2 Mathematical preliminaries

2.1 Matrices

A matrix is a rectangular array of numbers arranged in rows and columns. The numbers in a matrix are called entries or elements and we can refer to the entries with a pair of indices, the row index and and column index. The row are indexed by positive integers from top to the bottom starting with 1. Columns are indexed by positive integers from left to right starting with 1. Matrices are often denoted by capital latin letters, while their entries are denoted by the corresponding lower case letter with a pair of subscript indices. First we write the row index and then the column index separated by a comma. An example of a matrix is the following:

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

The matrix A above has 3 rows and 4 columns. We can refer to the entry in the 2nd row and 3rd column as $a_{2,3}$ and it equals to 0, while the entry in the 3rd row and 2nd column is referred as $a_{3,2}$ and it equals to 1. We note here that if the matrix is not too large and it makes no confusion we can omit the comma between the row and column indices, thus we can refer to the above entries as a_{23} and a_{32} too. In a larger matrix we refer to the the entry in the i-th row and j-th column as $a_{i,j}$ or a_{ij}. The notation

$$A = (a_{ij})$$

is used when we would like to refer to or denote the element of the matrix A in the i-th row and j-th column as a_{ij}. This notation is used only when it's clear what are the ranges of the indices i and j (i.e. what is the number of rows and the number of columns). Otherwise we can specify a matrix with

m rows and n columns as

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

If a matrix has m rows and n columns then we say the size (or dimension) of the matrix is $m \times n$. The sample matrix A above has size 3×4. The set of all real matrices (i.e. matrices whose entries are real numbers) with m rows and n columns is denoted by $\mathcal{M}_{m\times n}(\mathbb{R})$. A matrix with the same number of rows and columns is called a square matrix. The set of all real square matrices with n rows and n columns is denoted by $\mathcal{M}_n(\mathbb{R})$.

2.1.1 Matrix operations

Now we define the basic matrix operations such as addition, scalar multiplication and transposition.

Definition 1 *Let a $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ be two matrices of size $m \times n$. The **sum of** A **and** B is the matrix $C \in \mathcal{M}_{m\times n}(\mathbb{R})$ of size $m \times n$ which satisfies*

$$c_{ij} = a_{ij} + b_{ij}$$

if $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$. The sum of the matrices A and B is denoted by $A + B$.

Note that the sum of two matrices can be defined only if they have the same size.

Definition 2 *Let a $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ be a matrix of size $m \times n$ and let $\lambda \in \mathbb{R}$ (lambda) be a real number. The **product of the matrix** A **with the number** λ is the matrix $C \in \mathcal{M}_{m\times n}(\mathbb{R})$ of size $m \times n$ which satisfies*

$$c_{ij} = \lambda \cdot a_{ij}$$

if $A = (a_{ij})$ and $C = (c_{ij})$. The product of the matrix A with the number λ is denoted by $\lambda \cdot A$ or shortly λA if it makes no confusion.

The above operation is called **scalar multiplication**, but make sure to not confuse it with scalar product which is a different operation defined for vectors. The real number λ in the scalar multiplication is often called scalar (thus the name).

Theorem 1 *The addition and scalar multiplication of matrices have the following properties.*

1. $(A+B)+C = A+(B+C)$ *for any matrices* $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{R})$.

2. *There exists a* ***zero matrix*** *of size* $m \times n$ *denoted by* $0_{m\times n}$ *which satisfies* $A + 0_{m\times n} = A$ *for any matrix* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$.

3. *For any matrix* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ *there exists an opposite matrix of size* $m \times n$ *denoted by* $-A$ *which satisfies* $A + (-A) = 0_{m\times n}$.

4. $A + B = B + A$ *for any matrices* $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$.

5. $\lambda(A+B) = \lambda A + \lambda B$ *for any matrices* $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ *and any scalar* $\lambda \in \mathbb{R}$.

6. $(\lambda + \mu)A = \lambda A + \mu B$ *for any matrix* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ *and any scalars* $\lambda, \mu \in \mathbb{R}$ *(lambda and mu).*

7. $(\lambda \cdot \mu)A = \lambda(\mu A) = \mu(\lambda A)$.

Examples.

The zero matrix of size 2×3 is

$$0_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Furthermore let

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

Then

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

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

$$3 \cdot A = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot (-1) \\ 3 \cdot 3 & 3 \cdot 4 & 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3 & 6 & -3 \\ 9 & 12 & 0 \end{pmatrix}$$

$$(-2) \cdot B = \begin{pmatrix} (-2) \cdot 2 & (-2) \cdot (-4) & (-2) \cdot 0 \\ (-2) \cdot (-1) & (-2) \cdot 1 & (-2) \cdot 3 \end{pmatrix} = \begin{pmatrix} -4 & 8 & 0 \\ 2 & -2 & -6 \end{pmatrix}$$

Definition 3 *Let a $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ be a matrix of size $m \times n$ and let $\lambda \in \mathbb{R}$ (lambda) be a real number. The* ***transpose of the matrix*** *A the matrix $B \in \mathcal{M}_{n \times m}(\mathbb{R})$ of size $n \times m$ which satisfies*

$$b_{ij} = \lambda \cdot a_{ji}$$

if $A = (a_{ij})$ and $B = (c_{ij})$. The transpose of the matrix A is denoted by $A^\top$.

As an example let

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

Then

$$A^\top = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ -1 & 0 \end{pmatrix}$$

A square matrix $A \in \mathcal{M}_n(\mathbb{R})$ is called symmetrical if $A = A^\top$. For example the matrix

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

is a symmetrical matrix.

Theorem 2 *Let $A, B \in \mathcal{M}_{m \times n}(\mathbb{R})$ be arbitrary matrices of size $m \times n$ and let $\lambda \in \mathbb{R}$ be an arbitrary scalar. Then*

1. $(A + B)^\top = A^\top + B^\top$,
2. $(\lambda A)^\top = \lambda A^\top$,
3. $\left(A^\top\right)^\top = A$.

