Introduction to Linear Algebra

for Earth Science Students

Draft Version   Benjamin C. Loi

Preface

This is a Linear Algebra textbook specifically designed for students who study any Earth Science related major like Geophysics and Atmospheric Sciences. With these target readers in mind, we set out to provide them with a foundation in Linear Algebra that is adequate to deal with relevant Earth Science problems. We have avoided pedantic mathematical details in the book and are instead focusing on practical methods. Therefore, this book is not suitable for training mathematicians. In each chapter, we first discuss a selected Linear Algebra topic. Then we will move on to some Earth Science examples about that topic if possible. It is followed up by another section which demonstrates how coding in Python may help us solve these Linear Algebra problems. At the end of each chapter, a number of exercises are given for practices, and it can be done either by hand or programming. It is suggested that the readers install the newest version of Python via Anaconda.

Benjamin Loi

Contents

6

Introduction to Matrices and Linear Systems

Although the Earth System is well-known to be filled with non-linear processes, we still benefit from learning how to work with linear systems, as which many Earth Science problems can be approximated. This actually works well in a number of cases. For example, in Atmosphere Sciences, we often consider what is called a *perturbation equation*, which assumes that variations from the mean state are small enough to neglect quadratic perturbed terms. *Matrices* are one of the most fundamental objects in Linear Algebra, and we are going to address the basic aspects related to them in the first chapter.

1.1 Definition and Operations of Matrices

1.1.1 Basic Structure of Matrices

Matrices are rectangular arrays of numbers, which can be real or complex. For now we will work with real numbers, and defer the treatment about complex matrices to later chapters. A matrix having m rows and n columns is called an $m \times n$ matrix. A matrix with the same number of rows and columns, i.e. $m = n$, is called a ***square matrix***. Below shows some examples of matrices.

$$\begin{bmatrix} 1 & 2 & -2 & 5 \\ 1 & 3 & \sqrt{3} & 7 \end{bmatrix}$$

A 2×4 real matrix.

$$\begin{bmatrix} 3 & 2 & 9 \\ -4 & 0 & 4 \\ 5 & 2 & -3 \end{bmatrix}$$

A 3×3 real, square matrix.

$$\begin{bmatrix} 2 - \iota \\ 0 \\ 1 \\ 3\iota \end{bmatrix}$$

A 4×1 complex matrix.

Given any matrix A, its entry at row i and column j will be denoted as A_{ij}. For example,

$$A = \begin{bmatrix} \overset{\text{Col 1}}{2} & 1 & 9 \\ 5 & -3 & 7 \end{bmatrix} \text{Row 2} \qquad A_{21} = 5$$

Short Exercise: Find A_{13} and A_{22}.[1]

1.1.2 Matrix Operations

Addition and Subtraction

Addition and subtraction between two matrices A and B are carried out *element-wise*, which means that if $C = A \pm B$, then $C_{ij} = A_{ij} \pm B_{ij}$. This implies that the two matrix operands must be of the same shape, and addition/subtraction is not possible for two matrices with different extents. For instance, if we have

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 8 \\ 1 & 7 \end{bmatrix}$$

---

[1] $A_{13} = 9$, $A_{22} = -3$.

Then

$$A + B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 8 \\ 1 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 1+1 & 2+1 \\ 3+0 & 4+8 \\ 5+1 & 6+7 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3 \\ 3 & 12 \\ 6 & 13 \end{bmatrix}$$

Short Exercise: Find $A - B$.[2]

Scalar Multiplication

Multiplying a matrix by a constant constitutes a ***scalar multiplication***, in which all entries are multiplied by the scalar. It is illustrated in the example below.

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

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

$$= \begin{bmatrix} 3 \times 2 & 3 \times (-5) & 3 \times 6 \\ 3 \times (-1) & 3 \times 4 & 3 \times (-3) \end{bmatrix}$$

$$= \begin{bmatrix} 6 & -15 & 18 \\ -3 & 12 & -9 \end{bmatrix}$$

Short Exercise: Find $\frac{1}{4}A$.[3]

---

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

[3] $A/4 = \begin{bmatrix} 1/2 & -5/4 & 3/2 \\ -1/4 & 1 & -3/4 \end{bmatrix}$

Matrix Multiplication/Matrix Product

Meanwhile, multiplication between two matrices, commonly referred to as *matrix multiplication/matrix product*, is not element-wise. It can be only carried out if the number of columns of the first matrix equals to the number of rows of the second matrix, let's say r, in other words, they are of the shape $m \times r$ and $r \times n$ respectively. The resulting matrix is of the shape $m \times n$, which means that the number of rows/columns of the output matrix follows the first/second input matrix respectively. The following two examples explain this requirement.

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}$$

Since the shapes of A and B are 2×3 and 3×1, the number of columns in A and the number of rows in B are both 3, and hence matrix product between them is possible. The resulting matrix will be of size 2×1.

$$C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 6 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 3 & 1 \\ 0 & 5 \\ 3 & 2 \end{bmatrix}$$

As the number of columns in C is 4, which is not equal to the number of rows in D, 3, matrix product is undefined in this case. Now we are ready to see how the entries in matrix product is exactly computed.

Definition 1.1.1 (Matrix Product). Given an $m \times r$ matrix A and an $r \times n$ matrix B, if we denote the matrix product between A and B as $C = AB$, then to calculate any entry in C at row i and column j, we select row i from the first matrix A and column j from the second matrix B. Subsequently, take the products between the r pairs of numbers from that row and column. Their sum will then be the required value of the entry, i.e.

$$C_{ij} = (AB)_{ij} = A_{i1}B_{ij} + A_{i2}B_{2j} + A_{i3}B_{3j} + ... + A_{ir}B_{rj}$$

$$= \sum_{k=1}^{r} A_{ik} B_{kj}$$

again, r is the number of columns/rows in the first/second matrix.

Example 1.1.1. Calculate the matrix product $C = AB$, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Solution. The output will be a 2×2 matrix. Using the definition, we have

$$\begin{aligned}
C_{11} = (AB)_{11} &= A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} \\
&= (1)(1) + (3)(2) + (5)(3) = 22 \\
C_{12} = (AB)_{12} &= A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\
&= (1)(4) + (3)(5) + (5)(6) = 49
\end{aligned}$$

Hence the entries along the first row of C will be 22 and 49. The remaining entries at the second row are found in a similar way, and the readers are encouraged to do this themselves. You should be able to get

$$C = \begin{bmatrix} 22 & 49 \\ 28 & 64 \end{bmatrix}$$

□

Matrix product has some important properties, listed as follows.

Properties 1.1.2. If A, B, C are some matrices having compatible shapes so that the matrix multiplication operations below are valid, then

$$\underbrace{A \cdots A}_{k \text{ times}} = A^k \qquad\qquad k\text{-th power of a (square) matrix}$$

$$(AB)C = A(BC) = ABC \qquad \text{Associative Property}$$
$$(A \pm B)C = AC \pm BC \qquad \text{Distributive Property}$$
$$A(B \pm C) = AB \pm AC \qquad \text{Distributive Property}$$

Another important observation is that, in general $AB \neq BA$ even if the matrix products AB and BA are both well-defined, so they are not *commutative*. However, as we will see later, there are some distinct exceptions of this.

Example 1.1.2. Calculate $-2A + 3B$, where

$$A = \begin{bmatrix} 1 & 6 & 9 \\ 4 & 4 & 6 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 4 & 8 & 6 \\ -5 & 0 & 3 \end{bmatrix}$$

Solution.

$$
\begin{aligned}
-2A + 3B &= -2 \begin{bmatrix} 1 & 6 & 9 \\ 4 & 4 & 6 \end{bmatrix} + 3 \begin{bmatrix} 4 & 8 & 6 \\ -5 & 0 & 3 \end{bmatrix} \\
&= \begin{bmatrix} -2 & -12 & -18 \\ -8 & -8 & -12 \end{bmatrix} + \begin{bmatrix} 12 & 24 & 18 \\ -15 & 0 & 9 \end{bmatrix} \\
&= \begin{bmatrix} 10 & 12 & 0 \\ -23 & -8 & -3 \end{bmatrix}
\end{aligned}
$$

$\square$

Example 1.1.3. Compute $(A + 3B)(2A - B)$, where

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

Solution. Using distributive property in Properties 1.1.2, the expression is equivalent to

$$(A + 3B)(2A - B) = A(2A - B) + (3B)(2A - B)$$

$$= A(2A) + A(-B) + (3B)(2A) + (3B)(-B)$$
$$= 2A^2 - AB + 6BA - 3B^2$$

Bear in mind that $AB \neq BA$. We calculate each term, which gives

$$A^2 = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$$
$$= \begin{bmatrix} (1)(1) + (2)(3) & (1)(2) + (2)(5) \\ (3)(1) + (5)(3) & (3)(2) + (5)(5) \end{bmatrix}$$
$$= \begin{bmatrix} 7 & 12 \\ 18 & 31 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 4 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} (1)(-2) + (2)(4) & (1)(0) + (2)(-1) \\ (3)(-2) + (5)(4) & (3)(0) + (5)(-1) \end{bmatrix}$$
$$= \begin{bmatrix} 6 & -2 \\ 14 & -5 \end{bmatrix}$$

Similarly, it is not difficult to obtain

$$BA = \begin{bmatrix} -2 & -4 \\ 1 & 3 \end{bmatrix} \qquad\qquad B^2 = \begin{bmatrix} 4 & 0 \\ -12 & 1 \end{bmatrix}$$

Hence the final answer will be

$$2A^2 - AB + 6BA - 3B^2 = 2\begin{bmatrix} 7 & 12 \\ 18 & 31 \end{bmatrix} - \begin{bmatrix} 6 & -2 \\ 14 & -5 \end{bmatrix} + 6\begin{bmatrix} -2 & -4 \\ 1 & 3 \end{bmatrix} - 3\begin{bmatrix} 4 & 0 \\ -12 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 14 & 24 \\ 36 & 62 \end{bmatrix} - \begin{bmatrix} 6 & -2 \\ 14 & -5 \end{bmatrix} + \begin{bmatrix} -12 & -24 \\ 6 & 18 \end{bmatrix} - \begin{bmatrix} 12 & 0 \\ -36 & 3 \end{bmatrix}$$
$$= \begin{bmatrix} -16 & 2 \\ 64 & 82 \end{bmatrix}$$

$\square$

Alternatively, one can evaluate $C = A + 3B$ and $D = 2A - B$ first, and subsequently calculate the matrix dot product CD. (This is actually somehow easier.) The readers should try this as an exercise.

Matrix Equation Manipulation

For any matrix equation, one can do addition, subtraction and multiplication on both sides of the equation. However, one important note is that multiplying a matrix to an equation requires the matrix to be inserted to the left (or right) on both sides, respecting the order. So, for a matrix equation like (assuming the shapes are compatible),

$$AB - C = DE + F$$

if we want to take the product with a matrix G, then it can be

$$G(AB - C) = G(DE + F)$$
$$(AB - C)G = (DE + F)G$$

but we have, in general

$$G(AB - C) \neq (DE + F)G$$
$$(AB - C)G \neq G(DE + F)$$

Taking successive matrix products follows the same principle, step by step. Using the example above, for another matrix H, we note some possible outcomes.

$$HG(AB - C) = HG(DE + F)$$
$$(AB - C)GH = (DE + F)GH$$
$$G(AB - C)H = G(DE + F)H$$

However, be careful that cancellation at both sides may not be correct. If $AB = AC$, then we cannot conclude that $B = C$ for sure, although it is not impossible (and happens quite often).

1.2 Definition of Linear Systems of Equations

The prime application of matrices is to deal with ***linear systems (of equations)***. To understand what a linear system is, we first have to know the definition of

a ***linear equation*** (possibly in multiple variables, let's say $x_1, x_2, \ldots$). For a linear equation, in any additive term, there is at most one variable or unknown, with a power of one, like $3x_1, -x_2$. This means that there are no cross-product terms such as $2x_1x_2$, variables with a power that is not one, like x_1^3, or non-linear functions, including $\sin x_1, e^{x_2}$. For n variables, a linear equation has the following definition.

Definition 1.2.1 (Linear Equation). A linear equation has the form of

$$\sum_{j=1}^{n} a_j x_j = a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + a_n x_n = h$$

where $x_1, x_2, \ldots$ are the unknowns, while $a_1, a_2, \ldots$ and h are some constants. If $h = 0$, then it is known as a ***homogeneous linear equation***.

Short Exercise: Determine whether the equations below are (a) linear, and if they are linear, then (b) homogeneous or not.[4]

1. $3x + 4y = 5$

2. $\cos x + \ln y = 0$

3. $7x - 5z = 2$

4. $x^2 + y^{-3/2} = 1$

5. $x + 3y + 6z = 0$

6. $xyz = 8$

A system of linear equations are then simply a family of m linear equations, $m \geq 1$.

Definition 1.2.2 (Linear System of Equations). A linear system of size $m \times n$

---

[4]Linear/Inhomogeneous, Non-linear, Linear/Inhomogeneous, Non-linear, Linear/Homogeneous, Non-linear.

has the form of

$$
\begin{cases}
\sum_{j=1}^{n} a_j^{(1)} x_j = a_1^{(1)} x_1 + a_2^{(1)} x_2 + a_3^{(1)} x_3 + \cdots + a_n^{(1)} x_n & = h^{(1)} \\
\sum_{j=1}^{n} a_j^{(2)} x_j = a_1^{(2)} x_1 + a_2^{(2)} x_2 + a_3^{(2)} x_3 + \cdots + a_n^{(2)} x_n & = h^{(2)} \\
\vdots \\
\sum_{j=1}^{n} a_j^{(m)} x_j = a_1^{(m)} x_1 + a_2^{(m)} x_2 + a_3^{(m)} x_3 + \cdots + a_n^{(m)} x_n & = h^{(m)}
\end{cases}
$$

If $h^{(1)}, h^{(2)}, \ldots$ at the right hand side are all zeros, then the system is called a ***homogeneous linear system (of equations)***. It is not hard to see that for any homogeneous linear system, it always has a trivial solution of $x_j = 0$ for $j = 1, \ldots, n$, or expressed as $\vec{x} = \mathbf{0}$. However, such trivial solution may not be the only solution to the system, as we shall see in !???.

Below shows some examples of linear systems.

$$
\begin{cases}
3x + 4y & = 5 \\
7x + 9y & = 13
\end{cases}
$$

A 2×2 linear system with two equations, two unknowns.

$$
\begin{cases}
x + 2y - 4z & = 3 \\
x - y + 3z & = -4
\end{cases}
$$

A 2×3 linear system with two equations, three unknowns.

$$
\begin{cases}
x + 2y + 3z & = 0 \\
2x + 3z & = 0 \\
4x - 5y & = 0
\end{cases}
$$

A 3×3 homogeneous linear system (homogeneous as the R.H.S. are all zeros), notice the coefficients for some unknowns in some equations are zeros as well, e.g. in the second equation y has a coefficient of zero and does not appear.

The above formulation of a linear system closely resembles a tabular structure. Therefore, we are motivated to represent such systems with the language of

matrices. Indeed, it is possible to rewrite an $m \times n$ linear system as $A\vec{x} = \vec{h}$, where A is an $m \times n$ matrix with entries copied from the coefficients in front of the variables arranged like those in Definition 1.2.2. In this book sometimes we will call it a *coefficient matrix*. Meanwhile, $\vec{x}$ is a column vector (an $n \times 1$ matrix) holding the n unknowns, and $\vec{h}$ is another column vector (an $m \times 1$ matrix) that contains the m constants at the right hand side of the linear system.

Properties 1.2.3. For a linear system like that in Definition 1.2.2, it can be rewritten as $A\vec{x} = \vec{h}$, where $A_{ij} = a_j^{(i)}$, $\vec{x} = x_j$, and $\vec{h} = h^{(i)}$.

Using the second example above as an illustration, we can easily verify that

$$\begin{cases} x + 2y - 4z & = 3 \\ x - y + 3z & = -4 \end{cases}$$

can be expressed as (you should check it yourself)

$$\begin{bmatrix} 1 & 2 & -4 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

An even simpler representation is the ***augmented matrix*** which omits the unknowns and concatenates the remaining matrices.

$$\left[\begin{array}{ccc|c} 1 & 2 & -4 & 3 \\ 1 & -1 & 3 & -4 \end{array} \right]$$

1.3 Elementary Row Operations

When we construct a matrix, it is natural to think about how to utilize its structure. *Elementary row operations* provide such possibility in three ways, outlined in the following definition.

Definition 1.3.1 (Elementary Row Operations). The three types of elementary row operations are

1. Multiplying a row R_p by any constant $c \neq 0$.

2. Adding a row R_q times any constant $c \neq 0$, to another row R_p, such that the new row becomes $R_p + cR_q$

3. Swapping a row R_p with another row R_q.

To facilitate the operations, we mark these three actions using the following notations.

1. $cR_p \rightarrow R_p$,

2. $R_p + cR_q \rightarrow R_p$,

3. $R_p \leftrightarrow R_q$

For example, the matrix A

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 11 \end{bmatrix}$$

can be transformed to a new matrix A'

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 5 \end{bmatrix}$$

if we apply the elementary row operation, subtracting $2R_1$ from R_2 (i.e. $R_2 - 2R_1 \rightarrow R_2$).

