2. Matrix Algebra

In chapter 1 we found it convenient to represent linear systems of equations by augmented matrices. Matrix algebra leads to other applications, such as geometrical transformations (essential to image processing, one of many engineering applications) and evolutionary models (economics, probability - Markov chains, etc.).

2.1 Simple Matrix Algebra

A matrix with m rows and n columns has **dimensions** or **size** ($\boldsymbol{m}\times\boldsymbol{n}$) and is said to be an "$m$ by n matrix". The number of rows is always written first and the number of columns second.

An example of a 2×3 matrix is $A = \begin{bmatrix} 3 & 2 & 1 \\ 0 & -1 & -2 \end{bmatrix}$.

A $1\times n$ matrix is a **row matrix**. $R = \begin{bmatrix} 12 & 2 & -5 & 10 \end{bmatrix}$ is a row matrix (of size 1×4).
(also known as a **row vector**).

An $n\times 1$ matrix is a **column matrix**. $C = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ is a column matrix (of size 3×1).

(also known as a **column vector**).
A matrix with equal numbers of rows and columns is a **square matrix**.

$D = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 5 \end{bmatrix}$ is a square matrix of dimensions (3×3).

The entry in row i and column j of matrix D is d_{ij}.
In matrix D above, $d_{23} = 3$.

The **main diagonal** of a matrix extends down and right from the top left corner; the elements of the main diagonal of matrix $A = [\, a_{ij} \,]$ are a_{ii}.

For the four matrices above, the main diagonals are highlighted here:

$$A = \begin{bmatrix} \boxed{3} & 2 & 1 \\ 0 & \boxed{-1} & -2 \end{bmatrix}, \quad R = \begin{bmatrix} \boxed{12} & 2 & -5 & 10 \end{bmatrix}, \quad C = \begin{bmatrix} \boxed{-2} \\ 0 \\ 1 \end{bmatrix},$$

$$D = \begin{bmatrix} \boxed{-1} & 0 & 2 \\ 0 & \boxed{1} & 3 \\ 0 & 0 & \boxed{5} \end{bmatrix}.$$

Equality

Two matrices are equal if and only if they are the same size *and* all corresponding pairs of entries are equal.
In other words, $A = B$ iff $a_{ij} = b_{ij}$ for all i and for all j.
Example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} \Leftrightarrow \quad a = 2,\ b = 0,\ c = 1 \text{ and } d = 3$$

Addition is defined only for matrices of the same size.

Example 2.1.01

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 5 \\ 4 & 3 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad A + B = \begin{bmatrix} 6 & 6 \\ 7 & 5 \\ 1 & 1 \end{bmatrix}$$

Example 2.1.02

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 6 & 5 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad A + C \text{ is undefined.}$$

Matrix addition is commutative and associative.
For any matrices A, B, C of the same size,

$$A + B = B + A \quad \text{and}$$
$$A + (B + C) = (A + B) + C$$

The identity matrix under addition is the **zero matrix**:

All entries of any zero matrix are zero. The $(m \times n)$ zero matrix is O_{mn} (or just O if the size is obvious from the situation).

For all matrices X,

$X + O = X$ (where the zero matrix is the same size as X)

The inverse matrix of an $(m \times n)$ matrix A under addition is its **negative** $-A$, whose entries are all $-a_{ij}$.

For all matrices X,

$$X + (-X) = O \qquad \text{(where the zero matrix is the same size as } X\text{)}$$

The **difference** of two matrices A, B of the same size is

$$A - B = A + (-B)\,, \text{ whose elements are } [a_{ij} - b_{ij}]$$

Example 2.1.03

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 5 \\ 4 & 3 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad B - A = \begin{bmatrix} 6 & 4 \\ 1 & 1 \\ 3 & 1 \end{bmatrix}$$

Example 2.1.04

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 6 & 5 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad A - C \text{ is undefined.}$$

Scalar Multiplication

Multiplication of a matrix A by a scalar k causes every element of A to be multiplied by k.

$$A = \left[a_{ij}\right] \quad \Rightarrow \quad k\,A = \left[k a_{ij}\right]$$

Example 2.1.05

$$A = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 5 \\ 4 & 3 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad 2A = \begin{bmatrix} 0 & 2 \\ 6 & 4 \\ -2 & 0 \end{bmatrix}$$

$$\text{and} \quad 3A - 2B = \begin{bmatrix} 3(0)-2(6) & 3(1)-2(5) \\ 3(3)-2(4) & 3(2)-2(3) \\ 3(-1)-2(2) & 3(0)-2(1) \end{bmatrix} = \begin{bmatrix} -12 & -7 \\ 1 & 0 \\ -7 & -2 \end{bmatrix}$$

The **distributive** laws for matrices of the same size follow:

$$k(A+B) = kA + kB$$
$$(p+q)A = pA + qA$$
and
$$(pq)A = p(qA)$$

The **transpose** of a matrix $A = [\, a_{ij} \,]$ is $A^{\mathrm{T}} = [\, a_{ji} \,]$.

