row of A and the 2-nd column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ -1 & 0 & 7 & 6 \end{pmatrix} \begin{pmatrix} \times & 4 \\ \times & 3 \\ \times & -2 \\ \times & 1 \end{pmatrix} = \begin{pmatrix} \times & \times \\ \times & \times \\ \times & q_{32} \end{pmatrix}.$$

From the definition, we have

$$q_{32} = (-1) \cdot 4 + 0 \cdot 3 + 7 \cdot (-2) + 6 \cdot 1 = -4 + 0 + -14 + 6 = -12.$$

We therefore conclude that

$$AB = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & -2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 13 \\ 11 & 7 \\ 17 & -12 \end{pmatrix}.$$

EXAMPLE 2.1.6. Consider again the matrices

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & -2 \\ 3 & 1 \end{pmatrix}.$$

Note that B is a 4×2 matrix and A is a 3×4 matrix, so that we do not have a definition for the "product" BA.

We leave the proofs of the following results as exercises for the interested reader.

PROPOSITION 2C. (ASSOCIATIVE LAW) *Suppose that A is an $m \times n$ matrix, B is an $n \times p$ matrix and C is an $p \times r$ matrix. Then $A(BC) = (AB)C$.*

PROPOSITION 2D. (DISTRIBUTIVE LAWS)
(a) Suppose that A is an $m \times n$ matrix and B and C are $n \times p$ matrices. Then $A(B + C) = AB + AC$.
(b) Suppose that A and B are $m \times n$ matrices and C is an $n \times p$ matrix. Then $(A + B)C = AC + BC$.

PROPOSITION 2E. *Suppose that A is an $m \times n$ matrix, B is an $n \times p$ matrix, and that $c \in \mathbb{R}$. Then $c(AB) = (cA)B = A(cB)$.*

2.2. Systems of Linear Equations

Note that the system (2) of linear equations can be written in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where the matrices A, $\mathbf{x}$ and $\mathbf{b}$ are given by (3) and (4). In this section, we shall establish the following important result.

PROPOSITION 2F. *Every system of linear equations of the form (2) has either no solution, one solution or infinitely many solutions.*

PROOF. Clearly the system (2) has either no solution, exactly one solution, or more than one solution. It remains to show that if the system (2) has two distinct solutions, then it must have infinitely many solutions. Suppose that $\mathbf{x} = \mathbf{u}$ and $\mathbf{x} = \mathbf{v}$ represent two distinct solutions. Then

$$A\mathbf{u} = \mathbf{b} \qquad \text{and} \qquad A\mathbf{v} = \mathbf{b},$$

so that

$$A(\mathbf{u} - \mathbf{v}) = A\mathbf{u} - A\mathbf{v} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

where $\mathbf{0}$ is the zero $m \times 1$ matrix. It now follows that for every $c \in \mathbb{R}$, we have

$$A(\mathbf{u} + c(\mathbf{u} - \mathbf{v})) = A\mathbf{u} + A(c(\mathbf{u} - \mathbf{v})) = A\mathbf{u} + c(A(\mathbf{u} - \mathbf{v})) = \mathbf{b} + c\mathbf{0} = \mathbf{b},$$

so that $\mathbf{x} = \mathbf{u} + c(\mathbf{u} - \mathbf{v})$ is a solution for every $c \in \mathbb{R}$. Clearly we have infinitely many solutions. ○

2.3. Inversion of Matrices

For the remainder of this chapter, we shall deal with square matrices, those where the number of rows equals the number of columns.

DEFINITION. The $n \times n$ matrix

$$I_n = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

where

$$a_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

is called the identity matrix of order n.

REMARK. Note that

$$I_1 = (\,1\,) \qquad \text{and} \qquad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The following result is relatively easy to check. It shows that the identity matrix I_n acts as the identity for multiplication of $n \times n$ matrices.

PROPOSITION 2G. *For every $n \times n$ matrix A, we have $AI_n = I_nA = A$.*

This raises the following question: Given an $n \times n$ matrix A, is it possible to find another $n \times n$ matrix B such that $AB = BA = I_n$?

We shall postpone the full answer to this question until the next chapter. In Section 2.5, however, we shall be content with finding such a matrix B if it exists. In Section 2.6, we shall relate the existence of such a matrix B to some properties of the matrix A.

DEFINITION. An $n \times n$ matrix A is said to be invertible if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, we say that B is the inverse of A and write $B = A^{-1}$.

PROPOSITION 2H. *Suppose that A is an invertible $n \times n$ matrix. Then its inverse A^{-1} is unique.*

PROOF. Suppose that B satisfies the requirements for being the inverse of A. Then $AB = BA = I_n$. It follows that

$$A^{-1} = A^{-1}I_n = A^{-1}(AB) = (A^{-1}A)B = I_nB = B.$$

Hence the inverse A^{-1} is unique. ◯

PROPOSITION 2J. *Suppose that A and B are invertible $n \times n$ matrices. Then $(AB)^{-1} = B^{-1}A^{-1}$.*

PROOF. In view of the uniqueness of inverse, it is sufficient to show that $B^{-1}A^{-1}$ satisfies the requirements for being the inverse of AB. Note that

$$(AB)(B^{-1}A^{-1}) = A(B(B^{-1}A^{-1})) = A((BB^{-1})A^{-1}) = A(I_nA^{-1}) = AA^{-1} = I_n$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}(AB)) = B^{-1}((A^{-1}A)B) = B^{-1}(I_nB) = B^{-1}B = I_n$$

as required. ◯

PROPOSITION 2K. *Suppose that A is an invertible $n \times n$ matrix. Then $(A^{-1})^{-1} = A$.*

PROOF. Note that both $(A^{-1})^{-1}$ and A satisfy the requirements for being the inverse of A^{-1}. Equality follows from the uniqueness of inverse. ◯

2.4. Application to Matrix Multiplication

In this section, we shall discuss an application of invertible matrices. Detailed discussion of the technique involved will be covered in Chapter 7.

DEFINITION. An $n \times n$ matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

where $a_{ij} = 0$ whenever $i \neq j$, is called a diagonal matrix of order n.

EXAMPLE 2.4.1. The 3×3 matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

are both diagonal.

Given an $n \times n$ matrix A, it is usually rather complicated to calculate

$$A^k = \underbrace{A \ldots A}_{k}.$$

However, the calculation is rather simple when A is a diagonal matrix, as we shall see in the following example.

EXAMPLE 2.4.2. Consider the 3×3 matrix

$$A = \begin{pmatrix} 17 & -10 & -5 \\ 45 & -28 & -15 \\ -30 & 20 & 12 \end{pmatrix}.$$

Suppose that we wish to calculate A^{98}. It can be checked that if we take

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix},$$

then

$$P^{-1} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & 4/3 & 1 \\ 3 & -5/3 & -1 \end{pmatrix}.$$

Furthermore, if we write

$$D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

then it can be checked that $A = PDP^{-1}$, so that

$$A^{98} = \underbrace{(PDP^{-1}) \ldots (PDP^{-1})}_{98} = PD^{98}P^{-1} = P \begin{pmatrix} 3^{98} & 0 & 0 \\ 0 & 2^{98} & 0 \\ 0 & 0 & 2^{98} \end{pmatrix} P^{-1}.$$

This is much simpler than calculating A^{98} directly. Note that this example is only an illustration. We have not discussed here how the matrices P and D are found.

2.5. Finding Inverses by Elementary Row Operations

In this section, we shall discuss a technique by which we can find the inverse of a square matrix, if the inverse exists. Before we discuss this technique, let us recall the three elementary row operations we discussed in the previous chapter. These are: (1) interchanging two rows; (2) adding a multiple of one row to another row; and (3) multiplying one row by a non-zero constant.

Let us now consider the following example.

EXAMPLE 2.5.1. Consider the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Let us interchange rows 1 and 2 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Let us interchange rows 2 and 3 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Let us add 3 times row 1 to row 2 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{11}+a_{21} & 3a_{12}+a_{22} & 3a_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{11}+a_{21} & 3a_{12}+a_{22} & 3a_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Let us add -2 times row 3 to row 1 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} -2a_{31}+a_{11} & -2a_{32}+a_{12} & -2a_{33}+a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} -2a_{31}+a_{11} & -2a_{32}+a_{12} & -2a_{33}+a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Let us multiply row 2 of A by 5 and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 5a_{21} & 5a_{22} & 5a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 5a_{21} & 5a_{22} & 5a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Let us multiply row 3 of A by -1 and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Let us now consider the problem in general.

DEFINITION. By an elementary $n \times n$ matrix, we mean an $n \times n$ matrix obtained from I_n by an elementary row operation.

We state without proof the following important result. The interested reader may wish to construct a proof, taking into account the different types of elementary row operations.