Short Exercise: Find out the resulting matrix A'' if we multiply the first row of A' by 3 and then subtract the second row from the first row.[5]

Attentive readers may have noticed that these three operations are what we have been always doing to equations when solving a linear system, as taught in high

---

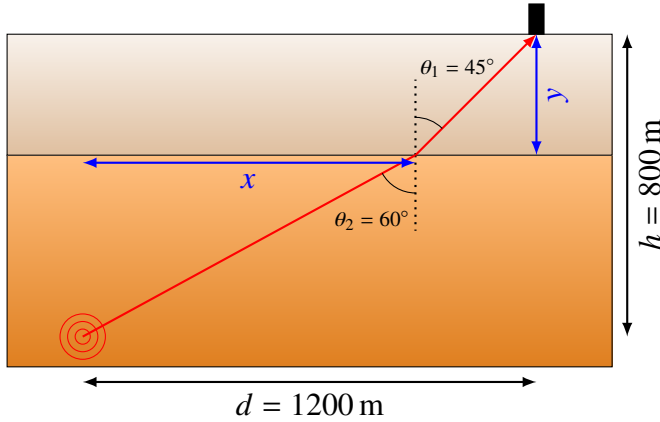[5] $\begin{bmatrix} 0 & 3 & 4 \\ 3 & 3 & 5 \end{bmatrix}$

Figure 1.1: The underground schematic for the seismic ray in Example 1.4.1.

school. We re-introduce them as elementary row operations here first as they are fundamental to the treatment of later chapters.

1.4 Earth Science Applications

Example 1.4.1. Seismic wave follows *Snell's Law* like a light ray when it comes to refraction. Assuming the ground can be modelled as a two-layer system (see Figure 1.1), and we know a particular train of seismic wave generated from an underground source that reaches the ground receiver travels at an angle of $\theta_1 = 45°/\theta_2 = 60°$ to the vertical at the top/bottom layer. Given that the horizontal and vertical distance between the seismic source and the surface receiver are $d = 1200$ m and $h = 800$ m. Construct a linear system for this situation in two unknowns: the depth of the top layer y and the horizontal displacement x (in meters) where the wave reaches at the interface relative to the source.

Solution. We can deduce two equations from the given information. Consider

the upper portion of the seismic ray, from basic trigonometry, we know that

$$\frac{d - x}{y} = \tan \theta_1$$
$$d - x = (\tan \theta_1)y$$
$$x + (\tan \theta_1)y = d$$

Similarly, for the lower portion of the seismic ray, we have

$$\frac{x}{h - y} = \tan \theta_2$$
$$x = (\tan \theta_2)h - (\tan \theta_2)y$$
$$x + (\tan \theta_2)y = (\tan \theta_2)h$$

The corresponding linear system is

$$\begin{cases} x + (\tan \theta_1)y & = d \\ x + (\tan \theta_2)y & = (\tan \theta_2)h \end{cases}$$

where x and y are the unknowns to be solved. d, h, θ_1 and θ_2 (and hence $\tan \theta_1$ and $\tan \theta_2$) are constants. Expressing the system in matrix form, we have

$$\begin{bmatrix} 1 & \tan \theta_1 \\ 1 & \tan \theta_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} d \\ (\tan \theta_2)h \end{bmatrix}$$

Substituting the provided values for the constants ($\tan \theta_1 = \tan(45°) = 1$, $\tan \theta_2 = \tan(60°) = \sqrt{3}$), we have

$$\begin{bmatrix} 1 & 1 \\ 1 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1200 \\ 800\sqrt{3} \end{bmatrix}$$

$\square$

Example 1.4.2. The radiation transfer across the atmosphere of any planet (including the Earth) in the Solar system can be compared to a *multi-layer model* with fully absorbing layers (note that it is just a simplistic approach). Assume

there are N such layers and the total rate of incident Solar radiation reaching the surface is E_{in}. Each of the layers also emits radiation to the other layers directly above/below itself. The rate of emission for the j-th layer that has a temperature T_j is $E_j = \sigma T_j^4$ according to the *Stefan–Boltzmann Law*, with $\sigma = 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$. The overall scenario can be seen in Figure 1.2. Formulate a linear system that represents the energy equilibrium (incoming radiation = outgoing radiation) of all layers and the surface, with E_j being the unknowns, over $j = 1, 2, \ldots, N, N + 1$.

Solution. Considering the energy equilibrium for the first (topmost) layer, we have

$$-2E_1 + E_2 = 0$$

Going down to the second layer, it is

$$E_1 - 2E_2 + E_3 = 0$$

In general, for the j-th layer in the middle, where j runs from 2 to N, we can similarly obtain

$$E_{j-1} - 2E_j + E_{j+1} = 0$$

Finally, for the surface (the $N + 1$-th layer), we have

$$E_N - E_{N+1} + E_{in} = 0$$
$$E_N - E_{N+1} = -E_{in}$$

Summarizing all the $N + 1$ equations, they can be expressed in matrix form as

$$
\begin{bmatrix}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 \\
1 & -2 & 1 & & 0 & 0 & 0 \\
0 & 1 & -2 & & 0 & 0 & 0 \\
\vdots & & & \ddots & & & \vdots \\
0 & 0 & 0 & & -2 & 1 & 0 \\
0 & 0 & 0 & & 1 & -2 & 1 \\
0 & 0 & 0 & \cdots & 0 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
E_1 \\
E_2 \\
E_3 \\
\vdots \\
E_{N-1} \\
E_N \\
E_{N+1}
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
0 \\
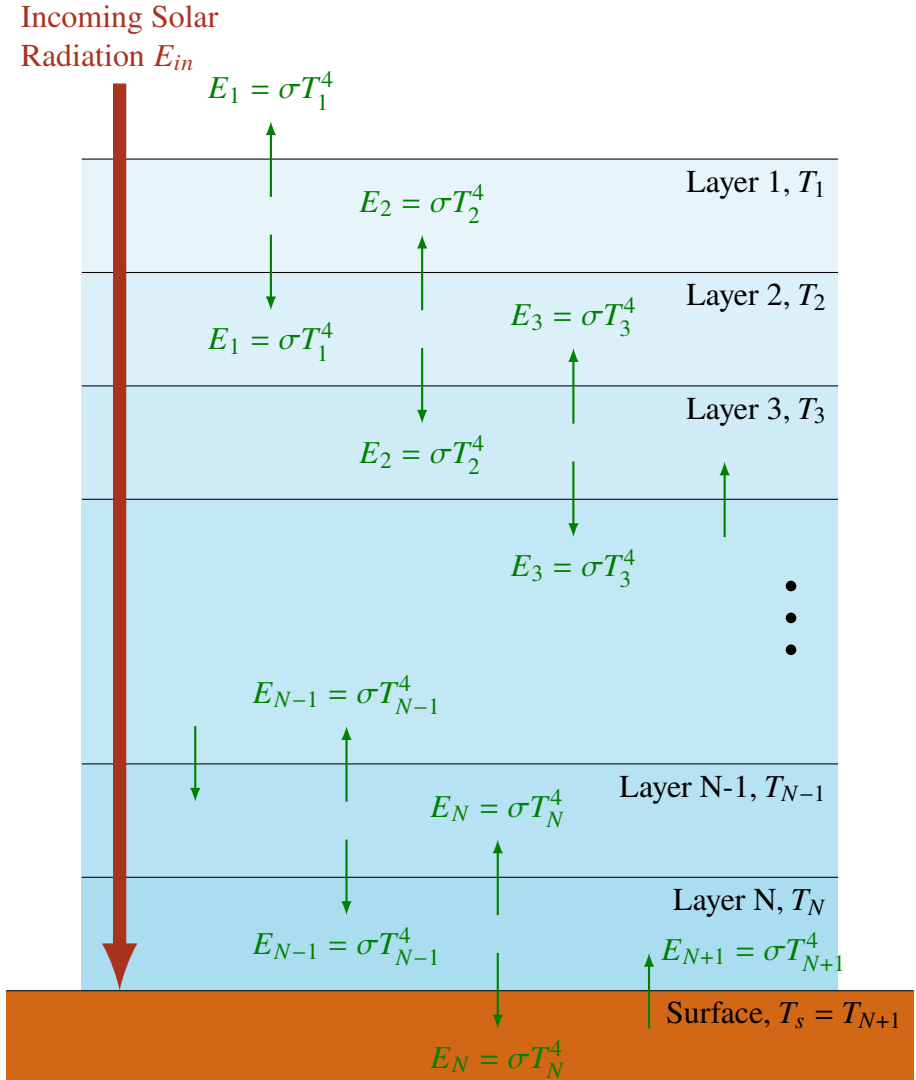\vdots \\
0 \\
0 \\
-E_{in}
\end{bmatrix}
$$

21

Figure 1.2: The atmospheric profile with multiple (N) absorbing layers in Example 1.4.2. The surface is treated as an extra $N+1$-th layer.

Particularly, for $N = 4$, it is

$$\begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -E_{in} \end{bmatrix}$$

$\square$

We will talk about how to solve the linear systems in these two examples in Section 3.3.

1.5 Python Programming

We will use the package `numpy` and `scipy` throughout the book to solve linear algebra problems via *Python* programming. First, we can define a 2D `numpy` array that works as a matrix.

```python
import numpy as np
myMatrix1 = np.array([[1, 4], [5, 3]])
print(myMatrix1)
```

which gives

```
[[1 4]
 [5 3]]
```

representing the matrix

$$\begin{bmatrix} 1 & 4 \\ 5 & 3 \end{bmatrix}$$

We can similarly define another matrix

```python
myMatrix2 = np.array([[1, 3], [5, 6]])
```

Addition, subtraction, and scalar multiplication are straight-forward.

```
myMatrix3 = 3*myMatrix1 - 4*myMatrix2
print(myMatrix3)
```

The above code produces

```
[[ -1   0]
 [ -5 -15]]
```

and you can verify the answer by hand. Meanwhile, matrix product is done by the function `np.matmul()`.

```
myMatrix4 = np.matmul(myMatrix1, myMatrix2) # or equivalently
    myMatrix1 @ myMatrix2
print(myMatrix4)
```

gives

```
[[21 27]
 [20 33]]
```

To select a specfic entry, use indexing by square brackets. The first index/second index represents row/column. Beware that each index starts at zero in *Python*. So putting the number 1 in the first/second index actually means the second row/column. So

```
print(myMatrix4[1,0])
```

refers to the entry at row 2, column 1 of `myMatrix4` which is 20. Also, we can select the *i*-th row (or the *j*-th column) by `<Matrix>[i-1, :]` (`<Matrix>[:, j-1]`), where the colon `:` implies selecting along the entire row (column). For example,

```
print(myMatrix3[0,:])
print(myMatrix4[:,1])
```

gives `[-1   0]` and `[27 33]` respectively. Now let's see how to perform elementary row operations. It will be easier and less error-prone if we copy the array before performing the operations.

```
myMatrix5 = np.copy(myMatrix4)
myMatrix5[0,:] = myMatrix5[:,:]/3
print(myMatrix5)
```

The lines above, when executed, divide the second row of `myMatrix5` (which is a copy of `myMatrix4`) by 3, and give

```
[[ 7  9]
 [20 33]]
```

Meanwhile, the subsequent lines below

```
myMatrix5[1,:] = myMatrix5[1,:] - 2*myMatrix5[0,:]
print(myMatrix5)
```

proceed to subtract 2 times the first row from the second row, and produce

```
[[ 7  9]
 [ 6 15]]
```

Row interchange is a bit more tricky.

```
myMatrix6 = np.copy(myMatrix4)
myMatrix6[[0, 1],:] = myMatrix6[[1, 0],:]
```

This swaps the first and second row. (You can swap columns in a similar way.) Printing out the new matrix by `print(myMatrix6)` shows

```
[[20 33]
 [21 27]]
```

An important pitfall is that, since our inputs to `np.array` are all integers, the previous arrays will automatically have a data type of `int` (integer). This may produce unexpected errors when the calculation leads to decimals/fractions. If it is the case, then we can avoid such bugs by declaring the array with the keyword `dtype=float` to use *floating point numbers*, like

```
myMatrix1 = np.array([[1, 4], [5, 3]], dtype=float)
```

when printed out via `print(myMatrix1)` it gives

```
[[1. 4.]
 [5. 3.]]
```

Notice the newly appeared decimal points after the original integers. Alternatively, we can add decimal points to the integer entries during the array declaration, as

```
myMatrix1 = np.array([[1., 4.], [5., 3.]])
```

1.6 Exercises

Exercise 1.1 Let

$$A = \begin{bmatrix} 1 & 2 \\ 5 & -1 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} -4 & 3 \\ -2 & 7 \end{bmatrix}$$

Find:

(a) $A + B$,

(b) $2A - \frac{3}{2}B$,

(c) AB,

(d) BA.

Exercise 1.2 Let

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

Find:

(a) AB,

(b) BA.

Exercise 1.3 Let

$$A = \begin{bmatrix} 4 & 6 \\ 3 & 3 \end{bmatrix}$$

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

$$C = \begin{bmatrix} 3 & 9 & 1 \\ 4 & 3 & -1 \end{bmatrix}$$

Find:

(a) $(A + B)C$,

(b) $AC + BC$,

(c) $(AB)C$,

(d) $A(BC)$.

Exercise 1.4 Let

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 7 & 2 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 5 & -2 \\ 4 & 3 & 1 \\ 0 & 2 & 3 \end{bmatrix}$$

Find:

(a) $(A + B)(2A - B)$,

(b) $(\frac{3}{2}A - B)(-A + \frac{1}{2}B)$.

Exercise 1.5 Let

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 6 \\ 5 & 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 3 & 8 \\ 4 & 0 & 7 \end{bmatrix}$$

Find:

(a) A^2,

(b) B^2,

(c) AB,

(d) BA.

Exercise 1.6 Let

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$$

Show that $AB = BA$ in this case.

Exercise 1.7 Rewrite the following system of linear equations into matrix form.

$$\begin{cases} 3y - 4z & = 6 \\ 5x - y + 2z & = 13 \\ 6x + z & = 8 \end{cases}$$

Exercise 1.8 For the following matrix,

$$\begin{bmatrix} 2 & 3 & 5 & 7 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 6 & 10 \end{bmatrix}$$

Find the results if the following elementary row operations are applied on it:

(a) Multiplying the third row by 2, and then subtracting the third row by the second row,

(b) Adding the first row by 3 times the third row, and then interchanging the first and second row.

Exercise 1.9 For the following matrix,

$$\begin{bmatrix} 3 & 0 & 4 & 6 & 9 \\ 5 & 3 & 8 & -1 & 3 \\ 2 & 5 & 4 & 3 & -7 \end{bmatrix}$$

Find the elementary row operations needed to reduce the matrix to

$$\begin{bmatrix} 1 & \frac{5}{2} & 2 & \frac{3}{2} & -\frac{7}{2} \\ 5 & \frac{21}{2} & 10 & -\frac{5}{2} & -\frac{33}{2} \\ 3 & 0 & 4 & 6 & 9 \end{bmatrix}$$

Exercise 1.10 The *dry adiabatic lapse rate*, which is the rate of decrease in air temperature when an unsaturated air parcel rises, is about $\Gamma_{dry} = 9.8\,°\text{C}\,\text{km}^{-1}$. When the temperature of the air parcel falls below the *dew point*, the air saturates and condensation occurs. Typically, dew point temperature of an air parcel will decrease at a rate of roughly $\Gamma_{dew} = 2\,°\text{C}\,\text{km}^{-1}$. Now, an air parcel with an initial air temperature/dew point temperature of $T_{a,ini} = 25.4\,°\text{C}\,/\,T_{dew,ini} = 17.8\,°\text{C}$ at the ground starts to rise. Let z_{cd} and T_{cd} be the height above the ground (in km) and temperature (in °C) of the air parcel when condensation occurs. Construct a linear system with z_{cd} and T_{cd} as the unknowns to represent this situation.

Exercise 1.11 In some ancient Chinese Mathematics texts, the problem of *Chickens and Rabbits in the Same Cage* was posed. *"Now there are some chickens and rabbits placed in the same cage, with a total number of* 35 *heads and* 94 *legs. How many chickens and rabbits are there respectively?"* Given the fact that a chicken (rabbit) has two (four) legs (and obviously only one head), write down the corresponding linear system in terms of the numbers of chickens *x* and rabbits *y*.

Inverses and Determinants

In this chapter, we are going to discuss two important concepts about matrices, which are the *Inverse* and *Determinant*. To derive them, we need to introduce some prerequisite ideas first, including *Identity*, *Transpose*, and the methods of *Gaussian Elimination* and *Laplace Expansion*.

2.1 Identity Matrices and Transpose

2.1.1 Identity Matrices

One important type of matrices is the ***Identity Matrices***. They are $n \times n$ square matrices, where n can be any positive integer, with entries along the ***Main Diagonal*** (where index of row = column) being 1 and other off-diagonal elements being 0. Usually, they are denoted by I_n, or simply I.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad\qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Identity matrices of size 2×2 and 3×3 with the main diagonal 1s highlighted.

Definition 2.1.1 (Identity Matrix). An identity matrix of the square shape $n \times n$ I_n is defined as $[I_n]_{ij} = 1$, for $i = j$, and $[I_n]_{ij} = 0$, for $i \neq j$, where $1 \leq i, j \leq n$.

Short Exercise: Explicitly write down I_5.[1]

One important property of identity matrices is

Properties 2.1.2. Matrix product between any matrix A with an identity matrix I always produces A whenever the matrix product is defined. If A is of the shape $m \times n$, then $AI_n = I_m A = A$. If A is now a square matrix such that $m = n$ and we let $I_m = I_n = I$, then $AI = IA = A$.

