A Review of Matrix Analysis

Part A – Matrix Operations

1.1 Matrix Notation

Matrices are simply rectangular arrays of quantities. Each quantity in the array is called an *element* of the matrix and an element can be either a numerical value, a symbolic expression or another matrix. The number of rows and columns in the matrix indicate the size of the matrix.

Definition 1.1: The number of rows in a matrix is called the *row dimension* of the matrix and the number of columns is called the *column dimension*. A matrix with M rows and N columns is called an $M \times N$ matrix and has MN elements.

Notice that the row dimension is always stated first, followed by the column dimension. Throughout this review, symbolic matrices will always be denoted using boldface type, while the elements of matrices will be put in square brackets. For example, a 3×4 matrix of ones would be denoted by

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

If a matrix is used in a purely symbolic fashion, it may be helpful to indicate its dimensions. For example if **B** is a 2×4 matrix and **C** is a 4×4, then the product of **B** and **C** might be written as $\mathbf{B}_{2 \times 4}\mathbf{C}_{4 \times 4}$.

Note 1.1: A matrix with only one row and one column is *equivalent to a scalar.*

Definition 1.2: If we turn every row in a given matrix, **A**, into a column in a new matrix, **B**, the new matrix, **B**, is called the *transpose* of **A**, denoted by $\mathbf{B} = \mathbf{A}^T$. If **A** is $M \times N$, then **B** will be $N \times M$.

Example 1.1: If

$$\mathbf{A} = \begin{bmatrix} 3 & -2 & 0 \\ 1 & 5 & -4 \end{bmatrix} \text{ then } \mathbf{A}^T = \begin{bmatrix} 3 & 1 \\ -2 & 5 \\ 0 & -4 \end{bmatrix}.$$

Definition 1.3: A matrix with M rows and one column is called a *vector.*, while a matrix with one row and M columns is the *transpose of a vector.*, but is *not* a vector itself.

Definition 1.4: Any ordered array of numbers or expressions can be considered a vector. For example the array $(x, 2, 3, y, z)$ can be written as the vector

$$\mathbf{v} = \begin{bmatrix} x \\ 2 \\ 3 \\ y \\ z \end{bmatrix} \text{ and } \mathbf{v}^T = \begin{bmatrix} x & 2 & 3 & y & z \end{bmatrix}.$$

Note 1.2: Vectors, by default, are *always* matrices with *one column*. Matrices with only one row are the *transpose* of vectors, but are usually not referred to as vectors.

By convention, we will always "try" to use lower case letters for vectors and their transposes and upper case letters for matrices whose row and column dimensions are both greater than one. Notice that if **a** is an $1 \times M$ vector and **b** is an $M \times 1$ vector with the same elements as **a**, they are not equal to one another, since they have different row and column dimensions.

A subscript notation is used to refer to the elements of a matrix. For example, if **A** is a 5×3 matrix, the element in row 4 and column 2 would be indicated by A_{42}. Again, notice that the row number is the first subscript, the column number is the second.

Example 1.2: For the vector **v** in the last example we have

$$v_{41} = y = (v^T)_{14},$$

however the element v_{14} does not exist.

1.2 Matrix Addition and Subtraction

Only matrices with equal numbers of rows and columns can be added or subtracted. The operation of addition or subtraction is carried out by adding or subtracting the corresponding elements of each matrix.

Definition 1.5: If $\mathbf{A}_{M \times N}$ has elements A_{ij} and $\mathbf{B}_{M \times N}$ has elements B_{ij}, then their *sum* or *difference* is given by

$$\mathbf{C}_{M \times N} = \mathbf{A}_{M \times N} \pm \mathbf{B}_{M \times N} = \begin{cases} \mathbf{A}_{M \times N} + \mathbf{B}_{M \times N} \\ \mathbf{A}_{M \times N} - \mathbf{B}_{M \times N} \end{cases}, \tag{1.1}$$

and the elements C_{ij} of $\mathbf{C}_{M \times N}$ are computed using

$$C_{ij} = A_{ij} \pm B_{ij} = \begin{cases} A_{ij} + B_{ij} \\ A_{ij} - B_{ij} \end{cases}, i = 1, \ldots, M, j = 1, \ldots, N. \tag{1.2}$$

Matrix addition and subtraction have the additional property that they are associative, i.e., they satisfy

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \text{ and } \mathbf{C} = \mathbf{A} - \mathbf{B} = -\mathbf{B} + \mathbf{A}. \tag{1.3}$$

Example 1.3: Find the sum and difference of **A** and **B** if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & a & z \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} z^2 & -1 & 2 \\ 4 & 5 & 6 \end{bmatrix}.$$

Solution: The sum is

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = \begin{bmatrix} z^2+1 & 1 & 5 \\ 8 & a+5 & z+6 \end{bmatrix}_{2 \times 3},$$

while the difference is

$$\mathbf{C} = \mathbf{A} - \mathbf{B} = \begin{bmatrix} -z^2+1 & 3 & 1 \\ 0 & a-5 & z-6 \end{bmatrix}_{2 \times 3}.$$

1.3 Multiplication

A matrix can be multiplied by either a scalar or another matrix and each of these operations are discussed below.

1.3.1 Multiplying a Matrix by a Scalar

When a matrix is multiplied by a scalar, the result is a new matrix of the same size as the original, where each element of the new matrix is multiplied by the scalar.

Definition 1.6: If x is a scalar and **A** is an $M \times N$ matrix, then the *product* of x with **A** is defined by

$$\mathbf{B}_{M \times N} = x\mathbf{A}_{M \times N} = \mathbf{A}_{M \times N}x, \tag{1.4}$$

where

$$B_{ij} = xA_{ij} = A_{ij}x, i = 1, \quad , M, j = 1, \quad , N. \tag{1.5}$$

Example 1.4: If $x = 4$ and $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then

$$\mathbf{B} = x\mathbf{A} = 6\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}6 = \begin{bmatrix} 6 & 12 & 18 \\ 24 & 30 & 36 \end{bmatrix}.$$

1.3.2 Multiplying Two Matrices

Multiplying two matrices is a bit more complicated than multiplying a matrix by a scalar. The basic rules for this type of multiplication are outlined in the definition below.

Definition 1.7: Two matrices, $\mathbf{A}_{M \times N}$ and $\mathbf{B}_{P \times Q}$ can be multiplied if and only if the column dimension of the first equal to the row dimension of the second (i.e., $N = P$). The result is a new matrix, $\mathbf{C}_{R \times S}$ whose row dimension is equal to the row dimension of **A** (i.e., $R = M$) and whose column dimension is equal to the column dimension of **B** (i.e., $S = Q$). Thus

$$\mathbf{A}_{M \times N}\mathbf{B}_{N \times P} = \mathbf{C}_{M \times P}. \tag{1.6}$$

Notice that in general, matrix multiplication is not commutative, i.e.,

$$\mathbf{C}_{M \times P} = \mathbf{A}_{M \times N}\mathbf{B}_{N \times P} \neq \mathbf{B}_{N \times P}\mathbf{A}_{M \times N}, \tag{1.7}$$

except in the special case when $M = P$.