2.1.2 Matrix multiplication

The addition and scalar multiplication of matrices defined in the previous section are calculated elementwise. This means that if we want to add the matrices A and B, then we only need to add the corresponding elements. Similarly to multiply a matrix by a scalar we need to multiply each element of the matrix with that scalar. Now we introduce the product of two matrices, which is defined in a different manner (not elementwise).

Definition 4 *Let a $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ be a matrix of size $m \times n$ and $B \in \mathcal{M}_{n\times l}(\mathbb{R})$ be a matrix of size $n \times l$. The **product of** A **and** B is the matrix $C \in \mathcal{M}_{m\times l}(\mathbb{R})$ of size $m \times l$ which satisfies*

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

if $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$. The product of the matrices A and B is denoted by $A \cdot B$ or shortly AB.

Note that the product of A and B is defined only if A has the same number of columns as the the number of rows of B. Note also that the element of the product matrix in the i-th row and j-th column is computed as the product of the i-th row of A and the j-th column of B in the same manner as the dot product of two n-dimensional vectors are defined. To see an example let

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

Then the product matrix is $A \cdot B = C = (c_{ij})$, where

$$\begin{aligned} c_{11} &= \sum_{k=1}^{4} a_{1k} \cdot b_{k1} = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} + a_{14} \cdot b_{41} \\ &= 1 \cdot 2 + 2 \cdot (-1) + (-1) \cdot 0 + 0 \cdot 1 = 0 \\ c_{12} &= \sum_{k=1}^{4} a_{1k} \cdot b_{k2} = a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} + a_{14} \cdot b_{42} \\ &= 1 \cdot (-1) + 2 \cdot 1 + (-1) \cdot 3 + 0 \cdot 2 = -2 \end{aligned}$$

$$
\begin{aligned}
c_{21} &= \sum_{k=1}^{4} a_{2k} \cdot b_{k1} = a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} + a_{24} \cdot b_{41} \\
&= 3 \cdot 2 + 4 \cdot (-1) + 0 \cdot 0 + (-2) \cdot 1 = 0 \\
c_{22} &= \sum_{k=1}^{4} a_{2k} \cdot b_{k2} = a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} + a_{24} \cdot b_{42} \\
&= 3 \cdot (-1) + 4 \cdot 1 + 0 \cdot 3 + (-2) \cdot 2 = -3 \\
c_{31} &= \sum_{k=1}^{4} a_{3k} \cdot b_{k1} = a_{31} \cdot b_{11} + a_{32} \cdot b_{21} + a_{33} \cdot b_{31} + a_{34} \cdot b_{41} \\
&= 0 \cdot 2 + 1 \cdot (-1) + 1 \cdot 0 + 5 \cdot 1 = 4 \\
c_{32} &= \sum_{k=1}^{4} a_{3k} \cdot b_{k2} = a_{31} \cdot b_{12} + a_{32} \cdot b_{22} + a_{33} \cdot b_{32} + a_{44} \cdot b_{42} \\
&= 0 \cdot (-1) + 1 \cdot 1 + 1 \cdot 3 + 5 \cdot 2 = 14
\end{aligned}
$$

Thus

$$
A \cdot B = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 0 & -3 \\ 4 & 14 \end{pmatrix}
$$

Until you are not experienced in the multiplication of matrices it makes easier to calculate the product if you arrange the factors A and B not next to each other, but A to the bottom left and B to the top right position of a 2-by-2 arrangement. Then the product matrix is written in the bottom right position as the following formula shows.

$$
\begin{array}{cc}
 & \begin{pmatrix} 2 & -1 \\ -1 & 1 \\ 0 & 3 \\ 1 & 2 \end{pmatrix} \\
\begin{pmatrix} 1 & 2 & -1 & 0 \\ 3 & 4 & 0 & -2 \\ 0 & 1 & 1 & 5 \end{pmatrix} & \begin{pmatrix} 0 & -2 \\ 0 & -3 \\ 4 & 14 \end{pmatrix}
\end{array}
$$

Theorem 3 *The multiplication of matrices has the following properties.*

1. *$A \cdot (B \cdot C) = (A \cdot B) \cdot C$ for any matrices A, B, C with appropriate sizes.*

2. *There exists an identity matrix of size $n \times n$ denoted by I_n, which satisfies $I_n \cdot A = A \cdot I_n = A$ for any square matrix A of size $n \times n$.*

3. *$A \cdot (B + C) = A \cdot B + A \cdot C$ for any matrices A, B, C with appropriate sizes.*

4. *$(A + B) \cdot C = A \cdot C + B \cdot C$ for any matrices A, B, C with appropriate sizes.*

5. *$\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$ for any scalar $\lambda \in \mathbb{R}$ and any matrices A, B with appropriate sizes.*

Examples:

The identity matrices of size 2×2 and 3×3 are

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The identity matrix of size $n \times n$ is

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

We can also define the identity matrix with the help of the Kronecker-delta, which is

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

Then the identity matrix of size $n \times n$ is $I_n = (\delta_{ij})$.

Consider now the following matrices.