In other words, the identity I can be regarded to be the "1" in the world of matrices. This is one of the cases that $AB = BA$ commutes (if A is a square matrix and $B = I$). Using the matrix

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

as an example, the readers can try to computed AI_2 and $I_2 A$ to see if the results are A itself.

2.1.2 Transpose

Transpose of a matrix, denoted by adding the superscript ^{T}, is formed by interchanging its rows and columns. In the special case of square matrix, this operation can be viewed as flipping the elements about the main diagonal.

Definition 2.1.3 (Transpose)**.** The transpose of an $m \times n$ matrix A, denoted as A^T, is formed according to the relation $A^T_{pq} = A_{qp}$, $1 \le p \le m$, $1 \le q \le n$, i.e. switching the row and column indices. Now A^T is an $n \times m$ matrix.

$$[1] I_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Two examples are given below to show the outcome of applying transpose on matrices.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad\qquad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & -4 & 3 \\ 2 & -2 & 0 \\ -3 & 1 & 4 \end{bmatrix} \qquad\qquad B^T = \begin{bmatrix} 1 & 2 & -3 \\ -4 & -2 & 1 \\ 3 & 0 & 4 \end{bmatrix}$$

Particularly, in the second example, we have highlighted the main diagonal of B (as well as B^T) and how the elements flip about it when transpose is carried out. Some useful properties about transpose are listed as follows.

Properties 2.1.4. For two matrices A and B, we have

1. $(cA)^T = cA^T$, where c is any constant,

2. $(A^T)^T = A$, i.e. transposing twice returns the original matrix (quite obvious),

3. $(A \pm B)^T = A^T \pm B^T$, if A and B have the same shape,

4. $(AB)^T = B^T A^T$, if their matrix product is valid,

5. $A_{kk} = A^T_{kk}$ for any k that A_{kk} is defined, i.e. the main diagonal is unaffected by transpose.

Short Exercise: Show that $(ABC)^T = C^T B^T A^T$ if the matrices have compatible shapes for the matrix product.[2]

2.1.3 Symmetric Matrices

A **Symmetric Matrix** has its elements mirrored about the main diagonal. Taking transpose of such a matrix will leave it unchanged. Implicitly, it is required to

---

[2]By (4), $(ABC)^T = ((AB)(C))^T = C^T (AB)^T = C^T B^T A^T$

be a square matrix.

Definition 2.1.5 (Symmetric Matrix). If an $n \times n$ square matrix A and its transpose A^T are the same, i.e. $A_{pq} = A_{pq}^T = A_{qp}$ for all $1 \le p, q \le n$, or simply $A = A^T$, then A, and also A^T, are symmetric.

As an example,

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

is a 3×3 symmetric matrix.

Short Exercise: Show that $Y = XX^T$ and $Z = X^T X$ are symmetric for any matrix X.[3]

In contrast, we have ***Skew-symmetric Matrices*** such that $A^T = -A$. This automatically requires that elements along the main diagonal to be all zeros.

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

A 3×3 skew-symmetric matrix.

2.2 Inverses

2.2.1 Definition and Properties of Inverses

Inverse of a square matrix, denoted by adding the superscript $^{-1}$, is another square matrix such that their matrix product yields an identity matrix.

---

[3]By Properties 2.1.4, $Y^T = (XX^T)^T = (X^T)^T (X)^T = XX^T = Y$, similar goes for $Z = X^T X$.

Definition 2.2.1 (Inverse). An $n \times n$ square matrix B is said to be the inverse of another $n \times n$ square matrix A if $AB = BA = I_n$. Henceforth, we write $B = A^{-1}$, and $AA^{-1} = A^{-1}A = I$. The opposite direction also holds, i.e. A is the inverse of $B = A^{-1}$. We say that $A = B^{-1}$ and $B = A^{-1}$ are the inverse of each other.

If there exists an inverse A^{-1} for the square matrix A, then both A and A^{-1} are called **invertible**. Otherwise, A is said to be **singular**. This is another situation in which a matrix product $AB = BA$ (if $B = A^{-1}$) can commute. We can show why $AA^{-1} = I$ implies $A^{-1}A = I$, or vice versa.

Proof. Assume only $AA^{-1} = I$ is true, then multiplying A to the right on both sides of the equation leads to

$$AA^{-1}A = IA$$
$$A(A^{-1}A) = A \qquad \text{(Properties 1.1.2 and Properties 2.1.2)}$$
$$AP = A$$

where we write $P = A^{-1}A$. The above implies that multiplying A by P returns A itself. We are tempted to claim that $P = I$ by observing Properties 2.1.2. However, it is not trivial to show that P cannot be any matrix other than I, and to do so we have to wait until later chapters. Nevertheless, it is indeed true as long as A is invertible. Therefore, $P = A^{-1}A = I$. The opposite direction is proved similarly. $\square$

The inverse operation can be somehow viewed as taking the reciprocal in the world of matrices. In addition,

---

Properties 2.2.2 (Uniqueness of Inverse). If A has an inverse A^{-1}, it is unique.

Proof. This property can be proved easily by first assuming that the invertible matrix A has two different inverses, B and C. Subsequently, by Definition 2.2.1, we have $BA = I$ (and also $AC = I$). Multiplying by C to the right on both sides gives

$$BAC = IC$$

$$B(AC) = C \qquad \text{(Properties 1.1.2 and 2.1.2)}$$
$$B(I) = C \qquad (AC = I \text{ from assumption})$$
$$B = C \qquad \text{(Properties 2.1.2)}$$

So, B and C are actually the same matrix, implying that the inverse of A is unique. □

Example 2.2.1. Let

$$A = \begin{bmatrix} 4 & 6 \\ 3 & 5 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} \frac{5}{2} & -3 \\ -\frac{3}{2} & 2 \end{bmatrix}$$

Show that A and B are inverse to each other.

Solution.

$$
\begin{aligned}
AB &= \begin{bmatrix} 4 & 6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \frac{5}{2} & -3 \\ -\frac{3}{2} & 2 \end{bmatrix} \\
&= \begin{bmatrix} (4)(\frac{5}{2}) + (6)(-\frac{3}{2}) & (4)(-3) + (6)(2) \\ (3)(\frac{5}{2}) + (5)(-\frac{3}{2}) & (3)(-3) + (5)(2) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2
\end{aligned}
$$

It is easy to verify that $BA = I$ and the readers are invited to do so. Hence $AB = BA = I$, A and B are indeed the inverse of each other. □

The followings are some properties of inverses.

Properties 2.2.3. If a square matrix A is invertible and has an inverse A^{-1}, then

1. $(cA)^{-1} = \frac{1}{c}A^{-1}$, for any constant $c \neq 0$,

2. $(A^{-1})^{-1} = A$, i.e. the inverse of an inverse is the original matrix,

3. $(A^n)^{-1} = (A^{-1})^n$, for any positive integer n,

4. $(AB)^{-1} = B^{-1}A^{-1}$, provided that B is invertible too,

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

However, $(A \pm B)^{-1}$ may not be equal to $A^{-1} \pm B^{-1}$, or even may be singular. We shall briefly prove (4) here.

Proof. It is given that A and B is invertible, and by Definition 2.2.1, we have $AA^{-1} = I$ and

$$BB^{-1} = I$$

Multiplying by A and A^{-1} to the left and right on both sides yields

$$ABB^{-1}A^{-1} = AIA^{-1}$$
$$AB(B^{-1}A^{-1}) = (AI)A^{-1} = AA^{-1} \qquad \text{(Properties 1.1.2 and 2.1.2)}$$
$$= I \qquad\qquad\qquad \text{(Definition 2.2.1)}$$

This shows that multiplying AB by $B^{-1}A^{-1}$ produces an identity matrix, and therefore $(AB)^{-1} = B^{-1}A^{-1}$ is the unique inverse of AB by Definition 2.2.1 and Properties 2.2.2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Short Exercise: Show that $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ if A, B and C are invertible.[4]

(4) of Properties 2.2.3 explicitly shows that the product AB is invertible if A and B are themselves invertible. We will show that the converse is true as well. Hence

Properties 2.2.4. For two square matrices A and B, AB is invertible if and only if A and B are invertible.

---

[4]By (4), $(ABC)^{-1} = ((AB)(C))^{-1} = C^{-1}(AB)^{-1} = C^{-1}B^{-1}A^{-1}$

Proof. The "if" part is just (4) of Properties 2.2.3 and the "only if" part will be proved as follows. Let's assume AB is invertible and has an inverse $C = (AB)^{-1}$, hence we have $(AB)C = I$ by Definition 2.2.1, (notice that A, B, and C are all square matrices of the same extent) and by Properties 1.1.2, $A(BC) = I$. Using Definition 2.2.1 (as well as Properties 2.2.2) again, we immediately identify BC as the inverse of A and A is invertible. The case for B is similarly proved. $\quad\square$

2.2.2 (Reduced) Row Echelon Form

To find the inverse of any matrix, we have to understand a form of matrices called *(Reduced) Row Echelon Form* first. The requirements of a matrix being in such a form are shown below.

Definition 2.2.5 ((Reduced) Row Echelon Form)**.** A matrix is in row echelon form if

1. The first non-zero number in every row is 1, which is known as the *"Leading 1"*,

2. *"Leading 1"* of a lower row must appear farther to the right than that of any higher row,

3. Any row consisted of all zeros is placed at the bottom;

4. If additionally, any column containing a leading 1 have zeros elsewhere in that column, then it is in *reduced* row echelon form.

It is apparent that all identity matrices are in (reduced) row echelon form. Examples of row echelon form (but not *reduced*), with the leading 1s highlighted are

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 3 & 1 & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 4 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Meanwhile, examples of *reduced* row echelon form are

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The following matrices are *not* in row echelon form.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

Short Exercise: Decide if the following matrices are in (reduced) row echelon form or not.[5]

$$X = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

The next goal is to transform the matrix in question to reduced row echelon form by elementary row operations, whose importance will be demonstrated soon. The procedure is comprised of two major parts, *Forward Elimination*, converting the matrix into row echelon form, and *Backward Elimination*, further converting it into reduced row echelon form. The first phase is also named **Gaussian Elimination**, and together they are called **Gauss-Jordan Elimination**. We demonstrate the entire procedure using an example.

---

[5]Yes, Yes (reduced), No.

Example 2.2.2. Carry out Gauss-Jordan Elimination on the following matrix.

$$A = \begin{bmatrix} 2 & 0 & 4 & 6 \\ 3 & 3 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

Solution. At each step of the forward phase, the strategy is to look at on the leftmost column that has at least one non-zero entries first (any column consisting of full zeros is ignored). Along that column, we either find an existing leading 1, or create a leading 1 via multiplying a suitable row having a starting entry a (Leading entries selected by the algorithm are known as **Pivots**, and the process is called **Pivoting**) that is as large as possible in magnitude, by the constant $1/a$. The row holding the leading 1 is subsequently put at the top, by a row interchange if needed. Such a row is highlighted in red during this example.

$$\begin{bmatrix} 2 & 0 & 4 & 6 \\ 3 & 3 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 3 & 3 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix} \qquad \frac{1}{2}R_1 \rightarrow R_1$$

We have picked R_1 here but a leading 1 can be obtained from R_2 or R_3 as well. Subsequently, we make all the elements below the leading 1 along that column become zero, by adding the top row (holding the leading 1), times $-a_i$ (where a_i is the corresponding leading entry of row i) to other rows. Those zeros will be highlighted in blue.

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 3 & 3 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 3 & -5 & -9 \\ 1 & 2 & 3 & 4 \end{bmatrix} \qquad R_2 - 3R_1 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 3 & -5 & -9 \\ 0 & 2 & 1 & 1 \end{bmatrix} \qquad R_3 - R_1 \rightarrow R_3$$

The first iteration is finished. We now repeat the same process over the remaining

submatrix made up of elements that are not yet highlighted in colour, recursively.

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 3 & -5 & -9 \\ 0 & 2 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -\frac{5}{3} & -3 \\ 0 & 2 & 1 & 1 \end{bmatrix} \qquad \frac{1}{3}R_2 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -\frac{5}{3} & -3 \\ 0 & 0 & \frac{13}{3} & 7 \end{bmatrix} \qquad R_3 - 2R_2 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -\frac{5}{3} & -3 \\ 0 & 0 & 1 & \frac{21}{13} \end{bmatrix} \qquad \frac{3}{13}R_3 \rightarrow R_3$$

Now, below every leading 1, all entries are zeros, indicating the forward phase is completed. We have obtained the row echelon form as an intermediate. The backward phase is done similarly but in a bottom up fashion. By adding some multiples of lower rows to higher rows, we turn all the non-zero elements above any leading 1 into zeros.

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -\frac{5}{3} & -3 \\ 0 & 0 & 1 & \frac{21}{13} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & -\frac{4}{13} \\ 0 & 0 & 1 & \frac{21}{13} \end{bmatrix} \qquad R_2 + \frac{5}{3}R_3 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{13} \\ 0 & 1 & 0 & -\frac{4}{13} \\ 0 & 0 & 1 & \frac{21}{13} \end{bmatrix} \qquad R_1 - 2R_3 \rightarrow R_1$$

The matrix is now in reduced row echelon form as required. $\qquad \square$

Short Exercise: Repeat the example but start by interchanging R_1 and R_3.[6]

---

[6]For checking, after the first iteration, it will be

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -8 & -12 \\ 0 & -4 & -2 & -2 \end{bmatrix}$$

and the end result will be the same.

From the short exercise above, we can see that even if we apply different elementary row operations (particularly for the creation of a leading 1) during Gauss-Jordan Elimination, we will acquire the same reduced echelon form in the end. In fact,

Theorem 2.2.6 (Uniqueness of Reduced Row Echelon Form)**.** Reduced row echelon form of a matrix is unique.

We shall omit the proof here. The following properties further reveal how elementary row operations are associated with reduced row echelon form.

Properties 2.2.7. If a matrix can be transformed into another matrix by elementary row operations, they are said to be ***Row Equivalent***.

Since for any pair of row equivalent matrices, either of them can be transformed into the other one by elementary row operations, and hence can be further transformed into the reduced row echelon form of the other matrix, by Theorem 2.2.6, the uniqueness of reduced row echelon form implies that

Properties 2.2.8. Row equivalent matrices have the same reduced row echelon form. Particularly, they are row equivalent to their own reduced row echelon form. If two matrices have different reduced row echelon forms, then they are not row equivalent, and vice versa.

Let's go through one more simple example about Gauss-Jordan Elimination.

Example 2.2.3. Transform the following matrix into reduced row echelon form.

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 6 & 4 & 1 \\ 2 & 3 & 2 \\ 2 & 1 & 0 \end{bmatrix}$$

Solution. One possible way to do the forward elimination is

$$
\begin{bmatrix} 2 & 2 & 1 \\ 6 & 4 & 1 \\ 2 & 3 & 2 \\ 2 & 1 & 0 \end{bmatrix}
\rightarrow
\begin{bmatrix} 6 & 4 & 1 \\ 2 & 2 & 1 \\ 2 & 3 & 2 \\ 2 & 1 & 0 \end{bmatrix}
\qquad\qquad R_1 \leftrightarrow R_2
$$

$$
\rightarrow
\begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} \\ 2 & 2 & 1 \\ 2 & 3 & 2 \\ 2 & 1 & 0 \end{bmatrix}
\qquad\qquad \frac{1}{6}R_1 \rightarrow R_1
$$

$$
\rightarrow
\begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} \\ 0 & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{5}{3} & \frac{5}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}
\qquad\qquad R_i - 2R_1 \rightarrow R_i, \text{ for } i = 2, 3, 4
$$

$$
\rightarrow
\begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} \\ 0 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{5}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}
\qquad\qquad \frac{3}{2}R_2 \rightarrow R_2
$$

$$
\rightarrow
\begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\qquad R_3 - \frac{5}{3}R_2 \rightarrow R_3, R_4 + \frac{1}{3}R_2 \rightarrow R_4
$$

The backward elimination is simple.

$$
\begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\rightarrow
\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\qquad R_1 - \frac{2}{3}R_2 \rightarrow R_1
$$

□

2.2.3 Finding Inverses by Gaussian Elimination

With Gaussian Elimination, obtaining the inverse A^{-1} of any invertible matrix A is straight-forward. We start by writing out an identity matrix I of the same extent and concatenate this identity matrix to the right of A, leading to a form of $[A|I]$. Then we carry out elementary row operations simultaneously on both sides of $[A|I]$ such that the matrix to the left, originally as A, is reduced to the identity matrix I by Gaussian Elimination. The identity matrix to the right will then be transformed into the desired inverse by the same set of elementary operations, such that the concatenated matrix will appear as $[I|A^{-1}]$,

Example 2.2.4. Find the inverse of

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

Solution. Appending an 3×3 identity matrix to the right, we have

$$\left[\begin{array}{ccc|ccc} 1 & 4 & 5 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \qquad R_2 - 2R_3 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right] \qquad R_2 \leftrightarrow R_3$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right] \qquad R_2 - R_3 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 0 & 1 & -5 & 10 \\ 0 & 1 & 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right] \qquad R_1 - 5R_3 \rightarrow R_1$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right] \qquad R_1 - 4R_2 \rightarrow R_1$$

Hence the required inverse is

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

□

Short Exercise: Find the inverse of A^{-1} above by the same elimination method but in an opposite fashion. [7]

The underlying reason why the above procedure can produce the inverse matrix is the equivalence between elementary row operations and multiplication by appropriate ***Elementary Matrices***.