Thus the rows of the transpose are the columns of the original matrix and vice versa.
The transpose of an $(m \times n)$ matrix is an $(n \times m)$ matrix.
In particular, the transpose of a row matrix is a column matrix
and the transpose of a column matrix is a row matrix.

Example 2.1.06

Write down the transpose of the following matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 9 \\ -6 \\ 1 \\ 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 & 6 \\ -4 & 0 & 5 \end{bmatrix}$$

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad B^{\mathrm{T}} = \begin{bmatrix} 9 & -6 & 1 & 2 \end{bmatrix}, \quad C^{\mathrm{T}} = \begin{bmatrix} 1 & -4 \\ 3 & 0 \\ 6 & 5 \end{bmatrix}$$

Further properties of transposition:
For all equal-size matrices A, B and all scalars k,

$$\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$$
$$(kA)^{\mathrm{T}} = k\left(A^{\mathrm{T}}\right)$$
$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$$

A matrix for which $A^{\mathrm{T}} = A$ is **symmetric**.
Symmetric matrices are necessarily square $(n \times n)$
and the main diagonal is a line of symmetry..

Example 2.1.07

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

Matrix A is symmetric because $A^{\mathrm{T}} = A$.

Matrix B is not symmetric because $B^{\mathrm{T}} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = B$.

Matrix C cannot be symmetric because it is not square.

Miscellaneous Examples

Example 2.1.08 Textbook exercises 2.1 page 34 question 1(b)

Find a, b, c and d if $\begin{bmatrix} a-b & b-c \\ c-d & d-a \end{bmatrix} = 2\begin{bmatrix} 1 & 1 \\ -3 & 1 \end{bmatrix}$.

This generates the system of simultaneous linear equations

$$\begin{aligned} a - b \qquad\qquad &= 2 \\ b - c \qquad &= 2 \\ c - d &= -6 \\ -a \qquad\qquad + d &= 2 \end{aligned}$$

Solving the linear system,

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 & -6 \\ -1 & 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow[R_4 + R_1]{} \left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 & -6 \\ 0 & -1 & 0 & 1 & 4 \end{array}\right]$$

$$\xrightarrow[R_4 + R_2]{} \left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 & -6 \\ 0 & 0 & -1 & 1 & 6 \end{array}\right] \xrightarrow[R_4 + R_3]{} \left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

which is row-echelon form.

d is a non-leading variable and is assigned a parametric value t (where t may be any real number).

Example 2.1.08 (continued)

The system is now

$$\begin{aligned} a - b \quad\quad\quad &= 2 \\ b - c \quad\quad &= 2 \\ c - d &= -6 \\ d &= t \end{aligned}$$

Using back-substitution,

$c = t - 6$
$b = c + 2 = t - 4$
$a = b + 2 = t - 2$

The values of a, b, c and d are therefore
$(a, b, c, d) = (t - 2, t - 4, t - 6, t)$ or equivalently
$(a, b, c, d) = (-2, -4, -6, 0) + t\,(1, 1, 1, 1)\,, \quad (t \in \mathbb{R})$.

Example 2.1.09

Find the transpose of $A = \begin{bmatrix} 0 & 5 & -2 \\ -5 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$.

$$A^{\mathrm{T}} = \begin{bmatrix} 0 & -5 & 2 \\ 5 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} = -A$$

Matrices which are such that $A^{\mathrm{T}} = -A$ are **skew-symmetric**.
In any skew-symmetric matrix A, the main diagonal elements $a_{ii} = 0$.

Example 2.1.10 Textbook exercises 2.1 page 35 question 15(a)

Find the matrix A that satisfies the equation

$$\left(A + 3\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix}\right)^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix}$$

Method 1.

$$A + 3\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 2 & 0 & 3 \\ 1 & 5 & 8 \end{bmatrix}$$

$$\Rightarrow \quad A = \begin{bmatrix} 2 & 0 & 3 \\ 1 & 5 & 8 \end{bmatrix} - 3\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix} = \underline{\underline{\begin{bmatrix} -1 & 3 & 3 \\ -2 & -1 & -4 \end{bmatrix}}}$$

Method 2.

$$A^{\mathrm{T}} + 3\begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix}$$

$$\Rightarrow \quad A^{\mathrm{T}} = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix} - 3\begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 3 & -1 \\ 3 & -4 \end{bmatrix}$$

$$\Rightarrow \quad A = \begin{bmatrix} -1 & -2 \\ 3 & -1 \\ 3 & -4 \end{bmatrix}^{\mathrm{T}} = \underline{\underline{\begin{bmatrix} -1 & 3 & 3 \\ -2 & -1 & -4 \end{bmatrix}}}$$

Example 2.1.11 Textbook exercises 2.1 page 35 question 17

Show that $A + A^T$ is symmetric for *any* square matrix A.

