Chapter 2

MATRICES

2.1 Matrix arithmetic

A matrix over a field F is a rectangular array of elements from F. The symbol $M_{m\times n}(F)$ denotes the collection of all $m \times n$ matrices over F. Matrices will usually be denoted by capital letters and the equation $A = [a_{ij}]$ means that the element in the i–th row and j–th column of the matrix A equals a_{ij}. It is also occasionally convenient to write $a_{ij} = (A)_{ij}$. For the present, all matrices will have rational entries, unless otherwise stated.

EXAMPLE 2.1.1 The formula $a_{ij} = 1/(i + j)$ for $1 \le i \le 3$, $1 \le j \le 4$ defines a 3×4 matrix $A = [a_{ij}]$, namely

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}.$$

DEFINITION 2.1.1 (Equality of matrices) Matrices A and B are said to be equal if A and B have the same size and corresponding elements are equal; that is A and $B \in M_{m\times n}(F)$ and $A = [a_{ij}]$, $B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \le i \le m$, $1 \le j \le n$.

DEFINITION 2.1.2 (Addition of matrices) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be of the same size. Then $A + B$ is the matrix obtained by adding corresponding elements of A and B; that is

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

DEFINITION 2.1.3 (Scalar multiple of a matrix) Let $A = [a_{ij}]$ and $t \in F$ (that is t is a *scalar*). Then tA is the matrix obtained by multiplying all elements of A by t; that is

$$tA = t[a_{ij}] = [ta_{ij}].$$

DEFINITION 2.1.4 (Additive inverse of a matrix) Let $A = [a_{ij}]$. Then $-A$ is the matrix obtained by replacing the elements of A by their additive inverses; that is

$$-A = -[a_{ij}] = [-a_{ij}].$$

DEFINITION 2.1.5 (Subtraction of matrices) Matrix subtraction is defined for two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ of the same size, in the usual way; that is

$$A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}].$$

DEFINITION 2.1.6 (The zero matrix) For each m, n the matrix in $M_{m\times n}(F)$, all of whose elements are zero, is called the *zero* matrix (of size $m \times n$) and is denoted by the symbol 0.

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, s and t will be arbitrary scalars and A, B, C are matrices of the same size.)

1. $(A + B) + C = A + (B + C)$;
2. $A + B = B + A$;
3. $0 + A = A$;
4. $A + (-A) = 0$;
5. $(s + t)A = sA + tA, \quad (s - t)A = sA - tA$;
6. $t(A + B) = tA + tB, \quad t(A - B) = tA - tB$;
7. $s(tA) = (st)A$;
8. $1A = A, \quad 0A = 0, \quad (-1)A = -A$;
9. $tA = 0 \Rightarrow t = 0$ or $A = 0$.

Other similar properties will be used when needed.

DEFINITION 2.1.7 (Matrix product) Let $A = [a_{ij}]$ be a matrix of size $m \times n$ and $B = [b_{jk}]$ be a matrix of size $n \times p$; (that is the number of columns of A equals the number of rows of B). Then AB is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i,\, k)$–th element is defined by the formula

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}.$$

EXAMPLE 2.1.2

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1\times 5 + 2\times 7 & 1\times 6 + 2\times 8 \\ 3\times 5 + 4\times 7 & 3\times 6 + 4\times 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$;

2. $\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$;

3. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$;

4. $\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \end{bmatrix}$;

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

Matrix multiplication obeys many of the familiar laws of arithmetic apart from the commutative law.

1. $(AB)C = A(BC)$ if A, B, C are $m \times n$, $n \times p$, $p \times q$, respectively;

2. $t(AB) = (tA)B = A(tB)$, $\quad A(-B) = (-A)B = -(AB)$;

3. $(A + B)C = AC + BC$ if A and B are $m \times n$ and C is $n \times p$;

4. $D(A + B) = DA + DB$ if A and B are $m \times n$ and D is $p \times m$.

We prove the associative law only:
First observe that $(AB)C$ and $A(BC)$ are both of size $m \times q$.
Let $A = [a_{ij}]$, $B = [b_{jk}]$, $C = [c_{kl}]$. Then

$$\begin{aligned} ((AB)C)_{il} &= \sum_{k=1}^{p}(AB)_{ik}c_{kl} = \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl} \\ &= \sum_{k=1}^{p}\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}. \end{aligned}$$

Similarly

$$(A(BC))_{il} = \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ij} b_{jk} c_{kl}.$$

However the double summations are equal. For sums of the form

$$\sum_{j=1}^{n} \sum_{k=1}^{p} d_{jk} \quad \text{and} \quad \sum_{k=1}^{p} \sum_{j=1}^{n} d_{jk}$$

represent the sum of the np elements of the rectangular array $[d_{jk}]$, by rows and by columns, respectively. Consequently

$$((AB)C)_{il} = (A(BC))_{il}$$

for $1 \le i \le m$, $1 \le l \le q$. Hence $(AB)C = A(BC)$.

The system of m linear equations in n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is equivalent to a single matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

that is $AX = B$, where $A = [a_{ij}]$ is the *coefficient matrix* of the system, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the *vector of unknowns* and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ is the *vector of constants.*

Another useful matrix equation equivalent to the above system of linear equations is

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

EXAMPLE 2.1.3 The system

$$\begin{aligned} x+y+z &= 1 \\ x-y+z &= 0. \end{aligned}$$

is equivalent to the matrix equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and to the equation

$$x \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

2.2 Linear transformations

An *n–dimensional column vector* is an $n \times 1$ matrix over F. The collection of all n–dimensional column vectors is denoted by F^n.

Every matrix is associated with an important type of function called a *linear transformation.*

DEFINITION 2.2.1 (Linear transformation) With $A \in M_{m\times n}(F)$, we associate the function $T_A : F^n \to F^m$ defined by $T_A(X) = AX$ for all $X \in F^n$. More explicitly, using components, the above function takes the form

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n, \end{aligned}$$

where $y_1, y_2, \cdots, y_m$ are the components of the column vector $T_A(X)$.