The elements of **C** are computed using

$$C_{ij} = \sum_{k=1}^{N} A_{ik}B_{kj}, i = 1, \quad , M, j = 1, \quad , P. \tag{1.8}$$

An equivalent way of thinking of Equation (1.8) is by taking the elements of row i of **A** and multiplying them by column j of **B** and adding the results. Hence if row i of **A** is $\mathbf{a}_i$ and column j of **B** is $\mathbf{b}_j$, then

$$C_{ij} = \mathbf{a}_i^T\mathbf{b}_j. \tag{1.9}$$

Example 1.5: Determine the product of the matrices **A** and **B** if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$$

Solution: The product is the 2×3 matrix $\mathbf{C} = \mathbf{AB}$ and since **A** and **B** satisfy the rules for matrix multiplication, then the elements of **C** are given by

$$C_{ij} = \sum_{k=1}^{3} A_{ik}B_{kj}, i = 1, \quad , 2, j = 1, \quad , 3.$$

Hence

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} = 1(9) + 2(6) + 3(3) = 30,$$
$$C_{12} = A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{23} = 1(8) + 2(5) + 3(2) = 24,$$
$$C_{13} = A_{11}B_{13} + A_{12}B_{23} + A_{13}B_{33} = 1(7) + 2(4) + 3(1) = 18,$$
$$C_{21} = A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} = 4(9) + 5(6) + 6(3) = 84,$$
$$C_{22} = A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{23} = 4(8) + 5(5) + 6(2) = 69,$$
$$C_{23} = A_{21}B_{13} + A_{22}B_{23} + A_{23}B_{33} = 4(7) + 5(4) + 6(1) = 54,$$

and

$$\mathbf{C} = \begin{bmatrix} 30 & 24 & 18 \\ 84 & 69 & 54 \end{bmatrix}.$$

Note 1.3: Although the expressions above are correct, it is much easier to think of matrix multiplication in terms of products of the rows of the first matrix times the columns of the second matrix.

Example 1.6: Since each row of a matrix is itself a matrix and each column of a matrix is also matrix then we can write any matrix using these concepts. For example, the matrix **B** from the previous example can be written in the following ways:

$$\mathbf{B} = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix},$$

where

$$\mathbf{r}_1 = \begin{bmatrix} 9 & 8 & 7 \end{bmatrix}, \ \mathbf{r}_2 = \begin{bmatrix} 6 & 5 & 4 \end{bmatrix}, \text{ and } \mathbf{r}_3 = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix},$$

are row matrices of **B** and

$$\mathbf{c}_1 = \begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix}, \ \mathbf{c}_2 = \begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix}, \text{ and } \mathbf{c}_3 = \begin{bmatrix} 7 \\ 4 \\ 1 \end{bmatrix},$$

are column matrices of **B**.

Example 1.7: If we write the matrices of Example(1.5) in the forms

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix},$$

then the product of **A** times **B** can be written in the form

$$\mathbf{AB} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1\mathbf{c}_1 & \mathbf{r}_1\mathbf{c}_2 & \mathbf{r}_1\mathbf{c}_3 \\ \mathbf{r}_2\mathbf{c}_1 & \mathbf{r}_2\mathbf{c}_2 & \mathbf{r}_3\mathbf{c}_3 \end{bmatrix},$$

where

$$C_{11} = \mathbf{r}_1\mathbf{c}_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 30 \end{bmatrix} = 30, \ C_{12} = \mathbf{r}_1\mathbf{c}_2 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 24 \end{bmatrix} = 24,$$

$$C_{13} = \mathbf{r}_1\mathbf{c}_3 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \end{bmatrix} = 18, \ C_{21} = \mathbf{r}_2\mathbf{c}_1 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 84 \end{bmatrix} = 84,$$

$$C_{22} = \mathbf{r}_2\mathbf{c}_2 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 69 \end{bmatrix} = 69, \ C_{23} = \mathbf{r}_2\mathbf{c}_3 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 54 \end{bmatrix} = 54.$$

Hence

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1\mathbf{c}_1 & \mathbf{r}_1\mathbf{c}_2 & \mathbf{r}_1\mathbf{c}_3 \\ \mathbf{r}_2\mathbf{c}_1 & \mathbf{r}_2\mathbf{c}_2 & \mathbf{r}_3\mathbf{c}_3 \end{bmatrix} = \begin{bmatrix} 30 & 24 & 18 \\ 84 & 69 & 54 \end{bmatrix}.$$

The two kinds of multiplication discussed so far are the only kinds defined in classical matrix analysis. However we should make note of the fact that some classical vector operations can be represented in matrix notation.

Definition 1.8: The *scalar (or dot or inner) product* between two vectors

$$\mathbf{a} = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + \quad + a_N\mathbf{i}_N \text{ and } \mathbf{b} = b_1\mathbf{i}_1 + b_2\mathbf{i}_2 + \quad + b_N\mathbf{i}_N,$$

is defined by

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1 + a_2b_2 + \quad + a_Nb_N = \mathbf{b}\cdot\mathbf{a},$$

in classical vector notation. In matrix notation, we have

$$\mathbf{a}_{N\times 1} = \begin{bmatrix} a_1 & a_2 & \quad & a_N \end{bmatrix}^T \text{ and } \mathbf{b}_{N\times 1} = \begin{bmatrix} b_1 & b_2 & \quad & b_N \end{bmatrix}^T,$$

and we see that

$$\mathbf{a}^T_{N\times 1}\mathbf{b}_{N\times 1} = \begin{bmatrix} a_1 & a_2 & \quad & a_N \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \\ b_N \end{bmatrix} = a_1b_1 + a_2b_2 + \quad + a_Nb_N = \mathbf{b}^T_{N\times 1}\mathbf{a}_{N\times 1},$$

hence

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{a}^T\mathbf{b} = \mathbf{b}\cdot\mathbf{a} = \mathbf{b}^T\mathbf{a}. \tag{1.10}$$

Definition 1.9: The *magnitude* of a vector is defined by

$$|\mathbf{a}| = \sqrt{\mathbf{a}\cdot\mathbf{a}} = \sqrt{\mathbf{a}^T\mathbf{a}}. \tag{1.11}$$

Definition 1.10: The *cross-product* in classical vector notation can only be carried out between two three dimensional vectors. The operation is defined by

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i}_1 + (a_3b_1 - a_1b_3)\mathbf{i}_2 + (a_1b_2 - a_2b_1)\mathbf{i}_3 = -(\mathbf{b}\times\mathbf{a}), \tag{1.12}$$

in classical vector notation, while in matrix notation we have

$$\mathbf{a}\times\mathbf{b} = \mathbf{A}\mathbf{b} = \mathbf{B}\mathbf{a},\ -(\mathbf{b}\times\mathbf{a}) = -\mathbf{A}^T\mathbf{b} = -\mathbf{B}^T\mathbf{a} \tag{1.13}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{bmatrix} = -\mathbf{A}^T,\ \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{1.14}$$

and

$$\mathbf{B} = \begin{bmatrix} 0 & b_3 & -b_2 \\ -b_3 & 0 & b_1 \\ b_2 & -b_1 & 0 \end{bmatrix} = -\mathbf{B}^T, \ \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}. \tag{1.15}$$

Definition 1.11: The *outter product* between two vectors

$$\mathbf{a} = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + \quad + a_N\mathbf{i}_N \text{ and } \mathbf{b} = b_1\mathbf{i}_1 + b_2\mathbf{i}_2 + \quad + b_N\mathbf{i}_N,$$

is defined by

$$\mathbf{a} \quad \mathbf{b} \equiv \mathbf{a}\mathbf{b}^T. \tag{1.16}$$

If

$$\mathbf{a}_{N \times 1} = \begin{bmatrix} a_1 & a_2 & \quad & a_N \end{bmatrix}^T \text{ and } \mathbf{b}_{N \times 1} = \begin{bmatrix} b_1 & b_2 & \quad & b_N \end{bmatrix}^T,$$

then the outer product of two vectors produces the matrix

$$\mathbf{a}_{N \times 1}\mathbf{b}^T_{N \times 1} = \begin{bmatrix} a_1 \\ a_2 \\ \\ a_N \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \quad & b_N \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & \quad & a_1b_N \\ a_2b_1 & a_2b_2 & \quad & a_2b_N \\ & & & \\ a_Nb_1 & a_Nb_2 & \quad & a_Nb_N \end{bmatrix} = \mathbf{C}_{N \times N}. \tag{1.17}$$