First note that if A is not square, then the dimensions of A and A^T will be different, so that $A + A^T$ is not defined at all.

$(A + A^T)^T = A^T + A = A + A^T$ (matrix addition is commutative).

Therefore the matrix $(A + A^T)$ is symmetric for all square matrices A.

Building on this example, any square matrix A can be written as the sum of a symmetric matrix S and a skew-symmetric matrix K: $A = S + K$

S is symmetric $\Rightarrow$ $S^T = S$.
K is skew-symmetric $\Rightarrow$ $K^T = -K$.
$A^T = (S + K)^T = S^T + K^T = S - K$

$$\Rightarrow \quad A + A^T = (S + K) + (S - K) = 2S \qquad \Rightarrow \quad S = \frac{A + A^T}{2}$$

and

$$A - A^T = (S + K) - (S - K) = 2K \qquad \Rightarrow \quad K = \frac{A - A^T}{2}$$

so that the symmetric matrix S and the skew-symmetric matrix K are uniquely determined for each square matrix A. [This is also question 20, exercise 2.1, on page 36 of the textbook.]

2.2 Matrix Multiplication

Dot product

The dot product of a row vector $R = \begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix}$

and a column vector $C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^{\mathrm{T}}$

is defined to be $R \bullet C = \begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = r_1 c_1 + r_2 c_2 + \cdots + r_n c_n = \sum_{k=1}^{n} r_k c_k$

Note that the dimensions of the row and column vectors must be $(1 \times n)$ and $(n \times 1)$ respectively, otherwise the sum $R \bullet C = \sum_{k=1}^{n} r_k c_k$ is not defined.

The order of multiplication in the dot product is important.

Example 2.2.01

The numbers of atoms of carbon, hydrogen and oxygen in each molecule of water, methanol and ethanol are represented in the matrix A:

$$A = \begin{array}{c} \\ C \\ H \\ O \end{array} \begin{array}{c} \begin{array}{ccc} \text{water} & \text{methanol} & \text{ethanol} \end{array} \\ \begin{bmatrix} 0 & 1 & 2 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

The composition of a form of denatured alcohol and a dilution of that alcohol in water is described by the numbers of molecules of water, methanol and ethanol per 20 molecules of the alcohol, as listed in matrix B:

$$B = \begin{array}{c} \\ \text{water} \\ \text{methanol} \\ \text{ethanol} \end{array} \begin{array}{c} \begin{array}{cc} \text{denatured} & \text{diluted} \end{array} \\ \begin{bmatrix} 0 & 10 \\ 2 & 1 \\ 18 & 9 \end{bmatrix} \end{array}$$

Find the ratio of carbon atoms to hydrogen atoms to oxygen atoms in the diluted alcohol.

Example 2.2.01 (continued)

Every 20 molecules on average of the diluted alcohol contains 10 molecules of water, 1 molecule of methanol and 9 molecules of ethanol. Reading the atomic contents of these three molecules from matrix A, we find the numbers of atoms per 20 molecules of diluted alcohol to be:

	Water + Meth. + Ethanol
Carbon:	$0 \times 10 + 1 \times 1 + 2 \times 9 = 19$ atoms
Hydrogen:	$2 \times 10 + 4 \times 1 + 6 \times 9 = 78$ atoms
Oxygen:	$1 \times 10 + 1 \times 1 + 1 \times 9 = 20$ atoms

On average, every 20 molecules of diluted alcohol contain 19 atoms of carbon, 78 atoms of hydrogen and 20 atoms of oxygen. The ratio is C:H:O: = **19 : 78 : 20**.

Note how the numbers of atoms were found.
The number of carbon atoms is the dot product of the first row of A with the second column of B.
The number of hydrogen atoms is the dot product of the second row of A with the second column of B.
The number of oxygen atoms is the dot product of the third row of A with the second column of B.