The function just defined has the property that

$$T_A(sX + tY) = sT_A(X) + tT_A(Y) \tag{2.1}$$

for all $s, t \in F$ and all n–dimensional column vectors X, Y. For

$$T_A(sX + tY) = A(sX + tY) = s(AX) + t(AY) = sT_A(X) + tT_A(Y).$$

REMARK 2.2.1 It is easy to prove that if $T : F^n \to F^m$ is a function satisfying equation 2.1, then $T = T_A$, where A is the $m \times n$ matrix whose columns are $T(E_1), \ldots, T(E_n)$, respectively, where $E_1, \ldots, E_n$ are the n–dimensional *unit vectors* defined by

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots \quad , E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

One well–known example of a linear transformation arises from rotating the (x, y)–plane in 2-dimensional Euclidean space, anticlockwise through θ radians. Here a point (x, y) will be transformed into the point (x_1, y_1), where

$$\begin{aligned} x_1 &= x\cos\theta - y\sin\theta \\ y_1 &= x\sin\theta + y\cos\theta. \end{aligned}$$

In 3–dimensional Euclidean space, the equations

$$x_1 = x\cos\theta - y\sin\theta,\ y_1 = x\sin\theta + y\cos\theta,\ z_1 = z;$$
$$x_1 = x,\ y_1 = y\cos\phi - z\sin\phi,\ z_1 = y\sin\phi + z\cos\phi;$$
$$x_1 = x\cos\psi - z\sin\psi,\ y_1 = y,\ z_1 = x\sin\psi + z\cos\psi;$$

correspond to rotations about the positive z, x, y–axes, anticlockwise through θ, ϕ, ψ radians, respectively.

The product of two matrices is related to the product of the corresponding linear transformations:

If A is $m\times n$ and B is $n\times p$, then the function $T_AT_B : F^p \to F^m$, obtained by first performing T_B, then T_A is in fact equal to the linear transformation T_{AB}. For if $X \in F^p$, we have

$$T_AT_B(X) = A(BX) = (AB)X = T_{AB}(X).$$

The following example is useful for producing rotations in 3–dimensional animated design. (See [27, pages 97–112].)

EXAMPLE 2.2.1 The linear transformation resulting from successively rotating 3–dimensional space about the positive z, x, y–axes, anticlockwise through θ, ϕ, ψ radians respectively, is equal to T_{ABC}, where

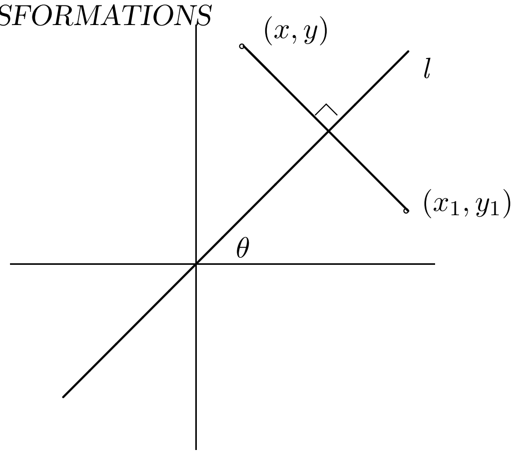

Figure 2.1: Reflection in a line.

$$C = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}.$$

$$A = \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix}.$$

The matrix ABC is quite complicated:

$$A(BC) = \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \cos\phi\sin\theta & \cos\phi\cos\theta & -\sin\phi \\ \sin\phi\sin\theta & \sin\phi\cos\theta & \cos\phi \end{bmatrix}$$

$$= \begin{bmatrix} \cos\psi\cos\theta - \sin\psi\sin\phi\sin\theta & -\cos\psi\sin\theta - \sin\psi\sin\phi\sin\theta & -\sin\psi\cos\phi \\ \cos\phi\sin\theta & \cos\phi\cos\theta & -\sin\phi \\ \sin\psi\cos\theta + \cos\psi\sin\phi\sin\theta & -\sin\psi\sin\theta + \cos\psi\sin\phi\cos\theta & \cos\psi\cos\phi \end{bmatrix}.$$

EXAMPLE 2.2.2 Another example of a linear transformation arising from geometry is reflection of the plane in a line l inclined at an angle θ to the positive x–axis.

We reduce the problem to the simpler case $\theta = 0$, where the equations of transformation are $x_1 = x$, $y_1 = -y$. First rotate the plane clockwise through θ radians, thereby taking l into the x–axis; next reflect the plane in the x–axis; then rotate the plane anticlockwise through θ radians, thereby restoring l to its original position.

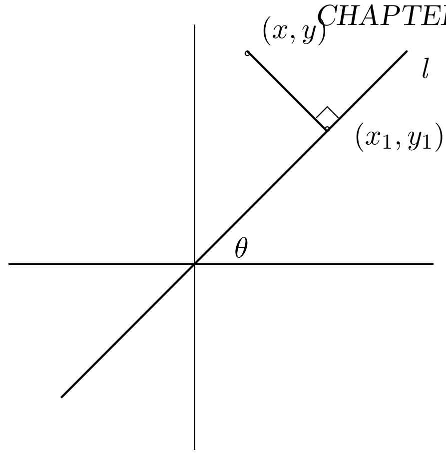

Figure 2.2: Projection on a line.

In terms of matrices, we get transformation equations

$$\begin{aligned}
\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.
\end{aligned}$$

The more general transformation

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = a \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix}, \quad a > 0,$$

represents a rotation, followed by a scaling and then by a translation. Such transformations are important in computer graphics. See [23, 24].

EXAMPLE 2.2.3 Our last example of a geometrical linear transformation arises from projecting the plane onto a line l through the origin, inclined at angle θ to the positive x–axis. Again we reduce that problem to the simpler case where l is the x–axis and the equations of transformation are $x_1 = x$, $y_1 = 0$.

In terms of matrices, we get transformation equations

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

2.3 Recurrence relations

DEFINITION 2.3.1 (The identity matrix) The $n \times n$ matrix $I_n = [\delta_{ij}]$, defined by $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ if $i \neq j$, is called the $n \times n$ *identity* matrix of order n. In other words, the columns of the identity matrix of order n are the unit vectors $E_1, \cdots, E_n$, respectively.

For example, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

THEOREM 2.3.1 If A is $m \times n$, then $I_m A = A = AI_n$.

DEFINITION 2.3.2 (k–th power of a matrix) If A is an $n \times n$ matrix, we define A^k recursively as follows: $A^0 = I_n$ and $A^{k+1} = A^k A$ for $k \geq 0$.

For example $A^1 = A^0 A = I_n A = A$ and hence $A^2 = A^1 A = AA$.