Theorem 2.2.9 (Elementary Matrices)**.** Any elementary row operation can be represented by multiplying to the left with a suitable elementary matrix. Such a matrix is essentially the one formed after applying that particular elementary row operation on an identity matrix. For the three types of elementary row operations in Definition 1.3.1

1. $cR_p \rightarrow R_p$, $c \neq 0$,

2. $R_p + cR_q \rightarrow R_p$,

3. $R_p \leftrightarrow R_q$

The corresponding elementary matrices E are square, and *invertible* (see the remark below) matrices with

1. $E_{kk} = 1$ for any k, except $E_{pp} = c$,

2. $E_{kk} = 1$ for all k, with $E_{pq} = c$,

3. $E_{kk} = 1$ for any k, except $E_{pp} = 0$ and $E_{qq} = 0$, with $E_{pq} = E_{qp} = 1$.

Entries not mentioned are all zeros.

---

[7]You should be able to retrieve the matrix A back.

Since it is quite abstract, it is useful to have some actual examples.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Multiplying } R_2 \text{ by a factor of } 2, 2R_2 \rightarrow R_2$$

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Adding 3 times } R_2 \text{ to } R_1, R_1 + 3R_2 \rightarrow R_1$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \text{Swapping } R_1 \text{ and } R_3, R_1 \leftrightarrow R_3$$

Any elementary row operation can be apparently undone by an inverse elementary row operation (addition vs subtraction, multiplication vs division ($c \neq 0$), swapping twice). Accordingly, any elementary matrix has another corresponding elementary matrix as its inverse, and the readers are invited to think about it in the exercise below.

Short Exercise: Write down the inverses of the three example elementary matrices above.[8]

For instance, consider a matrix

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

then subtracting R_2 from R_3, $R_3 - R_2 \rightarrow R_3$. can be written as

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

---

[8] $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

Short Exercise: Find out the 3×3 elementary matrix for subtracting 4 times the first row from the third row.[9]

Now we are ready to see why finding inverses by Gaussian Elimination works.

Theorem 2.2.10. If a matrix A can be converted to an identity matrix I as its reduced row echelon form by Gaussian Elimination, then it is invertible since the same steps can in turn be applied on I, producing its inverse A^{-1}.

Using the language of Properties 2.2.8, the matrix A has to be row equivalent to I for A^{-1} to exist. This also means if Gaussian Elimination fails to reduce A to I (i.e. the reduced row echelon form of A is some matrix other than the identity), then A^{-1} does not exist.

Proof. Assume A is invertible and hence $AA^{-1} = I$ (Definition 2.2.1). From Theorem 2.2.9, When doing Gaussian Elimination over A, the i-th elementary row operation executed can be represented by an elementary row matrix, denoted as E_i, $i = 1, 2, \ldots, n$ where n is the total number of steps. If we multiply these E_i successively to the left on both sides of the equation $AA^{-1} = I$, we have

$$(E_n \cdots E_3 E_2 E_1 A) A^{-1} = (E_n \cdots E_3 E_2 E_1) I$$
$$A^{-1} = (I) A^{-1} = (E_n \cdots E_3 E_2 E_1) I$$

$E_n \cdots E_3 E_2 E_1 A = I$ because the elementary row operations E_i reduces A to I during Gaussian Elimination as we demand. With $A^{-1} = E_n \cdots E_3 E_2 E_1 I$, we claim that the same set of elementary row operations, equivalent to E_i, $i = 1, 2, \ldots, n$, can also change I to A^{-1} and thus indeed A is invertible. $\square$

As a corollary, because we have $E_n \cdots E_3 E_2 E_1 A = I$ from above, and all E_i are invertible by Theorem 2.2.9, we can multiply their inverses $E_i' = E_i^{-1}$ (which are

---

[9] $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$

also elementary matrices), to the left on both sides successively, where i runs backwards from n to 1. This leads to

$$(E_1^{-1}E_2^{-1}E_3^{-1}\cdots E_n^{-1})E_n\cdots E_3E_2E_1A = (E_1^{-1}E_2^{-1}E_3^{-1}\cdots E_n^{-1})I$$
$$A = E_1'E_2'E_3'\cdots E_n'$$

as each of $E_n^{-1}E_n$, $E_{n-1}^{-1}E_{n-1}$, ..., $E_1^{-1}E_1$ cancels out to produce I and hence

Properties 2.2.11. All invertible matrices can be written as a product of some sequence of elementary matrices.

2.3 Determinants

2.3.1 Computing Determinants

Determinant of a *square* matrix A, denoted by $\det(A)$ or $|A|$, is a number linked to certain intrinsic properties of the matrix which can help us to find its inverse (determinant of non-square matrix is undefined). Determinant of a 1×1 matrix is equal to the matrix's only entry. Determinants of 2×2 and 3×3 matrices can be calculated by a trick called ***Sarrus' Rule***.

Sarrus' Rule

Properties 2.3.1 (Sarrus' Rule)**.** Determinants of size 2×2 and 3×3 matrices can be found by the Sarrus' Rule. For a 2×2 matrix

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

Its determinant is computed by

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$|A| = a_{11}a_{22} - a_{21}a_{12}$$

which is the product of elements crossed by the red arrow, minus the blue one. Similarly, for a 3×3 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Its determinant is found by

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

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

Example 2.3.1. Find the determinant of the following matrix.

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

Solution. By Sarrus's Rule, we have

$$|A| = \begin{vmatrix} 1 & 2 & 4 \\ -5 & 0 & -3 \\ 4 & 3 & 1 \end{vmatrix}$$
$$= ((1)(0)(1) + (2)(-3)(4) + (4)(-5)(3))$$
$$- ((4)(0)(4) + (3)(-3)(1) + (1)(-5)(2))$$
$$= (0 - 24 - 60) - (0 - 9 - 10)$$

$$= -65$$

□

Cofactor Expansion

Another commonly used method to calculate determinants is *Cofactor Expansion*, also known as *Laplace Expansion*. Before discussing cofactor expansion, it is necessary to know what *cofactors* are.

Definition 2.3.2 (Cofactor and Minor). ***Cofactor*** C_{ij} at (i, j) of a matrix A is simply the determinant of the submatrix formed by deleting the i-th row and j-th column of A, which is called ***Minor*** at (i, j) M_{ij}, times the factor of $(-1)^{i+j}$. Mathematically, $C_{ij} = (-1)^{i+j} M_{ij}$.

The $(-1)^{i+j}$ factor can be visualized as a checkerboard pattern like

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

So, for a matrix like

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 5 & 7 \end{bmatrix}$$

Its cofactor at $(2, 1)$ is

$$C_{21} = (-1)^{(2+1)} \begin{vmatrix} 3 & 5 \\ 5 & 7 \end{vmatrix} \qquad \text{(Definition 2.3.2)}$$

$$= (-1)((3)(7) - (5)(5)) \qquad \text{(Properties 2.3.1)}$$
$$= 4$$

Short Exercise: Find C_{13} and C_{32} for the matrix above.[10]

With ***Cofactor (Laplace) Expansion***, the determinant is computed as the sum of products between each entry and the corresponding cofactor along a picked row/column.

Properties 2.3.3 (Cofactor/Laplace Expansion)**.** The determinant of a $n \times n$ square matrix A, $|A|$, can be found by selecting either a fixed row i, or column j, and adding up the products of every element-cofactor pair along that row/column. For the former case, the determinant is computed as

$$|A| = A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in}$$
$$= \sum_{k=1}^{n} A_{ik}C_{ik}$$

For the latter case, the determinant is similarly found by

$$|A| = A_{1j}C_{1j} + A_{2j}C_{2j} + \cdots + A_{nj}C_{nj}$$
$$= \sum_{k=1}^{n} A_{kj}C_{kj}$$

where C is defined as in Definition 2.3.2. Regardless of the row or column chosen, the result is always the same.

Example 2.3.2. Again, for the matrix

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

---

[10]$C_{13} = -2$, $C_{32} = 4$

Find its determinant via cofactor expansion.

Solution. According to Properties 2.3.3, if we choose the first row, its determinant is

$$|A| = A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13}$$

$$= (1)((-1)^{1+1}\begin{vmatrix} 4 & 6 \\ 5 & 7 \end{vmatrix}) + (3)((-1)^{1+2}\begin{vmatrix} 2 & 6 \\ 3 & 7 \end{vmatrix})$$

$$+ (5)((-1)^{1+3}\begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix}) \qquad \text{(Definition 2.3.2)}$$

$$= (1)(-2) + (3)(4) + (5)(-2) = 0 \qquad \text{(Properties 2.3.1)}$$

$\square$

Short Exercise: Confirm the answer by carrying out cofactor expansion on another row or column.[11]

Example 2.3.3. Find the determinant of

$$A = \begin{bmatrix} 1 & 4 & 4 & 4 \\ 2 & 0 & 4 & 6 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 3 & 1 \end{bmatrix}$$

Solution. It is a 4×4 matrix and we have to apply cofactor expansion. We can choose row or column that contains zero to reduce the computation. Here we pick the second column and by Properties 2.3.3, we have

$$|A| = (-1)^{1+2}(4)\begin{vmatrix} 2 & 4 & 6 \\ 2 & 1 & 0 \\ 6 & 3 & 1 \end{vmatrix} + (-1)^{2+2}(0)\begin{vmatrix} 1 & 4 & 4 \\ 2 & 1 & 0 \\ 6 & 3 & 1 \end{vmatrix}$$

---

[11]You should able to get $|A| = 0$, no matter which row/column is selected.

$$+ (-1)^{3+2}(1)\begin{vmatrix} 1 & 4 & 4 \\ 2 & 4 & 6 \\ 6 & 3 & 1 \end{vmatrix} + (-1)^{4+2}(2)\begin{vmatrix} 1 & 4 & 4 \\ 2 & 4 & 6 \\ 2 & 1 & 0 \end{vmatrix}$$

By Sarrus' Rule (Properties 2.3.1), we have (the detailed calculations are omitted)

$$|A| = (-4)(-6) + 0 + (-1)(50) + (2)(18) = 10$$

□

Finally, we can derive two simple results about determinants from the perspective of cofactor expansion.

Properties 2.3.4. If a matrix have a row/column with full zeros, or two identical/proportional rows/columns, then it has a determinant of zero.

The first case is trivial (just do the expansion along the row/column with full zeros). We will prove the second case alongside the introduction of the properties of determinants coming up in the next subsection.

2.3.2 Properties of Determinants

There are some important results about determinants. First of all, it is very easy to see that determinants for any $n \times n$ identity matrix I_n is just 1. Second, there is a close relation between elementary row operations/elementary matrices and (their effects on) determinants.

Properties 2.3.5. The three types of elementary row operations in Definition 1.3.1, when applied on some matrix A,

1. $cR_p \rightarrow R_p$, $c \neq 0$,

2. $R_p + cR_q \rightarrow R_p$,

3. $R_p \leftrightarrow R_q$,

change the determinant of A by a factor of c, 1, and -1, respectively.

Properties 2.3.6. For the three types of elementary matrices E in Theorem 2.2.9 that correspond to the elementary row operations in Definition 1.3.1,

1. $E_{kk} = 1$ for any k, except $E_{pp} = c$ ($cR_p \to R_p$, $c \neq 0$),

2. $E_{kk} = 1$ for all k, with $E_{pq} = c$ ($R_p + cR_q \to R_p$),

3. $E_{kk} = 1$ for any k, except $E_{pp} = 0$ and $E_{qq} = 0$, with $E_{pq} = E_{qp} = 1$ ($R_p \leftrightarrow R_q$),

their determinants are c, 1, and -1, respectively.

We will prove the above properties for the second kind of elementary row operations/elementary matrices (corresponding to addition/subtraction). The properties for the two other types of elementary matrices are easy to show and we will take them for granted, such that we can complete the second case in Properties 2.3.4 first, which is in turn used for demonstrating the final result.

Proof. Consider an $n \times n$ square matrix A that has two identical and adjacent rows with indices i_1 and i_2, where $i_2 = i_1 + 1$ (hence one of the indices is odd and another is even), then cofactor expansion along the odd row (let's say i_1) will give

$$|A| = \sum_{k=1}^{n} A_{i_1 k} C_{i_1 k}$$

$$= \sum_{k=1}^{n} (-1)^{i_1 + k} A_{i_1 k} M_{i_1 k}$$

by Properties 2.3.3 and Definition 2.3.2. Similarly by considering the even row, we have

$$|A| = \sum_{k=1}^{n} A_{i_2 k} C_{i_2 k}$$

$$= \sum_{k=1}^{n} (-1)^{i_2 + k} A_{i_2 k} M_{i_2 k}$$

But since the i_1-th and i_2-th row are identical, $A_{i_1k} = A_{i_2k}$. Furthermore, as these two identical rows are also adjacent, the minors $M_{i_1k} = M_{i_2k}$ are also equal. The only difference between the two expressions for $|A|$ above is the $(-1)^{i+j}$ factor. And because i_1 is odd and i_2 is even, they are differed by a negative sign only. Explicitly, we have

$$
\begin{aligned}
|A| &= \sum_{k=1}^{n} (-1)^{i_1+k} A_{i_1k} M_{i_1k} \\
&= \sum_{k=1}^{n} (-1)^{(i_2-1)+k} A_{i_2k} M_{i_2k} \\
&= (-1) \sum_{k=1}^{n} (-1)^{i_2+k} A_{i_2k} M_{i_2k} \\
&= -|A|
\end{aligned}
$$

Therefore $|A| = -|A|$ and $|A| = 0$ must equal to zero. Now we can generalize the results to non-adjacent, proportional rows by doing the first and third kind (multiplication and swapping) of elementary row operations when appropriate, and the second case in Properties 2.3.4 is completed. Subsequently, for the addition/subtraction type of elementary row operations, let's say $R_p + cR_q \to R_p$ is applied on some matrix A (this A is not the same one in the first part) to produce A', then

$$
A' = \begin{bmatrix}
\vdots & \vdots & & \vdots \\
A_{p1} + cA_{q1} & A_{p2} + cA_{q2} & \cdots & A_{pn} + cA_{qn} \\
\vdots & \vdots & & \vdots \\
A_{q1} & A_{q2} & \cdots & A_{qn} \\
\vdots & \vdots & & \vdots
\end{bmatrix}
$$

where we have only written out the rows R_p and R_q. By applying cofactor expansion along R_p following Properties 2.3.3, we have

$$
|A'| = \sum_{k=1}^{n} [(A_{pk} + cA_{qk}) C_{pk}]
$$

$$= \sum_{k=1}^{n} A_{pk}C_{pk} + c \sum_{k=1}^{n} A_{qk}C_{pk}$$

We identify the first term with $|A|$ that is computed from using cofactor expansion on the row R_p of A. The second term can be thought as the determinant of a matrix $\tilde{A}$ that is formed by replacing R_p by R_q in A and subsequently expanded along that row. So $\tilde{A}$ practically has two identical rows $R_p = R_q$ and by the previous result the value of $|\tilde{A}|$ is zero. Therefore $|A'| = |A| + c|\tilde{A}| = |A| + c(0) = |A|$, implying that the addition/subtraction type of elementary row operations does not affect the value of determinant. □

Since the values of determinants for elementary matrices, by Properties 2.3.6, coincide exactly with the factors by how the determinant of some other matrix changes when the corresponding elementary row operations are applied on it (represented by multiplication to the left by these elementary matrices) as shown in Properties 2.3.5, we can conclude

Theorem 2.3.7. For any elementary matrix E and another arbitrary matrix A, we have

$$\det(EA) = \det(E)\det(A)$$

This theorem will be of use when we later prove other properties of determinant. However, before doing so, we will demonstrate how to utilize Properties 2.3.5 (or equivalently 2.3.6) to ease the calculation of determinants.

Example 2.3.4. Re-do Example 2.3.3 utilizing Properties 2.3.5.

Solution. We can factor out the 2 in second row and subtract 3 times the third row from the fourth row. By Properties 2.3.5, we have

$$|A| = \begin{vmatrix} 1 & 4 & 4 & 4 \\ 2 & 0 & 4 & 6 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 3 & 1 \end{vmatrix} = 2 \begin{vmatrix} 1 & 4 & 4 & 4 \\ 1 & 0 & 2 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 3 & 1 \end{vmatrix}$$

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

The determinant in the last line can be computed by doing cofactor expansion along the fourth row which now contains two zeros. With Properties 2.3.3 and 2.3.1, it is

$$\begin{vmatrix} 1 & 4 & 4 & 4 \\ 1 & 0 & 2 & 3 \\ 2 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{vmatrix} = 0 + (-1)^{4+2}(-1) \begin{vmatrix} 1 & 4 & 4 \\ 1 & 2 & 3 \\ 2 & 1 & 0 \end{vmatrix} + 0 + (-1)^{4+4}(1) \begin{vmatrix} 1 & 4 & 4 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \end{vmatrix}$$

$$= 0 + (-1)(9) + 0 + (1)(14) = 5$$

and hence $|A| = 2(5) = 10$. $\qquad\square$

Finally, it is the time to show the following properties about determinants.

Properties 2.3.8. An invertible matrix has a non-zero determinant. Otherwise, a singular matrix has a determinant of zero.