The product of the two matrices yields the number of atoms per 20 molecules of each of the two substances:

$$AB = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix} \bullet \begin{bmatrix} 0 & 10 \\ 2 & 1 \\ 18 & 9 \end{bmatrix} = \begin{bmatrix} (0\times 0 + 1\times 2 + 2\times 18) & (0\times 10 + 1\times 1 + 2\times 9) \\ (2\times 0 + 4\times 2 + 6\times 18) & (2\times 10 + 4\times 1 + 6\times 9) \\ (1\times 0 + 1\times 2 + 1\times 18) & (1\times 10 + 1\times 1 + 1\times 9) \end{bmatrix}$$

denatured diluted

$$\Rightarrow \quad AB = \begin{bmatrix} 38 & 19 \\ 116 & 78 \\ 20 & 20 \end{bmatrix} \begin{matrix} C \\ H \\ O \end{matrix}$$

The product of two general matrices follows.

The product of an $(m\times n)$ matrix A with a $(p\times q)$ matrix B (in that order) is not defined unless $p = n$.

The product $C = AB$ of an $(m\times n)$ matrix A with an $(n\times q)$ matrix B (in that order) is the $(m\times q)$ matrix $C = [\ c_{ij}\ ]$, where the entry in row i and column j of C is the dot product of the i^{th} row of A with the j^{th} column of B:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} c_{kj}$$

Example 2.2.02

Find the matrix products AB and BA where

$$A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}.$$

$$AB = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} (1\times 2+0\times 1) & (1\times 1+0\times -2) \\ (-2\times 2+1\times 1) & (-2\times 1+1\times -2) \\ (4\times 2+3\times 1) & (4\times 1+3\times -2) \end{bmatrix} = \underline{\underline{\begin{bmatrix} 2 & 1 \\ -3 & -4 \\ 11 & -2 \end{bmatrix}}}.$$

BA is **not defined** because B is $(2\times\mathbf{2})$ and A is $(\mathbf{3}\times 2)$. The number of columns of the left matrix does not match the number of rows of the right matrix.

Note that this example demonstrates that matrix multiplication is **not commutative** in general, that is $BA \neq AB$.

The **identity matrix** of order n is the square ($n \times n$) matrix whose main diagonal entries are one and whose other entries are all zero.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ etc.}$$

For any ($m \times n$) matrix A, $I_m A = A I_n = A$.

I is therefore the identity element for the operation of matrix multiplication.

Where it is obvious from the context, I_n is represented by just I.

Example 2.2.03

$$A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix} \quad \Rightarrow \quad IA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix}$$

and

$$AI = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix}$$

Where the product is defined, the product of the **zero matrix** with any other matrix is the zero matrix of the appropriate dimensions.

Example 2.2.04

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

but $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 4 & 3 \end{bmatrix}$ is not defined

---

Example 2.2.05

Find A^2, where $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

$$A^2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (1\times1+2\times3) & (1\times2+2\times4) \\ (3\times1+4\times3) & (3\times2+4\times4) \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$$

Example 2.2.06

Find A^2, where $A = \begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix}$ and k is any real number.

$$A^2 = \begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} (-1\times-1+k\times0) & (-1\times k+k\times1) \\ (0\times-1+1\times0) & (0\times k+1\times1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Note that in scalar arithmetic $x^2 = 1 \quad \Rightarrow \quad x = \pm1$, but in matrix multiplication
$A^2 = I \quad \not\Rightarrow \quad A = \pm I$

Example 2.2.07

Find A^2, where $A = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix}$.

$$A^2 = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} (2\times2+-4\times1) & (2\times-4+-4\times-2) \\ (1\times2+-2\times1) & (1\times-4+-2\times-2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O$$

Note that in scalar arithmetic $x^2 = 0 \quad \Rightarrow \quad x = 0$, but in matrix multiplication
$A^2 = O \quad \not\Rightarrow \quad A = O$

Some properties of matrix multiplication:

For any scalar k, matrices A, B, C of dimensions such that the matrix multiplications are defined, and identity and zero matrices of the appropriate dimensions,

$IA = AI = A$ [identity]
$OA = AO = O$ [zero]
$A(BC) = (AB)C$ [associative law]
$A(B+C) = AB + AC$ [distributive law]
$(B+C)A = BA + CA$ [distributive law]
$k(AB) = (kA)B = A(kB)$
but note that $AB \neq BA$ in general. Matrices for which $AB = BA$ are said to commute. Be very careful of the order of matrix multiplication.

$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$

As first seen in Chapter 1, any system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{p1}x_1 + a_{p2}x_2 + a_{p3}x_3 + \ldots + a_{pn}x_n &= b_p \end{aligned}$$

can be written more compactly as the matrix equation

$$AX = B$$

where $A = \left[a_{ij} \right] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \ldots & a_{pn} \end{bmatrix}$,

and X and B are the column vectors $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_p \end{bmatrix}$.