PROPOSITION 2L. *Suppose that A is an $n \times n$ matrix, and suppose that B is obtained from A by an elementary row operation. Suppose further that E is an elementary matrix obtained from I_n by the same elementary row operation. Then $B = EA$.*

We now adopt the following strategy. Consider an $n \times n$ matrix A. Suppose that it is possible to reduce the matrix A by a sequence $\alpha_1, \alpha_2, \ldots, \alpha_k$ of elementary row operations to the identity matrix I_n. If $E_1, E_2, \ldots, E_k$ are respectively the elementary $n \times n$ matrices obtained from I_n by the same elementary row operations $\alpha_1, \alpha_2 \ldots, \alpha_k$, then

$$I_n = E_k \ldots E_2 E_1 A.$$

We therefore must have

$$A^{-1} = E_k \ldots E_2 E_1 = E_k \ldots E_2 E_1 I_n.$$

It follows that the inverse A^{-1} can be obtained from I_n by performing the same elementary row operations $\alpha_1, \alpha_2, \ldots, \alpha_k$. Since we are performing the same elementary row operations on A and I_n, it makes sense to put them side by side. The process can then be described pictorially by

$$\begin{aligned} (A|I_n) &\xrightarrow{\alpha_1} (E_1 A | E_1 I_n) \\ &\xrightarrow{\alpha_2} (E_2 E_1 A | E_2 E_1 I_n) \\ &\xrightarrow{\alpha_3} \ldots \\ &\xrightarrow{\alpha_k} (E_k \ldots E_2 E_1 A | E_k \ldots E_2 E_1 I_n) = (I_n | A^{-1}). \end{aligned}$$

In other words, we consider an array with the matrix A on the left and the matrix I_n on the right. We now perform elementary row operations on the array and try to reduce the left hand half to the matrix I_n. If we succeed in doing so, then the right hand half of the array gives the inverse A^{-1}.

EXAMPLE 2.5.2. Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix}.$$

To find A^{-1}, we consider the array

$$(A|I_3) = \begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 3 & 0 & 3 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We now perform elementary row operations on this array and try to reduce the left hand half to the matrix I_3. Note that if we succeed, then the final array is clearly in reduced row echelon form. We therefore follow the same procedure as reducing an array to reduced row echelon form. Adding −3 times row 1 to row 2, we obtain

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

Adding 2 times row 1 to row 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 5 & 4 & 2 & 0 & 1 \end{pmatrix}.$$

Multiplying row 3 by 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 15 & 12 & 6 & 0 & 3 \end{pmatrix}.$$

Adding 5 times row 2 to row 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 1 by 3, we obtain

$$\begin{pmatrix} 3 & 3 & 6 & 3 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding 2 times row 3 to row 1, we obtain

$$\begin{pmatrix} 3 & 3 & 0 & -15 & 10 & 6 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding −1 times row 3 to row 2, we obtain

$$\begin{pmatrix} 3 & 3 & 0 & -15 & 10 & 6 \\ 0 & -3 & 0 & 6 & -4 & -3 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding 1 times row 2 to row 1, we obtain

$$\begin{pmatrix} 3 & 0 & 0 & -9 & 6 & 3 \\ 0 & -3 & 0 & 6 & -4 & -3 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 1 by 1/3, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & -3 & 0 & 6 & -4 & -3 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 2 by $-1/3$, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & 1 & 0 & -2 & 4/3 & 1 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 3 by $-1/3$, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & 1 & 0 & -2 & 4/3 & 1 \\ 0 & 0 & 1 & 3 & -5/3 & -1 \end{pmatrix}.$$

Note now that the array is in reduced row echelon form, and that the left hand half is the identity matrix I_3. It follows that the right hand half of the array represents the inverse A^{-1}. Hence

$$A^{-1} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & 4/3 & 1 \\ 3 & -5/3 & -1 \end{pmatrix}.$$

Example 2.5.3. Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 4 & 5 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

To find A^{-1}, we consider the array

$$(A|I_4) = \begin{pmatrix} 1 & 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 2 & 2 & 4 & 5 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We now perform elementary row operations on this array and try to reduce the left hand half to the matrix I_4. Adding -2 times row 1 to row 2, we obtain

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

Adding 1 times row 2 to row 4, we obtain

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

Interchanging rows 2 and 3, we obtain

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

At this point, we observe that it is impossible to reduce the left hand half of the array to I_4. For those who remain unconvinced, let us continue. Adding 3 times row 3 to row 1, we obtain

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

Adding -1 times row 4 to row 3, we obtain

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

Multiplying row 1 by 6 (here we want to avoid fractions in the next two steps), we obtain

$$\begin{pmatrix} 6 & 6 & 12 & 0 & -30 & 18 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -2 & 1 & 0 & 1 \end{pmatrix}.$$

Adding -15 times row 4 to row 1, we obtain

$$\begin{pmatrix} 6 & 6 & 12 & 0 & 0 & 3 & 0 & -15 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -2 & 1 & 0 & 1 \end{pmatrix}.$$

Adding -2 times row 2 to row 1, we obtain

$$\begin{pmatrix} 6 & 0 & 12 & 0 & 0 & 3 & -2 & -15 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -2 & 1 & 0 & 1 \end{pmatrix}.$$

Multiplying row 1 by 1/6, multiplying row 2 by 1/3, multiplying row 3 by -1 and multiplying row 4 by $-1/2$, we obtain

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

Note now that the array is in reduced row echelon form, and that the left hand half is not the identity matrix I_4. Our technique has failed. In fact, the matrix A is not invertible.

2.6. Criteria for Invertibility

Examples 2.5.2–2.5.3 raise the question of when a given matrix is invertible. In this section, we shall obtain some partial answers to this question. Our first step here is the following simple observation.

PROPOSITION 2M. *Every elementary matrix is invertible.*

PROOF. Let us consider elementary row operations. Recall that these are: (1) interchanging two rows; (2) adding a multiple of one row to another row; and (3) multiplying one row by a non-zero constant.

These elementary row operations can clearly be reversed by elementary row operations. For (1), we interchange the two rows again. For (2), if we have originally added c times row i to row j, then we can reverse this by adding $-c$ times row i to row j. For (3), if we have multiplied any row by a non-zero constant c, we can reverse this by multiplying the same row by the constant $1/c$. Note now that each elementary matrix is obtained from I_n by an elementary row operation. The inverse of this elementary matrix is clearly the elementary matrix obtained from I_n by the elementary row operation that reverses the original elementary row operation. ◯

Suppose that an $n \times n$ matrix B can be obtained from an $n \times n$ matrix A by a finite sequence of elementary row operations. Then since these elementary row operations can be reversed, the matrix A can be obtained from the matrix B by a finite sequence of elementary row operations.

Definition. An $n \times n$ matrix A is said to be row equivalent to an $n \times n$ matrix B if there exist a finite number of elementary $n \times n$ matrices $E_1, \ldots, E_k$ such that $B = E_k \ldots E_1 A$.

Remark. Note that $B = E_k \ldots E_1 A$ implies that $A = E_1^{-1} \ldots E_k^{-1} B$. It follows that if A is row equivalent to B, then B is row equivalent to A. We usually say that A and B are row equivalent.

The following result gives conditions equivalent to the invertibility of an $n \times n$ matrix A.

PROPOSITION 2N. *Suppose that*

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

and that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{and} \qquad \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

are $n \times 1$ matrices, where $x_1, \ldots, x_n$ are variables.

(a) Suppose that the matrix A is invertible. Then the system $A\mathbf{x} = \mathbf{0}$ of linear equations has only the trivial solution.

(b) Suppose that the system $A\mathbf{x} = \mathbf{0}$ of linear equations has only the trivial solution. Then the matrices A and I_n are row equivalent.

(c) Suppose that the matrices A and I_n are row equivalent. Then A is invertible.

Proof. (a) Suppose that $\mathbf{x}_0$ is a solution of the system $A\mathbf{x} = \mathbf{0}$. Then since A is invertible, we have

$$\mathbf{x}_0 = I_n \mathbf{x}_0 = (A^{-1}A)\mathbf{x}_0 = A^{-1}(A\mathbf{x}_0) = A^{-1}\mathbf{0} = \mathbf{0}.$$

It follows that the trivial solution is the only solution.

(b) Note that if the system $A\mathbf{x} = \mathbf{0}$ of linear equations has only the trivial solution, then it can be reduced by elementary row operations to the system

$$x_1 = 0, \qquad \ldots, \qquad x_n = 0.$$

This is equivalent to saying that the array

$$\left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} & 0 \end{array} \right)$$

can be reduced by elementary row operations to the reduced row echelon form

$$\left(\begin{array}{ccc|c} 1 & \ldots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{array}\right).$$

Hence the matrices A and I_n are row equivalent.

(c) Suppose that the matrices A and I_n are row equivalent. Then there exist elementary $n \times n$ matrices $E_1, \ldots, E_k$ such that $I_n = E_k \ldots E_1 A$. By Proposition 2M, the matrices $E_1, \ldots, E_k$ are all invertible, so that

$$A = E_1^{-1} \ldots E_k^{-1} I_n = E_1^{-1} \ldots E_k^{-1}$$

is a product of invertible matrices, and is therefore itself invertible. ◯

2.7. Consequences of Invertibility

Suppose that the matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

is invertible. Consider the system $A\mathbf{x} = \mathbf{b}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

are $n \times 1$ matrices, where $x_1, \ldots, x_n$ are variables and $b_1, \ldots, b_n \in \mathbb{R}$ are arbitrary. Since A is invertible, let us consider $\mathbf{x} = A^{-1}\mathbf{b}$. Clearly

$$A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b},$$

so that $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution of the system. On the other hand, let $\mathbf{x}_0$ be any solution of the system. Then $A\mathbf{x}_0 = \mathbf{b}$, so that

$$\mathbf{x}_0 = I_n\mathbf{x}_0 = (A^{-1}A)\mathbf{x}_0 = A^{-1}(A\mathbf{x}_0) = A^{-1}\mathbf{b}.$$

It follows that the system has unique solution. We have proved the following important result.

PROPOSITION 2P. *Suppose that*

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

and that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

are $n \times 1$ matrices, where $x_1, \ldots, x_n$ are variables and $b_1, \ldots, b_n \in \mathbb{R}$ are arbitrary. Suppose further that the matrix A is invertible. Then the system $A\mathbf{x} = \mathbf{b}$ of linear equations has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

We next attempt to study the question in the opposite direction.