1.4 Transposes and the Identity Matrix

Some basic rules for the transpose operation are listed below. In this list **A** and **B** are real matrices and x is a real scalar.

$$(\mathbf{A}^T)^T = \mathbf{A}. \tag{1.18}$$

$$(\mathbf{A} \pm \mathbf{B})^T = \mathbf{A}^T \pm \mathbf{B}^T = \pm\mathbf{B}^T + \mathbf{A}^T. \tag{1.19}$$

$$(x\mathbf{A})^T = x\mathbf{A}^T = \mathbf{A}^T x. \tag{1.20}$$

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T \neq \mathbf{A}^T\mathbf{B}^T. \tag{1.21}$$

The identity matrix is a special square matrix, usually denoted by **I**, which has the following properties:

$$\mathbf{I}_{M \times M}\mathbf{A}_{M \times N} = \mathbf{A}_{M \times N}. \tag{1.22}$$

$$\mathbf{A}_{M \times N}\mathbf{I}_{N \times N} = \mathbf{A}_{M \times N}. \tag{1.23}$$

Definition 1.12: The elements of the $N \times N$ *identity matrix*, I are given by

$$I_{ij} \equiv \delta_{ij} \equiv \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \ i, j = 1, \quad, N, \tag{1.24}$$

where δ_{ij} is called the *Kronecker delta.*

1.5 Some Special Matrices

1.5.1 The Zero Matrix

The symbol **0** is used for the $M \times N$ matrix whose elements are all zero and this matrix is sometimes called the *null matrix*. For example, the matrix equation $\mathbf{A} - \mathbf{B} = \mu\mathbf{C} - \mathbf{D}$, can be written in the alternative form

$$\mathbf{A} - \mathbf{B} - \mu\mathbf{C} + \mathbf{D} = \mathbf{0},$$

where **0** is an appropriately sized zero matrix.

1.5.2 Diagonal Matrices

A square matrix is diagonal if non–zero elements occur only on the main diagonal.

Definition 1.13: The $N \times N$ matrix **A** with elements $A_{ij}, i, j = 1, \quad , N$, is **diagonal** iff

$$A_{ij} \equiv 0, i \neq j. \qquad (1.25)$$

Definition 1.14: The *main diagonal* of the $N \times N$ matrix **A** are the elements

$$A_{ii}, i = 1, \quad , N, \qquad (1.26)$$

and is often displayed as a row matrix in the form

$$\mathbf{D} = \begin{bmatrix} A_{11} & & A_{NN} \end{bmatrix}. \qquad (1.27)$$

The *first sub–diagonal* of **A** are the elements

$$A_{i, i-1}, i = 2, \quad , N, \qquad (1.28)$$

the *second sub–diagonal* of **A** are the elements

$$A_{i, i-2}, i = 3, \quad , N, \qquad (1.29)$$

and the j^{th} *sub–diagonal* of **A** are the elements

$$A_{i, i-j}, i = j+1, \quad , N. \qquad (1.30)$$

The *first super–diagonal* of **A** are the elements

$$A_{i, i+1}, i = 1, \quad , N-1, \qquad (1.31)$$

the *second super–diagonal* of **A** are the elements

$$A_{i, i+2}, i = 1, \quad , N-2, \qquad (1.32)$$

and the j^{th} *super–diagonal* of **A** are the elements

$$A_{i, i+j}, i = 1, \quad , N-j. \qquad (1.33)$$

1.5.3 Triangular Matrices

A matrix is *lower–triangular* if all elements above the main diagonal are zero and is *upper–triangular* if all elements below the main diagonal are zero.

Definition 1.15: The $M \times N$ matrix **A** with elements A_{ij}, is *lower–triangular* if

$$A_{ij} = 0, \quad i < j, \tag{1.34}$$

and is *upper–triangular* if

$$A_{ij} = 0, \quad i > j. \tag{1.35}$$

Definition 1.16: The $N \times N$ matrix **A** with elements $A_{ij}, i, j = 1, \ldots, N$, is *unit lower–triangular* if

$$A_{ij} = 0, \quad i < j, \text{ and } A_{ii} = 1, i = 1, \ldots, N \tag{1.36}$$

and is *unit upper–triangular* if

$$A_{ij} = 0, \quad i > j, \text{ and } A_{ii} = 1, i = 1, \ldots, N \tag{1.37}$$

1.5.4 Banded Matrices

Banded matrices are zero outside a certain band on either side of the main diagonal. If the number of nonzero sub–diagonals is h_L and the number of nonzero super–diagonals is h_U, then the bandwidth of the matrix is

$$b = h_L + h_U + 1. \tag{1.38}$$

The number of nonzero sub–diagonals is h_L is called the *lower bandwidth* of the matrix and the number of nonzero super–diagonals is h_U, is called the *upper bandwidth* of the matrix. If $h_L = h_U = h$, then h is called the *half–bandwidth* of the matrix. Banded matrices are often referred to using the notation (h_L, h_U).

Example 1.8: The matrix given by

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 4 & 1 & 0 & 0 & 0 \\ 6 & 5 & 3 & 1 & 0 & 0 \\ 0 & 6 & 5 & 3 & 1 & 0 \\ 0 & 0 & 6 & 5 & 3 & 1 \\ 0 & 0 & 0 & 6 & 4 & 2 \end{bmatrix},$$

has a lower bandwidth of $h_L = 2$ and an upper bandwidth of $b_U = 1$. Hence this is an $(2, 1)$–band matrix with a bandwidth of $b = 2 + 1 + 1 = 4$.

An $N \times N$ matrix with an (h_L, h_U)–band can be stored more compact as a matrix with dimensions $(h_L + h_U + 1) \times N = b \times N$. Hence instead of requiring N^2 storage locations, it can be stored in bN locations. For example if **A** is a 100×100 matrix with lower and upper band-widths of 5 and 7, then instead of requiring $100^2 = 10000$ storage locations, **A** can be stored in

$$(5 + 7 + 1)100 = 1300$$

storage locations. Special techniques exist for solving banded matrix systems.

Definition 1.17: A *tridiagonal matrix* is a (1,1)–band matrix with the property that

$$A_{ij} = 0 \text{ if } |i - j| \geq 2 . \tag{1.39}$$

Definition 1.18: A *pentadiagonal matrix* is a (2,2)–band matrix with the property that

$$A_{ij} = 0 \text{ if } |i - j| \geq 3 . \tag{1.40}$$

1.6 Inverse of a Matrix

In classical matrix analysis, there is no such operation as division between matrices. In algebra, if we have the scalar equation $ax = b$, then we can easily solve this to get

$$x = \frac{b}{a} = a^{-1}b = ba^{-1},$$

where $a^{-1} \equiv 1/a$. However if we have the matrix equation, $\mathbf{Ax} = \mathbf{b}$, then $\mathbf{x} = \mathbf{b}/\mathbf{A}$ has no meaning in classical matrix analysis. In fact, the solution to $\mathbf{Ax} = \mathbf{b}$ is usually written as

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \neq \mathbf{b}\mathbf{A}^{-1},$$

where $\mathbf{A}^{-1}$ ($\neq 1/\mathbf{A}$) is called the *inverse* of **A**.