Given an inhomogeneous linear system $AX = B$, there is an **associated homogeneous system**

$$AX = O$$

If the column vector X_1 is any one solution to $AX = B$ and
the column vector X_0 is any one solution to $AX = O$, then
$(X_0 + X_1)$ is also a solution to $AX = B$. [This requires $AX = B$ to be consistent.]

Thus the general solution to the system $AX = B$ may be expressed as the sum of the general solution to the associated homogeneous system and a **particular solution** of the inhomogeneous system.

Proof:

Let X_2 be any solution to $AX = B$ (so that $AX_2 = B$)
and let X_1 be a known particular solution to $AX = B$ (so that $AX_1 = B$).
Let $X_0 = X_2 - X_1$.

Then $AX_0 = A(X_2 - X_1) = AX_2 - AX_1 = B - B = O$
$\Rightarrow$ X_0 is a solution to the associated homogeneous system $AX = O$.

Occasionally it is easier to find a particular solution and to solve the associated homogeneous system than it is to solve the original inhomogeneous system all at once.

We will see this concept of partitioning a solution into a particular solution and the solution of the associated homogeneous system again when we study ordinary differential equations in a future course (MATH 3260 or ENGI 3424 or ENGI 3425/4425).

If A is an $(m\times n)$ matrix of rank r, then the homogeneous linear system of m equation in n variables $AX = O$ has exactly $(n-r)$ basic solutions, one for each parameter and every solution is a linear combination of these basic solutions.

Example 2.2.08

Find basic solutions of $AX = O$, where $A = \begin{bmatrix} 1 & 2 & 0 & 1 & 1 \\ 2 & 4 & 1 & 1 & -2 \\ 3 & 6 & 1 & 2 & -1 \\ 1 & 2 & 1 & 0 & -3 \end{bmatrix}$

Show that $X = \begin{bmatrix} 1 & 2 & 0 & 2 & 1 \end{bmatrix}^{\mathrm{T}}$ is a solution to $AX = B$, where
$B = \begin{bmatrix} 8 & 10 & 18 & 2 \end{bmatrix}^{\mathrm{T}}$. Hence find the complete solution to $AX = B$.

Example 2.2.08 (continued)

Reducing the augmented matrix to row-echelon form:

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 1 & 0 \\ 2 & 4 & 1 & 1 & -2 & 0 \\ 3 & 6 & 1 & 2 & -1 & 0 \\ 1 & 2 & 1 & 0 & -3 & 0 \end{array}\right] \xrightarrow[R_3 - 3R_1,\; R_4 - R_1]{R_2 - 2R_1} \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -4 & 0 \\ 0 & 0 & 1 & -1 & -4 & 0 \\ 0 & 0 & 1 & -1 & -4 & 0 \end{array}\right]$$

$$\xrightarrow[R_4 - R_2]{R_3 - R_2} \left[\begin{array}{ccccc|c} \boxed{1} & 2 & 0 & 1 & 1 & 0 \\ 0 & 0 & \boxed{1} & -1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

which is equivalent to

$x_1 + 2x_2 + 0x_3 + x_4 + x_5 = 0$ and

$x_3 - x_4 - 4x_5 = 0$

The leading variables are x_1 and x_3.

Assign parameters $x_2 = r,\ x_4 = s,\ x_5 = t,$ so that the general solution is

$x_1 = -2r - s - t\,,\quad x_3 = s + 4t$

Then

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2r - s - t \\ r \\ s + 4t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} -1 \\ 0 \\ 4 \\ 0 \\ 1 \end{bmatrix} t$$

The basic solutions are therefore

$X_1 = [\,-2\ 1\ 0\ 0\ 0\,]^{\mathrm{T}}$, $X_2 = [\,-1\ 0\ 1\ 1\ 0\,]^{\mathrm{T}}$, $X_3 = [\,-1\ 0\ 4\ 0\ 1\,]^{\mathrm{T}}$

and the general solution to $AX = O$ is $X = r\,X_1 + s\,X_2 + t\,X_3$.

$$AX = \begin{bmatrix} 1 & 2 & 0 & 1 & 1 \\ 2 & 4 & 1 & 1 & -2 \\ 3 & 6 & 1 & 2 & -1 \\ 1 & 2 & 1 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 10 \\ 18 \\ 2 \end{bmatrix} = B$$