Proof. Let's denote the matrix in question as A. Assume that A is invertible and hence by Properties 2.2.11 it can be written as the product of some elementary matrices $E_1, E_2, \ldots, E_{n-1}, E_n$, i.e.

$$A = E_1 E_2 \cdots E_{n-1} E_n$$

Taking the determinant of both sides, we have

$$\det(A) = \det(E_1 E_2 \cdots E_{n-1} E_n)$$

By repetitively using Theorem 2.3.7, we have

$$\det(A) = \det(E_1 (E_2 \cdots E_{n-1} E_n))$$

$$= \det(E_1) \det(E_2 \cdots E_{n-1} E_n)$$
$$= \det(E_1) \det(E_2) \det(\cdots E_{n-1} E_n)$$
$$= \det(E_1) \det(E_2) \cdots \det(E_{n-1}) \det(E_n)$$

Since by Properties 2.3.6, all elementary matrices have a non-zero determinant (particularly we have required $c \neq 0$ in Properties 2.3.6), and all $\det(E_i) \neq 0$, we have $\det(A) \neq 0$. To finish the part about singular matrices, we note that by Theorem 2.2.10, singular matrices have reduced row echelon forms that are not the identity. Furthermore, we have the observation that all other square reduced row echelon forms that are not the identity has a determinant of zero. (Why?)[12] With these two pieces of information, we leave the remaining proof to interested readers. □

Properties 2.3.9. For any $n \times n$ square matrices A and B, we have

1. $\det(A^T) = \det(A)$,

2. $\det(kA) = k^n \det(A)$, for any constant k,

3. $\det(AB) = \det(A) \det(B)$, and

4. $\det(A^{-1}) = \frac{1}{\det(A)}$, if A is invertible.

By extension, $\det(A_1 A_2 \cdots A_n) = \det(A_1) \det(A_2) \cdots \det(A_n)$.

For instance, if

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 9 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 4 & 5 \\ 1 & 0 \end{bmatrix}$$

then

$$|A| = (2)(9) - (3)(5) = 3 \qquad |B| = (4)(0) - (5)(1) = -5$$

$$AB = \begin{bmatrix} 2 & 3 \\ 5 & 9 \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 1 & 0 \end{bmatrix}$$

---

[12]Such matrices must have at least one row of full zeros, and by Properties 2.3.4 we are done.

$$= \begin{bmatrix} (2)(4) + (3)(1) & (2)(5) + (3)(0) \\ (5)(4) + (9)(1) & (5)(5) + (9)(0) \end{bmatrix}$$

$$= \begin{bmatrix} 11 & 10 \\ 29 & 25 \end{bmatrix}$$

$$|AB| = (11)(25) - (10)(29)$$

$$= -15 = (3)(-5) = |A||B|$$

So we can see in this case, $\det(AB) = \det(A)\det(B)$ indeed. Now we will formally prove this ((3) of Properties 2.3.9).

Proof. There are two cases to consider, A being invertible or singular. If A is singular, then by Properties 2.2.4, AB is also singular. And by Properties 2.3.8, both $\det(A)$ and $\det(AB)$ will be zero, and the equality holds trivially. Otherwise, if A is invertible, then we can follow the idea in the proof of Properties 2.3.8, and let $A = E_1 E_2 \cdots E_{n-1} E_n$ as a product of elementary matrices in sequence. By using Theorem 2.3.7 back and forth, we can readily show

$$\det(AB) = \det(E_1 E_2 \cdots E_{n-1} E_n B)$$
$$= \det(E_1)\det(E_2) \cdots \det(E_{n-1})\det(E_n)\det(B)$$
$$= (\det(E_1)\det(E_2) \cdots \det(E_{n-1})\det(E_n))\det(B) = \det(A)\det(B)$$

So the equality is true in both cases. $\qquad\square$

Short Exericse: Prove (4) of Properties 2.3.9.[13]

---

[13]Consider $A^{-1}A = I$, and take determinant on both sides. By (3), we have

$$\det\left(A^{-1}A\right) = \det(I)$$

$$\det\left(A^{-1}\right)\det(A) = 1 \quad \text{(The identity always has a determinant of 1)}$$

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}$$

2.3.3 Finding Inverses by Adjugate

An alternative method to compute the inverse of a matrix is by using its ***Adjugate***, which is the transpose of its associated cofactor matrix.

Definition 2.3.10 (Adjugate)**.** For a matrix A, its adjugate is defined as

$$[\text{adj}(A)]_{pq} = (C_{pq})^T = C_{qp}$$

where C is formulated in Definition 2.3.2.

Properties 2.3.11. The inverse of a matrix A can be computed from its adjugate by

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$$

From this formula, obviously, a singular matrix that has a determinant of zero does not has an inverse.

Example 2.3.5. For a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

It is not difficult to see that the determinant is $ad - bc$, and the adjugate matrix is

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

So the inverse, if $ad - bc \neq 0$, is

$$\frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Example 2.3.6. Find the inverse of the following matrix by evaluating its adjugate.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 4 & 11 \end{bmatrix}$$

Solution. First of all, by Sarrus' Rule (Properties 2.3.1)

$$\begin{aligned} |A| &= ((1)(3)(11) + (2)(5)(1) + (3)(1)(4)) \\ &\quad - ((3)(3)(1) + (1)(5)(4) + (2)(1)(11)) \\ &= (33 + 10 + 12) - (9 + 20 + 22) \\ &= 4 \end{aligned}$$

The adjugate matrix is

$$\text{adj}(A) = \begin{bmatrix} \begin{vmatrix} 3 & 5 \\ 4 & 11 \end{vmatrix} & -\begin{vmatrix} 1 & 5 \\ 1 & 11 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 4 & 11 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 11 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} \end{bmatrix}^T$$

$$= \begin{bmatrix} 13 & -6 & 1 \\ -10 & 8 & -2 \\ 1 & -2 & 1 \end{bmatrix}^T$$

$$= \begin{bmatrix} 13 & -10 & 1 \\ -6 & 8 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

(be careful not to forget the transpose!) Putting the pieces together according to the formula in Properties 2.3.11, we have

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

$$= \frac{1}{4} \begin{bmatrix} 13 & -10 & 1 \\ -6 & 8 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{13}{4} & -\frac{5}{2} & \frac{1}{4} \\ -\frac{3}{2} & 2 & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

$\square$

A summarizing point to be emphasized is that

Theorem 2.3.12 (Equivalence Statement)**.** For a square matrix A, the followings are equivalent:

(a) A is invertible, i.e. A^{-1} exists,

(b) $\det(A) \neq 0$,

(c) The reduced row echelon form of A is I.

---

which is just a rephrasing of Properties 2.3.8 and Theorem 2.2.10. Particularly, invertibility is equivalent to a non-zero determinant. We will see the expansion of this equivalence statement in later chapters.

2.4 Python Programming

To create an identity matrix of size n, we use `np.identity(n)`. For example,

```python
import numpy as np
I4 = np.identity(4)
print(I4)
```

returns

```
[[1. 0. 0. 0.]
 [0. 1. 0. 0.]
 [0. 0. 1. 0.]
 [0. 0. 0. 1.]]
```

Applying transpose on a matrix is simple where we just add `.T` after the array variable, like

```
myMatrix1 = np.array([[1., 0., 3.],
                      [1., 4., 1.],
                      [-1., 2., 4.]])
print(myMatrix1)
print(myMatrix1.T)
```

yields

```
[[ 1.  0.  3.]
 [ 1.  4.  1.]
 [-1.  2.  4.]]
[[ 1.  1. -1.]
 [ 0.  4.  2.]
 [ 3.  1.  4.]]
```

Finding the inverse of a matrix requires the `scipy.linalg` library and call the `inv` function.

```
from scipy import linalg
myMatrix2 = linalg.inv(myMatrix1)
print(myMatrix2)
print(myMatrix1@myMatrix2) # Check: should give the identity
```

gives the expected results of

```
[[ 0.4375   0.1875  -0.375  ]
 [-0.15625  0.21875  0.0625 ]
 [ 0.1875  -0.0625   0.125  ]]
[[1. 0. 0.]
 [0. 1. 0.]
 [0. 0. 1.]]
```

Meanwhile, we can use the `det` function to calculate the determinant of a matrix as follows. First,

```
print(linalg.det(myMatrix1))
```

gives the expected output of `32.0`. As another example,

```
myMatrix3 = np.array([[3., 1., 3., 2.],
                      [0., -1., -3., 1.],
                      [1., -1., -2., 0.],
                      [2., 0., 1., 0.]])
print(linalg.det(myMatrix3))
```

produces an extremely small value of `1.1102230246251562e-16`. In fact, the matrix

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

has a determinant of exactly zero. It is an artifact of numerical error when using floating point numbers. If we keep going ahead and computes its inverse by `linalg.inv(myMatrix3)`, we will obtain an absurd output of

```
[[ 1.200959e+15  -2.401919e+15   3.602879e+15  -3.602879e+15]
 [ 6.004799e+15  -1.200959e+16   1.801439e+16  -1.801439e+16]
 [-2.401919e+15   4.803839e+15  -7.205759e+15   7.205759e+15]
 [-1.200959e+15   2.401919e+15  -3.602879e+15   3.602879e+15]]
```

that have entries of extremely large magnitude. This phenomenon is due to the extremely small "determinant", through Properties 2.3.11, magnifies the adjugate by being in the denominator. (The actual computation does not use Properties 2.3.11 directly but this is one perspective to view the problem.) To prevent this, we can add a `if` condition to look for singularity, defining a function like

```
def safe_inv(matrix):
    if np.abs(linalg.det(matrix)) < np.finfo(float).eps:
        print("Warning: The matrix is highly singular!")
        return(np.nan)
    else:
        return(linalg.inv(matrix))
```

where `np.finfo(float).eps` gives the so-called *machine epsilon* ϵ (the order of relative round-off error) of `float` and we want the absolute value of the determinant be larger than that. Subsequently, calling `safe_inv(myMatrix3)` will print a warning. Finally, we note that we can use `sympy` to acquire the reduced row echelon form of a matrix. Let's use the matrix in Example 2.2.3 for demonstration.

```python
import sympy

myMatrix4 = np.array([[2., 2., 1.],
                      [6., 4., 1.],
                      [2., 3., 2.],
                      [2., 1., 0.]])
myMatrix4_sympy = sympy.Matrix(myMatrix4) # Convert the numpy
    array to a sympy matrix
print(myMatrix4_sympy.rref())
```

then returns two objects

```
(Matrix([
[1, 0, -0.5],
[0, 1,  1.0],
[0, 0,    0],
[0, 0,    0]]), (0, 1))
```

The first one is the reduced row echelon form we want, and the second is a tuple which keeps the column indices of the pivots. `sympy` also does *zero testing* such that

```python
myMatrix3_sympy = sympy.Matrix(myMatrix3)
print(myMatrix3_sympy**(-1))
```

raises properly the error of

```
NonInvertibleMatrixError("Matrix det == 0; not invertible.")
    sympy.matrices.common.NonInvertibleMatrixError: Matrix det
    == 0; not invertible.
```

2.5 Exercises

Exercise 2.1 Find the determinant of the matrix below by inspection.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 0 & 10 & 11 & 12 \\ 0 & 0 & 0 & 13 & 14 \\ 0 & 0 & 0 & 0 & 15 \end{bmatrix}$$

Exercise 2.2 Let

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 4 & 6 \\ 0 & 1 \end{bmatrix}$$

Verify:

(a) $(AB)^T = B^T A^T$,

(b) $(AB)^{-1} = B^{-1} A^{-1}$, and

(c) $\det(AB) = \det(A)\det(B)$.

for this particular case.

Exercise 2.3 If

$$A = \begin{bmatrix} 3 & 2 & 9 \\ 1 & 2 & 3 \\ 4 & 0 & 4 \end{bmatrix}$$

Find its inverse by

(a) Gaussian Elimination, and

(b) Determinant and adjugate.

66

Exercise 2.4 Let

$$A = \begin{bmatrix} 0 & 2 & 5 \\ 0 & 4 & 9 \\ 1 & 2 & 1 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 4 & 6 \\ 3 & 5 & 8 \end{bmatrix}$$

Verify:

(a) $(AB)^T = B^T A^T$,

(b) $(AB)^{-1} = B^{-1} A^{-1}$, and

(c) $\det(AB) = \det(A)\det(B)$.

for this particular case.

Exercise 2.5 Show that

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

is singular.

Exercise 2.6 Given

$$A = \begin{bmatrix} 1 & 9 & 1 & 4 \\ 0 & 6 & 2 & 8 \\ 1 & 9 & 3 & 9 \\ 0 & 9 & 0 & 1 \end{bmatrix}$$

Find its determinant, inverse, and determinant of the inverse.

Exercise 2.7 For the following matrix,

$$A = \begin{bmatrix} p & 1 & 2 \\ 0 & 2 & p \\ 4 & -2 & 0 \end{bmatrix}$$

Find the values of p such that A is invertible.

Exercise 2.8 Show that for any square matrix A, $A + A^T$ is symmetric and $A - A^T$ is skew-symmetric. Hence show with an explicit formula that any square matrix A can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

Exercise 2.9 Prove that if A is an invertible $n \times n$ matrix, $|A| \neq 0$, then we have

$$\det(\text{adj}(A)) = (\det(A))^{n-1}$$

using Properties 2.3.9 and 2.3.11.

Solutions for Linear Systems

The last chapter has introduced the necessary machinery for solving linear systems and now we are going to see how to apply them under suitable circumstances. Remember, in the first chapter, we have formulated some problems about linear systems of equations appearing in the Earth System, and they will be solved accordingly.

3.1 Number of Solutions for Linear Systems

Before tackling any linear system, we may like to know there are how many solutions. In fact, there are only three possibilities.

Theorem 3.1.1 (Number of Solutions for a Linear System)**.** For a system of linear equations $A\vec{x} = \vec{h}$ (recall Definition 1.2.2 and Properties 1.2.3), it has either:

1. No solution,

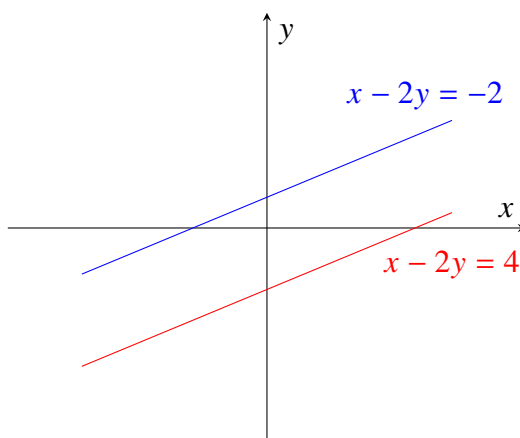2. An unique solution, or

3. Infinitely many solutions.

for the unknowns $\vec{x}$.

This can be illustrated by considering a linear system with two equations and two unknowns, with each equation representing a line. There are three types of scenarios.
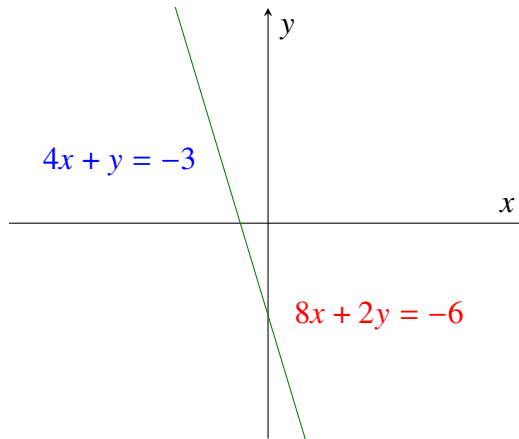
$$\begin{cases} a_1 x + b_1 y & = h_1 \\ a_2 x + b_2 y & = h_2 \end{cases}$$
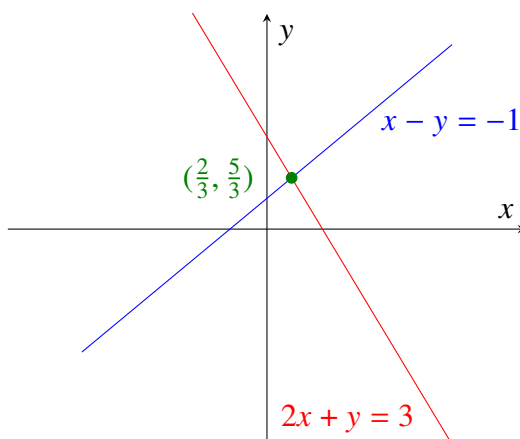
One Solution: Two non-parallel lines (red/blue) intersecting at one point (green).

No Solution: Two parallel lines never touch each other.

Infinitely Many Solutions: Two parallel lines overlap each other.

It goes similarly for any linear system of three unknowns in which equations represent planes instead. The readers can try to imagine and visualize the possibilities. (The intersection of two non-parallel planes will be a line.) In fact, this theorem about the existence of solutions is true for any number of variables and equations. Some readers may think if there can be finitely many solutions only. Unfortunately, it is impossible. Assume there are at least two distinct solutions $\vec{x}_1, \vec{x}_2$ to the system $A\vec{x} = \vec{h}$, then it is easy to show by construction all $\vec{x}_t = t\vec{x}_1 + (1 - t)\vec{x}_2$ for any t will be valid solutions which are infinitely many.

Naturally, the next question is about how to find out which case the linear system belongs to. The following theorem reveals the relation between the number of solutions for a *square* linear system and the determinant of its coefficient matrix.

Theorem 3.1.2. For a square linear system $A\vec{x} = \vec{h}$, if the coefficient matrix A is invertible, i.e. $\det(A) \neq 0$, there is always only one unique solution. However, if A is singular, $\det(A) = 0$, then it has either no solution, or infinitely many solutions.

As a consequence, if the homogeneous linear system $A\vec{x} = \mathbf{0}$ is singular with $\det(A) = 0$, since it always has a trivial solution of $\vec{x} = \mathbf{0}$, the above theorem

implies that the homogeneous system must have infinitely many solutions (since it does not have no solution). We defer the proof of Theorem 3.1.2, as well as the discussion about non-square systems, until we start actually solving linear systems in the next subsection.