The usual index laws hold provided $AB = BA$:

1. $A^m A^n = A^{m+n}$, $(A^m)^n = A^{mn}$;

2. $(AB)^n = A^n B^n$;

3. $A^m B^n = B^n A^m$;

4. $(A + B)^2 = A^2 + 2AB + B^2$;

5. $(A + B)^n = \sum_{i=0}^{n} \binom{n}{i} A^i B^{n-i}$;

6. $(A + B)(A - B) = A^2 - B^2$.

We now state a basic property of the natural numbers.

AXIOM 2.3.1 (PRINCIPLE OF MATHEMATICAL INDUCTION) *If for each $n \geq 1$, $\mathcal{P}_n$ denotes a mathematical statement and*

(i) $\mathcal{P}_1$ *is true,*

(ii) *the truth of* $\mathcal{P}_n$ *implies that of* $\mathcal{P}_{n+1}$ *for each* $n \geq 1$,

then $\mathcal{P}_n$ *is true for all* $n \geq 1$.

EXAMPLE 2.3.1 Let $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$. Prove that

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix} \quad \text{if } n \geq 1.$$

Solution. We use the principle of mathematical induction.

Take $\mathcal{P}_n$ to be the statement

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix}.$$

Then $\mathcal{P}_1$ asserts that

$$A^1 = \begin{bmatrix} 1+6\times 1 & 4\times 1 \\ -9\times 1 & 1-6\times 1 \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix},$$

which is true. Now let $n \geq 1$ and assume that $\mathcal{P}_n$ is true. We have to deduce that

$$A^{n+1} = \begin{bmatrix} 1+6(n+1) & 4(n+1) \\ -9(n+1) & 1-6(n+1) \end{bmatrix} = \begin{bmatrix} 7+6n & 4n+4 \\ -9n-9 & -5-6n \end{bmatrix}.$$

Now

$$\begin{aligned} A^{n+1} &= A^n A \\ &= \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix} \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} \\ &= \begin{bmatrix} (1+6n)7+(4n)(-9) & (1+6n)4+(4n)(-5) \\ (-9n)7+(1-6n)(-9) & (-9n)4+(1-6n)(-5) \end{bmatrix} \\ &= \begin{bmatrix} 7+6n & 4n+4 \\ -9n-9 & -5-6n \end{bmatrix}, \end{aligned}$$

and "the induction goes through".

The last example has an application to the solution of a system of *recurrence relations*:

EXAMPLE 2.3.2 The following system of recurrence relations holds for all $n \geq 0$:

$$\begin{aligned} x_{n+1} &= 7x_n + 4y_n \\ y_{n+1} &= -9x_n - 5y_n. \end{aligned}$$

Solve the system for x_n and y_n in terms of x_0 and y_0.

Solution. Combine the above equations into a single matrix equation

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix},$$

or $X_{n+1} = AX_n$, where $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$ and $X_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$.
We see that

$$\begin{aligned} X_1 &= AX_0 \\ X_2 &= AX_1 = A(AX_0) = A^2X_0 \\ &\vdots \\ X_n &= A^nX_0. \end{aligned}$$

(The truth of the equation $X_n = A^nX_0$ for $n \geq 1$, strictly speaking follows by mathematical induction; however for simple cases such as the above, it is customary to omit the strict proof and supply instead a few lines of motivation for the inductive statement.)

Hence the previous example gives

$$\begin{aligned} \begin{bmatrix} x_n \\ y_n \end{bmatrix} = X_n &= \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\ &= \begin{bmatrix} (1+6n)x_0 + (4n)y_0 \\ (-9n)x_0 + (1-6n)y_0 \end{bmatrix}, \end{aligned}$$

and hence $x_n = (1+6n)x_0 + 4ny_0$ and $y_n = (-9n)x_0 + (1-6n)y_0$, for $n \geq 1$.

2.4 PROBLEMS

1. Let A, B, C, D be matrices defined by

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 1 & 0 \\ -4 & 1 & 3 \end{bmatrix},$$

$$C = \begin{bmatrix} -3 & -1 \\ 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & -1 \\ 2 & 0 \end{bmatrix}.$$

Which of the following matrices are defined? Compute those matrices which are defined.

$$A+B,\ A+C,\ AB,\ BA,\ CD,\ DC,\ D^2.$$

[Answers: $A+C$, BA, CD, D^2;

$$\begin{bmatrix} 0 & -1 \\ 1 & 3 \\ 5 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 12 \\ -4 & 2 \\ -10 & 5 \end{bmatrix}, \quad \begin{bmatrix} -14 & 3 \\ 10 & -2 \\ 22 & -4 \end{bmatrix}, \quad \begin{bmatrix} 14 & -4 \\ 8 & -2 \end{bmatrix}.]$$

2. Let $A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Show that if B is a 3×2 such that $AB = I_2$, then

$$B = \begin{bmatrix} a & b \\ -a-1 & 1-b \\ a+1 & b \end{bmatrix}$$

for suitable numbers a and b. Use the associative law to show that $(BA)^2 B = B$.

3. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, prove that $A^2 - (a+d)A + (ad-bc)I_2 = 0$.

4. If $A = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$, use the fact $A^2 = 4A - 3I_2$ and mathematical induction, to prove that

$$A^n = \frac{(3^n-1)}{2}A + \frac{3-3^n}{2}I_2 \quad \text{if } n \geq 1.$$

5. A sequence of numbers $x_1, x_2, \ldots, x_n, \ldots$ satisfies the recurrence relation $x_{n+1} = ax_n + bx_{n-1}$ for $n \geq 1$, where a and b are constants. Prove that

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = A \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix},$$

where $A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$ and hence express $\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix}$ in terms of $\begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$.
If $a = 4$ and $b = -3$, use the previous question to find a formula for x_n in terms of x_1 and x_0.

[Answer:

$$x_n = \frac{3^n - 1}{2} x_1 + \frac{3 - 3^n}{2} x_0.]$$

6. Let $A = \begin{bmatrix} 2a & -a^2 \\ 1 & 0 \end{bmatrix}$.

   (a) Prove that

   $$A^n = \begin{bmatrix} (n+1)a^n & -na^{n+1} \\ na^{n-1} & (1-n)a^n \end{bmatrix} \quad \text{if } n \geq 1.$$

   (b) A sequence $x_0, x_1, \ldots, x_n, \ldots$ satisfies the recurrence relation $x_{n+1} = 2ax_n - a^2 x_{n-1}$ for $n \geq 1$. Use part (a) and the previous question to prove that $x_n = na^{n-1}x_1 + (1-n)a^n x_0$ for $n \geq 1$.

7. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and suppose that λ_1 and λ_2 are the roots of the quadratic polynomial $x^2 - (a+d)x + ad - bc$. (λ_1 and λ_2 may be equal.) Let k_n be defined by $k_0 = 0$, $k_1 = 1$ and for $n \geq 2$

   $$k_n = \sum_{i=1}^{n} \lambda_1^{n-i} \lambda_2^{i-1}.$$

   Prove that

   $$k_{n+1} = (\lambda_1 + \lambda_2)k_n - \lambda_1 \lambda_2 k_{n-1},$$

   if $n \geq 1$. Also prove that

   $$k_n = \begin{cases} (\lambda_1^n - \lambda_2^n)/(\lambda_1 - \lambda_2) & \text{if } \lambda_1 \neq \lambda_2, \\ n\lambda_1^{n-1} & \text{if } \lambda_1 = \lambda_2. \end{cases}$$

   Use mathematical induction to prove that if $n \geq 1$,

   $$A^n = k_n A - \lambda_1 \lambda_2 k_{n-1} I_2,$$

   [Hint: Use the equation $A^2 = (a+d)A - (ad - bc)I_2$.]

8. Use Question 7 to prove that if $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, then

$$A^n = \frac{3^n}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{(-1)^{n-1}}{2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

if $n \geq 1$.

9. The Fibonacci numbers are defined by the equations $F_0 = 0$, $F_1 = 1$ and $F_{n+1} = F_n + F_{n-1}$ if $n \geq 1$. Prove that

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right)$$

if $n \geq 0$.

10. Let $r > 1$ be an integer. Let a and b be arbitrary positive integers. Sequences x_n and y_n of positive integers are defined in terms of a and b by the recurrence relations

$$\begin{aligned} x_{n+1} &= x_n + r y_n \\ y_{n+1} &= x_n + y_n, \end{aligned}$$

for $n \geq 0$, where $x_0 = a$ and $y_0 = b$.

Use Question 7 to prove that

$$\frac{x_n}{y_n} \to \sqrt{r} \quad \text{as } n \to \infty.$$

2.5 Non–singular matrices

DEFINITION 2.5.1 (Non–singular matrix)

A square matrix $A \in M_{n\times n}(F)$ is called *non–singular* or *invertible* if there exists a matrix $B \in M_{n\times n}(F)$ such that

$$AB = I_n = BA.$$

Any matrix B with the above property is called an *inverse* of A. If A does not have an inverse, A is called *singular*.

THEOREM 2.5.1 (Inverses are unique)

If A has inverses B and C, then $B = C$.

Proof. Let B and C be inverses of A. Then $AB = I_n = BA$ and $AC = I_n = CA$. Then $B(AC) = BI_n = B$ and $(BA)C = I_nC = C$. Hence because $B(AC) = (BA)C$, we deduce that $B = C$.

REMARK 2.5.1 If A has an inverse, it is denoted by A^{-1}. So

$$AA^{-1} = I_n = A^{-1}A.$$

Also if A is non–singular, it follows that A^{-1} is also non–singular and

$$(A^{-1})^{-1} = A.$$

THEOREM 2.5.2 If A and B are non–singular matrices of the same size, then so is AB. Moreover

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

Proof.

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

Similarly

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

REMARK 2.5.2 The above result generalizes to a product of m non–singular matrices: If $A_1, \ldots, A_m$ are non–singular $n \times n$ matrices, then the product $A_1 \ldots A_m$ is also non–singular. Moreover

$$(A_1 \ldots A_m)^{-1} = A_m^{-1} \ldots A_1^{-1}.$$

(Thus the inverse of the product equals the product of the inverses *in the reverse order*.)

EXAMPLE 2.5.1 If A and B are $n \times n$ matrices satisfying $A^2 = B^2 = (AB)^2 = I_n$, prove that $AB = BA$.

Solution. Assume $A^2 = B^2 = (AB)^2 = I_n$. Then A, B, AB are non–singular and $A^{-1} = A$, $B^{-1} = B$, $(AB)^{-1} = AB$.

But $(AB)^{-1} = B^{-1}A^{-1}$ and hence $AB = BA$.

EXAMPLE 2.5.2 $A = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$ is singular. For suppose $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is an inverse of A. Then the equation $AB = I_2$ gives

$$\begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and equating the corresponding elements of column 1 of both sides gives the system

$$\begin{aligned} a + 2c &= 1 \\ 4a + 8c &= 0 \end{aligned}$$

which is clearly inconsistent.

THEOREM 2.5.3 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\Delta = ad - bc \neq 0$. Then A is non–singular. Also

$$A^{-1} = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

REMARK 2.5.3 The expression $ad - bc$ is called the *determinant* of A and is denoted by the symbols $\det A$ or $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Proof. Verify that the matrix $B = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ satisfies the equation $AB = I_2 = BA$.

EXAMPLE 2.5.3 Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & 0 & 0 \end{bmatrix}.$$

Verify that $A^3 = 5I_3$, deduce that A is non–singular and find A^{-1}.

Solution. After verifying that $A^3 = 5I_3$, we notice that

$$A\left(\frac{1}{5}A^2\right) = I_3 = \left(\frac{1}{5}A^2\right)A.$$

Hence A is non–singular and $A^{-1} = \frac{1}{5}A^2$.

THEOREM 2.5.4 If the coefficient matrix A of a system of n equations in n unknowns is non–singular, then the system $AX = B$ has the unique solution $X = A^{-1}B$.

Proof. Assume that A^{-1} exists.

1. (Uniqueness.) Assume that $AX = B$. Then

$$\begin{aligned}(A^{-1}A)X &= A^{-1}B,\\ I_nX &= A^{-1}B,\\ X &= A^{-1}B.\end{aligned}$$

2. (Existence.) Let $X = A^{-1}B$. Then

$$AX = A(A^{-1}B) = (AA^{-1})B = I_nB = B.$$

THEOREM 2.5.5 (Cramer's rule for 2 equations in 2 unknowns)

The system

$$\begin{aligned}ax + by &= e\\ cx + dy &= f\end{aligned}$$

has a unique solution if $\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, namely

$$x = \frac{\Delta_1}{\Delta}, \qquad y = \frac{\Delta_2}{\Delta},$$

where

$$\Delta_1 = \begin{vmatrix} e & b \\ f & d \end{vmatrix} \quad \text{and} \quad \Delta_2 = \begin{vmatrix} a & e \\ c & f \end{vmatrix}.$$

Proof. Suppose $\Delta \neq 0$. Then $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has inverse

$$A^{-1} = \Delta^{-1}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

and we know that the system

$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix}$$

has the unique solution

$$\begin{aligned}\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} e \\ f \end{bmatrix} &= \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} e \\ f \end{bmatrix} \\ &= \frac{1}{\Delta}\begin{bmatrix} de-bf \\ -ce+af \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} = \begin{bmatrix} \Delta_1/\Delta \\ \Delta_2/\Delta \end{bmatrix}.\end{aligned}$$

Hence $x = \Delta_1/\Delta$, $y = \Delta_2/\Delta$.