Therefore the complete solution to $AX = B$ is

$$X = [\,1\ 2\ 0\ 2\ 1\,]^{\mathrm{T}} + r\,X_1 + s\,X_2 + t\,X_3\,.$$

Block Multiplication

Example 2.2.09

Suppose that matrices A, B, P, X and Y are defined as

$$A = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 3 & 1 \end{array}\right] = \begin{bmatrix} I_2 & O_{23} \\ O_{22} & P \end{bmatrix} \quad \text{and} \quad B = \left[\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 2 & 0 \\ 0 & 3 \\ 1 & 2 \end{array}\right] = \begin{bmatrix} X \\ Y \end{bmatrix}$$

then

$$AB = \begin{bmatrix} I_2 & O_{23} \\ O_{22} & P \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} (I_2X + O_{23}Y) \\ (O_{22}X + PY) \end{bmatrix} = \begin{bmatrix} X \\ PY \end{bmatrix}$$

$$PY = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ 1 & 11 \end{bmatrix}$$

$$\Rightarrow \quad AB = \begin{bmatrix} X \\ PY \end{bmatrix} = \left[\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 7 & 6 \\ 1 & 11 \end{array}\right]$$

This is somewhat faster than the direct evaluation of

$$AB = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 2 & 0 \\ 0 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 7 & 6 \\ 1 & 11 \end{bmatrix}$$

The partitioning of the matrices in a matrix multiplication must be such that all matrix products are defined.

Additional Examples

Example 2.2.10

Find the complete set of (2×2) matrices that commute with $P = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Let the general (2×2) matrix be $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$AP = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix}$$

and

$$PA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix}$$

$AP = PA$ if and only if $c = 0$ and $d = a$.

Therefore the complete set of (2×2) matrices that commute with $P = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is

$A = \left\{\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}\right\}$, where a and b are any real numbers.

Example 2.2.11

For the matrix $A = \begin{bmatrix} I & X \\ O & -I \end{bmatrix}$, where X, I and O are all square matrices of the same size ($k \times k$), find an expression for any natural number power of A, A^n.

$$A^2 = \begin{bmatrix} I_k & X \\ O_{kk} & -I_k \end{bmatrix}\begin{bmatrix} I_k & X \\ O_{kk} & -I_k \end{bmatrix} = \begin{bmatrix} I_k & O_{kk} \\ O_{kk} & I_k \end{bmatrix} = I_{(2k)}$$

$\Rightarrow \quad A^3 = A^2 A = IA = A, \qquad A^4 = A^2 A^2 = I, \qquad A^5 = A^4 A = A,$ etc.

Therefore

$$A^n = \begin{cases} I & (n \text{ even}) \\ A & (n \text{ odd}) \end{cases}$$

The topic of adjacency matrices for directed graphs (textbook page 46) will be explored in an assignment.

Example 2.2.12 (Textbook, exercises 2.2, page 48, question 11)

Given that $A\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = O = A\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$ and that $X_0 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ is a solution to $AX = B$, find a two-parameter family of solutions to $AX = B$.

The homogeneous system $AX = O$ has at least a two-parameter family of solutions

$$X_h = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} s + \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} t, \qquad (s, t \in \mathbb{R})$$

We have a particular solution to the inhomogeneous system $AX = B$,

$$X_p = X_0 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

Therefore a two-parameter family of solutions to $AX = B$ is

$$X = X_p + X_h = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} s + \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} t, \qquad (s, t \in \mathbb{R})$$

2.3 - Matrix Inverses

For $(n \times n)$ matrices A, B, if

$$AB = BA = I$$

then $B = A^{-1}$ is the **inverse matrix** of A.

A matrix that possesses an inverse is **invertible**. A non-invertible matrix is **singular**.

Example 2.3.1

Show that $B = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$ is the inverse of $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

and

$$BA = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Therefore B is the inverse matrix of A.

If the inverse to a matrix A exists, then it is unique.

Proof:
Suppose that matrices B and C are both inverses of A. Then
$AB = BA = I$ and $AC = CA = I$.
$\Rightarrow \quad C = IC = (BA)C = B(AC) = BI = B$
The inverse matrix, if it exists, is therefore unique.

From Example 2.2.6 above, $A = \begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix}$ is its own inverse for all values of the real number k: $\quad A^2 = \begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$.
Therefore in this case $A^{-1} = A$, even though A is not $\pm I$.

The uniqueness of the inverse allows us to check just one of $A^{-1}A = I$ or $AA^{-1} = I$.

Inverse of a (2×2) Matrix

The **adjugate** (or **adjoint**) of a (2×2) matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\operatorname{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

$$A \operatorname{adj}(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (ad-bc)I$$

The **determinant** of A is defined to be $\det A = ad - bc$.

For all (2×2) matrices such that $\det A \neq 0$, it is clear that

$$\boxed{A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}$$

A matrix whose determinant is zero is singular (has no inverse).