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

$$C = \begin{pmatrix} 2 & 0 \\ 1 & 3 \\ 0 & -1 \end{pmatrix} \quad D = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

Then

$$A \cdot B = \begin{pmatrix} -2 & -2 & -1 \\ 4 & 4 & 2 \end{pmatrix}$$

but the product $B \cdot A$ doesn't exist as B has 3 columns and A has 2 rows. Similarly the product $A \cdot C$ doesn't exist as A has 2 columns and C has 3 rows, but the product $C \cdot A$ exists and

$$C \cdot A = \begin{pmatrix} 2 & -2 \\ -5 & 5 \\ 2 & -2 \end{pmatrix}$$

Both of the products $B \cdot C$ and $C \cdot B$ exist

$$B \cdot C = \begin{pmatrix} 4 & 7 \\ 10 & 12 \end{pmatrix} \quad \text{and} \quad C \cdot B = \begin{pmatrix} 2 & 4 & -2 \\ 10 & 14 & -1 \\ -3 & -4 & 0 \end{pmatrix}$$

but $B \cdot C \neq C \cdot B$ as these product don't even have the same size. Similarly both of the products $A \cdot D$ and $D \cdot A$ exist

$$A \cdot D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad D \cdot A = \begin{pmatrix} 3 & -3 \\ 3 & -3 \end{pmatrix}$$

but $A \cdot D \neq D \cdot A$ even though these products have the same size. Finally we see that

$$A^2 = A \cdot A = \begin{pmatrix} 3 & -3 \\ -6 & 6 \end{pmatrix} \quad \text{and} \quad D^2 = D \cdot D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

An important note is that if we interchange the factors in the multiplication of matrices, then the product may be undefined (see $A \cdot B$ and $B \cdot A$, or $A \cdot C$ and $C \cdot A$ above), but even if both products are defined it may happen that they don't equal (see $B \cdot C$ and $C \cdot B$, or $A \cdot D$ and $D \cdot A$). In fact if we randomly take two square matrices of size $n \times n$, A and B (this ensures, that both $A \cdot B$ and $B \cdot A$ are defined and have the same size), then it's very likely that $A \cdot B \neq B \cdot A$.

Another interesting property of the multiplication of matrices is that it may happen that none of the matrices A and B is the zero matrix, but their product is the zero matrix (see $A \cdot D$ above). Note that this is a property which is not valid for the multiplication of real numbers, since if $a, b \in \mathbb{R}$ and $a \cdot b = 0$, then $a = 0$ or $b = 0$ (this is called the zero-product property). We can go even further as there exist a nonzero matrix whose square (i.e. the product with itself) is the zero matrix (see D^2 above).

2.1.3 Matrices of special shape

We already introduced square matrices (i.e. matrices with the same number of rows and columns) in the above sections. Now we define further matrices of special shape. First of all lets mention that a single column matrix (which has only one column) with n rows, or a single row matrix (which has only one row) with n columns can be identified as an element of the n-dimensional coordinate space $\mathbb{R}^n$, whose elements are often called vectors. Thus a single column matrix is called a **column vector**, while a single row matrix is called a **row vector**. Note that the transpose of a column vector is a row vector and the transpose of a row vector is a column vector. In the rest of the text we apply the convention that if nothing else said then by a vector we always mean a column vector. Now let's see further matrices of special shape.

Definition 5 *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, $A = (a_{ij})$ be a matrix of size n. The* ***diagonal of*** *A is the sequence of its elements with equal row and column indices, i.e. $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$. The matrix A is called* ***diagonal matrix*** *if all elements of A not contained in the diagonal equal to zero, i.e. $a_{ij} = 0$ if $i \neq j$. The matrix A is called* ***upper triangular matrix*** *if all elements of A below the diagonal equal to zero, i.e. $a_{ij} = 0$ if $i > j$. The matrix A is called* ***lower triangular matrix*** *if all elements of A above the diagonal equal to zero, i.e. $a_{ij} = 0$ if $i < j$.*

We note that a diagonal matrix may contain zeros in the diagonal too. Similarly an upper triangular matrix may have zeros in and above the diagonal, just as a lower triangular matrix may have zeros in and below the diagonal. As examples the zero matrix and the identity matrix can be considered as diagonal, upper triangular or lower triangular matrix. Further instances are the following.

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

Above the matrix A is diagonal (and also upper and lower triangular), B is upper triangular, C is lower triangular.

Definition 6 *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $A = (a_{ij})$ be a matrix. If the i-th row of A contains at least one nonzero element, then the leftmost nonzero element is called the* ***pivot element*** *of the i-th row. We say that the matrix is A in a* ***row echelon form*** *if all rows consisting of only zeros are at the bottom and the pivot element of any row (except the first) is strictly to the right from the pivot element of the previous row. In other words A in a row echelon form if considering the pivot elements from the top to the bottom, the column indices of the pivot elements form a strictly increasing sequence. Furthermore we say A is in* ***reduced row echelon form*** *if it's in row echelon form, all pivot elements equal to 1, and each column containing a pivot element has zeros in all other entries.*

Consider the following matrices.

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

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

Here A in not in row echelon form since the pivot element of the second row has column index 3, but the pivot element of the third row has column index 2. B in not in row echelon form also, since the column indices of the pivot elements in the second and third rows equal. C is in row echelon form, but not reduced echelon form, since none of the pivot elements equal to 1, and furthermore the second column and fourth column contain nonzero entries besides the pivot element. The matrix D is in reduced row echelon form.

2.1.4 Elementary row operations

We will see in the later chapters that matrices of reduced row echelon form have an important role in the solution of systems of linear equations. Thus now we introduce elementary row operations that can be applied to transform a matrix into reduced row echelon form.

Definition 7 *The elementary row operations are the following.*

1. *Interchange two rows of a matrix.*
2. *Multiply each element of a row in a matrix with a nonzero scalar $\lambda \in \mathbb{R}$.*
3. *Add each element of a row of a matrix to the corresponding elements of another row.*
4. *As a combination of the above two we can add each element of a row multiplied by $\lambda \in \mathbb{R}$ to the corresponding elements of another row.*

Note that during the last operation if we add λ-times the i-th row the j-th row, then only the entries of the j-th row are modified, while the i-t row remain unchanged. Note also that if we multiply a row with the multiplicative inverse nonzero scalar $\lambda \in \mathbb{R}$, then it has the same result as if we divide with $\lambda \in \mathbb{R}$. Thus the second elementary row operation means that we can also divide each element of a row in a matrix with a nonzero scalar $\lambda \in \mathbb{R}$.