COROLLARY 2.5.1 The homogeneous system

$$\begin{aligned} ax+by &= 0 \\ cx+dy &= 0 \end{aligned}$$

has only the trivial solution if $\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$.

EXAMPLE 2.5.4 The system

$$\begin{aligned} 7x+8y &= 100 \\ 2x-9y &= 10 \end{aligned}$$

has the unique solution $x = \Delta_1/\Delta$, $y = \Delta_2/\Delta$, where

$$\Delta = \begin{vmatrix} 7 & 8 \\ 2 & -9 \end{vmatrix} = -79,\ \Delta_1 = \begin{vmatrix} 100 & 8 \\ 10 & -9 \end{vmatrix} = -980,\ \Delta_2 = \begin{vmatrix} 7 & 100 \\ 2 & 10 \end{vmatrix} = -130.$$

So $x = \frac{980}{79}$ and $y = \frac{130}{79}$.

THEOREM 2.5.6 Let A be a square matrix. If A is non–singular, the homogeneous system $AX = 0$ has only the trivial solution. Equivalently, if the homogenous system $AX = 0$ has a non–trivial solution, then A is singular.

Proof. If A is non–singular and $AX = 0$, then $X = A^{-1}0 = 0$.

REMARK 2.5.4 If $A_{*1}, \ldots, A_{*n}$ denote the columns of A, then the equation

$$AX = x_1A_{*1} + \ldots + x_nA_{*n}$$

holds. Consequently theorem 2.5.6 tells us that if there exist scalars $x_1, \ldots, x_n$, *not all zero*, such that

$$x_1A_{*1} + \ldots + x_nA_{*n} = 0,$$

that is, if the columns of A are *linearly dependent*, then A is singular. An equivalent way of saying that the columns of A are linearly dependent is that one of the columns of A is expressible as a sum of certain scalar multiples of the remaining columns of A; that is one column is a *linear combination* of the remaining columns.

EXAMPLE 2.5.5

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

is singular. For it can be verified that A has reduced row–echelon form

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and consequently $AX = 0$ has a non–trivial solution $x = -1$, $y = -1$, $z = 1$.

REMARK 2.5.5 More generally, if A is row–equivalent to a matrix containing a zero row, then A is singular. For then the homogeneous system $AX = 0$ has a non–trivial solution.

An important class of non–singular matrices is that of the *elementary row matrices.*

DEFINITION 2.5.2 (Elementary row matrices) There are three types, E_{ij}, $E_i(t)$, $E_{ij}(t)$, corresponding to the three kinds of elementary row operation:

1. E_{ij}, $(i \neq j)$ is obtained from the identity matrix I_n by interchanging rows i and j.
2. $E_i(t)$, $(t \neq 0)$ is obtained by multiplying the i–th row of I_n by t.
3. $E_{ij}(t)$, $(i \neq j)$ is obtained from I_n by adding t times the j–th row of I_n to the i–th row.

EXAMPLE 2.5.6 $(n = 3.)$

$$E_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \; E_2(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; E_{23}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The elementary row matrices have the following distinguishing property:

THEOREM 2.5.7 If a matrix A is pre–multiplied by an elementary row–matrix, the resulting matrix is the one obtained by performing the corresponding elementary row–operation on A.

EXAMPLE 2.5.7

$$E_{23}\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ e & f \\ c & d \end{bmatrix}.$$

COROLLARY 2.5.2 The three types of elementary row–matrices are non–singular. Indeed

1. $E_{ij}^{-1} = E_{ij}$;
2. $E_i^{-1}(t) = E_i(t^{-1})$;
3. $(E_{ij}(t))^{-1} = E_{ij}(-t)$.

Proof. Taking $A = I_n$ in the above theorem, we deduce the following equations:

$$\begin{aligned} E_{ij}E_{ij} &= I_n \\ E_i(t)E_i(t^{-1}) &= I_n = E_i(t^{-1})E_i(t) \quad \text{if } t \neq 0 \\ E_{ij}(t)E_{ij}(-t) &= I_n = E_{ij}(-t)E_{ij}(t). \end{aligned}$$

EXAMPLE 2.5.8 Find the 3×3 matrix $A = E_3(5)E_{23}(2)E_{12}$ explicitly. Also find A^{-1}.

Solution.

$$A = E_3(5)E_{23}(2)\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_3(5)\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

To find A^{-1}, we have

$$\begin{aligned} A^{-1} &= (E_3(5)E_{23}(2)E_{12})^{-1} \\ &= E_{12}^{-1}\,(E_{23}(2))^{-1}\,(E_3(5))^{-1} \\ &= E_{12}E_{23}(-2)E_3(5^{-1}) \end{aligned}$$

$$
\begin{aligned}
&= E_{12}E_{23}(-2)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix} \\
&= E_{12}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -\frac{2}{5} \\ 1 & 0 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}.
\end{aligned}
$$

REMARK 2.5.6 Recall that A and B are row–equivalent if B is obtained from A by a sequence of elementary row operations. If $E_1, \ldots, E_r$ are the respective corresponding elementary row matrices, then

$$B = E_r(\ldots(E_2(E_1A))\ldots) = (E_r \ldots E_1)A = PA,$$

where $P = E_r \ldots E_1$ is non–singular. Conversely if $B = PA$, where P is non–singular, then A is row–equivalent to B. For as we shall now see, P is in fact a product of elementary row matrices.

THEOREM 2.5.8 Let A be non–singular $n \times n$ matrix. Then

(i) A is row–equivalent to I_n,

(ii) A is a product of elementary row matrices.