Example 2.3.1 (again)

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \quad \Rightarrow \quad \det(A) = 1\times 5 - 2\times 3 = -1 \quad \text{and} \quad \operatorname{adj}(A) = \begin{bmatrix} 5 & -2 \\ -3 & 1 \end{bmatrix}$$

$$\Rightarrow \quad A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)} = -\begin{bmatrix} 5 & -2 \\ -3 & 1 \end{bmatrix} = +\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

Example 2.2.6 (again)

$$A = \begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad \det(A) = 1\times(-1) - (-k)\times 0 = -1 \quad \text{and} \quad \operatorname{adj}(A) = \begin{bmatrix} 1 & -k \\ 0 & -1 \end{bmatrix}$$

$$\Rightarrow \quad A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)} = -\begin{bmatrix} 1 & -k \\ 0 & -1 \end{bmatrix} = +\begin{bmatrix} -1 & k \\ 0 & 1 \end{bmatrix} = A$$

In a square linear system (n equations in n unknowns), if the coefficient matrix A has rank n, then it is invertible and
$AX = B \quad \Rightarrow \quad A^{-1}AX = A^{-1}B \quad \Rightarrow \quad IX = A^{-1}B \quad \Rightarrow$
the solution to the linear system is

$$X = A^{-1}B$$

and the solution is [necessarily] unique.

If rank $A < n$, then A^{-1} does not exist and the system is either inconsistent or has infinitely many solutions, but not a unique solution.

Example 2.3.2

Solve the linear system

$$\begin{aligned} 3x_1 + 2x_2 &= 10 \\ 5x_1 + 4x_2 &= 8 \end{aligned}$$

$$A = \begin{bmatrix} 3 & 2 \\ 5 & 4 \end{bmatrix} \quad \Rightarrow \quad \operatorname{adj}(A) = \begin{bmatrix} 4 & -2 \\ -5 & 3 \end{bmatrix} \quad \text{and} \quad \det(A) = 12 - 10 = 2$$

$$\Rightarrow \quad A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)} = \frac{1}{2}\begin{bmatrix} 4 & -2 \\ -5 & 3 \end{bmatrix}$$

The unique solution to the linear system is

$$X = A^{-1}B = \frac{1}{2}\begin{bmatrix} 4 & -2 \\ -5 & 3 \end{bmatrix}\begin{bmatrix} 10 \\ 8 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (40-16) \\ (-50+24) \end{bmatrix} = \underline{\underline{\begin{bmatrix} 12 \\ -13 \end{bmatrix}}}$$

Check by substituting the solution into the left side of the linear system:

$$\begin{aligned} 3x_1 + 2x_2 &= 3\times 12 + 2\times(-13) = 36 - 26 = 10 \\ 5x_1 + 4x_2 &= 5\times 12 + 4\times(-13) = 60 - 52 = 8 \end{aligned}$$

Matrix Inversion by Gaussian Elimination

Iff matrix A is invertible, then the reduced row-echelon form of $[\,A \;\; I\,]$ is $[\,I \;\; A^{-1}\,]$. The details are on page 54 of the textbook.

Example 2.3.3 (Textbook, page 59, exercises 2.3, question 2(c), modified)

Find the inverse of $A = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 2 & 0 \\ -1 & -1 & 0 \end{bmatrix}$ **and** hence solve the linear system

$$\begin{aligned} x - z &= 1 \\ 3x + 2y &= -3 \\ -x - y &= 2 \end{aligned}$$

$$[A\,|\,I] = \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 + R_1]{R_2 - 3R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 3 & -3 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 & 1 \end{array}\right] \xrightarrow{R_2 \times \frac{1}{2}} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} & 0 \\ 0 & -1 & -1 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 + R_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{array}\right] \xrightarrow[R_3 \times 2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -1 & 1 & 2 \end{array}\right]$$

$$\xrightarrow[R_2 - \frac{3}{2}R_3]{R_1 + R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & -1 & 1 & 2 \end{array}\right]$$

Therefore $A^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & -1 & -3 \\ -1 & 1 & 2 \end{bmatrix}$

One can easily verify that $AA^{-1} = A^{-1}A = I$.