Theorem 4 *We can transform any matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ into a matrix of reduced row echelon form with the help of finitely many elementary row operations.*

Now we give a sketch of the proof of the above theorem by telling what steps are required to achieve the reduced row echelon form. The procedure, that transforms an arbitrary matrix to a matrix in reduced row echelon from with the help of elementary row operations, is called **Gaussian elimination**.

First let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ be an arbitrary matrix. Let k be the counter of how many pivot elements we found. Set the initial value of k to zero. Let B_k denote the part of A below the k-th row. At the beginning, when $k = 0$ let $B_0 = A$. Now follow the steps below to transform A into a matrix of reduced row echelon form.

1. If B_k is the zero matrix, then A is in reduced row echelon form and the procedure terminates.

2. Otherwise let j be the column index of the leftmost nonzero column of B_k.

3. If the first element of the j-th column of B_k is zero, then look for a row in B_k which contains a nonzero element in the j-th column and interchange it with the first row of B_k (i.e. with the $k+1$-th row of A).

4. Divide each element of the $k+1$-th row of A with $a_{k+1,j}$.

5. Add $-a_{ij}$-times the $k+1$-th row to the i-th row of A for all $i \in \{1, 2, \ldots, m\}$, $i \neq k+1$.

6. Increase the value of the counter k by 1.

7. Repeat steps (1)-(6) until the procedure terminates in step (1) or k is increased to n.

The above procedure transforms any matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ into a matrix of reduced row echelon form. Let's illustrate the procedure on the following matrix.

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

First $k = 0$ and $B_0 = A$ is not the zero matrix. The leftmost nonzero column is the second column. The first element of the second column is zero, thus we need to choose a row whose entry in the second column is nonzero. Let's choose the second row. Now we interchange the first and second row of A. This step is denoted as follows.

$$A = \begin{pmatrix} 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & 3 & 2 & 3 & 0 & 4 \\ 0 & -2 & 1 & 5 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix} \xrightarrow{R1 \leftrightarrow R2} \begin{pmatrix} 0 & 3 & 2 & 3 & 0 & 4 \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & -2 & 1 & 5 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix}$$

Note that we don't write equality between the above matrices, because they are not the same, but they are similar in some sense. In the above procedure we refer to this new matrix as A, since we assume that the new matrix overwrites A. Now we apply the same notation, i.e. when we refer to an element of A, then we always mean the latest version of the matrix. The next step is to divide each element of the first row of A with $a_{12} = 3$. This step is denoted as follows.

$$\begin{pmatrix} 0 & 3 & 2 & 3 & 0 & 4 \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & -2 & 1 & 5 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix} \xrightarrow{\frac{1}{3}\cdot R1} \begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & -2 & 1 & 5 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix}$$

Now we add $-a_{22} = 0$-times the first row to the second row, $-a_{32} = 2$-times the first row to the third row, $-a_{42} = -1$-times the first row to the fourth row, and $-a_{52} = 0$-times the first row to the fifth row. However note that if we add 0-times a row to another row, then nothing changes, thus it's enough to work with those rows, which contain a nonzero element in the second column. Considering this we only add 2-times the first row to the third row, and (-1)-times the first row to the fourth row. This step is denoted as follows.

$$\begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & -2 & 1 & 5 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix} \xrightarrow[R4-1\cdot R1]{R3+2\cdot R1} \begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & 0 & \frac{7}{3} & 7 & 1 & -\frac{1}{3} \\ 0 & 0 & -\frac{2}{3} & -2 & 0 & \frac{2}{3} \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix}$$

Then we set $k = 1$. Here B_1 is not the zero matrix and its leftmost nonzero column is the third column. The first element of the third column in $B1$ is nonzero, thus we divide the second row by $a_{23} = 3$.

$$\begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 3 & 9 & 1 & -1 \\ 0 & 0 & \frac{7}{3} & 7 & 1 & -\frac{1}{3} \\ 0 & 0 & -\frac{2}{3} & -2 & 0 & \frac{2}{3} \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix} \xrightarrow{\frac{1}{3}\cdot R2} \begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & \frac{-1}{3} \\ 0 & 0 & \frac{7}{3} & 7 & 1 & -\frac{1}{3} \\ 0 & 0 & -\frac{2}{3} & -2 & 0 & \frac{2}{3} \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix}$$

Then we add $-a_{13} = -\frac{2}{3}$-times the second row to the first row, $-a_{33} = -\frac{7}{3}$-times the second row to the third row, $-a_{43} = \frac{2}{3}$-times the second row to the fourth row, and $-a_{53} = 1$-times the second row to the fifth row (the later means only that we add the second row to the fifth row).

$$\begin{pmatrix} 0 & 1 & \frac{2}{3} & 1 & 0 & \frac{4}{3} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & \frac{-1}{3} \\ 0 & 0 & \frac{7}{3} & 7 & 1 & -\frac{1}{3} \\ 0 & 0 & -\frac{2}{3} & -2 & 0 & \frac{2}{3} \\ 0 & 0 & -1 & -3 & 1 & 3 \end{pmatrix} \xrightarrow[\substack{R4+\frac{2}{3}\cdot R2 \\ R5+R2}]{\substack{R1-\frac{2}{3}\cdot R2 \\ R3-\frac{7}{3}\cdot R2}} \begin{pmatrix} 0 & 1 & 0 & -1 & -\frac{2}{9} & \frac{14}{9} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & \frac{-1}{3} \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{4}{3} & \frac{8}{3} \end{pmatrix}$$

Now set $k = 2$. Here B_2 is not the zero matrix and its leftmost nonzero column is the fifth column. The first element of the fifth column in B_2 is nonzero, thus we divide the third row by $a_{35} = \frac{2}{9}$.