PROPOSITION 2Q. *Suppose that*

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

and that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

are $n \times 1$ matrices, where $x_1, \ldots, x_n$ are variables. Suppose further that for every $b_1, \ldots, b_n \in \mathbb{R}$, the system $A\mathbf{x} = \mathbf{b}$ of linear equations is soluble. Then the matrix A is invertible.

PROOF. Suppose that

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \ldots, \qquad \mathbf{b}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

In other words, for every $j = 1, \ldots, n$, $\mathbf{b}_j$ is an $n \times 1$ matrix with entry 1 on row j and entry 0 elsewhere. Now let

$$\mathbf{x}_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \qquad \ldots, \qquad \mathbf{x}_n = \begin{pmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{pmatrix}$$

denote respectively solutions of the systems of linear equations

$$A\mathbf{x} = \mathbf{b}_1, \qquad \ldots, \qquad A\mathbf{x} = \mathbf{b}_n.$$

It is easy to check that

$$A\,(\,\mathbf{x}_1 \quad \ldots \quad \mathbf{x}_n\,) = (\,\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n\,);$$

in other words,

$$A \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \ldots & x_{nn} \end{pmatrix} = I_n,$$

so that A is invertible. ◯

We can now summarize Propositions 2N, 2P and 2Q as follows.

PROPOSITION 2R. *In the notation of Proposition 2N, the following four statements are equivalent:*
(a) The matrix A is invertible.
(b) The system $A\mathbf{x} = \mathbf{0}$ of linear equations has only the trivial solution.
(c) The matrices A and I_n are row equivalent.
(d) The system $A\mathbf{x} = \mathbf{b}$ of linear equations is soluble for every $n \times 1$ matrix $\mathbf{b}$.

2.8. Application to Economics

In this section, we describe briefly the Leontief input-output model, where an economy is divided into n sectors.

For every $i = 1, \ldots, n$, let x_i denote the monetary value of the total output of sector i over a fixed period, and let d_i denote the output of sector i needed to satisfy outside demand over the same fixed period. Collecting together x_i and d_i for $i = 1, \ldots, n$, we obtain the vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \qquad \text{and} \qquad \mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \in \mathbb{R}^n,$$

known respectively as the production vector and demand vector of the economy.

On the other hand, each of the n sectors requires material from some or all of the sectors to produce its output. For $i, j = 1, \ldots, n$, let c_{ij} denote the monetary value of the output of sector i needed by sector j to produce one unit of monetary value of output. For every $j = 1, \ldots, n$, the vector

$$\mathbf{c}_j = \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix} \in \mathbb{R}^n$$

is known as the unit consumption vector of sector j. Note that the column sum

$$c_{1j} + \ldots + c_{nj} \le 1 \tag{5}$$

in order to ensure that sector j does not make a loss. Collecting together the unit consumption vectors, we obtain the matrix

$$C = (\,\mathbf{c}_1 \quad \ldots \quad \mathbf{c}_n\,) = \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \ldots & c_{nn} \end{pmatrix},$$

known as the consumption matrix of the economy.

Consider the matrix product

$$C\mathbf{x} = \begin{pmatrix} c_{11}x_1 + \ldots + c_{1n}x_n \\ \vdots \\ c_{n1}x_1 + \ldots + c_{nn}x_n \end{pmatrix}.$$

For every $i = 1, \ldots, n$, the entry $c_{i1}x_1 + \ldots + c_{in}x_n$ represents the monetary value of the output of sector i needed by all the sectors to produce their output. This leads to the production equation

$$\mathbf{x} = C\mathbf{x} + \mathbf{d}. \tag{6}$$

Here $C\mathbf{x}$ represents the part of the total output that is required by the various sectors of the economy to produce the output in the first place, and $\mathbf{d}$ represents the part of the total output that is available to satisfy outside demand.

Clearly $(I - C)\mathbf{x} = \mathbf{d}$. If the matrix $I - C$ is invertible, then

$$\mathbf{x} = (I - C)^{-1}\mathbf{d}$$

represents the perfect production level. We state without proof the following fundamental result.

PROPOSITION 2S. *Suppose that the entries of the consumption matrix C and the demand vector $\mathbf{d}$ are non-negative. Suppose further that the inequality (5) holds for each column of C. Then the inverse matrix $(I-C)^{-1}$ exists, and the production vector $\mathbf{x}=(I-C)^{-1}\mathbf{d}$ has non-negative entries and is the unique solution of the production equation (6).*

Let us indulge in some heuristics. Initially, we have demand $\mathbf{d}$. To produce $\mathbf{d}$, we need $C\mathbf{d}$ as input. To produce this extra $C\mathbf{d}$, we need $C(C\mathbf{d}) = C^2\mathbf{d}$ as input. To produce this extra $C^2\mathbf{d}$, we need $C(C^2\mathbf{d}) = C^3\mathbf{d}$ as input. And so on. Hence we need to produce

$$\mathbf{d}+C\mathbf{d}+C^2\mathbf{d}+C^3\mathbf{d}+\ldots=(I+C+C^2+C^3+\ldots)\mathbf{d}$$

in total. Now it is not difficult to check that for every positive integer k, we have

$$(I-C)(I+C+C^2+C^3+\ldots+C^k)=I-C^{k+1}.$$

If the entries of C^{k+1} are all very small, then

$$(I-C)(I+C+C^2+C^3+\ldots+C^k)\approx I,$$

so that

$$(I-C)^{-1}\approx I+C+C^2+C^3+\ldots+C^k.$$

This gives a practical way of approximating $(I-C)^{-1}$, and also suggests that

$$(I-C)^{-1}=I+C+C^2+C^3+\ldots.$$

EXAMPLE 2.8.1. An economy consists of three sectors. Their dependence on each other is summarized in the table below:

	To produce one unit of monetary value of output in sector 1	2	3
monetary value of output required from sector 1	0.3	0.2	0.1
monetary value of output required from sector 2	0.4	0.5	0.2
monetary value of output required from sector 3	0.1	0.1	0.3

Suppose that the final demand from sectors 1, 2 and 3 are respectively 30, 50 and 20. Then the production vector and demand vector are respectively

$$\mathbf{x}=\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix} \qquad\text{and}\qquad \mathbf{d}=\begin{pmatrix} d_1\\ d_2\\ d_3\end{pmatrix}=\begin{pmatrix} 30\\ 50\\ 20\end{pmatrix},$$

while the consumption matrix is given by

$$C=\begin{pmatrix} 0.3 & 0.2 & 0.1\\ 0.4 & 0.5 & 0.2\\ 0.1 & 0.1 & 0.3\end{pmatrix}, \qquad\text{so that}\qquad I-C=\begin{pmatrix} 0.7 & -0.2 & -0.1\\ -0.4 & 0.5 & -0.2\\ -0.1 & -0.1 & 0.7\end{pmatrix}.$$

The production equation $(I-C)\mathbf{x}=\mathbf{d}$ has augmented matrix

$$\left(\begin{array}{ccc|c} 0.7 & -0.2 & -0.1 & 30\\ -0.4 & 0.5 & -0.2 & 50\\ -0.1 & -0.1 & 0.7 & 20\end{array}\right), \qquad\text{equivalent to}\qquad \left(\begin{array}{ccc|c} 7 & -2 & -1 & 300\\ -4 & 5 & -2 & 500\\ -1 & -1 & 7 & 200\end{array}\right),$$

and which can be converted to reduced row echelon form

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3200/27 \\ 0 & 1 & 0 & 6100/27 \\ 0 & 0 & 1 & 700/9 \end{array}\right).$$

This gives $x_1 \approx 119$, $x_2 \approx 226$ and $x_3 \approx 78$, to the nearest integers.

2.9. Matrix Transformation on the Plane

Let A be a 2×2 matrix with real entries. A matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ can be defined as follows: For every $\mathbf{x} = (x_1, x_2) \in \mathbb{R}$, we write $T(\mathbf{x}) = \mathbf{y}$, where $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$ satisfies

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Such a transformation is linear, in the sense that $T(\mathbf{x}' + \mathbf{x}'') = T(\mathbf{x}') + T(\mathbf{x}')$ for every $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^2$ and $T(c\mathbf{x}) = cT(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^2$ and every $c \in \mathbb{R}$. To see this, simply observe that

$$A \begin{pmatrix} x_1' + x_1'' \\ x_2' + x_2'' \end{pmatrix} = A \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} + A \begin{pmatrix} x_1'' \\ x_2'' \end{pmatrix} \qquad \text{and} \qquad A \begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix} = cA \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We shall study linear transformations in greater detail in Chapter 8. Here we confine ourselves to looking at a few simple matrix transformations on the plane.