Definition 1.19: If **A** is an $N \times N$ matrix, then the *inverse* of **A** is the $N \times N$ matrix **B** which satisfies

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}, \tag{1.41}$$

where **I** is the $N \times N$ identity matrix. Symbolically, $\mathbf{B} = \mathbf{A}^{-1}$, hence we normally write

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{1.42}$$

Some properties of the inverse are shown below.

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A} \quad . \tag{1.43}$$

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T. \tag{1.44}$$

$$(x\mathbf{A})^{-1} = x^{-1}\mathbf{A}^{-1} = \frac{\mathbf{A}^{-1}}{x}. \tag{1.45}$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{1.46}$$

$$(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}. \tag{1.47}$$

Note 1.4: Not all square matrices *have* an inverse.

Definition 1.20: If the $N \times N$ matrix **A** does not have an inverse, **A** is said to be *singular*. If **A** does have an inverse, the matrix is said to be *non–singular*.

In the following sections, we will see the conditions necessary for an inverse to exist.

1.7 Determinant of a Matrix

The determinant is an operation on a square matrix which produces a scalar. Later on we will see that the determinant is related to the question of invertibility.

Definition 1.21: If **A** is a 1×1 matrix whose *only* element is A_{11} then the *determinant* of **A** is defined by

$$\det(\mathbf{A}) \equiv A_{11}\,. \tag{1.48}$$

If **A** is a 2×2 matrix given by

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then the *determinant* of **A** is defined by

$$\det(\mathbf{A}) \equiv A_{11}A_{22} - A_{21}A_{12}\,. \tag{1.49}$$

If **A** is a 3×3 matrix given by

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \tag{1.50}$$

then the *determinant* of **A** is defined by

$$\det(\mathbf{A}) \equiv (A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{32}A_{21}) - (A_{11}A_{32}A_{23} + A_{12}A_{21}A_{33} + A_{13}A_{22}A_{31})\,. \tag{1.51}$$

For matrices beyond 3×3, the determinant is defined in terms of other operations, as shown below.

Definition 1.22: If **A** is an $N \times N$ matrix with elements A_{ij}, then the *minor* of A_{ij} is the determinant of the $(N-1) \times (N-1)$ matrix formed by deleting the i^{th} row and j^{th} column of **A** and is denoted by M_{ij}. Every element of **A** has a minor.

Definition 1.23: If **A** is an $N \times N$ matrix with minors M_{ij}, then the *cofactor* of A_{ij} is the number defined by

$$A^C_{ij} \equiv (-1)^{i+j} M_{ij} \,. \qquad (1.52)$$

The cofactors themselves form an $N \times N$ matrix, $\mathbf{A}^C$ called the *cofactor matrix* of **A**.

Definition 1.24: If **A** is an $N \times N$ matrix with elements A_{ij} and cofactors A^C_{ij} then the determinant of **A** is defined by

$$\det(\mathbf{A}) \equiv \begin{cases} \sum_{j=1}^{N} A_{ij} A^C_{ij}, & \text{expansion using row i} \\ \sum_{i=1}^{N} A_{ij} A^C_{ij}, & \text{expansion using column j} \end{cases} \,. \qquad (1.53)$$

Example 1.9: Compute the determinant of the matrix

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

by expanding row 3.

Solution: If we expand using row then

$$\det(\mathbf{A}) = A_{31}A^C_{31} + A_{32}A^C_{32} + A_{33}A^C_{33}$$

The minors for row 3 are

$$M_{31} = \det\begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix} = -6\,,\ M_{32} = \det\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} = 12\,,\ M_{33} = \det\begin{bmatrix} 4 & -2 \\ 0 & 1 \end{bmatrix} = 4\,,$$

and the corresponding cofactors are

$$A^C_{31} = (-1)^{3+1} M_{31} = -6\,,\ A^C_{32} = (-1)^{3+2} M_{32} = -12\,,\ A^C_{33} = (-1)^{3+3} M_{33} = 4\,.$$

Hence the determinant is

$$\det(\mathbf{A}) = 2(-6) + 4(-12) + (-3)4 = -72\,.$$

Definition 1.25: If **A** is an $N \times N$ matrix with cofactor matrix $\mathbf{A}^C$ then the *adjoint* of **A** is the $N \times N$ matrix defined by

$$\text{adj}(\mathbf{A}) \equiv (\mathbf{A}^C)^T \,. \qquad (1.54)$$

Definition 1.26: *Cramer's rule* for determining the *inverse* of **A** is defined by

$$\mathbf{A}^{-1} \equiv \frac{\text{adj}(\mathbf{A})}{\det(\mathbf{A})}. \tag{1.55}$$

Although Cramer's rule produces an exact inverse in infinite precision arithmetic, it is probably the worst method that you can use in finite precision arithmetic (the arithmetic of computers and calculators). However it does show us that if $\det(\mathbf{A}) = 0$, then the inverse of **A** does not exist. Hence we see that

1. If $\det(\mathbf{A}) = 0$ then **A** is *singular.*

2. If $\det(\mathbf{A}) \neq 0$ then **A** is *non–singular.*

If **A** and **B** are real $N \times N$ matrices and x is a real scalar, the determinant has the following properties:

$$\det(\mathbf{A}^T) = \det(\mathbf{A}). \tag{1.56}$$

$$\det(x\mathbf{A}) = x^N \det(\mathbf{A}). \tag{1.57}$$

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}). \tag{1.58}$$

$$\det(\mathbf{A} \pm \mathbf{B}) \neq \det(\mathbf{A}) \pm \det(\mathbf{B}). \tag{1.59}$$

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}. \tag{1.60}$$

1.8 Condition of a Matrix

When solving systems of linear equations of the form $\mathbf{Ax} = \mathbf{b}$, the symbolic solution is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ and we know that if $\det(\mathbf{A}) = 0$, then $\mathbf{A}^{-1}$ does not exist and the system either has no solution at all, or it has no unique solution. Although the determinant gives the condition for an exact singularity, it does not give a good indication of how close a matrix is to being singular. For example, consider the system $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 600 & 800 \\ 30001 & 40002 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 200 \\ 10001 \end{bmatrix}.$$

Since $\det(\mathbf{A}) = 400$ "seems" to be very large, we might conclude that this matrix is not very close to being singular. If we use Cramer's rule and exact arithmetic, the inverse of **A** is

$$\mathbf{A}^{-1} = \frac{\text{adj}(\mathbf{A})}{\det(\mathbf{A})} = \frac{\begin{bmatrix} 40002 & -800 \\ -30001 & 600 \end{bmatrix}}{400} = \begin{bmatrix} \frac{20001}{200} & -2 \\ -\frac{30001}{400} & \frac{3}{2} \end{bmatrix},$$

and the corresponding "exact" solution to this system is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} \frac{20001}{200} & -2 \\ -\frac{30001}{400} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 200 \\ 10001 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

However, if we change b_1 from 200 to 201, we find that the solution changes to

$$\mathbf{x} = \begin{bmatrix} \frac{20001}{200} & -2 \\ -\frac{30001}{400} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 201 \\ 10001 \end{bmatrix} = \begin{bmatrix} 99.005 \\ -74.0025 \end{bmatrix}.$$

Definition 1.27: For the system $\mathbf{Ax} = \mathbf{b}$, whenever "small" changes in either **A** or **b** lead to "large" changes in **x**, the system is said to be *ill–conditioned* and even though the determinant "seems" large, the matrix **A** is "nearly" singular.

Clearly the determinant is not a good measure of how close a matrix is to being singular. However, there is another measure of how close a matrix is to being singular called the *condition number* of the matrix.