$$\begin{pmatrix} 0 & 1 & 0 & -1 & -\frac{2}{9} & \frac{14}{9} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & \frac{-1}{3} \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{4}{3} & \frac{8}{3} \end{pmatrix} \xrightarrow{\frac{9}{2}\cdot R3} \begin{pmatrix} 0 & 1 & 0 & -1 & -\frac{2}{9} & \frac{14}{9} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & \frac{-1}{3} \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{4}{3} & \frac{8}{3} \end{pmatrix}$$

Then we add $-a_{15} = \frac{2}{9}$-times the third row to the first row, $-a_{25} = -\frac{1}{3}$-times the third row to the second row, $-a_{45} = -\frac{2}{9}$-times the third row to the fourth row, and $-a_{55} = -\frac{4}{3}$-times the third row to the fifth row.

$$\begin{pmatrix} 0 & 1 & 0 & -1 & -\frac{2}{9} & \frac{14}{9} \\ 0 & 0 & 1 & 3 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & \frac{2}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 0 & \frac{4}{3} & \frac{8}{3} \end{pmatrix} \xrightarrow[\substack{R4-\frac{2}{9}\cdot R3 \\ R5-\frac{4}{3}\cdot R3}]{\substack{R1+\frac{2}{9}\cdot R3 \\ R2-\frac{1}{3}\cdot R3}} \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Finally set $k = 3$. Now B_3 is the zero matrix, thus the procedure terminates and we claim that the resulting matrix is in reduced row echelon form, and it clearly is.

2.2 Systems of linear equations

A linear equation for the unknowns (or variables) $x_1, x_2, x_3, \dots x_n$ is an equation of the form where a linear combination of $x_1, x_2, \dots x_n$ equals to a constant. If the coefficients in the linear combination are $a_1, a_2, \dots, a_n$, and the constant is b, then the linear equation is

$$a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_nx_n = b$$

Sometimes the same unknowns must satisfy not just one, but several linear equations. Then we talk about a system of linear equations. As the combining coefficients in different equations vary, it's better to use double indexing form these coefficients. For example the coefficient of the unknown x_j is the i-th equation can be denoted by a_{ij}. The constants may be different in different equations, hence these should be also indexed. Let b_i denote the constant on the right in the i-th equation. Then a system of m equations for the unknowns $x_1, x_2, x_3, \dots x_n$ is of the following form.

$$\left.\begin{array}{r}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\vdots \qquad\qquad\qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\end{array}\right\}$$

The coefficients in the above system naturally define a matrix of size $m \times n$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

which is called the **coefficient matrix** of the system. The coefficient matrix together with the multiplication of matrices gives a good opportunity to make a short notation of systems of linear equalities. If $\underline{x}$ denotes the column vector (i.e. single column matrix) containing the unknowns and $\underline{b}$ denotes the column vector (i.e. single column matrix) containing the the constants on the right, then the system can be written as

$$A \cdot \underline{x} = \underline{b}$$

which is called the matrix form of the system. Clearly the product on the left has m rows and only one column (as $\underline{x}$ has only one column). The column vector $\underline{b}$ has the same size as $A \cdot \underline{x}$, and the two column vectors equal to each other if all corresponding components equal to each other, which means exactly that the unknowns must satisfy the system of linear equalities above. We note here that the elements of the n-dimensional coordinate space $\mathbb{R}^n$ as vectors are usually denoted by underlined lowercase letters in order to emphasize that they are vectors and not to confuse them with real numbers, which are that scalars. Thus the notation $\underline{x}$ and $\underline{b}$ of column vectors.

Definition 8 *If $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $\underline{b}$ is column vector of m elements, then the **extended coefficient matrix** of the system of linear equalities $A \cdot \underline{x} = \underline{b}$ is the matrix of size $m \times (n+1)$ whose first n columns are exactly the same as the columns of A and the last column equals to the column vector $\underline{b}$. The extended coefficient matrix is denoted by $(A|\underline{b})$*

The last column of the extended coefficient matrix is often separated by a vertical line because of it's special role. Now let's see the following example of a system of linear equations.

$$\left.\begin{aligned} 3x_1 - 2x_2 + 4x_3 - x_4 &= 8 \\ x_1 + x_2 - 5x_4 &= -1 \\ 2x_1 + 3x_3 + x_4 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 &= 1 \\ x_2 - x_3 - 2x_4 &= 6 \end{aligned}\right\}$$

The coefficient matrix and the column vector of right-hand-side constants are

$$A = \begin{pmatrix} 3 & -2 & 4 & -1 \\ 1 & 1 & 0 & -5 \\ 2 & 0 & 3 & 1 \\ -1 & -1 & 2 & -3 \\ 0 & 1 & -1 & -2 \end{pmatrix} \qquad \underline{b} = \begin{pmatrix} 8 \\ -1 \\ 0 \\ 1 \\ 6 \end{pmatrix}$$

and the extended coefficient matrix is

$$\left(\begin{array}{cccc|c} 3 & -2 & 4 & -1 & 8 \\ 1 & 1 & 0 & -5 & -1 \\ 2 & 0 & 3 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & -2 & 6 \end{array}\right)$$

Note that whenever the coefficient of an unknown seems to be missing in the system (see for example x_1 and x_2 in the second equation) then it only means that the coefficient is 1. It's also possible that some of the unknowns are missing is some equations (see for example x_3 in the second equation or x_2 is the third equation). Then the corresponding coefficient is zero.