Short Exercise: By inspection, determine the number of solutions for the following linear systems.[1]

$$\begin{bmatrix} 2 & 1 & 6 \\ 3 & 0 & 4 \\ 1 & 1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 4 & 3 \\ 1 & 5 & 2 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3.2 Solving Linear Systems

Finally it is the time to get down to solving linear systems (preferably written in form of matrices), and we have two methods to choose.

1. By Gaussian Elimination, for linear system in any shape, or

2. By Inverse, which is apparently only applicable for square, invertible coefficient matrices.

3.2.1 Solving Linear Systems by Gaussian Elimination

Like in Section 2.2.3, applying Gaussian Elimination on the augmented matrix (introduced at the end of Section 1.2) of a linear system can yield the solution at right hand side. The principles involving elementary row operations are the same as stated in Theorems 2.2.9 and 2.2.10, but with $A\vec{x} = \vec{h}$ instead of $AA^{-1} = I$. In addition, the coefficient matrix A can be non-square, but we will look at the easier case of a coefficient matrix A first.

---

[1]These two homogeneous linear system has a determinant of -1 and 0, and hence by Theorem 3.1.2 the first system has a unique solution and the second one has infinitely many solutions.

Square Systems

Example 3.2.1. Solve the following linear system by Gaussian Elimination.

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 4 \\ 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \\ 8 \end{bmatrix}$$

Solution. We re-write the system in augmented form and apply Gaussian Elimination, aiming to reduce the matrix to the left into the identity.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 1 & 1 & 4 & 10 \\ 2 & 0 & 3 & 8 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad R_2 - R_1 \rightarrow R_2, R_3 - 2R_1 \rightarrow R_3$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad R_2 - 2R_3 \rightarrow R_2, R_1 - R_3 \rightarrow R_1$$

which translates to

$$\begin{cases} x = 1 \\ y = 1 \quad \text{or} \\ z = 2 \end{cases} \qquad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Note that we have successfully converted the coefficient matrix to the identity along the way, which by Theorem 2.3.12 the coefficient matrix is invertible. This explains the first part of Theorem 3.1.2 as in this case every unknown is associated only to a leading 1 in the corresponding column and a unique solution can always be derived. □

Example 3.2.2. Solve the linear system of

$$\begin{bmatrix} 3 & 7 & 2 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \\ 2 \end{bmatrix}$$

Solution. Again, we apply Gaussian Elimination on the augmented matrix to obtain

$$\left[\begin{array}{ccc|c} 3 & 7 & 2 & 8 \\ 1 & 1 & 0 & 2 \\ 0 & 2 & 1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 3 & 7 & 2 & 8 \\ 0 & 2 & 1 & 2 \end{array}\right] \qquad R_1 \leftrightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 4 & 2 & 2 \\ 0 & 2 & 1 & 2 \end{array}\right] \qquad R_2 - 3R_1 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 1 & 2 \end{array}\right] \qquad \frac{1}{4}R_2 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{array}\right] \qquad R_3 - 2R_2 \rightarrow R_3$$

The last row corresponds to $0 = 1$ which is contradictory. As a consequence, the system is inconsistent, no solution exists. □

Example 3.2.3. Find the solution for the following linear system.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

Solution. Gaussian Elimination leads to

$$
\begin{bmatrix} 1 & 2 & 1 & | & 1 \\ 2 & 5 & 3 & | & 2 \\ 0 & 1 & 1 & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & | & 1 \\ 0 & 1 & 1 & | & 0 \\ 0 & 1 & 1 & | & 0 \end{bmatrix} \qquad R_2 - 2R_1 \rightarrow R_2
$$

$$
\rightarrow \begin{bmatrix} 1 & 0 & -1 & | & 1 \\ 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \qquad R_3 - R_2 \rightarrow R_3, R_1 - 2R_2 \rightarrow R_1
$$

Now, the last row corresponds to $0 = 0$, implying one equation is spurious. This also means that it has a ***Free Variable***, which means that we can assign one unknown as a parameter for expressing other variables. We choose such unknowns according to the rule that they should not be fixed to a pivot in the reduced coefficient matrix. As the variables x and y already correspond to the two pivots in the first/second columns, we can only let $z = t$. From the first row and second row, we obtain $x = 1 + t$, $y = -t$ respectively. Therefore,

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

where $-\infty < t < \infty$ is any scalar. The first column vector

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

is the so-called ***Particular Solution***. When it is complemented by the second column vector which is multiplied by the free parameter t

$$
t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}
$$

they constitute the entire set of ***General Solution***. $\qquad\qquad$ □

Short Exercise: Try plugging in any number t to the general solution and verify the consistency.[2]

The general solution encompasses all possible solutions to the linear system. For broader situations, it can contain more than one pairs of free parameter and column vector (or none, for the rather trivial cases of zero or a unique solution). The amount of free variables can be seen to be the number of columns in the coefficient matrix, minus the number of pivots in the reduced row echelon form. In case of multiple free variables, we assign the corresponding amount of free parameters to the non-pivots and apply the same procedure to get a set of general solution. Any column vector of a free parameter can be scaled as we desire.[3]

Meanwhile, the particular solution can be set to any valid solution to the system (the choice does not affect the structure of any column vector that comes along with a free parameter, see the footnote to the short exercise above). If the linear system is homogeneous, then the zero vector will always be a possible particular solution.

---

[2]Let's say $t = 1$ and $\tilde{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$, then clearly $A\tilde{x} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. It can become a new particular solution by noting that the original solution can be rewritten as

$$\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + (t - 1) \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + t' \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \tilde{x} + t' \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

where we "extract" $\tilde{x}$ from generating a shifted free parameter $t' = t - 1$ and according to this relation, it represents the same set of general solution as the original expression.

[3]Using the last example,

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{t}{2} \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

where we use $s = \frac{t}{2}$ as a new free parameter and the column vector is now scaled by a factor of 2.

We have seen in the previous two examples that if the reduced row echelon form of the square coefficient matrix has some row of full zeros, then it either leads to no solution (if inconsistent) or infinitely many solutions (if consistent). Since such a matrix at the same time has a determinant of zero (by Properties 2.3.4) and is singular. This establishes the second part of Theorem 3.1.2.

For non-square coefficient matrices, two cases occur.

1. There are more equations (rows) than unknowns (columns). The system is **Overdetermined**. The reduced row echelon form then must have at least one row of full zeros. If any one of them is inconsistent, then contradiction will arise just like in Example 3.2.2 and there will be no solution. However, if all zero rows are consistent (i.e. $0 = 0$), then there still can be a unique solution or infintely many of them.

2. There are fewer equations (rows) than unknowns (columns). The system is said to be **Underdetermined**. There must be unknowns that are non-pivots in the reduced row echelon form of the coefficient matrix. Hence free variables, and infinitely many solutions ensue if there is no *inconsistent* row of full zeros (then there is no solution). The calculation is similar to that in Example 3.2.3.

Let's see some examples for non-square linear systems.

Overdetermined Systems

Example 3.2.4. Find the solution to the following overdetermined system, if any.

$$\begin{bmatrix} 1 & 4 & 0 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 3 \\ 5 \end{bmatrix}$$

Solution.

$$
\begin{bmatrix}
1 & 4 & 0 & 4 \\
2 & 2 & 3 & 8 \\
1 & 1 & 2 & 3 \\
0 & 3 & 1 & 5
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 4 & 0 & 4 \\
0 & -6 & 3 & 0 \\
0 & -3 & 2 & -1 \\
0 & 3 & 1 & 5
\end{bmatrix}
\qquad R_2 - 2R_1 \rightarrow R_2,\, R_3 - R_1 \rightarrow R_3
$$

$$
\rightarrow
\begin{bmatrix}
1 & 4 & 0 & 4 \\
0 & 1 & -\frac{1}{2} & 0 \\
0 & -3 & 2 & -1 \\
0 & 3 & 1 & 5
\end{bmatrix}
\qquad -\frac{1}{6}R_2 \rightarrow R_2
$$

$$
\rightarrow
\begin{bmatrix}
1 & 4 & 0 & 4 \\
0 & 1 & -\frac{1}{2} & 0 \\
0 & 0 & \frac{1}{2} & -1 \\
0 & 0 & \frac{5}{2} & 5
\end{bmatrix}
\qquad R_3 + 3R_2 \rightarrow R_3,\, R_4 - 3R_2 \rightarrow R_4
$$

$$
\rightarrow
\begin{bmatrix}
1 & 4 & 0 & 4 \\
0 & 1 & -\frac{1}{2} & 0 \\
0 & 0 & 1 & -2 \\
0 & 0 & \frac{5}{2} & 5
\end{bmatrix}
\qquad 2R_3 \rightarrow R_3
$$

$$
\rightarrow
\begin{bmatrix}
1 & 4 & 0 & 4 \\
0 & 1 & -\frac{1}{2} & 0 \\
0 & 0 & 1 & -2 \\
0 & 0 & 0 & 10
\end{bmatrix}
\qquad R_4 - \frac{5}{2}R_3 \rightarrow R_4
$$

The last row is inconsistent and hence the overdetermined system has no solution. □

Example 3.2.5. Show that there are infinitely many solution to the following overdetermined system.

$$
\begin{bmatrix}
1 & 1 & 2 \\
1 & 2 & 5 \\
2 & 1 & 1 \\
1 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
=
\begin{bmatrix}
2 \\
3 \\
3 \\
1
\end{bmatrix}
$$

Solution.

$$
\begin{bmatrix} 1 & 1 & 2 & | & 2 \\ 1 & 2 & 5 & | & 3 \\ 2 & 1 & 1 & | & 3 \\ 1 & 0 & -1 & | & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 2 & | & 2 \\ 0 & 1 & 3 & | & 1 \\ 0 & -1 & -3 & | & -1 \\ 0 & -1 & -3 & | & -1 \end{bmatrix} \quad \begin{array}{c} R_2 - R_1 \rightarrow R_2, R_3 - 2R_1 \rightarrow R_3 \\ R_4 - R_1 \rightarrow R_4 \end{array}
$$

$$
\rightarrow \begin{bmatrix} 1 & 1 & 2 & | & 2 \\ 0 & 1 & 3 & | & 1 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \quad R_3 + R_2 \rightarrow R_3, R_4 + R_2 \rightarrow R_4
$$

$$
\rightarrow \begin{bmatrix} 1 & 0 & -1 & | & 1 \\ 0 & 1 & 3 & | & 1 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \quad R_1 - R_2 \rightarrow R_1
$$

Two out of the four equations are redundant and there are effectively two constraints only, over the three variables. We can let the non-pivot unknown $z = t$ be a free variable like in Example 3.2.3, and derive $x = 1 + t$, $y = 1 - 3t$ from the first two rows. Thus the general solution is

$$
\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 + t \\ 1 - 3t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}
$$

where $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is a particular solution. □

Underdetermined Systems

Example 3.2.6. Solve the following underdetermined system.

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 3 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Solution.

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 2 & 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 2 & 1 \end{array}\right] \qquad R_2 - 2R_1 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & -1 & -1 & 0 & 1 \end{array}\right] \qquad R_2 \leftrightarrow R_3$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{array}\right] \qquad R_3 + R_2 \rightarrow R_3$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad \frac{1}{2}R_3 \rightarrow R_3$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & -1 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad R_2 - 2R_3 \rightarrow R_2, R_1 - R_3 \rightarrow R_1$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad R_1 - R_2 \rightarrow R_1$$

From the third row, we have $v = 1$ immediately. The only unknown that is not associated to a pivot is u and we can let $u = t$ be a free variable. From the first two equations, we retrieve $y = -1 - t$ and $x = -t$, and therefore the general

solution is

$$\vec{x} = \begin{bmatrix} x \\ y \\ u \\ v \end{bmatrix} = \begin{bmatrix} -t \\ -1-t \\ t \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

with $\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$ as a particular solution. $\qquad\qquad\qquad\square$

3.2.2 Solving Linear Systems by Inverse

For a square linear system $A\vec{x} = \vec{h}$, if A has a non-zero determinant and is invertible, then we can utilize its inverse to recover the solution. Remember that multiplying the inverse to a matrix returns an identity matrix, it is possible to multiply the inverse A^{-1} to the left on both sides of the equation $A\vec{x} = \vec{h}$ to cancel out the A at the L.H.S., which leads to

$$A^{-1}A\vec{x} = (A^{-1}A)\vec{x} = A^{-1}\vec{h}$$
$$\vec{x} = I\vec{x} = A^{-1}\vec{h} \qquad \text{(Definition 2.2.1 and Properties 2.1.2)}$$

This solution is unique, guaranteed by Theorem 3.1.2.

Example 3.2.7. Given a linear system $A\vec{x} = \vec{h}$

$$\begin{bmatrix} 1 & -1 & -2 \\ 0 & 3 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$$

It can be checked that the inverse of the coefficient matrix is

$$\begin{bmatrix} 1 & -1 & -2 \\ 0 & 3 & 1 \\ 1 & 0 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{3}{2} & -\frac{1}{2} & \frac{3}{2} \end{bmatrix}$$

The readers are encouraged to verify the inverse. Subsequently, we have the solution to the linear system as $\vec{x} = A^{-1}\vec{h}$

$$\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{3}{2} & -\frac{1}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

Doing Gaussian Elimination to find the inverse and then compute the solution by $\vec{x} = A^{-1}\vec{h}$ in Section 3.2.2 is somehow the same as using Gaussian Elimination directly to solve the linear system suggested by Section 3.2.1. Hypothetically, if there are a large amount of linear systems which all share the same coefficient matrix A, but different $\vec{h}_k$ to be solved, then the former approach may be more efficient at first sight. However, in computer, calculation of inverse can be unstable (see Section 2.4) and there are some other practical reasons not to do so, as we shall see in Section 3.4. Besides, Theorem 2.3.12 can be extended as below by incorporating Theorem 3.1.2:

Theorem 3.2.1. [Equivalence Statement, ver. 2] For a square matrix A, the followings are equivalent:

(a) A is invertible, i.e. A^{-1} exists,

(b) $\det(A) \neq 0$,

(c) The reduced row echelon form of A is I,

(d) The linear system $A\vec{x} = \vec{h}$ has a unique solution, particularly $A\vec{x} = \mathbf{0}$ has only the trivial solution $\vec{x} = \mathbf{0}$.

3.3 Earth Science Applications

Now we are going to revisit and find the solutions to the two linear system problems in Section 1.4.

Example 3.3.1. Solve for the horizontal displacement x and depth of top layer y in the seismic ray problem of Example 1.4.1.

Solution. The linear system is

$$\begin{bmatrix} 1 & 1 \\ 1 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1200 \\ 800\sqrt{3} \end{bmatrix}$$

Since it is just a 2×2 coefficient matrix, we can directly use the expression in Example 2.3.5 to find its inverse, which is

$$\frac{1}{\sqrt{3} - 1} \begin{bmatrix} \sqrt{3} & -1 \\ -1 & 1 \end{bmatrix} = \frac{1 + \sqrt{3}}{2} \begin{bmatrix} \sqrt{3} & -1 \\ -1 & 1 \end{bmatrix}$$

and solve the system by multiplying the inverse following the method demonstrated in Section 3.2.2, leading to

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1 + \sqrt{3}}{2} \begin{bmatrix} \sqrt{3} & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1200 \\ 800\sqrt{3} \end{bmatrix} = \begin{bmatrix} 600 + 200\sqrt{3} \\ 600 - 200\sqrt{3} \end{bmatrix}$$

Therefore the required horizontal displacement and depth of top layer are about 946.4 m and 253.6 m respectively. □

Example 3.3.2. Find the radiative loss E_j and hence temperature T_j in each layer of the multi-layer model in Example 1.4.2. In particular, what is the temperature at the surface ($j = N + 1$)?

