Matrix Theory — Lecture notes
MAT 335, Fall 2022 — D. Ivanšić

1.1 Matrices and Vectors

A matrix is a principal object of our study — it is simply a rectangular array of numbers.

Definition. An $m \times n$ ("m by n") matrix is a rectangular array of numbers (scalars) that has m rows and n columns. The scalar in the i-th row and j-th column is typically denoted a_{ij} and called the (i,j)-entry of the matrix. For example:

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

$$\text{Or, in general,} \quad \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Two matrices of the same dimensions $m \times n$ can be added, subtracted and multiplied by a scalar, all done component-wise.

Example. Let $A = \begin{bmatrix} 2 & 4 & 5 \\ -4 & -3 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 & 8 \\ 7 & 0 & 1 \end{bmatrix}$. Compute $A+B$, $2A$, $4A-3B$.

Another operation on matrices is transposing, which can be imagined as flipping the matrix over its main diagonal (the downwards-going one). The transpose of the $m \times n$ matrix A is the $n \times m$ matrix A^T.

Rows of A become columns in A^T
Columns of A become rows of A^T
The (i,j)-element of A becomes the (j,i)-element of A^T

For example,

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

A special matrix is the *zero matrix*, whose entries are all zeroes. It will be denoted by 0, just like the scalar (context tells us if it is a matrix or a scalar).

Theorems 1.1, 1.2: Properties of matrix addition, scalar multiplication and the transpose. Let A, B and C be $m \times n$ matrices, s and t scalars. Let $-A$ denote $(-1)A$. Then

$$\begin{array}{ccc} A+B=B+A & A+0=A & s(tA)=(st)A \\ (A+B)+C=A+(B+C) & A+(-A)=0 & s(A+B)=sA+SB \\ \text{(commutativity and associativity)} & & (s+t)A=sA+tA \end{array}$$

$$(A+B)^T = A^T + B^T \qquad (sA)^T = sA^T \qquad (A^T)^T = A$$

Definition. A *vector* is a matrix with either exactly one column *(column vector)*, or one row *(row vector)*, for example

$$\begin{bmatrix} 4 \\ -7 \\ 3 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -5 & 2 & -6 \end{bmatrix}.$$

The *components* of the vector are its entries. If a vector has n components, then the set of all vectors is denoted by $\mathbf{R}^n$. We usually work with column vectors.
Vectors can be visualized as arrows in the plane or in space, for example

$$\mathbf{R}^3 \longleftrightarrow \text{points in space}$$

Note. The properties of matrix operations apply to vectors as well. Actually, when you first studied vectors, you were given the same list of properties as the ones above (without the transpose properties). The reason that the properties first introduced for vectors hold for matrices is that an $m \times n$ matrix can be thought of as list of mn numbers — that is, a vector with mn components — that was arranged in a rectangular array. So, for purposes of addition and multiplication by scalar, matrices act like vectors. The usefulness of putting those mn numbers in rectangular arrays will become clear as we progress through the course.

It is often convenient to think of an $m \times n$ matrix as a collection of n column vectors with m components.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & \dots & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ | & | & \dots & | \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \text{ where } \mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Geometry of vectors

Geometrically, adding vectors can be done via:

parallelogram law triangle law scalar multiplication

Multiplication by a scalar is essentially lenghtening the vector by a factor, reversing direction if scalar is negative.

In this course, we will mainly be thinking of vectors as all emanating from the origin — the parallelogram law then applies to any two vectors.

In general, a vector can be represented by an arrow with any starting point. Then, two such arrows are considered the same if they have the same direction (are parallel and point the same way) and magnitude. Alternatively, if there is a coordinate system set up in space, the arrows are the same if, for both arrows, the terminal point is reached from the starting point by going a_1 units in the direction of the x-axis, a_2 units in the direction of the y axis, and a_3 units in direction of the z axis. In this case, a_1, a_2 and a_3 are components of the vector. Under this interpretation, a vector is a "forest" of arrows of same direction and magnitude, while an arrow is a "tree" in this forest, said to *represent* the vector.

1.2 Linear Combinations and Matrix-Vector Products

Definition. A *linear combination* of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is any vector of the form

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n,$$

where $c_1, c_2, \ldots, c_n$ are scalars (called *coefficients* of the linear combination).

Identify all possible linear combinations of the vectors shown in space.