Proof. Assume that A is non–singular and let B be the reduced row–echelon form of A. Then B has no zero rows, for otherwise the equation $AX = 0$ would have a non–trivial solution. Consequently $B = I_n$.

It follows that there exist elementary row matrices $E_1, \ldots, E_r$ such that $E_r(\ldots(E_1A)\ldots) = B = I_n$ and hence $A = E_1^{-1} \ldots E_r^{-1}$, a product of elementary row matrices.

THEOREM 2.5.9 Let A be $n \times n$ and suppose that A is row–equivalent to I_n. Then A is non–singular and A^{-1} can be found by performing the same sequence of elementary row operations on I_n as were used to convert A to I_n.

Proof. Suppose that $E_r \ldots E_1A = I_n$. In other words $BA = I_n$, where $B = E_r \ldots E_1$ is non–singular. Then $B^{-1}(BA) = B^{-1}I_n$ and so $A = B^{-1}$, which is non–singular.

Also $A^{-1} = \left(B^{-1}\right)^{-1} = B = E_r((\ldots(E_1I_n)\ldots)$, which shows that A^{-1} is obtained from I_n by performing the same sequence of elementary row operations as were used to convert A to I_n.

REMARK 2.5.7 It follows from theorem 2.5.9 that if A is singular, then A is row–equivalent to a matrix whose last row is zero.

EXAMPLE 2.5.9 Show that $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ is non–singular, find A^{-1} and express A as a product of elementary row matrices.

Solution. We form the *partitioned* matrix $[A|I_2]$ which consists of A followed by I_2. Then any sequence of elementary row operations which reduces A to I_2 will reduce I_2 to A^{-1}. Here

$$[A|I_2] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

$$R_2 \to R_2 - R_1 \quad \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right]$$

$$R_2 \to (-1)R_2 \quad \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array}\right]$$

$$R_1 \to R_1 - 2R_2 \quad \left[\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right].$$

Hence A is row–equivalent to I_2 and A is non–singular. Also

$$A^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

We also observe that

$$E_{12}(-2)E_2(-1)E_{21}(-1)A = I_2.$$

Hence

$$\begin{aligned} A^{-1} &= E_{12}(-2)E_2(-1)E_{21}(-1) \\ A &= E_{21}(1)E_2(-1)E_{12}(2). \end{aligned}$$

The next result is the converse of Theorem 2.5.6 and is useful for proving the non–singularity of certain types of matrices.

THEOREM 2.5.10 Let A be an $n \times n$ matrix with the property that the homogeneous system $AX = 0$ has only the trivial solution. Then A is non–singular. Equivalently, if A is singular, then the homogeneous system $AX = 0$ has a non–trivial solution.

Proof. If A is $n \times n$ and the homogeneous system $AX = 0$ has only the trivial solution, then it follows that the reduced row–echelon form B of A cannot have zero rows and must therefore be I_n. Hence A is non–singular.

COROLLARY 2.5.3 Suppose that A and B are $n \times n$ and $AB = I_n$. Then $BA = I_n$.

Proof. Let $AB = I_n$, where A and B are $n \times n$. We first show that B is non–singular. Assume $BX = 0$. Then $A(BX) = A0 = 0$, so $(AB)X = 0$, $I_nX = 0$ and hence $X = 0$.

Then from $AB = I_n$ we deduce $(AB)B^{-1} = I_nB^{-1}$ and hence $A = B^{-1}$. The equation $BB^{-1} = I_n$ then gives $BA = I_n$.

Before we give the next example of the above criterion for non-singularity, we introduce an important matrix operation.

DEFINITION 2.5.3 (The transpose of a matrix) Let A be an $m \times n$ matrix. Then A^t, the *transpose* of A, is the matrix obtained by interchanging the rows and columns of A. In other words if $A = [a_{ij}]$, then $\left(A^t\right)_{ji} = a_{ij}$. Consequently A^t is $n \times m$.

The transpose operation has the following properties:

1. $\left(A^t\right)^t = A$;
2. $(A \pm B)^t = A^t \pm B^t$ if A and B are $m \times n$;
3. $(sA)^t = sA^t$ if s is a scalar;
4. $(AB)^t = B^tA^t$ if A is $m \times n$ and B is $n \times p$;
5. If A is non–singular, then A^t is also non–singular and
$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t;$$
6. $X^tX = x_1^2 + \ldots + x_n^2$ if $X = [x_1, \ldots, x_n]^t$ is a column vector.

We prove only the fourth property. First check that both $(AB)^t$ and B^tA^t have the same size $(p \times m)$. Moreover, corresponding elements of both matrices are equal. For if $A = [a_{ij}]$ and $B = [b_{jk}]$, we have

$$\begin{aligned} \left((AB)^t\right)_{ki} &= (AB)_{ik} \\ &= \sum_{j=1}^{n} a_{ij}b_{jk} \end{aligned}$$

$$\begin{aligned} &= \sum_{j=1}^{n} \left(B^t\right)_{kj} \left(A^t\right)_{ji} \\ &= \left(B^t A^t\right)_{ki}. \end{aligned}$$

There are two important classes of matrices that can be defined concisely in terms of the transpose operation.

DEFINITION 2.5.4 (Symmetric matrix) A real matrix A is called *symmetric* if $A^t = A$. In other words A is square ($n \times n$ say) and $a_{ji} = a_{ij}$ for all $1 \le i \le n$, $1 \le j \le n$. Hence

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

is a general 2×2 symmetric matrix.

DEFINITION 2.5.5 (Skew–symmetric matrix) A real matrix A is called *skew–symmetric* if $A^t = -A$. In other words A is square ($n \times n$ say) and $a_{ji} = -a_{ij}$ for all $1 \le i \le n$, $1 \le j \le n$.

REMARK 2.5.8 Taking $i = j$ in the definition of skew–symmetric matrix gives $a_{ii} = -a_{ii}$ and so $a_{ii} = 0$. Hence

$$A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$$

is a general 2×2 skew–symmetric matrix.

We can now state a second application of the above criterion for non–singularity.

COROLLARY 2.5.4 Let B be an $n \times n$ skew–symmetric matrix. Then $A = I_n - B$ is non–singular.

Proof. Let $A = I_n - B$, where $B^t = -B$. By Theorem 2.5.10 it suffices to show that $AX = 0$ implies $X = 0$.