Solution. The linear system is

$$\begin{bmatrix} -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & & 0 & 0 & 0 \\ 0 & 1 & -2 & & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & & -2 & 1 & 0 \\ 0 & 0 & 0 & & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_{N-1} \\ E_N \\ E_{N+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -E_{in} \end{bmatrix}$$

where N is any positive integer. Since N can be arbitrarily large, we may wish to avoid the direct computation of a massive inverse. Instead, we resort to a tactful way of row reduction to reveal the pattern of R_j. Rather than starting the reduction at the top as usual, we build up at the bottom, subtracting the lower row from the row directly above it and then moving up a row, repeated until we reach the top.

$$
\left[\begin{array}{ccccccc|c}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -2 & 1 & & 0 & 0 & 0 & 0 \\
0 & 1 & -2 & & 0 & 0 & 0 & 0 \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -2 & 1 & 0 & 0 \\
0 & 0 & 0 & & 1 & -2 & 1 & 0 \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -E_{in}
\end{array}\right]
$$

$$
\rightarrow
\left[\begin{array}{ccccccc|c}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -2 & 1 & & 0 & 0 & 0 & 0 \\
0 & 1 & -2 & & 0 & 0 & 0 & 0 \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -2 & 1 & 0 & 0 \\
0 & 0 & 0 & & 1 & -1 & 0 & -E_{in} \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -E_{in}
\end{array}\right]
\qquad R_N + R_{N+1} \to R_N
$$

$$
\rightarrow
\left[\begin{array}{ccccccc|c}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -2 & 1 & & 0 & 0 & 0 & 0 \\
0 & 1 & -2 & & 0 & 0 & 0 & 0 \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -1 & 0 & 0 & -E_{in} \\
0 & 0 & 0 & & 1 & -1 & 0 & -E_{in} \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -E_{in}
\end{array}\right]
\qquad R_{N-1} + R_N \to R_{N-1}
$$

$$\rightarrow \quad \vdots \qquad\qquad\qquad\qquad\qquad \text{(Keep going up)}$$

$$\rightarrow \left[\begin{array}{ccccccc|c}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -2 & 1 & & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & & 0 & 0 & 0 & -E_{in} \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -1 & 0 & 0 & -E_{in} \\
0 & 0 & 0 & & 1 & -1 & 0 & -E_{in} \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -R_{in}
\end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccccc|c}
-2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & & 0 & 0 & 0 & -E_{in} \\
0 & 1 & -1 & & 0 & 0 & 0 & -E_{in} \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -1 & 0 & 0 & -E_{in} \\
0 & 0 & 0 & & 1 & -1 & 0 & -E_{in} \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -E_{in}
\end{array}\right] \qquad R_2 + R_3 \rightarrow R_2$$

$$\rightarrow \left[\begin{array}{ccccccc|c}
-1 & 0 & 0 & \cdots & 0 & 0 & 0 & -E_{in} \\
1 & -1 & 0 & & 0 & 0 & 0 & -E_{in} \\
0 & 1 & -1 & & 0 & 0 & 0 & -E_{in} \\
\vdots & & & \ddots & & & \vdots & \vdots \\
0 & 0 & 0 & & -1 & 0 & 0 & -E_{in} \\
0 & 0 & 0 & & 1 & -1 & 0 & -E_{in} \\
0 & 0 & 0 & \cdots & 0 & 1 & -1 & -E_{in}
\end{array}\right] \qquad R_1 + R_2 \rightarrow R_1$$

From the first row, we readily obtain $E_1 = E_{in}$. The second row yields the equation

$$E_1 - E_2 = -E_{in}$$
$$E_2 = E_1 + E_{in} = E_{in} + E_{in} = 2E_{in}$$

Similarly, the subsequent rows are all in the form of $E_{j+1} = E_j + E_{in}$, and inductively we have $E_j = jE_{in}$. $E_1 = E_{in}$ is the emission of radiation from Earth as a whole as viewed from the space, and the *emission temperature* is $T_1 = \sqrt[4]{E_1/\sigma} = \sqrt[4]{E_{in}/\sigma}$ by the Stefan–Boltzmann Law. The surface releases terrestrial radiation at the rate of $E_{N+1} = (N+1)E_{in}$ and has a temperature of $T_{N+1} = \sqrt[4]{E_{N+1}/\sigma} = \sqrt[4]{(N+1)E_{in}/\sigma} = (N+1)^{1/4}\sqrt[4]{E_{in}/\sigma} = (N+1)^{1/4}T_1$, i.e.

the surface temperature is $(N + 1)^{1/4}$ times the emission temperature. Our earth has an emission temperature of 255 K and a surface temperature of 288 K on average (notice that we have to use Kelvin instead of degree Celsius!), which leads to an effective number of layers $N = (288/255)^4 - 1 = 0.627$. □

3.4 Python Programming

For solving square linear systems in the form of $A\vec{x} = \vec{h}$, we can again import the `scipy.linalg` library and call the `solve` function with the coefficient matrix A as the first argument and $\vec{h}$ placed in the second one.

```python
import numpy as np
from scipy import linalg

A = np.array([[1., 0., 1.],
              [2., 2., 3.],
              [1., 2., 0.]])
h = np.array([0., -1., 1.])
x = linalg.solve(A,h)
```

This corresponds to the linear system

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 3 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

which has a solution of

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

`print(x)` then gives the correct output of `[ 1. -0. -1.]`. However, if A is a singular matrix like the one shown in Section 2.4

```
A = np.array([[3., 1., 3., 2.],
              [0., -1., -3., 1.],
              [1., -1., -2., 0.],
              [2., 0., 1., 0.]]) # "myMatrix3" in the last
                  chapter
h = np.array([0., 1., 1., -1.])
x = linalg.solve(A,h)
print(x)
```

raises a warning and an unreasonable output of

```
LinAlgWarning: Ill-conditioned matrix
(rcond=3.42661e-18): result may not be accurate.
  x = linalg.solve(A,h)
[ 4.803839e+15  2.401919e+16 -9.607679e+15 -4.803839e+15]
```

Again, we can use the `sympy` package for the rescue as follows.

```
import sympy

A_sympy = sympy.Matrix(A)
h_sympy = sympy.Matrix(h)
A_sympy.solve(h_sympy)
```

which raises the same "not invertible" error as in Section 2.4. We note that, unfortunately, there is no simple way to deal with over/under-determined systems using either `scipy` or `sympy`. Moreover, there are two questions that may come to the curious readers when reading the programming sections of these two chapters. First, which of `scipy` and `sympy` should we choose over another? Second, why we don't compute the inverse of A and solve the system by something along the line of `x = linalg.inv(A) @ h`? For the first question, we note that `scipy` is numerical while `sympy` is symbolic, which means that if we are dealing with real data we may find `scipy` adequate and more efficient, while if we are focusing on the theoretical part of Mathematics we can obtain a more analytical solution with `sympy`. To the second question, we refer the readers to this excellent Stack Overflow post (31256252).

3.5 Exercises

Exercise 3.1 Solve the following linear system.

$$\begin{cases} 5x + y + 3z & = 6 \\ 2x - y + z & = \frac{7}{2} \\ 3x + 2y - 4z & = -\frac{13}{2} \end{cases}$$

Exercise 3.2 Solve $A\vec{x} = \vec{h}_k$, where

$$A = \begin{bmatrix} 6 & 7 & 7 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \end{bmatrix} \qquad\qquad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\vec{h}_1 = \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix} \qquad\qquad \vec{h}_2 = \begin{bmatrix} 19/4 \\ 1 \\ 5/4 \end{bmatrix}$$

Exercise 3.3 Derive the solution to the following linear system.

$$\begin{cases} 3x + 4z & = 2 \\ x + y + 2z & = -1 \\ x - 2y & = 0 \end{cases}$$

Exercise 3.4 Solve the following linear system.

$$\begin{cases} m + n - p - 3q & = 2 \\ m - q & = 5 \\ 3m + 2n - 2p - 7q & = 9 \end{cases}$$

How about if the R.H.S. of the third equation is equal to 3 instead?

Exercise 3.5 For the following linear system,

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

Find the values of α so that the system has no solution, or infinitely many solutions.

Exercise 3.6 In a geology field trip, an outcrop is examined. It is observed that the rock mainly consists of crystals of three distinct colors (gray/pink/black). Assume that crystal of each color corresponds to exactly one type of mineral. Three samples are gathered, have their densities measured and composition percentages of the three types of crystal analyzed. The data are as follows:

	gray	pink	black	density (g/cm^3)
Sample A	40%	50%	10%	2.645
Sample B	55%	40%	5%	2.6325
Sample C	45%	45%	10%	2.65

From the data, infer the densities of the constituent minerals.

Exercise 3.7 *Ohm's law* relates voltage drop of a current due to resistance by $V = IR$. In addition, *Kirchhoff's Second Law* states that: The voltage gain balances the voltage drop around any closed loop (net voltage change must be zero). The clockwise convention is adopted, i.e. around a loop, a battery with its positive terminal facing the clockwise direction is considered a voltage gain, and clockwise current passing through a resistor is deemed as a voltage drop. Together with the knowledge that current at a junction must conserve (*Kirchhoff's First Law*), find I_1, I_2, I_3 (assumed flowing in the direction as indicated) for the circuit in Figure 3.1.
You will obtain two equations by considering any two loops with Kirchhoff's Second Law, and one from Kirchhoff's First Law. So, there are three equations, for the three unknown currents.
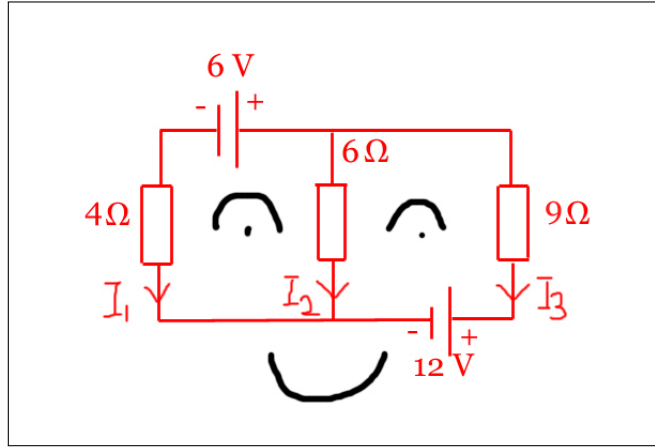
Figure 3.1: The circuit for Exercise 3.7

Exercise 3.8 The *shallow water equations* (see Figure 3.2) describe the evolution of gravity wave under some approximations such as *hydrostatic balance* and a sufficiently shallow fluid depth, and has the form of

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + H(\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y}) & = 0 \\[2ex] \dfrac{\partial u}{\partial t} & = -g\dfrac{\partial \eta}{\partial x} \\[2ex] \dfrac{\partial v}{\partial t} & = -g\dfrac{\partial \eta}{\partial y} \end{cases}$$

when the Coriolis effect is ignored. By assuming a travelling wave solution

$$u = \tilde{U} \cos(kx + ly - \omega t)$$
$$v = \tilde{V} \cos(kx + ly - \omega t)$$
$$\eta = \tilde{\eta} \cos(kx + ly - \omega t)$$

where $\tilde{U}$, $\tilde{V}$, $\tilde{\eta}$ are some constants to be determined, show that the equations become

$$\begin{cases} \omega\tilde{\eta} - kH\tilde{U} - lH\tilde{V} & = 0 \\ \omega\tilde{U} - kg\tilde{\eta} & = 0 \\ \omega\tilde{V} - lg\tilde{\eta} & = 0 \end{cases}$$
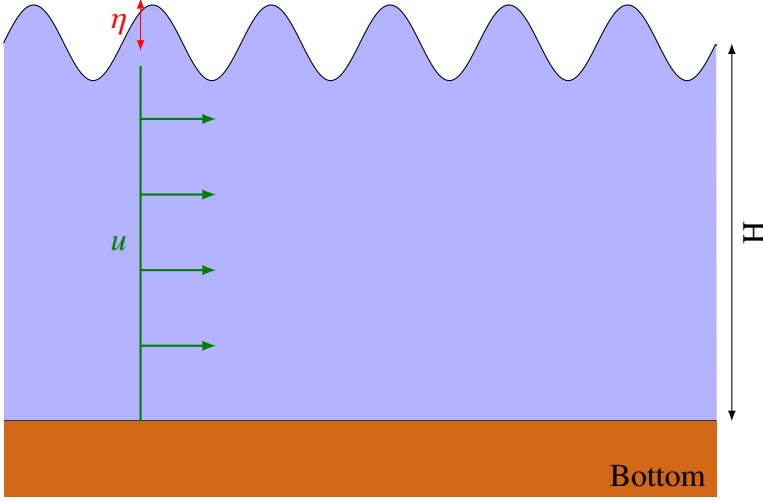
Figure 3.2: The *x*-*z* cross-section of shallow water system in Exercise 3.8. η is the height of free surface, H is the mean depth of the fluid, and u is the fluid velocity along *x*-axis.

By requiring that $\tilde{U}, \tilde{V}, \tilde{\eta}$ have a non-trivial solution so that they are not all zeros, derive the dispersion relation of gravity wave, which is

$$\omega^2 = gH(k^2 + l^2)$$
$$\omega = c\kappa$$

where $c = \sqrt{gH}$ is the wave speed, and $\kappa = \sqrt{k^2 + l^2}$ is the total wavenumber.

Exercise 3.9 Solve for the condensation height and temperature z_{cd} and T_{cd} in Exercise 1.10.

Exercise 3.10 Solve the *Chickens and Rabbits in the Same Cage* problem in Exercise 1.11. If we now introduce a new type of mystical creature who has one head and three legs, and put them in another cage along with some chickens and rabbits, find all possible numbers of the three species if the cage now has 48 heads and 122 legs.

Answers to Exercises

Exercise 1.1

(a) $\begin{bmatrix} -3 & 5 \\ 3 & 6 \end{bmatrix}$

(b) $\begin{bmatrix} 8 & -\frac{1}{2} \\ 13 & -\frac{25}{2} \end{bmatrix}$

(c) $\begin{bmatrix} -8 & 17 \\ -18 & 8 \end{bmatrix}$

(d) $\begin{bmatrix} 11 & -11 \\ 33 & -11 \end{bmatrix}$

Exercise 1.2

(a) $\begin{bmatrix} -2 & 1 & 3 \\ -1 & -1 & -9 \\ -8 & 2 & -2 \end{bmatrix}$

(b) $\begin{bmatrix} -8 & -5 \\ 15 & 3 \end{bmatrix}$

Exercise 1.3

(a) $\begin{bmatrix} 42 & 72 & 0 \\ 32 & 51 & -1 \end{bmatrix}$

(b) Same as above

(c) $\begin{bmatrix} 90 & 162 & 2 \\ 51 & 99 & 3 \end{bmatrix}$

(d) Same as above

Exercise 1.4

(a) $\begin{bmatrix} 16 & 23 & 129 \\ 133 & 33 & 102 \\ 27 & 9 & 128 \end{bmatrix}$

(b) $\begin{bmatrix} -\frac{233}{4} & -\frac{19}{4} & \frac{69}{2} \\ -\frac{339}{4} & -16 & 31 \\ \frac{109}{4} & \frac{33}{4} & -\frac{289}{4} \end{bmatrix}$

Exercise 1.5

(a) $\begin{bmatrix} 16 & 6 & 3 \\ 34 & 13 & 12 \\ 9 & 2 & 27 \end{bmatrix}$

(b) $\begin{bmatrix} 27 & 15 & 69 \\ 37 & 12 & 85 \\ 36 & 12 & 69 \end{bmatrix}$

(c) $\begin{bmatrix} 14 & 3 & 26 \\ 29 & 9 & 60 \\ 12 & 21 & 41 \end{bmatrix}$

(d) $\begin{bmatrix} 33 & 13 & 24 \\ 47 & 19 & 21 \\ 39 & 14 & 12 \end{bmatrix}$

Exercise 1.6

$$AB = BA = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Exercise 1.7

$$\begin{bmatrix} 0 & 3 & -4 \\ 5 & -1 & 2 \\ 6 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 13 \\ 8 \end{bmatrix}$$

or

$$\left[\begin{array}{ccc|c} 0 & 3 & -4 & 6 \\ 5 & -1 & 2 & 13 \\ 6 & 0 & 1 & 8 \end{array} \right]$$

Exercise 1.8

(a) $\begin{bmatrix} 2 & 3 & 5 & 7 \\ 1 & 2 & 4 & 8 \\ 1 & 4 & 8 & 12 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 4 & 8 \\ 5 & 12 & 23 & 37 \\ 1 & 3 & 6 & 10 \end{bmatrix}$

Exercise 1.9

1. Interchange Row 1 and Row 3 ($R_1 \leftrightarrow R_3$),

2. Multiply Row 1 by $\frac{1}{2}$ ($\frac{1}{2}R_1 \rightarrow R_1$),

3. Subtract Row 3 from Row 2 ($R_2 - R_3 \rightarrow R_2$),

4. Add 3 times Row 1 to Row 2 ($R_2 + 3R_1 \rightarrow R_2$).

The order of step 1 and 2, as well as step 3 and 4, can be interchanged.

Exercise 1.10 The air temperature/dew point at any height z before saturation is $T_a = T_{a,ini} - (\Gamma_{dry})z$ and $T_{dew} = T_{dew,ini} - (\Gamma_{dew})z$ respectively. At the condensation level $z = z_{cd}$, the air temperature equals to the dew point temperature $T_a = T_{dew} = T_{cd}$, and hence we have

$$T_{a,ini} - \Gamma_{dry}(z_{cd}) = T_{dew,ini} - \Gamma_{dew}(z_{cd}) = T_{cd}$$

which can be separated into two equations

$$\begin{cases} T_{a,ini} - \Gamma_{dry}(z_{cd}) & = T_{cd} \\ T_{dew,ini} - \Gamma_{dew}(z_{cd}) & = T_{cd} \end{cases}$$

Rearranging to put the unknowns z_{cd} and T_{cd} to the L.H.S., we obtain

$$\begin{cases} T_{cd} + \Gamma_{dry}(z_{cd}) & = T_{a,ini} \\ T_{cd} + \Gamma_{dew}(z_{cd}) & = T_{dew,ini} \end{cases}$$

or, in matrix form

$$\begin{bmatrix} 1 & \Gamma_{dry} \\ 1 & \Gamma_{dew} \end{bmatrix} \begin{bmatrix} T_{cd} \\ z_{cd} \end{bmatrix} = \begin{bmatrix} T_{a,ini} \\ T_{dew,ini} \end{bmatrix}$$

Plugging in the lapse rates, we have

$$\begin{bmatrix} 1 & 9.8 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} T_{cd} \\ z_{cd} \end{bmatrix} = \begin{bmatrix} 25.4 \\ 17.8 \end{bmatrix}$$

Exercise 1.11 Obviously, there are 35 chickens and rabbits in total, and $x+y = 35$. Considering the total amount of legs, we also have $2x + 4y = 94$. Hence the required linear system is

$$\begin{cases} x + y & = 35 \\ 2x + 4y & = 94 \end{cases}$$

In matrix form, it is

$$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 35 \\ 94 \end{bmatrix}$$

Exercise 2.1 (Applying cofactor expansion along the leftmost column recursively) The determinant is just the product of the diagonal elements = $(1)(6)(10)(13)(15) = 11700$.