Example 2.9.1. The matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{satisfies} \qquad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents reflection across the x_1-axis, whereas the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{satisfies} \qquad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents reflection across the x_2-axis. On the other hand, the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{satisfies} \qquad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents reflection across the origin, whereas the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{satisfies} \qquad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents reflection across the line $x_1 = x_2$. We give a summary in the table below:

Transformation	Equations	Matrix
Reflection across x_1-axis	$\begin{cases} y_1 = x_1 \\ y_2 = -x_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
Reflection across x_2-axis	$\begin{cases} y_1 = -x_1 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Reflection across origin	$\begin{cases} y_1 = -x_1 \\ y_2 = -x_2 \end{cases}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
Reflection across $x_1 = x_2$	$\begin{cases} y_1 = x_2 \\ y_2 = x_1 \end{cases}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

EXAMPLE 2.9.2. Let k be a fixed positive real number. The matrix

$$A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \qquad \text{satisfies} \qquad A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} kx_1 \\ kx_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents a dilation if $k > 1$ and a contraction if $0 < k < 1$. On the other hand, the matrix

$$A = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{satisfies} \qquad A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} kx_1 \\ x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents an expansionn in the x_1-direction if $k > 1$ and a compression in the x_1-direction if $0 < k < 1$, whereas the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \qquad \text{satisfies} \qquad A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ kx_2 \end{pmatrix}$$

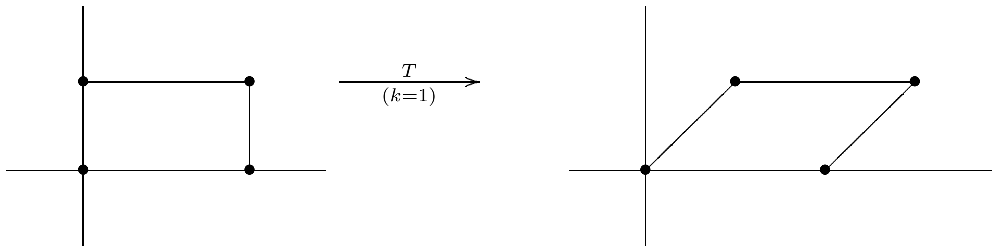
for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents a expansion in the x_2-direction if $k > 1$ and a compression in the x_2-direction if $0 < k < 1$. We give a summary in the table below:

Transformation	Equations	Matrix
Dilation or contraction by factor $k > 0$	$\begin{cases} y_1 = kx_1 \\ y_2 = kx_2 \end{cases}$	$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$
Expansion or compression in x_1-direction by factor $k > 0$	$\begin{cases} y_1 = kx_1 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$
Expansion or compression in x_2-direction by factor $k > 0$	$\begin{cases} y_1 = x_1 \\ y_2 = kx_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$

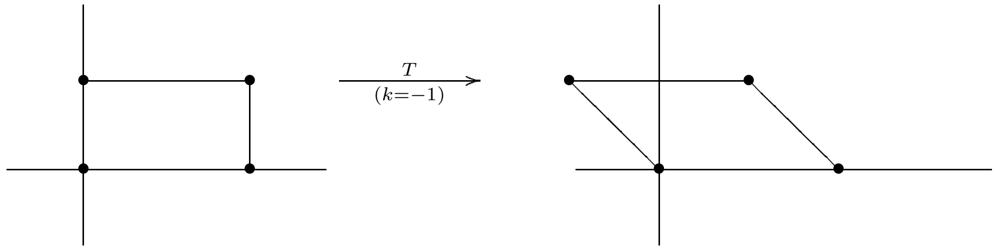
EXAMPLE 2.9.3. Let k be a fixed real number. The matrix

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \qquad \text{satisfies} \qquad A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + kx_2 \\ x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents a shear in the x_1-direction. For the case $k = 1$, we have the following:



For the case $k = -1$, we have the following:



Similarly, the matrix

$$A = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} \qquad \text{satisfies} \qquad A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ kx_1 + x_2 \end{pmatrix}$$

for every $(x_1, x_2) \in \mathbb{R}^2$, and so represents a shear in the x_2-direction. We give a summary in the table below:

Transformation	Equations	Matrix
Shear in x_1-direction	$\begin{cases} y_1 = x_1 + kx_2 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$
Shear in x_2-direction	$\begin{cases} y_1 = x_1 \\ y_2 = kx_1 + x_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$

EXAMPLE 2.9.4. For anticlockwise rotation by an angle θ, we have $T(x_1, x_2) = (y_1, y_2)$, where

$$y_1 + \mathrm{i}y_2 = (x_1 + \mathrm{i}x_2)(\cos\theta + \mathrm{i}\sin\theta),$$

and so

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It follows that the matrix in question is given by

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We give a summary in the table below:

Transformation	Equations	Matrix
Anticlockwise rotation by angle θ	$\begin{cases} y_1 = x_1\cos\theta - x_2\sin\theta \\ y_2 = x_1\sin\theta + x_2\cos\theta \end{cases}$	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

We conclude this section by establishing the following result which reinforces the linearity of matrix transformations on the plane.

PROPOSITION 2T. *Suppose that a matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is given by an invertible matrix A. Then*
(a) the image under T of a straight line is a straight line;
(b) the image under T of a straight line through the origin is a straight line through the origin; and
(c) the images under T of parallel straight lines are parallel straight lines.

PROOF. Suppose that $T(x_1, x_2) = (y_1, y_2)$. Since A is invertible, we have $\mathbf{x} = A^{-1}\mathbf{y}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

The equation of a straight line is given by $\alpha x_1 + \beta x_2 = \gamma$ or, in matrix form, by

$$(\,\alpha \quad \beta\,)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (\,\gamma\,).$$

Hence

$$\begin{pmatrix} \alpha & \beta \end{pmatrix} A^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (\gamma).$$

Let

$$\begin{pmatrix} \alpha' & \beta' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \end{pmatrix} A^{-1}.$$

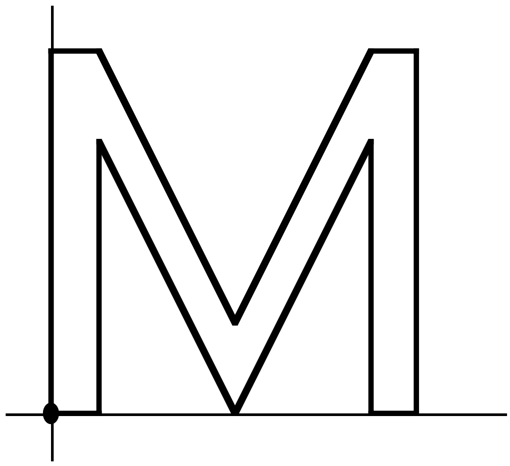
Then

$$\begin{pmatrix} \alpha' & \beta' \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (\gamma).$$

In other words, the image under T of the straight line $\alpha x_1 + \beta x_2 = \gamma$ is $\alpha' y_1 + \beta' y_2 = \gamma$, clearly another straight line. This proves (a). To prove (b), note that straight lines through the origin correspond to $\gamma = 0$. To prove (c), note that parallel straight lines correspond to different values of γ for the same values of α and β. ◯

2.10. Application to Computer Graphics

EXAMPLE 2.10.1. Consider the letter M in the diagram below:



Following the boundary in the anticlockwise direction starting at the origin, the 12 vertices can be represented by the coordinates

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 6 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 0 \end{pmatrix},$$
$$\begin{pmatrix} 8 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 8 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 8 \end{pmatrix}.$$

Let us apply a matrix transformation to these vertices, using the matrix

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix},$$

representing a shear in the x_1-direction with factor 0.5, so that

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \frac{1}{2} x_2 \\ x_2 \end{pmatrix} \quad \text{for every } (x_1, x_2) \in \mathbb{R}^2.$$

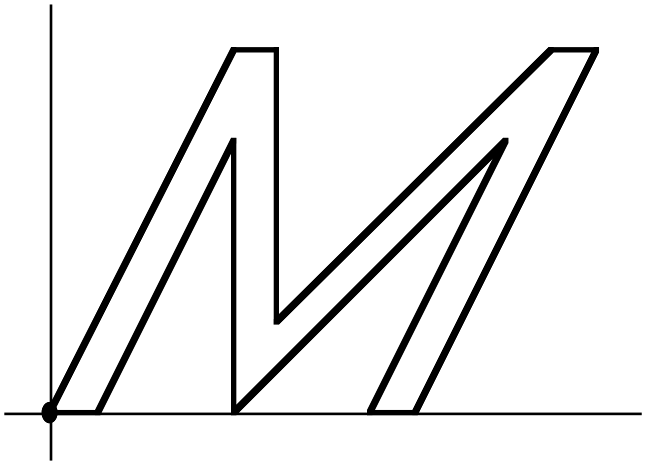
Then the images of the 12 vertices are respectively

$$\begin{pmatrix}0\\0\end{pmatrix}, \quad \begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}4\\6\end{pmatrix}, \quad \begin{pmatrix}4\\0\end{pmatrix}, \quad \begin{pmatrix}10\\6\end{pmatrix}, \quad \begin{pmatrix}7\\0\end{pmatrix},$$
$$\begin{pmatrix}8\\0\end{pmatrix}, \quad \begin{pmatrix}12\\8\end{pmatrix}, \quad \begin{pmatrix}11\\8\end{pmatrix}, \quad \begin{pmatrix}5\\2\end{pmatrix}, \quad \begin{pmatrix}5\\8\end{pmatrix}, \quad \begin{pmatrix}4\\8\end{pmatrix},$$

noting that

$$\begin{pmatrix}1 & \frac{1}{2}\\0 & 1\end{pmatrix}\begin{pmatrix}0 & 1 & 1 & 4 & 7 & 7 & 8 & 8 & 7 & 4 & 1 & 0\\0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8\end{pmatrix} = \begin{pmatrix}0 & 1 & 4 & 4 & 10 & 7 & 8 & 12 & 11 & 5 & 5 & 4\\0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8\end{pmatrix}.$$

In view of Proposition 2T, the image of any line segment that joins two vertices is a line segment that joins the images of the two vertices. Hence the image of the letter M under the shear looks like the following:



Next, we may wish to translate this image. However, a translation is a transformation by vector $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ is of the form

$$\begin{pmatrix}y_1\\y_2\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix} + \begin{pmatrix}h_1\\h_2\end{pmatrix} \quad \text{for every } (x_1, x_2) \in \mathbb{R}^2,$$

and this cannot be described by a matrix transformation on the plane. To overcome this deficiency, we introduce homogeneous coordinates. For every point $(x_1, x_2) \in \mathbb{R}^2$, we identify it with the point $(x_1, x_2, 1) \in \mathbb{R}^3$. Now we wish to translate a point (x_1, x_2) to $(x_1, x_2) + (h_1, h_2) = (x_1 + h_1, x_2 + h_2)$, so we attempt to find a 3×3 matrix A^* such that

$$\begin{pmatrix}x_1 + h_1\\x_2 + h_2\\1\end{pmatrix} = A^* \begin{pmatrix}x_1\\x_2\\1\end{pmatrix} \quad \text{for every } (x_1, x_2) \in \mathbb{R}^2.$$

It is easy to check that

$$\begin{pmatrix}x_1 + h_1\\x_2 + h_2\\1\end{pmatrix} = \begin{pmatrix}1 & 0 & h_1\\0 & 1 & h_2\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\1\end{pmatrix} \quad \text{for every } (x_1, x_2) \in \mathbb{R}^2.$$

It follows that using homogeneous coordinates, translation by vector $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ can be described by the matrix

$$A^* = \begin{pmatrix}1 & 0 & h_1\\0 & 1 & h_2\\0 & 0 & 1\end{pmatrix}.$$

Remark. Consider a matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ on the plane given by a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Suppose that $T(x_1, x_2) = (y_1, y_2)$. Then

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Under homogeneous coordinates, the image of the point $(x_1, x_2, 1)$ is now $(y_1, y_2, 1)$. Note that

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}.$$

It follows that homogeneous coordinates can also be used to study all the matrix transformations we have discussed in Section 2.9. By moving over to homogeneous coordinates, we simply replace the 2×2 matrix A by the 3×3 matrix

$$A^* = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 2.10.2. Returning to Example 2.10.1 of the letter M, the 12 vertices are now represented by homogeneous coordinates, put in an array in the form

$$\begin{pmatrix} 0 & 1 & 1 & 4 & 7 & 7 & 8 & 8 & 7 & 4 & 1 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Then the 2×2 matrix

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

is now replaced by the 3×3 matrix

$$A^* = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{aligned} & A^* \begin{pmatrix} 0 & 1 & 1 & 4 & 7 & 7 & 8 & 8 & 7 & 4 & 1 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 4 & 7 & 7 & 8 & 8 & 7 & 4 & 1 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 4 & 4 & 10 & 7 & 8 & 12 & 11 & 5 & 5 & 4 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \end{aligned}$$

Next, let us consider a translation by the vector $(2, 3)$. The matrix under homogeneous coordinates for this translation is given by

$$B^* = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

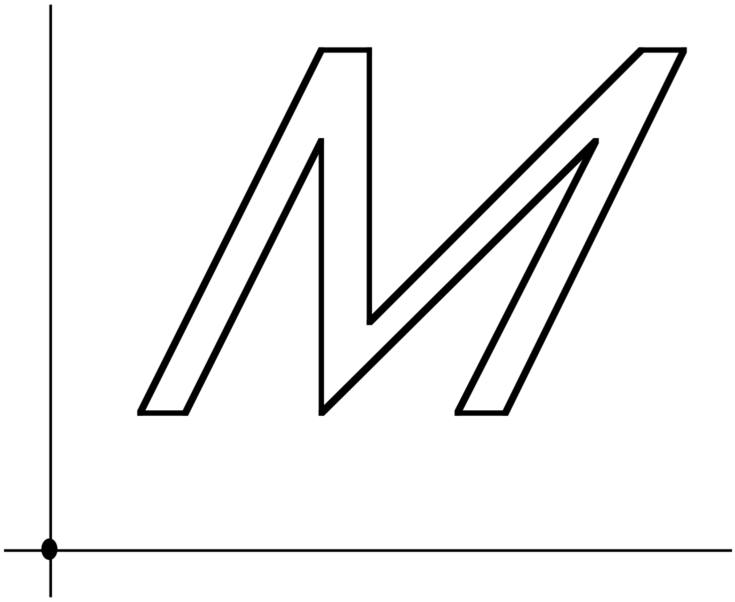
Note that

$$\begin{aligned}
B^*A^* &\begin{pmatrix} 0 & 1 & 1 & 4 & 7 & 7 & 8 & 8 & 7 & 4 & 1 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 4 & 4 & 10 & 7 & 8 & 12 & 11 & 5 & 5 & 4 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 8 & 8 & 2 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 3 & 6 & 6 & 12 & 9 & 10 & 14 & 13 & 7 & 7 & 6 \\ 3 & 3 & 9 & 3 & 9 & 3 & 3 & 11 & 11 & 5 & 11 & 11 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},
\end{aligned}$$

giving rise to coordinates in $\mathbb{R}^2$, displayed as an array

$$\begin{pmatrix} 2 & 3 & 6 & 6 & 12 & 9 & 10 & 14 & 13 & 7 & 7 & 6 \\ 3 & 3 & 9 & 3 & 9 & 3 & 3 & 11 & 11 & 5 & 11 & 11 \end{pmatrix}$$

Hence the image of the letter M under the shear followed by translation looks like the following:



EXAMPLE 2.10.3. Under homogeneous coordinates, the transformation representing a reflection across the x_1-axis, followed by a shear by factor 2 in the x_1-direction, followed by anticlockwise rotation by $90°$, and followed by translation by vector $(2, -1)$, has matrix

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.11. Complexity of a Non-Homogeneous System

Consider the problem of solving a system of linear equations of the form $A\mathbf{x} = \mathbf{b}$, where A is an $n \times n$ invertible matrix. We are interested in the number of operations required to solve such a system. By an operation, we mean interchanging, adding or multiplying two real numbers.

One way of solving the system $A\mathbf{x} = \mathbf{b}$ is to write down the augmented matrix

$$\left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} & b_n \end{array}\right), \tag{7}$$

and then convert it to reduced row echelon form by elementary row operations.

The first step is to reduce it to row echelon form:

(I) First of all, we may need to interchange two rows in order to ensure that the top left entry in the array is non-zero. This requires $n+1$ operations.

(II) Next, we need to multiply the new first row by a constant in order to make the top left pivot entry equal to 1. This requires $n+1$ operations, and the array now looks like

$$\left(\begin{array}{cccc|c} 1 & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_n \end{array}\right).$$

Note that we are abusing notation somewhat, as the entry a_{12} here, for example, may well be different from the entry a_{12} in the augmented matrix (7).

(III) For each row $i = 2, \ldots, n$, we now multiply the first row by $-a_{i1}$ and then add to row i. This requires $2(n-1)(n+1)$ operations, and the array now looks like

$$\left(\begin{array}{cccc|c} 1 & a_{12} & \ldots & a_{1n} & b_1 \\ 0 & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} & b_n \end{array}\right). \tag{8}$$

(IV) In summary, to proceed from the form (7) to the form (8), the number of operations required is at most $2(n+1) + 2(n-1)(n+1) = 2n(n+1)$.

(V) Our next task is to convert the smaller array

$$\left(\begin{array}{ccc|c} a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{n2} & \ldots & a_{nn} & b_n \end{array}\right)$$

to an array that looks like

$$\left(\begin{array}{cccc|c} 1 & a_{23} & \ldots & a_{2n} & b_2 \\ 0 & a_{33} & \ldots & a_{3n} & b_3 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n3} & \ldots & a_{nn} & b_n \end{array}\right).$$

These have one row and one column fewer than the arrays (7) and (8), and the number of operations required is at most $2m(m+1)$, where $m = n-1$. We continue in this way systematically to reach row echelon form, and conclude that the number of operations required to convert the augmented matrix (7) to row echelon form is at most

$$\sum_{m=1}^{n} 2m(m+1) \approx \frac{2}{3}n^3.$$

The next step is to convert the row echelon form to reduced row echelon form. This is simpler, as many entries are now zero. It can be shown that the number of operations required is bounded by something like $2n^2$ – indeed, by something like n^2 if one analyzes the problem more carefully. In any case, these estimates are insignificant compared to the estimate $\frac{2}{3}n^3$ earlier.

We therefore conclude that the number of operations required to solve the system $A\mathbf{x} = \mathbf{b}$ by reducing the augmented matrix to reduced row echelon form is bounded by something like $\frac{2}{3}n^3$ when n is large.

Another way of solving the system $A\mathbf{x} = \mathbf{b}$ is to first find the inverse matrix A^{-1}. This may involve converting the array

$$\left(\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1n} & 1 & & \\ \vdots & & \vdots & & \ddots & \\ a_{n1} & \ldots & a_{nn} & & & 1 \end{array}\right)$$

to reduced row echelon form by elementary row operations. It can be shown that the number of operations required is something like $2n^3$, so this is less efficient than our first method.

2.12. Matrix Factorization

In some situations, we may need to solve systems of linear equations of the form $A\mathbf{x} = \mathbf{b}$, with the same coefficient matrix A but for many different vectors $\mathbf{b}$. If A is an invertible square matrix, then we can find its inverse A^{-1} and then compute $A^{-1}\mathbf{b}$ for each vector $\mathbf{b}$. However, the matrix A may not be a square matrix, and we may have to convert the augmented matrix to reduced row echelon form.

In this section, we describe a way for solving this problem in a more efficient way. To describe this, we first need a definition.