Definition 1.28: The *condition number* of an $N \times N$ matrix **A** is defined as

$$\text{cond}(\mathbf{A}) \equiv ||\mathbf{A}||\,||\mathbf{A}^{-1}|| \geq 1, \tag{1.61}$$

where $||\mathbf{A}||$ is called the *norm* of **A** and is defined by

$$||\mathbf{A}|| \equiv \max_{j} \sum_{i=1}^{N} |A_{ij}| . \tag{1.62}$$

If we were able to perform all operations in infinite precision arithmetic, the every matrix would always have a condition number of exactly one. However since all calculator and computer operations are carried out in finite precision arithmetic, some matrices will have conditions number much larger than one.

If the condition number of a matrix is "small" then the matrix is well–conditioned and if it is "large" then the matrix is ill–conditioned. We can use the following rule–of–thumb:

1. if $\text{cond}(\mathbf{A}) \leq 1000$ then the matrix is "well–conditioned" (far from being singular).

2. if $\text{cond}(\mathbf{A}) \geq 10000$ then the matrix is "ill–conditioned" (close to being singular).

3. if $1000 < \text{cond}(\mathbf{A}) < 10000$ then the matrix is in the gray area and may be either.

As with all rules–of–thumb, it is NOT always valid. In practice, we do not use the formal definition of the condition number to compute its value, since this requires computing the inverse. There are special techniques for estimating the condition of a matrix.

1.9 Trace of a Matrix

The trace is another operation on a square matrix which produces a scalar.

Definition 1.29: If **A** is an $N \times N$ matrix, then the *trace* of **A** is simply the sum of all the diagonal elements of **A** and is defined by

$$\mathrm{tr}(\mathbf{A}) \equiv \sum_{i=1}^{N} A_{ii} \quad . \tag{1.63}$$

1.10 Partitioned Matrices

Earlier we discussed matrix multiplication by thinking of the rows of the first matrix as row matrices and the columns of the second matrix as column matrices. This process was a form of *matrix partitioning*. Whenever we break a matrix up into submatrices, this is referred to as partitioning.

As an example, consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}.$$

This matrix can be partitioned in the from

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{a}_2 \\ \mathbf{a}_3^T & \mathbf{A}_4 \end{bmatrix},$$

where

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{bmatrix}, \ \mathbf{a}_2 = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \ \mathbf{a}_3^T = \begin{bmatrix} 9 & 10 & 11 \end{bmatrix} \text{ and } \mathbf{A}_4 = \begin{bmatrix} 12 \end{bmatrix}.$$

Partitioning is advantageous when we want to isolate parts of a matrix to emphasize the roles those parts play in matrix equations. For example, it can be shown that if the 6×6 matrix **A** is partitioned in the form

$$\mathbf{A}_{6\times6} = \begin{bmatrix} \mathbf{P}_{4\times4} & \mathbf{Q}_{4\times2} \\ \mathbf{R}_{2\times4} & \mathbf{S}_{2\times2} \end{bmatrix},$$

and we define

$$\mathbf{W}_{2\times2} \equiv \mathbf{S}_{2\times2} - \mathbf{R}_{2\times4}\mathbf{P}_{4\times4}^{-1}\mathbf{Q}_{4\times2},$$

and

$$\mathbf{X}_{4\times4} \equiv \mathbf{P}_{4\times4}^{-1} + \mathbf{P}_{4\times4}^{-1}\mathbf{Q}_{4\times2}\mathbf{W}_{2\times2}^{-1}\mathbf{R}_{2\times4}\mathbf{P}_{4\times4}^{-1},$$

then the inverse of **A** is given by

$$\mathbf{A}_{6\times6}^{-1} = \begin{bmatrix} \mathbf{X}_{4\times4} & -\mathbf{P}_{4\times4}^{-1}\mathbf{Q}_{4\times2}\mathbf{W}_{2\times2}^{-1} \\ -\mathbf{W}_{2\times2}^{-1}\mathbf{R}_{2\times4}\mathbf{P}_{4\times4}^{-1} & \mathbf{W}_{2\times2}^{-1} \end{bmatrix}.$$

This means that we can invert a 6×6 matrix, **A** by inverting a 4×4 matrix, **P** and a 2×2 matrix, **W**.

Part B – Systems of Algebraic Equations

1.1 Basic Theory

We now consider *systems of linear algebraic equations*. For example, if we have a system with M equations in N unknowns, it can always be written in the form

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \quad + A_{1N}x_N &= b_1 \\ A_{21}x_1 + A_{22}x_2 + \quad + A_{2N}x_N &= b_2 \\ &. \\ A_{M1}x_1 + A_{M2}x_2 + \quad + A_{Mn}x_n &= b_M \end{aligned} \tag{1.1}$$

In order to simplify the task of dealing with such systems, we use matrix notation and define the $M \times N$ *coefficient matrix*

$$\mathbf{A} \equiv \begin{bmatrix} A_{11} & A_{12} & & A_{1N} \\ A_{21} & A_{22} & & A_{2N} \\ & & & \\ A_{M1} & A_{M2} & & A_{MN} \end{bmatrix}, \tag{1.2}$$

the $M \times 1$ *right-hand-side (rhs) vector* and the $N \times 1$ *vector of unknowns*

$$\mathbf{b} \equiv \begin{bmatrix} b_1 \\ b_2 \\ \\ b_M \end{bmatrix} \text{ and } \mathbf{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ \\ x_N \end{bmatrix}, \tag{1.3}$$

hence the system can be written in the compact form

$$\mathbf{Ax} = \mathbf{b}. \tag{1.4}$$

This represents 3 fundamentally different kinds of systems:

1. If $M < N$ then we have fewer equations than there are unknowns and the system is said to be *under-determined*. In this case, either no solution exists or an infinity of solutions exist.

2. If $M > N$ then we have more equations than there are unknowns and the system is said to be *over-determined.* In this case, either no solution exists or an infinity of solutions exist.

3. If $M = N$ then we have equal numbers of unknowns and equations and the system is said to be *determined* and is represented by a square matrix. In this case, either no solution exists or a unique solution exists.

In this chapter, the only case we consider is the case where $M = N$ and therefore if $\det(\mathbf{A}) \neq 0$, then $\mathbf{A}^{-1}$ exists and the unique solution is given by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{1.5}$$

If $\det(\mathbf{A}) = 0$, then $\mathbf{A}^{-1}$ does not exist and there is no solution to the system.

Computing the inverse of a matrix is computationally intensive and is often a very inefficient way to solve a system of equations. In the next sections we consider alternative methods of solving systems without ever computing the inverse of the matrix.

1.2 Gaussian Elimination

Gauss elimination is carried out by performing elementary row operations (EROs) on the matrix and its right hand side. There are three elementary row operations and the basic definitions relating to them are given below.

Definition 1.1: For an $M \times N$ matrix **A**, the three *elementary row operations* are:
ERO 1: interchange any two rows of **A**,
ERO 2: multiply any row of A by a nonzero scalar c,
ERO 3: replace the i^{th} row of **A** by the sum of the i^{th} row of **A** plus c times the j^{th} row of **A**, for $i \neq j$.

Theorem 1.1: Performing *any sequence* of EROs on the system $\mathbf{Ax} = \mathbf{b}$ does not change the solution **x**.

In order to perform EROs we define the following elementary vectors and matrices.

Definition 1.2: An $N \times 1$ *elementary vector* $\mathbf{e}_i$ is a column vector of zeros except in the i^{th} row, where there is a one.

Example 1.1: If $N = 6$, construct $\mathbf{e}_4$.

Solution: This vector is simply the 6×1 zero vector with a 1 in the 4th row, hence

$$\mathbf{e}_4 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T.$$

Definition 1.3: If **A** is an $M \times N$ matrix, then interchanging the i^{th} and j^{th} rows of **A** (ERO 1) is equivalent to $\mathbf{E}_{ij}\mathbf{A}$, where $\mathbf{E}_{ij}$ is the $M \times M$ matrix obtained by interchanging the i^{th} and j^{th} rows of the $M \times M$ identity matrix, **I**. The *elementary matrix* $\mathbf{E}_{ij}$ has the property that $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij}$.