Exercise 2.2

(a) $\begin{bmatrix} 8 & 20 \\ 15 & 37 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix}$

(b) $\begin{bmatrix} -\frac{37}{4} & \frac{15}{4} \\ 5 & -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & -\frac{3}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -7 & 3 \\ 5 & -2 \end{bmatrix}$

(c) $\begin{vmatrix} 8 & 15 \\ 20 & 37 \end{vmatrix} = -4 = (-1)(4) = \begin{vmatrix} 2 & 3 \\ 5 & 7 \end{vmatrix} \begin{vmatrix} 4 & 6 \\ 0 & 1 \end{vmatrix}$

Exercise 2.3

(a)

$$\begin{bmatrix} 3 & 2 & 9 & | & 1 & 0 & 0 \\ 1 & 2 & 3 & | & 0 & 1 & 0 \\ 4 & 0 & 4 & | & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 3 & | & 0 & 1 & 0 \\ 3 & 2 & 9 & | & 1 & 0 & 0 \\ 4 & 0 & 4 & | & 0 & 0 & 1 \end{bmatrix} \qquad R_1 \leftrightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 3 & | & 0 & 1 & 0 \\ 0 & -4 & 0 & | & 1 & -3 & 0 \\ 0 & -8 & -8 & | & 0 & -4 & 1 \end{bmatrix} \qquad R_2 - 3R_1 \rightarrow R_2, R_3 - 4R_1 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & 0 & | & -\frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 1 & 1 & | & 0 & \frac{1}{2} & -\frac{1}{8} \end{bmatrix} \qquad -\frac{1}{4}R_2 \rightarrow R_2, -\frac{1}{8}R_3 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & 0 & | & -\frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 1 & | & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{8} \end{bmatrix} \qquad R_3 - R_2 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left. \begin{matrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{8} \\ -\frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{8} \end{matrix} \right] \qquad R_1 - 3R_3 - 2R_2 \rightarrow R_1$$

(b) $\det(A) = -32$ and

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

$$= \begin{bmatrix} 8 & 8 & -8 \\ -8 & -24 & 8 \\ -12 & 0 & 4 \end{bmatrix}^T$$

$$= \begin{bmatrix} 8 & -8 & -12 \\ 8 & -24 & 0 \\ -8 & 8 & 4 \end{bmatrix}$$

Hence

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

$$= -\frac{1}{32} \begin{bmatrix} 8 & -8 & -12 \\ 8 & -24 & 0 \\ -8 & 8 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{8} \\ -\frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{8} \end{bmatrix}$$

Exercise 2.4

(a) $\begin{bmatrix} 19 & 35 & 9 \\ 33 & 61 & 16 \\ 52 & 96 & 24 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 3 \\ 3 & 4 & 5 \\ 4 & 6 & 8 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 2 & 4 & 2 \\ 5 & 9 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 18 & -10 & 1 \\ -6 & 3 & 1 \\ -\frac{11}{4} & \frac{7}{4} & -1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 2 & -2 \\ -1 & -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 7 & -4 & 1 \\ -\frac{9}{2} & \frac{5}{2} & 0 \\ 2 & -1 & 0 \end{bmatrix}$

(c) $\begin{vmatrix} 19 & 33 & 52 \\ 35 & 61 & 96 \\ 9 & 16 & 24 \end{vmatrix} = -4 = (-2)(2) = \begin{vmatrix} 0 & 2 & 5 \\ 0 & 4 & 9 \\ 1 & 2 & 1 \end{vmatrix} \begin{vmatrix} 2 & 3 & 4 \\ 2 & 4 & 6 \\ 3 & 5 & 8 \end{vmatrix}$

Exercise 2.5 Either by evaluating the determinant to show that $|A| = 0$, or find its reduced row echelon form which is

$$\begin{bmatrix} 1 & 0 & -\frac{1}{3} \\ 0 & 1 & \frac{5}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

and not equal to the identity.

Exercise 2.6

$$\det(A) = -42$$

$$\det\left(A^{-1}\right) = -\frac{1}{42}$$

$$A^{-1} = \begin{bmatrix} \frac{9}{7} & -\frac{3}{14} & -\frac{2}{7} & -\frac{6}{7} \\ -\frac{1}{21} & -\frac{1}{21} & \frac{1}{21} & \frac{1}{7} \\ -\frac{11}{7} & -\frac{15}{14} & \frac{11}{7} & \frac{5}{7} \\ \frac{3}{7} & \frac{3}{7} & -\frac{3}{7} & -\frac{2}{7} \end{bmatrix}$$

Exercise 2.7 By cofactor expansion along the first column, we can obtain the determinant of A as

$$|A| = 2p^2 + 4p - 16$$

which has two roots, $p = -4$ and $p = 2$ such that $|A| = 0$ and A is not invertible. All values of p other than $p = -4$ and $p = 2$ make A invertible.

Exercise 2.8 $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A^T + A,$
and $(A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T).$ We can split A into

$$A = A + \frac{1}{2}(A^T - A^T)$$

$$= \frac{1}{2}A + \frac{1}{2}A + \frac{1}{2}A^T - \frac{1}{2}A^T$$

$$= \frac{1}{2}A + \frac{1}{2}A^T + \frac{1}{2}A - \frac{1}{2}A^T$$

$$= \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

where the first term is symmetric and the second term is skew-symmetric.

Exercise 2.9

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$$

$$\det\left(A^{-1}\right) = \det\left(\frac{1}{\det(A)}\text{adj}(A)\right) \qquad \text{(Notice that } \frac{1}{\det(A)} \text{ is now a scalar)}$$

$$\frac{1}{\det(A)} = (\frac{1}{\det(A)})^n \det(\text{adj}(A))$$

$$\det(\text{adj}(A)) = (\det(A))^{n-1}$$

Exercise 3.1

$$A^{-1} = \begin{bmatrix} \frac{1}{21} & \frac{5}{21} & \frac{2}{21} \\ \frac{11}{42} & -\frac{29}{42} & \frac{1}{42} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \end{bmatrix}$$

$$\vec{x} = A^{-1}\vec{h} = \begin{bmatrix} \frac{1}{21} & \frac{5}{21} & \frac{2}{21} \\ \frac{11}{42} & -\frac{29}{42} & \frac{1}{42} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \end{bmatrix} \begin{bmatrix} 6 \\ \frac{7}{2} \\ -\frac{13}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -1 \\ \frac{3}{2} \end{bmatrix}$$

or

$$\begin{bmatrix} 5 & 1 & 3 & 6 \\ 2 & -1 & 1 & \frac{7}{2} \\ 3 & 2 & -4 & -\frac{13}{2} \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & \frac{1}{5} & \frac{3}{5} & \Big| & \frac{6}{5} \\ 2 & -1 & 1 & \Big| & \frac{7}{2} \\ 3 & 2 & -4 & \Big| & -\frac{13}{2} \end{bmatrix} \qquad \frac{1}{5}R_1 \rightarrow R_1$$

$$\rightarrow \begin{bmatrix} 1 & \frac{1}{5} & \frac{3}{5} & \Big| & \frac{6}{5} \\ 0 & -\frac{7}{5} & -\frac{1}{5} & \Big| & \frac{11}{10} \\ 0 & \frac{7}{5} & -\frac{29}{5} & \Big| & -\frac{101}{10} \end{bmatrix} \qquad R_2 - 2R_1 \rightarrow R_2, R_3 - 3R_1 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & \frac{1}{5} & \frac{3}{5} & \Big| & \frac{6}{5} \\ 0 & -\frac{7}{5} & -\frac{1}{5} & \Big| & \frac{11}{10} \\ 0 & 0 & -6 & \Big| & -9 \end{bmatrix} \qquad R_3 + R_2 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & \frac{1}{5} & \frac{3}{5} & \Big| & \frac{6}{5} \\ 0 & 1 & \frac{1}{7} & \Big| & -\frac{11}{14} \\ 0 & 0 & 1 & \Big| & \frac{3}{2} \end{bmatrix} \qquad -\frac{5}{7}R_2 \rightarrow R_2, -\frac{1}{6}R_3 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & \frac{1}{5} & 0 & \Big| & \frac{3}{10} \\ 0 & 1 & 0 & \Big| & -1 \\ 0 & 0 & 1 & \Big| & \frac{3}{2} \end{bmatrix} \qquad R_1 - \frac{3}{5}R_3 \rightarrow R_1, R_2 - \frac{1}{7}R_3 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & \Big| & \frac{1}{2} \\ 0 & 1 & 0 & \Big| & -1 \\ 0 & 0 & 1 & \Big| & \frac{3}{2} \end{bmatrix} \qquad R_1 - \frac{1}{5}R_2 \rightarrow R_1$$

Exercise 3.2

$$A^{-1} = \begin{bmatrix} -\frac{1}{8} & 0 & \frac{7}{8} \\ \frac{3}{16} & -\frac{1}{2} & -\frac{5}{16} \\ \frac{1}{16} & \frac{1}{2} & -\frac{7}{16} \end{bmatrix}$$

$$\vec{x}_1 = A^{-1}\vec{h}_1 \qquad\qquad \vec{x}_2 = A^{-1}\vec{h}_2$$

$$= \begin{bmatrix} -\frac{1}{8} & 0 & \frac{7}{8} \\ \frac{3}{16} & -\frac{1}{2} & -\frac{5}{16} \\ \frac{1}{16} & \frac{1}{2} & -\frac{7}{16} \end{bmatrix} \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix} \qquad = \begin{bmatrix} -\frac{1}{8} & 0 & \frac{7}{8} \\ \frac{3}{16} & -\frac{1}{2} & -\frac{5}{16} \\ \frac{1}{16} & \frac{1}{2} & -\frac{7}{16} \end{bmatrix} \begin{bmatrix} \frac{19}{4} \\ 1 \\ \frac{5}{4} \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} \qquad\qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{4} \end{bmatrix}$$

Exercise 3.3

$$\begin{bmatrix} 3 & 0 & 4 & 2 \\ 1 & 1 & 2 & -1 \\ 1 & -2 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 0 & 0 \\ 1 & 1 & 2 & -1 \\ 3 & 0 & 4 & 2 \end{bmatrix} \qquad R_1 \leftrightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 3 & 2 & -1 \\ 0 & 6 & 4 & 2 \end{bmatrix} \qquad R_2 - R_1 \rightarrow R_2, R_3 - 3R_1 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 6 & 4 & 2 \end{bmatrix} \qquad \frac{1}{3}R_2 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 4 \end{bmatrix} \qquad R_3 - 6R_2 \rightarrow R_3$$

The last row is inconsistent and the system has no solution.
Note: You may get, to the right of the last row, some number other than 4, but this is possible and not wrong. (Why?)

Exercise 3.4

$$\begin{bmatrix} 1 & 1 & -1 & -3 & 2 \\ 1 & 0 & 0 & -1 & 5 \\ 3 & 2 & -2 & -7 & 9 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & -1 & 5 \\ 1 & 1 & -1 & -3 & 2 \\ 3 & 2 & -2 & -7 & 9 \end{bmatrix} \qquad R_1 \leftrightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & -1 & 5 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 2 & -2 & -4 & -6 \end{bmatrix} \qquad R_2 - R_1 \rightarrow R_2, R_3 - 3R_1 \rightarrow R_3$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & -1 & 5 \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad R_3 - 2R_2 \rightarrow R_3$$

Let $p = s$, $q = t$ as the two free variables. Substituting them back into the equations, we have $m - t = 5$ and $n - s - 2t = -3$, hence $m = 5 + t$ and $n = -3 + s + 2t$, and

$$\begin{bmatrix} m \\ n \\ p \\ q \end{bmatrix} = \begin{bmatrix} 5 + t \\ -3 + s + 2t \\ s \\ t \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

Exercise 3.5 The determinant of the coefficient matrix can be found to be

$$\begin{vmatrix} 1 & 0 & \alpha \\ 0 & \alpha & 0 \\ \alpha & 0 & 1 \end{vmatrix} = -\alpha^3 + \alpha$$

$$= -\alpha(\alpha - 1)(\alpha + 1)$$

The system will have no solution or infinitely many of them only when the determinant equals to zero, which gives us three possible values of $\alpha = -1, 0, 1$. When $\alpha = -1$, the system is

$$\begin{bmatrix} 1 & 0 & -1 & | & -1 \\ 0 & -1 & 0 & | & 0 \\ -1 & 0 & 1 & | & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & | & -1 \\ 0 & -1 & 0 & | & 0 \\ 0 & 0 & 0 & | & -2 \end{bmatrix} \qquad R_3 + R_1 \rightarrow R_3$$

where the last row is inconsistent and there is no solution. When $\alpha = 0$, it becomes

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

It is obvious that $x = z = 0$, and $y = t$ is a free variable, so the solution is infinitely many and is in the form of

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

The last case, $\alpha = 1$, gives rise to the system of

$$
\left[\begin{array}{ccc|c}
1 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1
\end{array}\right]
\rightarrow
\left[\begin{array}{ccc|c}
1 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]
\qquad R_3 - R_1 \rightarrow R_3
$$

such that $y = 0$ and $z = t$ can be set to be a free variable and there are infinitely many solutions in the form of

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

Exercise 3.6 The system can be written as

$$
\begin{bmatrix}
0.4 & 0.5 & 0.1 \\
0.55 & 0.4 & 0.05 \\
0.45 & 0.45 & 0.1
\end{bmatrix}
\begin{bmatrix} gy \\ pk \\ bk \end{bmatrix} =
\begin{bmatrix} 2.645 \\ 2.6325 \\ 2.65 \end{bmatrix}
$$

and has a unique solution of $gy = 2.65$ (quartz), $pk = 2.55$ (feldspar), $bk = 3.1$ (biotite).

Exercise 3.7 The first two equations below come from the left inner loop and right inner loop, but one of them can be replaced by the outer loop as well.

$$
-4I_1 + 6I_2 = 6
$$
$$
-6I_2 + 9I_3 = -12
$$
$$
I_1 + I_2 + I_3 = 0
$$

and the solution is $I_1 = -\frac{3}{19}$, $I_2 = \frac{17}{19}$, $I_3 = -\frac{14}{19}$ (in Amperes).

Exercise 3.8 Substituting the given wave solution forms into the equation, we have

$$
\omega \tilde{\eta} \sin(kx + ly - \omega t) + H(-k\tilde{U} \sin(kx + ly - \omega t)
$$
$$
- l\tilde{V} \sin(kx + ly - \omega t)) = 0
$$

$$\omega \tilde{U} \sin(kx + ly - \omega t) = gk\tilde{\eta} \sin(kx + ly - \omega t)$$
$$\omega \tilde{V} \sin(kx + ly - \omega t) = gl\tilde{\eta} \sin(kx + ly - \omega t)$$

Cancelling out all the sine factors, we arrive at the linear system displayed in the question

$$\begin{cases} \omega\tilde{\eta} - kH\tilde{U} - lH\tilde{V} & = 0 \\ \omega\tilde{U} - kg\tilde{\eta} & = 0 \\ \omega\tilde{V} - lg\tilde{\eta} & = 0 \end{cases}$$

For $\tilde{U}, \tilde{V}, \tilde{\eta}$ to have a non-trivial solution other than all zeros, we require the determinant of the corresponding coefficient matrix to be zero according to Theorem 3.1.2, which leads to

$$\begin{vmatrix} \omega & -kH & -lH \\ -kg & \omega & 0 \\ -lg & 0 & \omega \end{vmatrix} = 0$$
$$\omega^3 - gHk^2\omega - gHl^2\omega = 0$$
$$\omega^2 - gH(k^2 + l^2) = 0$$

as the dispersion relation of gravity wave.

Exercise 3.9 $T_{cd} \approx 15.9\,°C$, $z_{cd} \approx 0.97\,\text{km}$.

Exercise 3.10 $x = 23$, $y = 12$. For the extra part, the new system of equations become (denote the number of third species as z)

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 48 \\ 122 \end{bmatrix}$$

By Gaussian Elimination, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 48 \\ 2 & 4 & 3 & 122 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 48 \\ 0 & 2 & 1 & 26 \end{array}\right] \qquad R_2 - 2R_1 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 48 \\ 0 & 1 & \frac{1}{2} & | & 13 \end{bmatrix} \qquad \frac{1}{2}R_2 \rightarrow R_2$$

$$\rightarrow \begin{bmatrix} 1 & 0 & \frac{1}{2} & | & 35 \\ 0 & 1 & \frac{1}{2} & | & 13 \end{bmatrix} \qquad R_1 - R_2 \rightarrow R_1$$

Let $z = t$ as the free variable, then we have $y = 13 - \frac{1}{2}t$ and $x = 35 - \frac{1}{2}t$, and hence

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 35 - \frac{1}{2}t \\ 13 - \frac{1}{2}t \\ t \end{bmatrix} = \begin{bmatrix} 35 \\ 13 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Since the numbers of species must be a non-negative integer, the solution can be expressed in a more good-looking form of

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 35 \\ 13 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

where $s = \frac{t}{2}$, and the range of s is $0, 1, \ldots, 13$ (when s reaches 13 there is no chicken remained).

Index

Bibliography

[1]   H. Anton and C. Rorres. *Elementary Linear Algebra: Applications Version, 11th Edition*. John Wiley & Sons Incorporated, 2013. ISBN: 9781118879160.

[2]   K.F. Riley, M.P. Hobson, and S.J. Bence. *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*. Cambridge University Press, 2006. ISBN: 9781139450997.