DEFINITION. A rectangular array of numbers is said to be in quasi row echelon form if the following conditions are satisfied:
(1) The left-most non-zero entry of any non-zero row is called a pivot entry. It is not necessary for its value to be equal to 1.
(2) All zero rows are grouped together at the bottom of the array.
(3) The pivot entry of a non-zero row occurring lower in the array is to the right of the pivot entry of a non-zero row occurring higher in the array.

In other words, the array looks like row echelon form in shape, except that the pivot entries do not have to be equal to 1.

We consider first of all a special case.

PROPOSITION 2U. *Suppose that an $m \times n$ matrix A can be converted to quasi row echelon form by elementary row operations but without interchanging any two rows. Then $A = LU$, where L is an $m \times m$ lower triangular matrix with diagonal entries all equal to* 1 *and U is a quasi row echelon form of A.*

SKETCH OF PROOF. Recall that applying an elementary row operation to an $m \times n$ matrix corresponds to multiplying the matrix on the left by an elementary $m \times m$ matrix. On the other hand, if we are aiming for quasi row echelon form and not row echelon form, then there is no need to multiply any row of the array by a non-zero constant. Hence the only elementary row operation we need to perform is to add a multiple of one row to another row. In fact, it is sufficient even to restrict this to adding a multiple of a row higher in the array to another row lower in the array, and it is easy to see that the corresponding elementary matrix is lower triangular, with diagonal entries all equal to 1. Let us call such elementary matrices unit lower triangular. If an $m \times n$ matrix A can be reduced in this way to quasi row echelon form U, then

$$U = E_k \ldots E_2 E_1 A,$$

where the elementary matrices $E_1, E_2, \ldots, E_k$ are all unit lower triangular. Let $L = (E_k \ldots E_2 E_1)^{-1}$. Then $A = LU$. It can be shown that products and inverses of unit lower triangular matrices are also unit lower triangular. Hence L is a unit lower triangular matrix as required. ◯

If $A\mathbf{x} = \mathbf{b}$ and $A = LU$, then $L(U\mathbf{x}) = \mathbf{b}$. Writing $\mathbf{y} = U\mathbf{x}$, we have

$$L\mathbf{y} = \mathbf{b} \qquad \text{and} \qquad U\mathbf{x} = \mathbf{y}.$$

It follows that the problem of solving the system $A\mathbf{x} = \mathbf{b}$ corresponds to first solving the system $L\mathbf{y} = \mathbf{b}$ and then solving the system $U\mathbf{x} = \mathbf{y}$. Both of these systems are easy to solve since both L and U have many zero entries. It remains to find L and U.

If we reduce the matrix A to quasi row echelon form by only performing the elementary row operation of adding a multiple of a row higher in the array to another row lower in the array, then U can be taken as the quasi row echelon form resulting from this. It remains to find L. However, note that $L = (E_k \ldots E_2 E_1)^{-1}$, where $U = E_k \ldots E_2 E_1 A$, and so

$$I = E_k \ldots E_2 E_1 L.$$

This means that the very elementary row operations that convert A to U will convert L to I. We therefore wish to create a matrix L such that this is satisfied. It is simplest to illustrate the technique by an example.

EXAMPLE 2.12.1. Consider the matrix

$$A = \begin{pmatrix} 2 & -1 & 2 & -2 & 3 \\ 4 & 1 & 6 & -5 & 8 \\ 2 & -10 & -4 & 8 & -5 \\ 2 & -13 & -6 & 16 & -5 \end{pmatrix}.$$

The entry 2 in row 1 and column 1 is a pivot entry, and column 1 is a pivot column. Adding -2 times row 1 to row 2, adding -1 times row 1 to row 3, and adding -1 times row 1 to row 4, we obtain

$$\begin{pmatrix} 2 & -1 & 2 & -2 & 3 \\ 0 & 3 & 2 & -1 & 2 \\ 0 & -9 & -6 & 10 & -8 \\ 0 & -12 & -8 & 18 & -8 \end{pmatrix}.$$

Note that the same three elementary row operations convert

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & * & 1 & 0 \\ 1 & * & * & 1 \end{pmatrix} \qquad \text{to} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & * & 1 & 0 \\ 0 & * & * & 1 \end{pmatrix}.$$

Next, the entry 3 in row 2 and column 2 is a pivot entry, and column 2 is a pivot column. Adding 3 times row 2 to row 3, and adding 4 times row 2 to row 4, we obtain

$$\begin{pmatrix} 2 & -1 & 2 & -2 & 3 \\ 0 & 3 & 2 & -1 & 2 \\ 0 & 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & 14 & 0 \end{pmatrix}.$$

Note that the same two elementary row operations convert

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & * & 1 \end{pmatrix} \qquad \text{to} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & * & 1 \end{pmatrix}.$$

Next, the entry 7 in row 3 and column 4 is a pivot entry, and column 4 is a pivot column. Adding -2 times row 3 to row 4, we obtain the quasi row echelon form

$$U = \begin{pmatrix} 2 & -1 & 2 & -2 & 3 \\ 0 & 3 & 2 & -1 & 2 \\ 0 & 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix},$$

where the entry 4 in row 4 and column 5 is a pivot entry, and column 5 is a pivot column. Note that the same elementary row operation converts

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \qquad \text{to} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Now observe that if we take

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & -4 & 2 & 1 \end{pmatrix},$$

then L can be converted to I_4 by the same elementary operations that convert A to U.

The strategy is now clear. Every time we find a new pivot, we note its value and the entries below it. The lower triangular entries of L are formed by these columns with each column divided by the value of the pivot entry in that column.

EXAMPLE 2.12.2. Let us examine our last example again. The pivot columns at the time of establishing the pivot entries are respectively

$$\begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \qquad \begin{pmatrix} * \\ 3 \\ -9 \\ -12 \end{pmatrix}, \qquad \begin{pmatrix} * \\ * \\ 7 \\ 14 \end{pmatrix}, \qquad \begin{pmatrix} * \\ * \\ * \\ 4 \end{pmatrix}.$$

Dividing them respectively by the pivot entries 2, 3, 7 and 4, we obtain respectively the columns

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

Note that the lower triangular entries of the matrix

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & -4 & 2 & 1 \end{pmatrix}$$

correspond precisely to the entries in these columns.

LU FACTORIZATION ALGORITHM.

(1) Reduce the matrix A to quasi row echelon form by only performing the elementary row operation of adding a multiple of a row higher in the array to another row lower in the array. Let U be the quasi row echelon form obtained.

(2) Record any new pivot column at the time of its first recognition, and modify it by replacing any entry above the pivot entry by zero and dividing every other entry by the value of the pivot entry.

(3) Let L denote the square matrix obtained by letting the columns be the pivot columns as modified in step (2).

EXAMPLE 2.12.3. We wish to solve the system of linear equations $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ -3 & 3 & -5 & 5 & -2 \\ 6 & -4 & 11 & -10 & 6 \\ -6 & 8 & -21 & 13 & -9 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ 9 \\ -15 \end{pmatrix}.$$

Let us first apply LU factorization to the matrix A. The first pivot column is column 1, with modified version

$$\begin{pmatrix} 1 \\ -1 \\ 2 \\ -2 \end{pmatrix}.$$

Adding row 1 to row 2, adding -2 times row 1 to row 3, and adding 2 times row 1 to row 4, we obtain

$$\begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ 0 & 2 & -3 & 1 & -1 \\ 0 & -2 & 7 & -2 & 4 \\ 0 & 6 & -17 & 5 & -7 \end{pmatrix}.$$

The second pivot column is column 2, with modified version

$$\begin{pmatrix} 0 \\ 1 \\ -1 \\ 3 \end{pmatrix}.$$

Adding row 2 to row 3, and adding -3 times row 2 to row 4, we obtain

$$\begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ 0 & 2 & -3 & 1 & -1 \\ 0 & 0 & 4 & -1 & 3 \\ 0 & 0 & -8 & 2 & -4 \end{pmatrix}.$$

The third pivot column is column 3, with modified version

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -2 \end{pmatrix}.$$

Adding 2 times row 3 to row 4, we obtain the quasi row echelon form

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

The last pivot column is column 5, with modified version

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

It follows that

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -2 & 3 & -2 & 1 \end{pmatrix} \qquad \text{and} \qquad U = \begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ 0 & 2 & -3 & 1 & -1 \\ 0 & 0 & 4 & -1 & 3 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

We now consider the system $L\mathbf{y} = \mathbf{b}$, with augmented matrix

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & -2 \\ 2 & -1 & 1 & 0 & 9 \\ -2 & 3 & -2 & 1 & -15 \end{array}\right).$$

Using row 1, we obtain $y_1 = 1$. Using row 2, we obtain $y_2 - y_1 = -2$, so that $y_2 = -1$. Using row 3, we obtain $y_3 + 2y_1 - y_2 = 9$, so that $y_3 = 6$. Using row 4, we obtain $y_4 - 2y_1 + 3y_2 - 2y_3 = -15$, so that $y_4 = 2$. Hence

$$\mathbf{y} = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix}.$$

We next consider the system $U\mathbf{x} = \mathbf{y}$, with augmented matrix

$$\left(\begin{array}{ccccc|c} 3 & -1 & 2 & -4 & 1 & 1 \\ 0 & 2 & -3 & 1 & -1 & -1 \\ 0 & 0 & 4 & -1 & 3 & 6 \\ 0 & 0 & 0 & 0 & 2 & 2 \end{array}\right).$$

Here the free variable is x_4. Let $x_4 = t$. Using row 4, we obtain $2x_5 = 2$, so that $x_5 = 1$. Using row 3, we obtain $4x_3 = 6 + x_4 - 3x_5 = 3 + t$, so that $x_3 = \frac{3}{4} + \frac{1}{4}t$. Using row 2, we obtain

$$2x_2 = -1 + 3x_3 - x_4 + x_5 = \tfrac{9}{4} - \tfrac{1}{4}t,$$

so that $x_2 = \frac{9}{8} - \frac{1}{8}t$. Using row 1, we obtain $3x_1 = 1 + x_2 - 2x_3 + 4x_4 - x_5 = \frac{27}{8}t - \frac{3}{8}$, so that $x_1 = \frac{9}{8}t - \frac{1}{8}$. Hence

$$(x_1, x_2, x_3, x_4, x_5) = \left(\frac{9t-1}{8}, \frac{9-t}{8}, \frac{3+t}{4}, t, 1 \right), \quad \text{where } t \in \mathbb{R}.$$

Remarks. (1) In practical situations, interchanging rows is usually necessary to convert a matrix A to quasi row echelon form. The technique here can be modified to produce a matrix L which is not unit lower triangular, but which can be made unit lower triangular by interchanging rows.