Definition 9 *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and let $\underline{b}$ be a column vector of m elements. The system of linear equations $A \cdot \underline{x} = \underline{b}$ is called **underdetermined** if the number of equations is less than the number of unknowns (i.e. $m < n$), and called **overdetermined** if the number of equations is larger than the number of unknowns (i.e. $n < m$). We say that the system is **solvable** if it has at least one solution, and otherwise we say it's **unsolvable**. Furthermore a solvable system $A \cdot \underline{x} = \underline{b}$ is called **determined** if it has exactly one solution, and called **undetermined** if it has more than one solution.*

Please note the difference between the properties underdetermined and undetermined. An underdeterminded system can be unsolvable, while a system is undetermined if it's solvable and has several solutions. However a solvable underdetermined system is undetermined. An overdetermined system can be solvable or unsolvable, and can be determined or undetermined when solvable.

An interesting property of systems of linear equations is that if a system has at least two solutions, then it has infinitely many solutions. To see this let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ be an aritrary matrix and let $\underline{b}$ be an arbitrary column vector of m elements. Assume that two different column vectors $\underline{u}$ and $\underline{v}$ of m elements are solutions of the system $A \cdot \underline{x} = \underline{b}$. This means that $A \cdot \underline{u} = \underline{b}$ and $A \cdot \underline{v} = \underline{b}$. Now choose and arbitrary scalar $t \in \mathbb{R}$ and construct the column vector $(1-t)\underline{u} + t\underline{v}$. Then

$$A \cdot \big((1-t)\underline{u} + t\underline{v}\big) = (1-t)\, A \cdot \underline{u} + t\, A \cdot \underline{v} =$$

$$(1-t)\,\underline{b} + t\,\underline{b} = (1-t+t)\,\underline{b} = 1 \cdot \underline{b} = \underline{b}$$

This means $A \cdot \big((1-t)\underline{u} + t\underline{v}\big) = \underline{b}$ and it implies that $(1-t)\underline{u} + t\underline{v}$ is also a solution of the system $A \cdot \underline{x} = \underline{b}$. This show the system has infinitely many solutions because $\underline{u}$ and $\underline{v}$ are different and we can choose infinitely many different scalars $t \in \mathbb{R}$.

Theorem 5 *Let $A \cdot \underline{x} = \underline{b}$ be a system of linear equations. If the we apply any of the elementary row operations to the extended coefficient matrix $(A|\underline{b})$, then the set of solutions of system defined by the new extended coefficient matrix is exactly the same as the set of solutions of $A \cdot \underline{x} = \underline{b}$.*

The consequence of the above theorem is that if the extended coefficient matrix of the system $A \cdot \underline{x} = \underline{b}$ is transformed into reduced row echelon form with the help of elementary row operations, then the set of solutions remains the same as for the original system. Since any matrix can be transformed into reduced row echelon form with the help of elementary row operations, it's enough to investigate the set of solutions of systems with extended coefficient matrix in reduced row echelon form.

Theorem 6 *Let $A \cdot \underline{x} = \underline{b}$ be a system of linear equations and assume that the extended coefficient matrix $(A|\underline{b})$ is in reduced row echelon form.*

1. *If $(A|\underline{b})$ has a row which contains at least one nonzero element, and the pivot element in that row is in the last column, then the system is unsolvable. Otherwise the system is solvable.*

2. *If $A \cdot \underline{x} = \underline{b}$ is solvable and the number of nonzero rows in $(A|\underline{b})$ equals to the number of unknowns, then the system is determined. The only solution is $\widehat{\underline{b}}$, which consists of those elements of $\underline{b}$, which are not included in any zero row of $(A|\underline{b})$.*

3. *If $A \cdot \underline{x} = \underline{b}$ is solvable and the number of nonzero rows in $(A|\underline{b})$ is less than the number of unknowns, then the system is undetermined.*

Note that if a matrix is in reduced row echelon form, then the number of nonzero rows can't be larger than the number of columns, since each nonzero row contains a pivot element, but in reduced row echelon form each column may contain at most one pivot element. By the above theorem we know what is the set of solutions if the system is unsolvable, or solvable and the number of nonzero rows in $(A|\underline{b})$ equals to the number of unknowns (in the former case it's the empty set). We also know that if a system is undetermined, then it has infinitely many solutions. Yet, we would like to give a method to characterize all the solutions in such a situation.

Let $A \cdot \underline{x} = \underline{b}$ be a system of linear equations whose extended coefficient matrix $(A|\underline{b})$ is in reduced row echelon form. Assume that the system is solvable and the number of nonzero rows in $(A|\underline{b})$ is less than the number of unknowns. Then at least one column of the coefficient matrix A contains no pivot element. Let $\widehat{\underline{x}}$ denote the column vector of those unknowns, whose corresponding columns don't contain a pivot element. The ordering of the elements in $\widehat{\underline{x}}$ is the same as in $\underline{x}$. The unknowns in the column vector $\widehat{\underline{x}}$ are called the **free variables** (since soon we will see that their values can be chosen freely). Let B be a matrix which consists of the nonzero rows of the coefficient matrix A and the opposites of those columns of A which contain no pivot element. The opposite of a column means that each element of that column is multiplied by -1. The ordering of the columns in B is the same as in A. The matrix B is called the **coefficient matrix of the free variables**.

Theorem 7 *Let $A \cdot \underline{x} = \underline{b}$ be a system of linear equations whose extended coefficient matrix $(A|\underline{b})$ is in reduced row echelon form. Assume that the system is solvable and the number of nonzero rows in $(A|\underline{b})$ is less than the number of unknowns. Then the system is undetermined and*

$$\underline{v} = B \cdot \underline{u} + \widehat{\underline{b}}$$

is a solution of the system, where B is the coefficient matrix of the free variables, $\widehat{\underline{b}}$ is the same as in the previous theorem, and $\underline{u}$ is a column vector of arbitrary real numbers with appropriate size. Moreover for all solution $\underline{v}$ of the system $A \cdot \underline{x} = \underline{b}$ there exists a column vector $\underline{u}$ such that $\underline{v} = B \cdot \underline{u} + \widehat{\underline{b}}$.