We have $(I_n - B)X = 0$, so $X = BX$. Hence $X^tX = X^tBX$.
Taking transposes of both sides gives

$$\begin{aligned} (X^tBX)^t &= (X^tX)^t \\ X^tB^t(X^t)^t &= X^t(X^t)^t \\ X^t(-B)X &= X^tX \\ -X^tBX &= X^tX = X^tBX. \end{aligned}$$

Hence $X^tX = -X^tX$ and $X^tX = 0$. But if $X = [x_1, \ldots, x_n]^t$, then $X^tX = x_1^2 + \ldots + x_n^2 = 0$ and hence $x_1 = 0, \ldots, x_n = 0$.

2.6 Least squares solution of equations

Suppose $AX = B$ represents a system of linear equations with real coefficients which may be inconsistent, because of the possibility of experimental errors in determining A or B. For example, the system

$$\begin{aligned} x &= 1 \\ y &= 2 \\ x + y &= 3.001 \end{aligned}$$

is inconsistent.

It can be proved that the associated system $A^tAX = A^tB$ is always consistent and that any solution of this system minimizes the sum $r_1^2 + \ldots + r_m^2$, where $r_1, \ldots, r_m$ (the *residuals*) are defined by

$$r_i = a_{i1}x_1 + \ldots + a_{in}x_n - b_i,$$

for $i = 1, \ldots, m$. The equations represented by $A^tAX = A^tB$ are called the *normal equations* corresponding to the system $AX = B$ and any solution of the system of normal equations is called a *least squares* solution of the original system.

EXAMPLE 2.6.1 Find a least squares solution of the above inconsistent system.

Solution. Here $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 2 \\ 3.001 \end{bmatrix}$.

Then $A^tA = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

Also $A^tB = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3.001 \end{bmatrix} = \begin{bmatrix} 4.001 \\ 5.001 \end{bmatrix}$.

So the normal equations are

$$\begin{aligned} 2x + y &= 4.001 \\ x + 2y &= 5.001 \end{aligned}$$

which have the unique solution

$$x = \frac{3.001}{3}, \quad y = \frac{6.001}{3}.$$

EXAMPLE 2.6.2 Points $(x_1, y_1), \ldots, (x_n, y_n)$ are experimentally determined and should lie on a line $y = mx + c$. Find a least squares solution to the problem.

Solution. The points have to satisfy

$$\begin{aligned} mx_1 + c &= y_1 \\ &\vdots \\ mx_n + c &= y_n, \end{aligned}$$

or $Ax = B$, where

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, X = \begin{bmatrix} m \\ c \end{bmatrix}, B = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

The normal equations are given by $(A^tA)X = A^tB$. Here

$$A^tA = \begin{bmatrix} x_1 & \ldots & x_n \\ 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} x_1^2 + \ldots + x_n^2 & x_1 + \ldots + x_n \\ x_1 + \ldots + x_n & n \end{bmatrix}$$

Also

$$A^tB = \begin{bmatrix} x_1 & \ldots & x_n \\ 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 + \ldots + x_ny_n \\ y_1 + \ldots + y_n \end{bmatrix}.$$

It is not difficult to prove that

$$\Delta = \det(A^tA) = \sum_{1 \le i < j \le n} (x_i - x_j)^2,$$

which is positive unless $x_1 = \ldots = x_n$. Hence if not all of $x_1, \ldots, x_n$ are equal, A^tA is non–singular and the normal equations have a unique solution. This can be shown to be

$$m = \frac{1}{\Delta} \sum_{1 \le i < j \le n} (x_i - x_j)(y_i - y_j), \; c = \frac{1}{\Delta} \sum_{1 \le i < j \le n} (x_iy_j - x_jy_i)(x_i - x_j).$$

REMARK 2.6.1 The matrix A^tA is symmetric.

2.7 PROBLEMS

1. Let $A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix}$. Prove that A is non–singular, find A^{-1} and express A as a product of elementary row matrices.

   [Answer: $A^{-1} = \begin{bmatrix} \frac{1}{13} & -\frac{4}{13} \\ \frac{3}{13} & \frac{1}{13} \end{bmatrix}$,

   $A = E_{21}(-3)E_2(13)E_{12}(4)$ is one such decomposition.]

2. A square matrix $D = [d_{ij}]$ is called *diagonal* if $d_{ij} = 0$ for $i \neq j$. (That is the *off–diagonal* elements are zero.) Prove that pre–multiplication of a matrix A by a diagonal matrix D results in matrix DA whose rows are the rows of A multiplied by the respective diagonal elements of D. State and prove a similar result for post–multiplication by a diagonal matrix.

   Let $\mathrm{diag}\,(a_1, \ldots, a_n)$ denote the diagonal matrix whose *diagonal* elements d_{ii} are $a_1, \ldots, a_n$, respectively. Show that

   $$\mathrm{diag}\,(a_1, \ldots, a_n)\mathrm{diag}\,(b_1, \ldots, b_n) = \mathrm{diag}\,(a_1b_1, \ldots, a_nb_n)$$

   and deduce that if $a_1 \ldots a_n \neq 0$, then $\mathrm{diag}\,(a_1, \ldots, a_n)$ is non–singular and

   $$(\mathrm{diag}\,(a_1, \ldots, a_n))^{-1} = \mathrm{diag}\,(a_1^{-1}, \ldots, a_n^{-1}).$$

   Also prove that $\mathrm{diag}\,(a_1, \ldots, a_n)$ is singular if $a_i = 0$ for some i.

3. Let $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 2 & 6 \\ 3 & 7 & 9 \end{bmatrix}$. Prove that A is non–singular, find A^{-1} and express A as a product of elementary row matrices.

   [Answers: $A^{-1} = \begin{bmatrix} -12 & 7 & -2 \\ \frac{9}{2} & -3 & 1 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}$,

   $A = E_{12}E_{31}(3)E_{23}E_3(2)E_{12}(2)E_{13}(24)E_{23}(-9)$ is one such decomposition.]

4. Find the rational number k for which the matrix $A = \begin{bmatrix} 1 & 2 & k \\ 3 & -1 & 1 \\ 5 & 3 & -5 \end{bmatrix}$ is singular. [Answer: $k = -3$.]

5. Prove that $A = \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$ is singular and find a non–singular matrix P such that PA has last row zero.

6. If $A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix}$, verify that $A^2 - 2A + 13I_2 = 0$ and deduce that $A^{-1} = -\frac{1}{13}(A - 2I_2)$.

7. Let $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$.

   (i) Verify that $A^3 = 3A^2 - 3A + I_3$.