Definition 1.4: If **A** is an $M \times N$ matrix, then multiplying the i^{th} row of **A** by a nonzero constant c (ERO 2) is equivalent to $\mathbf{Q}_i(c)\mathbf{A}$, where $\mathbf{Q}_i(c)$ is the $M \times M$ matrix

$$\mathbf{Q}_i(c) = \mathbf{I} + (c-1)\mathbf{e}_i\mathbf{e}_i^T. \qquad (1.6)$$

The *elementary matrix* $\mathbf{Q}_i(c)$ has the property that

$$[\mathbf{Q}_i(c)]^{-1} = \mathbf{Q}_i\left(\frac{1}{c}\right). \qquad (1.7)$$

Definition 1.5: If $\mathbf{A}$ is an $M \times N$ matrix, then adding c times the j^th row of $\mathbf{A}$ to the i^th row of $\mathbf{A}$ (ERO 3) is equivalent to $\mathbf{R}_{ij}(c)\mathbf{A}$, where $\mathbf{R}_{ij}(c)$ is the $M \times M$ matrix

$$\mathbf{R}_{ij}(c) = \mathbf{I} + c\mathbf{e}_i\mathbf{e}_j^T,\ i \neq j \tag{1.8}$$

The *elementary matrix* $\mathbf{R}_{ij}(c)$ has the property that

$$[\mathbf{R}_{ij}(c)]^{-1} = \mathbf{R}_{ij}(-c). \tag{1.9}$$

Summary: We have the following equivalences

1. $\mathbf{E}_{ij}\mathbf{A}$ is equivalent to interchanging the i^{th} and j^{th} rows of $\mathbf{A}$.

2. $\mathbf{Q}_i(c)\mathbf{A}$ is equivalent to multiplying the i^{th} row of $\mathbf{A}$ by $c \neq 0$

3. $\mathbf{R}_{ij}(c)\mathbf{A}$ is equivalent to adding $c \neq 0$ times the j^{th} row of $\mathbf{A}$ to the i^{th} row of $\mathbf{A}$.

Quick Tips: If $\mathbf{A}$ is an $M \times N$ matrix, we can quickly construct $\mathbf{Q}_i(c)$ by replacing the i^{th} diagonal element of the $M \times M$ identity matrix, $\mathbf{I}$, by c. Similarly, we can quickly construct $\mathbf{R}_{ij}(c)$ by replacing the ij^{th} element $(i \neq j)$ of the $M \times M$ identity matrix, $\mathbf{I}$, by c.

Example 1.2: Replace row 2 of the matrix

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

by row 2 minus 4.5 times row 4.

Solution: The elementary matrix $\mathbf{R}_{24}(-4.5)$ performs this operation and according to the quick tip

$$\mathbf{R}_{24}(-4.5) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Hence

$$\mathbf{A}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 4 & 2 & -1 \\ 1 & 0 & -2 & 3 \\ 5 & -1 & 0 & 2 \\ 1 & -1 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 2 & -1 \\ -3.5 & 4.5 & -15.5 & 12 \\ 5 & -1 & 0 & 2 \\ 1 & -1 & 3 & -2 \end{bmatrix},$$

which you can verify is the correct result.

Example 1.3: Multiply the 3rd row of the matrix $\mathbf{A}$ from the last example by 6.

Solution: The elementary matrix $\mathbf{Q}_3(6)$ performs this operation and according to the quick tip

$$\mathbf{Q}_3(6) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Hence

$$\mathbf{A}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 2 & -1 \\ 1 & 0 & -2 & 3 \\ 5 & -1 & 0 & 2 \\ 1 & -1 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 2 & -1 \\ 1 & 0 & -2 & 3 \\ 30 & -6 & 0 & 12 \\ 1 & -1 & 3 & -2 \end{bmatrix},$$

which you can verify is the correct result.

Definition 1.6: Simple *Gaussian Elimination* for the system $\mathbf{A}_{N \times N}\mathbf{x}_{N \times 1} = \mathbf{b}_{N \times 1}$ consists of two steps. Step 1, called *forward elimination*, consists of performing ERO's on the augmented matrix $\mathbf{A}_{N \times (N+1)}' = \begin{bmatrix} \mathbf{A}_{N \times N} & \mathbf{b}_{N \times 1} \end{bmatrix}$ until $\mathbf{A}'$ is upper-triangular. Step 2, called *backward elimination*, consists of solving the upper-triangular system obtained in Step 1 for the solution vector $\mathbf{x}_{N \times 1}$, starting with the last equation of the upper-triangular system and using simple substitution

Definition 1.7: Simple *Gauss-Jordan Elimination* for the system $\mathbf{A}_{N \times N}\mathbf{x}_{N \times 1} = \mathbf{b}_{N \times 1}$ consists of two steps. Step 1, also called *forward elimination*, is identical to Step 1 of simple Gaussian elimination. Step 2 consists of performing additional ERO's on the upper-triangular system obtained in Step 1, until the augmented matrix is the $N \times N$ identity matrix in the first N columns,

$$\mathbf{A}_{N \times (N+1)}' = \begin{bmatrix} \mathbf{I}_{N \times N} & \mathbf{x}_{N \times 1} \end{bmatrix}, \tag{1.10}$$

and is equivalent to the backward elimination of Gaussian elimination.

Definition 1.8: The *inverse* of the matrix $\mathbf{A}_{N \times N}$ can be obtained by a slightly modified version of Gauss-Jordan elimination. The process consists of performing ERO's on the augmented matrix

$$\mathbf{A}_{N \times 2N}' = \begin{bmatrix} \mathbf{A}_{N \times N} & \mathbf{e}_1 & \mathbf{e}_2 & & \mathbf{e}_N \end{bmatrix} \tag{1.11}$$

until the first N columns of $\mathbf{A}'$ is the $N \times N$ identity matrix. At this point the last N columns of

$$\mathbf{A}_{N \times 2N}' = \begin{bmatrix} \mathbf{I}_{N \times N} & \mathbf{A}_{N \times N}^{-1} \end{bmatrix}, \tag{1.12}$$

will be the $N \times N$ inverse of $\mathbf{A}_{N \times N}$.

Example 1.4: Solve the system

$$\begin{aligned} 2x + y - 2z &= 1 \\ 4x - y + 2z &= 5, \\ 2x - y + z &= 2 \end{aligned}$$

using elementary matrices and (a) Gaussian elimination. (b) Complete the operations necessary for Gauss–Jordan elimination. (c) Using the results of part (b) to invert the coefficient matrix for this system.