In the above theorem the column vector $\underline{u}$ determines the values of the free variables. The appropriate size means that the number of elements in $\underline{u}$ is the same as the number of columns in the matrix B. Then the product $B \cdot \underline{u}$ is well defined and its dimension is the same as the dimension of $\widehat{\underline{b}}$, thus the sum $B \cdot \underline{u} + \widehat{\underline{b}}$ is also well defined.

Now we are able to determine the set of solutions of any system of linear equations by transforming its extended coefficient matrix into reduced row echelon form with the help of elementary row operations, and then applying one of the above theorems.

Example 1.

$$\left.\begin{aligned} x_1 - 3x_2 + 2x_3 + x_4 &= 4 \\ 5x_3 - 2x_4 &= 12 \\ -x_1 + 2x_2 + x_3 - 4x_4 &= 5 \\ -2x_2 + 4x_3 + 3x_4 &= 5 \\ x_1 - x_2 - 11x_4 &= 12 \end{aligned}\right\}$$

This is an overdetermined system. The extended coefficient matrix of the above system is

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & -3 & 2 & 1 & 4 \\ 0 & 0 & 5 & -2 & 12 \\ -1 & 2 & 1 & -4 & 5 \\ 0 & -2 & 4 & 3 & 5 \\ 1 & -1 & 0 & -11 & 12 \end{array}\right)$$

Transform this into reduced row echelon form.

$$\left(\begin{array}{cccc|c} 1 & -3 & 2 & 1 & 4 \\ 0 & 0 & 5 & -2 & 12 \\ -1 & 2 & 1 & -4 & 5 \\ 0 & -2 & 4 & 3 & 5 \\ 1 & -1 & 0 & -11 & 12 \end{array}\right) \xrightarrow[R5-1\cdot R1]{R3+1\cdot R1} \left(\begin{array}{cccc|c} 1 & -3 & 2 & 1 & 4 \\ 0 & 0 & 5 & -2 & 12 \\ 0 & -1 & 3 & -3 & 9 \\ 0 & -2 & 4 & 3 & 5 \\ 0 & 2 & -2 & -12 & 8 \end{array}\right) \xrightarrow{R2\leftrightarrow R3}$$

$$\left(\begin{array}{cccc|c} 1 & -3 & 2 & 1 & 4 \\ 0 & -1 & 3 & -3 & 9 \\ 0 & 0 & 5 & -2 & 12 \\ 0 & -2 & 4 & 3 & 5 \\ 0 & 2 & -2 & -12 & 8 \end{array}\right) \xrightarrow{(-1)\cdot R2} \left(\begin{array}{cccc|c} 1 & -3 & 2 & 1 & 4 \\ 0 & 1 & -3 & 3 & -9 \\ 0 & 0 & 5 & -2 & 12 \\ 0 & -2 & 4 & 3 & 5 \\ 0 & 2 & -2 & -12 & 8 \end{array}\right) \xrightarrow[R5-2\cdot R2]{\substack{R1+3\cdot R2 \\ R4+2\cdot R2}}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & -7 & 10 & -23 \\ 0 & 1 & -3 & 3 & -9 \\ 0 & 0 & 5 & -2 & 12 \\ 0 & 0 & -2 & 9 & -13 \\ 0 & 0 & 4 & -18 & 26 \end{array}\right) \xrightarrow{\frac{1}{5}\cdot R3} \left(\begin{array}{cccc|c} 1 & 0 & -7 & 10 & -23 \\ 0 & 1 & -3 & 3 & -9 \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{12}{5} \\ 0 & 0 & -2 & 9 & -13 \\ 0 & 0 & 4 & -18 & 26 \end{array}\right) \xrightarrow[\substack{R4+2\cdot R3 \\ R5-4\cdot R3}]{\substack{R1+7\cdot R3 \\ R2+3\cdot R3}}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{36}{5} & -\frac{31}{5} \\ 0 & 1 & 0 & \frac{9}{5} & -\frac{9}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{12}{5} \\ 0 & 0 & 0 & \frac{41}{5} & -\frac{41}{5} \\ 0 & 0 & 0 & -\frac{82}{5} & \frac{82}{5} \end{array}\right) \xrightarrow{\frac{5}{41}\cdot R4} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{36}{5} & -\frac{31}{5} \\ 0 & 1 & 0 & \frac{9}{5} & -\frac{9}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{12}{5} \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -\frac{82}{5} & \frac{82}{5} \end{array}\right) \xrightarrow[\substack{R3+\frac{2}{5}\cdot R4 \\ R5+\frac{82}{5}\cdot R4}]{\substack{R1-\frac{36}{5}\cdot R4 \\ R2-\frac{9}{5}\cdot R4}}$$

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

Now this is in reduced row echelon form. There's no row whose pivot element is in the last column, thus the system is solvable. The number of nonzero rows is 4 just as the number of unknowns, thus the system is determined and the only solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}$$

Clearly the first row of the above matrix in reduced row echelon form gives the equation $1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 1$, that is $x_1 = 1$. The second row gives the equation $0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 0$, that is $x_2 = 0$. The third and fourth rows are similar.