(2) Computing an LU factorization of an $n \times n$ matrix takes approximately $\frac{2}{3}n^3$ operations. Solving the systems $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$ requires approximately $2n^2$ operations.

(3) LU factorization is particularly efficient when the matrix A has many zero entries, in which case the matrices L and U may also have many zero entries.

2.13. Application to Games of Strategy

Consider a game with two players. Player R, usually known as the row player, has m possible moves, denoted by $i = 1, 2, 3, \ldots, m$, while player C, usually known as the column player, has n possible moves, denoted by $j = 1, 2, 3, \ldots, n$. For every $i = 1, 2, 3, \ldots, m$ and $j = 1, 2, 3, \ldots, n$, let a_{ij} denote the payoff that player C has to make to player R if player R makes move i and player C makes move j. These numbers give rise to the payoff matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}.$$

The entries can be positive, negative or zero.

Suppose that for every $i = 1, 2, 3, \ldots, m$, player R makes move i with probability p_i, and that for every $j = 1, 2, 3, \ldots, n$, player C makes move j with probability q_j. Then

$$p_1 + \ldots + p_m = 1 \qquad \text{and} \qquad q_1 + \ldots + q_n = 1.$$

Assume that the players make moves independently of each other. Then for every $i = 1, 2, 3, \ldots, m$ and $j = 1, 2, 3, \ldots, n$, the number $p_i q_j$ represents the probability that player R makes move i and player C makes move j. Then the double sum

$$E_A(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} p_i q_j$$

represents the expected payoff that player C has to make to player R.

The matrices

$$\mathbf{p} = (\, p_1 \quad \ldots \quad p_m \,) \qquad \text{and} \qquad \mathbf{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$$

are known as the strategies of player R and player C respectively. Clearly the expected payoff

$$E_A(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} p_i q_j = (\, p_1 \quad \ldots \quad p_m \,) \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} = \mathbf{p}A\mathbf{q}.$$

Here we have slightly abused notation. The right hand side is a 1×1 matrix!

We now consider the following problem: Suppose that A is fixed. Is it possible for player R to choose a strategy $\mathbf{p}$ to try to maximize the expected payoff $E_A(\mathbf{p}, \mathbf{q})$? Is it possible for player C to choose a strategy $\mathbf{q}$ to try to minimize the expected payoff $E_A(\mathbf{p}, \mathbf{q})$?

FUNDEMENTAL THEOREM OF ZERO SUM GAMES. *There exist strategies* $\mathbf{p}^*$ *and* $\mathbf{q}^*$ *such that*

$$E_A(\mathbf{p}^*, \mathbf{q}) \geq E_A(\mathbf{p}^*, \mathbf{q}^*) \geq E_A(\mathbf{p}, \mathbf{q}^*)$$

for every strategy $\mathbf{p}$ *of player* R *and every strategy* $\mathbf{q}$ *of player* C.

Remark. The strategy $\mathbf{p}^*$ is known as an optimal strategy for player R, and the strategy $\mathbf{q}^*$ is known as an optimal strategy for player C. The quantity $E_A(\mathbf{p}^*, \mathbf{q}^*)$ is known as the value of the game. Optimal strategies are not necessarily unique. However, if $\mathbf{p}^{**}$ and $\mathbf{q}^{**}$ are another pair of optimal strategies, then $E_A(\mathbf{p}^*, \mathbf{q}^*) = E_A(\mathbf{p}^{**}, \mathbf{q}^{**})$.

Zero sum games which are strictly determined are very easy to analyze. Here the payoff matrix A contains saddle points. An entry a_{ij} in the payoff matrix A is called a saddle point if it is a least entry in its row and a greatest entry in its column. In this case, the strategies

$$\mathbf{p}^* = (\,0 \quad \ldots \quad 0 \quad 1 \quad 0 \quad \ldots \quad 0\,) \qquad \text{and} \qquad \mathbf{q} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the 1's occur in position i in $\mathbf{p}^*$ and position j in $\mathbf{q}^*$, are optimal strategies, so that the value of the game is a_{ij}.

REMARK. It is very easy to show that different saddle points in the payoff matrix have the same value.

EXAMPLE 2.13.1. In some sports mad school, the teachers require 100 students to each choose between rowing (R) and cricket (C). However, the students cannot make up their mind, and will only decide when the identities of the rowing coach and cricket coach are known. There are 3 possible rowing coaches and 4 possible cricket coaches the school can hire. The number of students who will choose rowing ahead of cricket in each scenario is as follows, where $R1$, $R2$ and $R3$ denote the 3 possible rowing coaches, and $C1$, $C2$, $C3$ and $C4$ denote the 4 possible cricket coaches:

	$C1$	$C2$	$C3$	$C4$
$R1$	75	50	45	60
$R2$	20	60	30	55
$R3$	45	70	35	30

[For example, if coaches $R2$ and $C1$ are hired, then 20 students will choose rowing, and so 80 students will choose cricket.] We first reset the problem by subtracting 50 from each entry and create a payoff matrix

$$A = \begin{pmatrix} 25 & 0 & -5 & 10 \\ -30 & 10 & -20 & 5 \\ -5 & 20 & -15 & -20 \end{pmatrix}.$$

[For example, the top left entry denotes that if each sport starts with 50 students, then 25 is the number cricket concedes to rowing.] Here the entry -5 in row 1 and column 3 is a saddle point, so the optimal strategy for rowing is to use coach $R1$ and the optimal strategy for cricket is to use coach $C3$.

In general, saddle points may not exist, so that the problem is not strictly determined. Then the solution for these optimal problems are solved by linear programming techniques which we do not discuss here. However, in the case of 2×2 payoff matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

which do not contain saddle points, we can write $p_2 = 1 - p_1$ and $q_2 = 1 - q_1$. Then

$$\begin{aligned} E_A(\mathbf{p},\mathbf{q}) &= a_{11}p_1q_1 + a_{12}p_1(1-q_1) + a_{21}(1-p_1)q_1 + a_{22}(1-p_1)(1-q_1) \\ &= ((a_{11} - a_{12} - a_{21} + a_{22})p_1 - (a_{22} - a_{21}))q_1 + (a_{12} - a_{22})p_1 + a_{22}. \end{aligned}$$

Let

$$p_1 = p_1^* = \frac{a_{22} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}.$$

Then

$$E_A(\mathbf{p}^*, \mathbf{q}) = \frac{(a_{12} - a_{22})(a_{22} - a_{21})}{a_{11} - a_{12} - a_{21} + a_{22}} + a_{22} = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}},$$

which is independent of $\mathbf{q}$. Similarly, if

$$q_1 = q_1^* = \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}},$$

then

$$E_A(\mathbf{p}, \mathbf{q}^*) = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}},$$

which is independent of $\mathbf{p}$. Hence

$$E_A(\mathbf{p}^*, \mathbf{q}) = E_A(\mathbf{p}^*, \mathbf{q}^*) = E_A(\mathbf{p}, \mathbf{q}^*) \quad \text{for all strategies } \mathbf{p} \text{ and } \mathbf{q}.$$

Note that

$$\mathbf{p}^* = \begin{pmatrix} \dfrac{a_{22} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}} & \dfrac{a_{11} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}} \end{pmatrix} \tag{9}$$

and

$$\mathbf{q}^* = \begin{pmatrix} \dfrac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}} \\ \dfrac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}} \end{pmatrix}, \tag{10}$$

with value

$$E_A(\mathbf{p}^*, \mathbf{q}^*) = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}.$$

Problems for Chapter 2

1. Consider the four matrices

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 4 \\ 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 7 & 2 & 9 \\ 9 & 2 & 7 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & 3 \\ 1 & 1 & 5 \\ 3 & 2 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 7 \\ 2 & 1 & 2 \\ 1 & 3 & 0 \end{pmatrix}.$$

Calculate all possible products.

2. In each of the following cases, determine whether the products AB and BA are both defined; if so, determine also whether AB and BA have the same number of rows and the same number of columns; if so, determine also whether $AB = BA$:

a) $A = \begin{pmatrix} 0 & 3 \\ 4 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$

b) $A = \begin{pmatrix} 1 & -1 & 5 \\ 3 & 0 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 \\ 3 & 6 \\ 1 & 5 \end{pmatrix}$

c) $A = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -4 \\ 12 & 1 \end{pmatrix}$

d) $A = \begin{pmatrix} 3 & 1 & -4 \\ -2 & 0 & 5 \\ 1 & -2 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

3. Evaluate A^2, where $A = \begin{pmatrix} 2 & -5 \\ 3 & 1 \end{pmatrix}$, and find $\alpha, \beta, \gamma \in \mathbb{R}$, not all zero, such that the matrix $\alpha I + \beta A + \gamma A^2$ is the zero matrix.