Example 2.3.3 (continued)

The linear system is $AX = B$, where $B = [\ 1\ \ -3\ \ 2\ ]^T$

$$\Rightarrow \quad X = A^{-1}B = \begin{bmatrix} 0 & 1 & 2 \\ 0 & -1 & -3 \\ -1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}$$

Therefore the unique solution is $(x, y, z) = \underline{(1, -3, 0)}$

Check of the solution:

$$AX = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 2 & 0 \\ -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = B$$

Example 2.3.4

Find the inverse of $A = \begin{bmatrix} 1 & -3 & 1 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$$[A \mid I] = \left[\begin{array}{cccc|cccc} 1 & -3 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 + 3R_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 7 & -1 & 1 & 3 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_2 - 2R_3]{R_1 - 7R_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -29 & 1 & 3 & -7 & 0 \\ 0 & 1 & 0 & -8 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

Example 2.3.4 (continued)

$$\xrightarrow[R_3-4R_4]{R_1+29R_4 \quad R_2+8R_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 3 & -7 & 29 \\ 0 & 1 & 0 & 0 & 0 & 1 & -2 & 8 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

Therefore

$$A^{-1} = \begin{bmatrix} 1 & 3 & -7 & 29 \\ 0 & 1 & -2 & 8 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

One can easily verify that $AA^{-1} = A^{-1}A = I$.

The following statements for an $(n \times n)$ matrix A are either all true or all false:

1) A^{-1} exists (that is, A is invertible).
2) The reduced row-echelon form of A is I_n .
3) $AX = O$ has only the trivial solution $X = O$.
4) $AX = B$ has a unique solution for every choice of B .

Example 2.3.5 (textbook, page 59, exercises 2.3, question 4(a))

Given $A^{-1} = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 5 \\ -1 & 1 & 0 \end{bmatrix}$, solve the system of equations $AX = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$.

The system has a unique solution because A^{-1} exists.

$$X = A^{-1}\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 0 & 5 \\ -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 17 \\ -2 \end{bmatrix}$$

Example 2.3.6 (textbook, page 61, exercises 2.3, question 24 modified)

Show that if the block matrix $M = \begin{bmatrix} A & X \\ O & B \end{bmatrix}$ is invertible, then the matrices A and B are invertible **and** find M^{-1}. Hence find M^{-1} when $M = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

Let $M^{-1} = \begin{bmatrix} C & Y \\ Z & D \end{bmatrix}$, where C and D are the same size as A and B respectively.

$$M M^{-1} = \begin{bmatrix} A & X \\ O & B \end{bmatrix}\begin{bmatrix} C & Y \\ Z & D \end{bmatrix} = \begin{bmatrix} AC+XZ & AY+XD \\ BZ & BD \end{bmatrix} = \begin{bmatrix} I & O \\ O & I \end{bmatrix}$$

$BD = I \;\Rightarrow\; D = B^{-1}$.
B is invertible and $BZ = O \;\Rightarrow\; Z = O$.
$\Rightarrow\; AC + XZ = AC + O = I \;\Rightarrow\; C = A^{-1}$.

Therefore if M is invertible then both A and B are invertible.

$AY + XD = AY + XB^{-1} = O \;\Rightarrow\; AY = -XB^{-1}$
$\Rightarrow\; Y = -A^{-1}XB^{-1}$

$$\therefore\; M^{-1} = \begin{bmatrix} A & X \\ O & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}XB^{-1} \\ O & B^{-1} \end{bmatrix}$$

Note that it follows from this result that, for any constant x and non-zero constants a, b,

$$\begin{bmatrix} a & x \\ 0 & b \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a} & \frac{-x}{ab} \\ 0 & \frac{1}{b} \end{bmatrix}.$$

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \;\Rightarrow\; A^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \;\Rightarrow\; B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and $X = O \;\Rightarrow\; -A^{-1}XB^{-1} = O$

$$\Rightarrow\; M^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Example 2.3.7

Given that $A = \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}$,

(a) Verify that $A^2 - A - 6I = O$; and

(b) Hence find A^{-1}.

(a) $A^2 = \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -1 \\ 0 & 4 \end{bmatrix}$

$$\Rightarrow \quad A^2 - A - 6I = \begin{bmatrix} 9 & -1 \\ 0 & 4 \end{bmatrix} - \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O$$

(b) $A^2 - A - 6I = O \quad \Rightarrow \quad A^{-1}\left(A^2 - A - 6I\right) = O$

$$\Rightarrow \quad A^{-1}AA - A^{-1}A - 6A^{-1}I = O \qquad \Rightarrow \quad A - I - 6A^{-1} = O$$

$$\Rightarrow \quad A^{-1} = \tfrac{1}{6}\left(A - I\right) = \frac{1}{6}\begin{bmatrix} 3-1 & -1 \\ 0 & -2-1 \end{bmatrix} = \underline{\underline{\frac{1}{6}\begin{bmatrix} 2 & -1 \\ 0 & -3 \end{bmatrix}}}$$

Check: $A^{-1} = \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}^{-1} = \dfrac{1}{3\times(-2)-(-1)\times 0}\begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} = \dfrac{1}{6}\begin{bmatrix} +2 & -1 \\ 0 & -3 \end{bmatrix}$