Example 2.

$$\left.\begin{aligned} x_1 - 2x_2 + 4x_3 &= 10 \\ x_1 - x_2 + 3x_3 + 3x_4 &= 10 \\ 2x_2 - x_3 - 3x_4 &= -7 \\ 3x_1 - 2x_2 + 9x_3 + 3x_4 &= 14 \end{aligned}\right\}$$

This system is neither underdetermined nor overdetermined. The extended coefficient matrix of the above system is

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & -2 & 4 & 0 & 10 \\ 1 & -1 & 3 & 3 & 10 \\ 0 & 2 & -1 & -3 & -7 \\ 3 & -2 & 9 & 3 & 14 \end{array}\right)$$

Transform this into reduced row echelon form.

$$\left(\begin{array}{cccc|c} 1 & -2 & 4 & 0 & 10 \\ 1 & -1 & 3 & 3 & 10 \\ 0 & 2 & -1 & -3 & -7 \\ 3 & -2 & 9 & 3 & 14 \end{array}\right) \xrightarrow[R2-1\cdot R1]{R4-3\cdot R1} \left(\begin{array}{cccc|c} 1 & -2 & 4 & 0 & 10 \\ 0 & 1 & -1 & 3 & 0 \\ 0 & 2 & -1 & -3 & -7 \\ 0 & 4 & -3 & 3 & -16 \end{array}\right) \xrightarrow[R4-4\cdot R2]{\substack{R1+2\cdot R2 \\ R3-2\cdot R2}}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 6 & 10 \\ 0 & 1 & -1 & 3 & 0 \\ 0 & 0 & 1 & -9 & -7 \\ 0 & 0 & 1 & -9 & -16 \end{array}\right) \xrightarrow[R4-1\cdot R3]{\substack{R1-2\cdot R3 \\ R2+1\cdot R3}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 24 & 24 \\ 0 & 1 & 0 & -6 & -7 \\ 0 & 0 & 1 & -9 & -7 \\ 0 & 0 & 0 & 0 & -9 \end{array}\right) \xrightarrow{-\frac{1}{9}\cdot R4}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 24 & 24 \\ 0 & 1 & 0 & -6 & -7 \\ 0 & 0 & 1 & -9 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R3-7\cdot R4]{\substack{R1-24\cdot R4 \\ R2+7\cdot R4}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 24 & 0 \\ 0 & 1 & 0 & -6 & 0 \\ 0 & 0 & 1 & -9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

Now this is in reduced row echelon form. The pivot element in the last row is in the last column, thus the system is unsolvable. Clearly the last row of the above matrix in reduced row echelon form gives the equation $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 1$, that is $0 = 1$, which is impossible.

Example 3.

$$\left.\begin{aligned} x_1 + 2x_3 + 3x_4 &= 0 \\ -x_1 + 2x_2 + x_3 - x_4 &= 3 \\ 2x_2 + 3x_3 + 2x_4 &= 3 \\ 3x_1 - 2x_2 + 3x_3 + 7x_4 &= -3 \\ -2x_1 + 6x_2 + 5x_3 &= 9 \end{aligned}\right\}$$

This is an overdetermined system. The extended coefficient matrix of the above system is

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 0 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 2 & 3 & 2 & 3 \\ 3 & -2 & 3 & 7 & -3 \\ -2 & 6 & 5 & 0 & 9 \end{array}\right)$$

Transform this into reduced row echelon form.

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 0 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 2 & 3 & 2 & 3 \\ 3 & -2 & 3 & 7 & -3 \\ -2 & 6 & 5 & 0 & 9 \end{array}\right) \xrightarrow[R5+2\cdot R1]{\substack{R2+1\cdot R1 \\ R4-3\cdot R1}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 0 \\ 0 & 2 & 3 & 2 & 3 \\ 0 & 2 & 3 & 2 & 3 \\ 0 & -2 & -3 & -2 & -3 \\ 0 & 6 & 9 & 6 & 9 \end{array}\right) \xrightarrow{\frac{1}{2}\cdot R2}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 0 \\ 0 & 1 & \frac{3}{2} & 1 & \frac{3}{2} \\ 0 & 2 & 3 & 2 & 3 \\ 0 & -2 & -3 & -2 & -3 \\ 0 & 6 & 9 & 6 & 9 \end{array}\right) \xrightarrow[R5-6\cdot R2]{\substack{R3-2\cdot R1 \\ R4+2\cdot R1}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 0 \\ 0 & 1 & \frac{3}{2} & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Now this is in reduced row echelon form. There's no row whose pivot element is in the last column, thus the system is solvable. The number of nonzero rows is 2, but the number of unknowns is 4, thus the system is undetermined. The free variables are x_3 and x_4 as the third and fourth columns of the coefficient matrix don't contain any pivot element. This means all unknowns can be given as a linear combination of the free variables plus and additive constant. Clearly the first row gives the equation $1 \cdot x_1 + 0 \cdot x_2 + 2 \cdot x_3 + 3 \cdot x_4 = 0$ and the second row gives the equation $0 \cdot x_1 + 1 \cdot x_2 + \frac{3}{2} \cdot x_3 + 1 \cdot x_4 = \frac{3}{2}$. The third, fourth and fifth rows give only the identity $0 = 0$. Thus we have

$$\left.\begin{aligned} x_1 + 2x_3 + 3x_4 &= 0 \\ x_2 + \tfrac{3}{2}x_3 + x_4 &= \tfrac{3}{2} \end{aligned}\right\}$$

which can be rearranged as

$$\left.\begin{aligned} x_1 &= -2x_3 - 3x_4 \\ x_2 &= -\tfrac{3}{2}x_3 - x_4 + \tfrac{3}{2} \end{aligned}\right\}$$

In matrix form this is

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ -\frac{3}{2} & -1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{3}{2} \end{pmatrix}$$

Here

$$B = \begin{pmatrix} -2 & -3 \\ -\frac{3}{2} & -1 \end{pmatrix}$$

is the coefficient matrix of the free variables and

$$\underline{\widehat{b}} = \begin{pmatrix} 0 \\ \frac{3}{2} \end{pmatrix}$$

is just as it's defined in Theorem 6. Now we can choose arbitrary real values for x_3 and x_4 and then computing $x_1 = -2x_3 - 3x_4$ and $x_2 = -\frac{3}{2}x_3 - x_4 + \frac{3}{2}$ we get a solution of the system. For example $x_3 = 0$ and $x_4 = 0$ gives $x_1 = 0$ and $x_2 = \frac{3}{2}$, thus it's a solution of the system. Another solution is $x_3 = 1$, $x_4 = 1$ which gives $x_1 = -5$, $x_2 = -1$.