   (ii) Express A^4 in terms of A^2, A and I_3 and hence calculate A^4 explicitly.

   (iii) Use (i) to prove that A is non–singular and find A^{-1} explicitly.

   [Answers: (ii) $A^4 = 6A^2 - 8A + 3I_3 = \begin{bmatrix} -11 & -8 & -4 \\ 12 & 9 & 4 \\ 20 & 16 & 5 \end{bmatrix}$;

   (iii) $A^{-1} = A^2 - 3A + 3I_3 = \begin{bmatrix} -1 & -3 & 1 \\ 2 & 4 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.]

8. (i) Let B be an $n \times n$ matrix such that $B^3 = 0$. If $A = I_n - B$, prove that A is non–singular and $A^{-1} = I_n + B + B^2$.
   Show that the system of linear equations $AX = b$ has the solution
   $$X = b + Bb + B^2b.$$

   (ii) If $B = \begin{bmatrix} 0 & r & s \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix}$, verify that $B^3 = 0$ and use (i) to determine $(I_3 - B)^{-1}$ explicitly.

$$[\text{Answer: } \begin{bmatrix} 1 & r & s+rt \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.]$$

9. Let A be $n \times n$.

   (i) If $A^2 = 0$, prove that A is singular.

   (ii) If $A^2 = A$ and $A \neq I_n$, prove that A is singular.

10. Use Question 7 to solve the system of equations

$$\begin{aligned} x + y - z &= a \\ z &= b \\ 2x + y + 2z &= c \end{aligned}$$

where a, b, c are given rationals. Check your answer using the Gauss–Jordan algorithm.

[Answer: $x = -a - 3b + c$, $y = 2a + 4b - c$, $z = b$.]

11. Determine explicitly the following products of 3×3 elementary row matrices.

(i) $E_{12}E_{23}$ (ii) $E_1(5)E_{12}$ (iii) $E_{12}(3)E_{21}(-3)$ (iv) $(E_1(100))^{-1}$
(v) E_{12}^{-1} (vi) $(E_{12}(7))^{-1}$ (vii) $(E_{12}(7)E_{31}(1))^{-1}$.

$$[\text{Answers: (i) } \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ (ii) } \begin{bmatrix} 0 & 5 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ (iii) } \begin{bmatrix} -8 & 3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{(iv) } \begin{bmatrix} \frac{1}{100} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ (v) } \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ (vi) } \begin{bmatrix} 1 & -7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ (vii) } \begin{bmatrix} 1 & -7 & 0 \\ 0 & 1 & 0 \\ -1 & 7 & 1 \end{bmatrix}.]$$

12. Let A be the following product of 4×4 elementary row matrices:

$$A = E_3(2)E_{14}E_{42}(3).$$

Find A and A^{-1} explicitly.

$$[\text{Answers: } A = \begin{bmatrix} 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; A^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 1 & -3 & 0 & 0 \end{bmatrix}.]$$

13. Determine which of the following matrices over $\mathbb{Z}_2$ are non–singular and find the inverse, where possible.

$$\text{(a)} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}.$$

[Answer: (a) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$.]

14. Determine which of the following matrices are non–singular and find the inverse, where possible.

$$\text{(a)} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 2 & 2 & 4 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 4 & 6 & -3 \\ 0 & 0 & 7 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\text{(d)} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 7 \end{bmatrix} \quad \text{(e)} \begin{bmatrix} 1 & 2 & 4 & 6 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{(f)} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

[Answers: (a) $\begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 1 & -1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} -\frac{1}{2} & 2 & 1 \\ 0 & 0 & 1 \\ \frac{1}{2} & -1 & -1 \end{bmatrix}$ (d) $\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}$

(e) $\begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$.]

15. Let A be a non–singular $n \times n$ matrix. Prove that A^t is non–singular and that $(A^t)^{-1} = (A^{-1})^t$.

16. Prove that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has no inverse if $ad - bc = 0$.

[Hint: Use the equation $A^2 - (a+d)A + (ad-bc)I_2 = 0$.]

17. Prove that the real matrix $A = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix}$ is non–singular by proving that A is row–equivalent to I_3.

18. If $P^{-1}AP = B$, prove that $P^{-1}A^nP = B^n$ for $n \geq 1$.

19. Let $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{3}{4} \end{bmatrix}$, $P = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}$. Verify that $P^{-1}AP = \begin{bmatrix} \frac{5}{12} & 0 \\ 0 & 1 \end{bmatrix}$ and deduce that

$$A^n = \frac{1}{7}\begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix} + \frac{1}{7}\left(\frac{5}{12}\right)^n \begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix}.$$

20. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a *Markov* matrix; that is a matrix whose elements are non–negative and satisfy $a+c = 1 = b+d$. Also let $P = \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix}$. Prove that if $A \neq I_2$ then

    (i) P is non–singular and $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & a+d-1 \end{bmatrix}$,

    (ii) $A^n \to \dfrac{1}{b+c}\begin{bmatrix} b & b \\ c & c \end{bmatrix}$ as $n \to \infty$, if $A \neq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

21. If $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ and $Y = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$, find XX^t, X^tX, YY^t, Y^tY.

    [Answers: $\begin{bmatrix} 5 & 11 & 17 \\ 11 & 25 & 39 \\ 17 & 39 & 61 \end{bmatrix}$, $\begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}$, $\begin{bmatrix} 1 & -3 & -4 \\ -3 & 9 & 12 \\ -4 & 12 & 16 \end{bmatrix}$, 26.]

22. Prove that the system of linear equations

$$\begin{aligned} x + 2y &= 4 \\ x + y &= 5 \\ 3x + 5y &= 12 \end{aligned}$$

    is inconsistent and find a least squares solution of the system.
    [Answer: $x = 6$, $y = -7/6$.]

23. The points (0, 0), (1, 0), (2, −1), (3, 4), (4, 8) are required to lie on a parabola $y = a + bx + cx^2$. Find a least squares solution for a, b, c. Also prove that no parabola passes through these points.

    [Answer: $a = \frac{1}{5}$, $b = -2$, $c = 1$.]

24. If A is a symmetric $n \times n$ real matrix and B is $n \times m$, prove that B^tAB is a symmetric $m \times m$ matrix.

25. If A is $m \times n$ and B is $n \times m$, prove that AB is singular if $m > n$.

26. Let A and B be $n \times n$. If A or B is singular, prove that AB is also singular.