Solution: (a) The augmented matrix is

$$\mathbf{A}_0 = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 4 & -1 & 2 & 5 \\ 2 & -1 & 1 & 2 \end{bmatrix},$$

and all the operations are shown below:

$$\mathbf{A}_1 = \mathbf{R}_{21}\left(\frac{A_{21}}{A_{11}}\right)\mathbf{A}_0 = \mathbf{R}_{21}(-2)\mathbf{A}_0 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 4 & -1 & 2 & 5 \\ 2 & -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 2 & -1 & 1 & 2 \end{bmatrix},$$

$$\mathbf{A}_2 = \mathbf{R}_{31}\left(\frac{A_{31}}{A_{11}}\right)\mathbf{A}_1 = \mathbf{R}_{31}(-1)\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 2 & -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 0 & -2 & 3 & 1 \end{bmatrix},$$

$$\mathbf{A}_3 = \mathbf{R}_{32}\left(\frac{A_{32}}{A_{22}}\right)\mathbf{A}_2 = \mathbf{R}_{32}\left(\frac{2}{3}\right)\mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 0 & -2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 0 & 0 & -1 & -1 \end{bmatrix},$$

which completes the forward elimination. The back–substitution then gives the gauss elimination solution

$$\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

(b) If we want to perform Gauss–Jordan reduction, we continue with

$$\mathbf{A}_4 = \mathbf{R}_{23}\left(\frac{A_{23}}{A_{33}}\right)\mathbf{A}_3 = \mathbf{R}_{23}(6)\mathbf{A}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 0 & 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix},$$

$$\mathbf{A}_5 = \mathbf{R}_{13}\left(\frac{A_{13}}{A_{33}}\right)\mathbf{A}_4 = \mathbf{R}_{13}(-2)\mathbf{A}_4 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix},$$

$$\mathbf{A}_6 = \mathbf{R}_{12}\left(\frac{A_{12}}{A_{22}}\right)\mathbf{A}_5 = \mathbf{R}_{12}\left(\frac{1}{3}\right)\mathbf{A}_5 = \begin{bmatrix} 1 & \frac{1}{3} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 0 & 3 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix},$$

$$\mathbf{A}_7 = \mathbf{Q}_1\left(\frac{1}{2}\right)\mathbf{Q}_2\left(-\frac{1}{3}\right)\mathbf{Q}_3(-1)\mathbf{A}_6 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{x} \end{bmatrix},$$

which completes the solution. Notice that we can do more than one operation at a time as in

$$\mathbf{Q}_1\left(\frac{1}{2}\right)\mathbf{Q}_2\left(\frac{1}{3}\right)\mathbf{Q}_3(-1) = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

For Gaussian elimination, the entire sequence of operations for the forward elimination can be summarized as

$$\mathbf{A}_3 = \mathbf{R}_{32}\left(\frac{2}{3}\right)\mathbf{R}_{31}(-1)\mathbf{R}_{21}(-2)\mathbf{A}_0 = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & -3 & 6 & 3 \\ 0 & 0 & -1 & -1 \end{bmatrix},$$

and for Gauss–Jordan reduction, the sequence is

$$\mathbf{A}_7 = \mathbf{Q}_1\left(\frac{1}{2}\right)\mathbf{Q}_2\left(\frac{1}{3}\right)\mathbf{Q}_3(-1)\mathbf{R}_{12}\left(\frac{1}{3}\right)\mathbf{R}_{13}(-2)\mathbf{R}_{23}(6) \times$$

$$\mathbf{R}_{32}\left(\frac{2}{3}\right)\mathbf{R}_{31}(-1)\mathbf{R}_{21}(-2)\mathbf{A}_0 = \mathbf{P}\mathbf{A}_0 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

where

$$\mathbf{P} = \mathbf{Q}_1\left(\frac{1}{2}\right)\mathbf{Q}_2\left(\frac{1}{3}\right)\mathbf{Q}_3(-1)\mathbf{R}_{12}\left(\frac{1}{3}\right)\mathbf{R}_{13}(-2)\mathbf{R}_{23}(6)\mathbf{R}_{32}\left(\frac{2}{3}\right)\mathbf{R}_{31}(-1)\mathbf{R}_{21}(-2)$$

$$= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{3} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & -2 \\ -\frac{1}{3} & \frac{2}{3} & -1 \end{bmatrix}.$$

Notice that

$$\mathbf{P}\mathbf{A}_0 = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & -2 \\ -\frac{1}{3} & \frac{2}{3} & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 \\ 4 & -1 & 2 & 5 \\ 2 & -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \mathbf{A}_7 = \begin{bmatrix} \mathbf{I} & \mathbf{x} \end{bmatrix}.$$

(c) We have also performed all of the operations necessary to invert **A**. If we form the augmented matrix

$$\mathbf{B}_0 = \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 & 0 & 0 \\ 4 & -1 & 2 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{bmatrix},$$

then

$$\mathbf{P}\mathbf{B}_0 = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & -2 \\ -\frac{1}{3} & \frac{2}{3} & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 & 1 & 0 & 0 \\ 4 & -1 & 2 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & -1 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1} \end{bmatrix},$$

hence

$$\mathbf{A}^{-1} = \frac{1}{6}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 6 & -12 \\ -2 & 4 & -6 \end{bmatrix},$$

which we can check with

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \frac{1}{6}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 6 & -12 \\ -2 & 4 & -6 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 \\ 4 & -1 & 2 \\ 2 & -1 & 1 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} = \mathbf{I}.$$

1.3 An Algorithm for Naive Gaussian Elimination

If we let $\mathbf{A}\mathbf{x} = \mathbf{b}$ be an $N \times N$ system of equations where **A** and **b** are known. To solve this system perform the following three steps.

I. Form augmented matrix:

$$\mathbf{A}' = [\mathbf{A}\ \mathbf{b}],$$

and designate the rows of $\mathbf{A}'$ as $\mathbf{r}_1, \quad, \mathbf{r}_N$, each $1 \times N$, then

$$\mathbf{A}' = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \\ \mathbf{r}_N \end{bmatrix}.$$

II. Perform Forward Elimination – perform the following elementary row operations on each column of $\mathbf{A}'$ in the order indicated:

Step 1. Eliminate column 1 of $\mathbf{A}'$ by performing the $N - 1$ elementary row operations

$$\mathbf{r}_2 - \frac{A_{21}}{A_{11}}\mathbf{r}_1, \mathbf{r}_3 - \frac{A_{31}}{A_{11}}\mathbf{r}_1, \quad , \mathbf{r}_N - \frac{A_{N1}}{A_{11}}\mathbf{r}_1.$$

Step 2. Eliminate column 2 of $\mathbf{A}'$ by performing the $N - 2$ elementary row operations

$$\mathbf{r}_3 - \frac{A_{32}}{A_{22}}\mathbf{r}_2, \mathbf{r}_4 - \frac{A_{42}}{A_{22}}\mathbf{r}_2, \quad , \mathbf{r}_N - \frac{A_{N2}}{A_{22}}\mathbf{r}_2.$$

Step 3. Eliminate column 3 of $\mathbf{A}'$ by performing the $N - 3$ elementary row operations

$$\mathbf{r}_4 - \frac{A_{43}}{A_{33}}\mathbf{r}_3, \mathbf{r}_5 - \frac{A_{53}}{A_{33}}\mathbf{r}_3, \quad , \mathbf{r}_N - \frac{A_{N3}}{A_{33}}\mathbf{r}_3.$$

...

Step N - 1. Eliminate column $N - 1$ of $\mathbf{A}'$ by performing the following $N - (N - 1) = 1$ elementary row operation

$$\mathbf{r}_N - \frac{A_{N, N-1}}{A_{N-1, N-1}}\mathbf{r}_{N-1}.$$

III. Backward Elimination – upon completing **II**, the matrix $\mathbf{A}'$ will be *upper triangular* and starting with the last equation, we perform the following operations in the order indicated:

Step 1. $x_N = \dfrac{A_{N, N+1}}{A_{NN}}.$

Step 2. $x_{N-1} = \dfrac{A_{N-1, N+1} - A_{N-1, N}x_N}{A_{N-1, N-1}}.$

Step 3. $x_{N-2} = \dfrac{A_{N-2, N+1} - A_{N-2, N-1}x_{N-1} - A_{N-2, N}x_N}{A_{N-2, N-2}}.$

...