4. a) Let $A = \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix}$. Show that A^2 is the zero matrix.

b) Find all 2×2 matrices $B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ such that B^2 is the zero matrix.

5. Prove that if A and B are matrices such that $I - AB$ is invertible, then the inverse of $I - BA$ is given by the formula $(I - BA)^{-1} = I + B(I - AB)^{-1}A$.
[Hint: Write $C = (I - AB)^{-1}$. Then show that $(I - BA)(I + BCA) = I$.]

6. For each of the matrices below, use elementary row operations to find its inverse, if the inverse exists:

a) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 & -2 \\ 1 & 5 & 3 \\ 2 & 6 & -1 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix}$

d) $\begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \\ 2 & 3 & 3 \end{pmatrix}$

e) $\begin{pmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{pmatrix}$

7. a) Using elementary row operations, show that the inverse of

$$\begin{pmatrix} 2 & 5 & 8 & 5 \\ 1 & 2 & 3 & 1 \\ 2 & 4 & 7 & 2 \\ 1 & 3 & 5 & 3 \end{pmatrix} \qquad \text{is} \qquad \begin{pmatrix} 3 & -2 & 1 & -5 \\ -2 & 5 & -2 & 3 \\ 0 & -2 & 1 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix}.$$

b) Without performing any further elementary row operations, use part (a) to solve the system of linear equations

$$\begin{aligned} 2x_1 + 5x_2 + 8x_3 + 5x_4 &= 0, \\ x_1 + 2x_2 + 3x_3 + x_4 &= 1, \\ 2x_1 + 4x_2 + 7x_3 + 2x_4 &= 0, \\ x_1 + 3x_2 + 5x_3 + 3x_4 &= 1. \end{aligned}$$

8. Consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 3 & 1 \\ 1 & 1 & 5 & 5 \\ 2 & 1 & 9 & 8 \\ 2 & 0 & 6 & 3 \end{pmatrix}.$$

a) Use elementary row operations to find the inverse of A.
b) Without performing any further elementary row operations, use your solution in part (a) to solve the system of linear equations

$$\begin{aligned} x_1 \phantom{{}+x_2} + 3x_3 + x_4 &= 1, \\ x_1 + x_2 + 5x_3 + 5x_4 &= 0, \\ 2x_1 + x_2 + 9x_3 + 8x_4 &= 0, \\ 2x_1 \phantom{{}+x_2} + 6x_3 + 3x_4 &= 0. \end{aligned}$$

9. In each of the following, solve the production equation $\mathbf{x} = C\mathbf{x} + \mathbf{d}$:

a) $C = \begin{pmatrix} 0.1 & 0.5 \\ 0.6 & 0.2 \end{pmatrix}$ and $\mathbf{d} = \begin{pmatrix} 50000 \\ 30000 \end{pmatrix}$

b) $C = \begin{pmatrix} 0 & 0.6 \\ 0.5 & 0.2 \end{pmatrix}$ and $\mathbf{d} = \begin{pmatrix} 36000 \\ 22000 \end{pmatrix}$

c) $C = \begin{pmatrix} 0.2 & 0.2 & 0 \\ 0.1 & 0 & 0.2 \\ 0.3 & 0.1 & 0.3 \end{pmatrix}$ and $\mathbf{d} = \begin{pmatrix} 4000000 \\ 8000000 \\ 6000000 \end{pmatrix}$

10. Consider three industries A, B and C. For industry A to manufacture \$1 worth of its product, it needs to purchase 25c worth of product from each of industries B and C. For industry B to manufacture \$1 worth of its product, it needs to purchase 65c worth of product from industry A and 5c worth of product from industry C, as well as use 5c worth of its own product. For industry C to manufacture \$1 worth of its product, it needs to purchase 55c worth of product from industry A and 10c worth of product from industry B. In a particular week, industry A receives \$500000 worth of outside order, industry B receives \$250000 worth of outside order, but industry C receives no outside order. What is the production level required to satisfy all the demands precisely?

11. Suppose that C is an $n \times n$ consumption matrix with all column sums less than 1. Suppose further that $\mathbf{x}'$ is the production vector that satisfies an outside demand $\mathbf{d}'$, and that $\mathbf{x}''$ is the production vector that satisfies an outside demand $\mathbf{d}''$. Show that $\mathbf{x}' + \mathbf{x}''$ is the production vector that satisfies an outside demand $\mathbf{d}' + \mathbf{d}''$.

12. Suppose that C is an $n \times n$ consumption matrix with all column sums less than 1. Suppose further that the demand vector $\mathbf{d}$ has 1 for its top entry and 0 for all other entries. Describe the production vector $\mathbf{x}$ in terms of the columns of the matrix $(I - C)^{-1}$, and give an interpretation of your observation.

13. Consider a pentagon in $\mathbb{R}^2$ with vertices $(1,1)$, $(3,1)$, $(4,2)$, $(2,4)$ and $(1,3)$. For each of the following transformations on the plane, find the 3×3 matrix that describes the transformation with respect to homogeneous coordinates, and use it to find the image of the pentagon:
a) reflection across the x_2-axis
b) reflection across the line $x_1 = x_2$
c) anticlockwise rotation by $90°$
d) translation by the fixed vector $(3, -2)$
e) shear in the x_2-direction with factor 2
f) dilation by factor 2
g) expansion in x_1-direction by factor 2
h) reflection across the x_2-axis, followed by anticlockwise rotation by $90°$
i) translation by the fixed vector $(3, -2)$, followed by reflection across the line $x_1 = x_2$
j) shear in the x_2-direction with factor 2, followed by dilation by factor 2, followed by expansion in x_1-direction by factor 2

14. In homogeneous coordinates, a 3×3 matrix that describes a transformation on the plane is of the form

$$A^* = \begin{pmatrix} a_{11} & a_{12} & h_1 \\ a_{21} & a_{22} & h_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that this transformation can be described by a matrix transformation on $\mathbb{R}^2$ followed by a translation in $\mathbb{R}^2$.

15. Consider the matrices

$$A_1^* = \begin{pmatrix} 1 & 0 & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad A_2^* = \begin{pmatrix} \sec\phi & -\tan\phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $\phi \in (0, \frac{1}{2}\pi)$ is fixed.
a) Show that A_1^* represents a shear in the x_2-direction followed by a compression in the x_2-direction.
b) Show that A_2^* represents a shear in the x_1-direction followed by an expansion in the x_1-direction.
c) What transformation on the plane does the matrix $A_2^* A_1^*$ describe?
[Remark: This technique is often used in computer graphics to speed up calculations.]

16. Consider the matrices

$$A_1^* = \begin{pmatrix} 1 & -\tan\theta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad A_2^* = \begin{pmatrix} 1 & 0 & 0 \\ \sin 2\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $\theta \in \mathbb{R}$ is fixed.
a) What transformation on the plane does the matrix A_1^* describe?
b) What transformation on the plane does the matrix A_2^* describe?
c) What transformation on the plane does the matrix $A_1^* A_2^* A_1^*$ describe?
[Remark: This technique is often used to reduce the number of multiplication operations.]

17. Show that the products and inverses of 3×3 unit lower triangular matrices are also unit lower triangular.

18. For each of the following matrices A and $\mathbf{b}$, find an LU factorization of the matrix A and use it to solve the system $A\mathbf{x} = \mathbf{b}$:

a) $A = \begin{pmatrix} 2 & 1 & 2 \\ 4 & 6 & 5 \\ 4 & 6 & 8 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 6 \\ 21 \\ 24 \end{pmatrix}$

b) $A = \begin{pmatrix} 3 & 1 & 3 \\ 9 & 4 & 10 \\ 6 & -1 & 5 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 5 \\ 18 \\ 9 \end{pmatrix}$

c) $A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 4 & 3 & 5 & 4 \\ 4 & 3 & 5 & 7 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 9 \\ 18 \end{pmatrix}$

d) $A = \begin{pmatrix} 3 & 1 & 1 & 5 \\ 9 & 3 & 4 & 19 \\ 6 & 2 & -1 & 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 10 \\ 35 \\ 7 \end{pmatrix}$

e) $A = \begin{pmatrix} 2 & -3 & 1 & 2 \\ -6 & 10 & -5 & -4 \\ 4 & -7 & 6 & -1 \\ 4 & -2 & -10 & 19 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ -6 \\ 28 \end{pmatrix}$

f) $A = \begin{pmatrix} 2 & -2 & 1 & 2 & 2 \\ 4 & -3 & 0 & 7 & 5 \\ -4 & 7 & -5 & 3 & 2 \\ 6 & -8 & 19 & -8 & 18 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 4 \\ 12 \\ 14 \\ 48 \end{pmatrix}$

19. Consider a payoff matrix

$$A = \begin{pmatrix} 4 & -1 & -6 & 4 \\ -6 & 2 & 0 & 8 \\ -3 & -8 & 7 & -5 \end{pmatrix}.$$

a) What is the expected payoff if $\mathbf{p} = (\,1/3 \quad 0 \quad 2/3\,)$ and $\mathbf{q} = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$?

b) Suppose that player R adopts the strategy $\mathbf{p} = (\,1/3 \quad 0 \quad 2/3\,)$. What strategy should player C adopt?

c) Suppose that player C adopts the strategy $\mathbf{q} = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$. What strategy should player R adopt?

20. Construct a simple example show that optimal strategies are not unique.

21. Show that the entries in the matrices in (9) and (10) are in the range $[0, 1]$.