Step N - 1. $x_2 = \dfrac{A_{2, N+1} - A_{2, 3}x_3 - A_{2, 4}x_4 - \quad - A_{2, N-1}x_{N-1} - A_{2, N}x_N}{A_{2, 2}}.$

Step N. $x_1 = \dfrac{A_{1, N+1} - A_{1, 2}x_2 - A_{1, 3}x_3 - \quad - A_{1, N-1}x_{N-1} - A_{1, N}x_N}{A_{1, 1}}.$

Forward Elimination of **A**: The following is pseudo-code for forward elimination of **A** *without* using the augmented matrix.

```
k = 1 to n - 1
   i = k + 1 to n
      L_ik ← A_ik / A_kk
      j = k to n
         A_ij ← A_ij − L_ik A_kj
      next j
   next i
next k
```

Forward Elimination of **b**: Pseudo-code for the forward elimination of **b**.

```
j = 1 to n - 1
   i = j + 1 to n
      b_i ← b_i − L_ij b_j
   next i
next j
```

Backward Elimination: Pseudo-code for the backward elimination of **b** to get **x** (in this form of the code, **x** is stored in **b**).

```
j = n downto 1
   b_j ← b_j / A_jj
   i = 1 to j - 1
      b_i ← b_i − A_ij b_j
   next i
next j
```

1.4 Pivoting

The Gauss elimination process described above fails immediately if any main diagonal element is zero or becomes zero during the forward elimination steps. In some cases this may mean that no solution to the problem exists, in others, it is simply an artifact of the solution process which can be avoided by interchanging rows of the system.

Definition 1.9: Any element on the main diagonal of a matrix is called the *pivot element* and if the row (or column) containing the pivot is interchanged with any other row (or column) the process is called *pivoting* and pivoting can always be represented by the elementary matrix E_{ij}.

If both rows and columns are pivoting, the process is called *full pivoting* and if only rows or only columns are pivoted, the process is called *partial pivoting*.

Definition 1.10: *Partial row pivoting* is carried out by first identifying the element below the main diagonal in the pivot row which is largest in absolute value. The row that this element is in and the pivot row are then interchanged to accomplish the pivoting.

Example 1.5: Solve the following system using gauss elimination with partial row pivoting.

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

Solution: First form the augmented matrix

$$\mathbf{A}_0' = [\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} 0 & 2 & 1 & 5 \\ 4 & 1 & -1 & -3 \\ -2 & 3 & -3 & 5 \end{bmatrix}$$

Since $A_{11} = 0$, then we must pivot before carrying out the elimination of the first column. Since row 2 has the largest element in absolute value below A_{11}, then we interchange rows 1 and 2 to get

$$\mathbf{A}_1' = \mathbf{E}_{12}\mathbf{A}_0' = \begin{bmatrix} 4 & 1 & -1 & -3 \\ 0 & 2 & 1 & 5 \\ -2 & 3 & -3 & 5 \end{bmatrix},$$

and then eliminate the rest of column 1:

$$\mathbf{A}_2' = \mathbf{R}_{31}(1/2)\mathbf{A}_1' = \begin{bmatrix} 4 & 1 & -1 & -3 \\ 0 & 2 & 1 & 5 \\ 0 & \frac{7}{2} & -\frac{7}{2} & \frac{7}{2} \end{bmatrix}.$$

Although $A_{22} \neq 0$, if pivoting is being used, it is used every time we eliminate a new column. Thus, since 7/2 (in row 3) is larger than the 2 (in row 2), we must interchange rows 2 and 3 to get

$$\mathbf{A}_3' = \mathbf{E}_{23}\mathbf{A}_2' = \begin{bmatrix} 4 & 1 & -1 & -3 \\ 0 & \frac{7}{2} & -\frac{7}{2} & \frac{7}{2} \\ 0 & 2 & 1 & 5 \end{bmatrix}.$$

Elimination of column 2 then yields

$$\mathbf{A}_4{}' = \mathbf{R}_{32}(4/7)\mathbf{A}_3{}' = \begin{bmatrix} 4 & 1 & -1 & -3 \\ 0 & \frac{7}{2} & -\frac{7}{2} & \frac{7}{2} \\ 0 & 0 & 3 & 3 \end{bmatrix} = \begin{bmatrix} \mathbf{U} & \mathbf{b}_4 \end{bmatrix},$$

and the solution to $\mathbf{Ux} = \mathbf{b}_4$ is

$$\mathbf{x} = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix}^T.$$

1.5 Scaling a Matrix

If the ratio of the largest element in absolute value to the smallest element in absolute value is a large number (> 100), then significant round-off errors may occur. In such cases it may be beneficial to scale the matrix

---

Definition 1.11: *Scaling* is carried out before pivoting, by dividing the elements of each row of the augmented matrix by the largest element in that row (excluding b_i).

Example 1.6: Solve the system

$$\begin{bmatrix} 3 & 2 & 105 \\ 2 & -3 & 103 \\ 1 & 1 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 104 \\ 98 \\ 3 \end{bmatrix},$$

(a) exactly. Using 3 significant figures in all calculations, solve the system (b) without scaling and (c) with scaling and compare the results.

Solution: (a) the exact solution is $\mathbf{x} = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}^T$. For (b) we form the augmented matrix and notice that no pivoting is necessary for the first column

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 3 & 2 & 105 & 104 \\ 2 & -3 & 103 & 98 \\ 1 & 1 & 3 & 3 \end{bmatrix} \begin{matrix} \\ R_2 - 0.667R_1 \\ R_3 - 0.333R_1 \end{matrix} \quad \begin{bmatrix} 3 & 2 & 105 & 104 \\ 0 & -4.33 & 33.0 & 28.6 \\ 0 & 0.334 & -32.0 & -31.6 \end{bmatrix}.$$

Since pivoting is not required for the second column either, then

$$\begin{bmatrix} 3 & 2 & 105 & 104 \\ 0 & -4.33 & 33.0 & 28.6 \\ 0 & 0.334 & -32.0 & -31.6 \end{bmatrix} \begin{matrix} \\ \\ R_3 + 0.0771R_2 \end{matrix} \quad \begin{bmatrix} 3 & 2 & 105 & 104 \\ 0 & -4.33 & 33.0 & 28.6 \\ 0 & 0 & -29.5 & -29.4 \end{bmatrix}.$$

Backward elimination yields the result

$$\mathbf{x} = \begin{bmatrix} -0.844 \\ 0.924 \\ 0.997 \end{bmatrix} \text{ with errors } \varepsilon = \begin{bmatrix} 15.6 \\ 7.60 \\ 0.3 \end{bmatrix}\%.$$

For (c), we first scale the matrix to get

$$\mathbf{A}' = \left[\mathbf{A}\ \mathbf{b}\right] = \begin{bmatrix} 3 & 2 & 105 & 104 \\ 2 & -3 & 103 & 98 \\ 1 & 1 & 3 & 3 \end{bmatrix} \begin{matrix} R_1/105 \\ R_2/103 \\ R_3/3 \end{matrix} \quad \begin{bmatrix} 0.0286 & 0.0190 & 1.00 & 0.991 \\ 0.0194 & -0.0291 & 1.00 & 0.952 \\ 0.333 & 0.333 & 1.00 & 1.00 \end{bmatrix},$$

then elimination of columns 1 and 2 gives

$$\begin{bmatrix} 0.333 & 0.333 & 1.00 & 1.00 \\ 0 & -0.0485 & 0.942 & 0.894 \\ 0 & 0 & 0.727 & 0.728 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} -0.990 \\ 0.990 \\ 1.00 \end{bmatrix} \quad \varepsilon = \begin{bmatrix} 1.00 \\ 1.00 \\ 0 \end{bmatrix} \%.$$

The scaling process significantly improves the results.