Example.

Is the vector $\begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$?

Is the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$?

Example.

Is the vector $\begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$?

Is the vector $\begin{bmatrix} 3 \\ 0 \\ 9 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$?

Note. Solving these problems boiled down to solving systems of linear equations.

Definition. The *standard basis vectors* of $\mathbf{R}^n$ is the collection of n vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

The $n \times n$ matrix $I_n = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix}$ is called the *identity matrix*

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

For any vector $\mathbf{v}$ in $\mathbf{R}^n$ whose components are $v_1, v_2, \dots, v_n$ it is clear that

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n,$$

in other words, every vector in $\mathbf{R}^n$ is a linear combination of the standard basis vectors.

Illustrating this for $\mathbf{R}^3$:

Definition of Matrix-Vector Multiplication. Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix given with columns $\mathbf{a}_1, \mathbf{a}_2, \dots \mathbf{a}_n$, and v be an $n \times 1$ vector with components $v_1, v_2, \dots, v_n$. The product of matrix A and vector v is

$$A\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + \dots + v_n\mathbf{a}_n$$

$$\text{Alternatively, } A\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n \end{bmatrix}$$

In the first interpretation, $A\mathbf{v}$ is an $m \times 1$ vector that is the linear combination of columns of A with coefficients the components of $\mathbf{v}$.

In the second interpretation, $A\mathbf{v}$ is an $m \times 1$ vector whose components are dot products of rows of A with the column $\mathbf{v}$.

Example. Multiply.

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

Example. Find the components of the vector obtained by rotating the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ by angle θ around the origin.

Therefore, rotating the vector $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ by angle θ results in vector $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.
The matrix A_θ shown is called a rotation matrix.

Rotate the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ around the origin by angle $\frac{\pi}{3}$.

Example. Below is data on nutritional value of various foods per 100g serving. We can use a matrix to represent this data, and we can use matrix-vector multiplication to find out the nutritional value of a meal of x servings of chicken, y servings of rice and z servings of lettuce contains.

Find the nutritional values for a meal with 2 servings of chicken, 1 servings of rice and 2 servings of lettuce.

	chicken	rice	lettuce
energy (kcal)	149	359	17
fat (g)	6	1	0
protein (g)	24	7	1
carbohydrates (g)	0	80	3

Start by arranging the data in a matrix:

$$\begin{bmatrix} 149 & 359 & 17 \\ 6 & 1 & 0 \\ 24 & 7 & 1 \\ 0 & 80 & 3 \end{bmatrix}$$

Theorems 1.3: Properties of matrix-vector multiplication. Let A, B be $m \times n$ matrices, and $\mathbf{u}$ and $\mathbf{v}$ vectors in $\mathbf{R}^n$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ standard basis vectors for $\mathbf{R}^n$, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ columns of A, c a scalar. Then

$$A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} \qquad A\mathbf{e}_j = \mathbf{a}_j, j = 1\ldots, n$$
$$A(c\mathbf{u}) = c(A\mathbf{u}) = (cA)\mathbf{u} \qquad A\mathbf{0} = \mathbf{0}$$
$$(A+B)\mathbf{u} = A\mathbf{u} + B\mathbf{u} \qquad 0\mathbf{v} = \mathbf{0}$$
$$I_n\mathbf{v} = \mathbf{v}$$

If $A\mathbf{u} = B\mathbf{u}$ for every vector $\mathbf{u}$ in $\mathbf{R}^n$ then $A = B$

Proof of some statements.

1.3 Systems of Linear Equations

Definition. A *linear equation* in variables $x_1, \ldots, x_n$ is an equation of form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where the real numbers $a_1, \ldots, a_n$ are called *coefficients* and b is called the *constant term.*

Example. What do linear equations in two or three variables represent in $\mathbf{R}^2$ or $\mathbf{R}^3$?

Definition. A *system of linear equations* is a set of m linear equations in n variables (like above). It can be written in form

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

A *solution* of a system of linear equations in variables $x_1, \ldots, x_n$ is the set of all vectors (n-tuples) $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ that make *all* equations true.

Example. Use geometric interpretation to see how many solutions a system of equations in two or three variables could have.

To solve a system of equations in an organized way, at every step we transform the system into an *equivalent* system, that is, a system of equations that has the same solutions as the original one. The goal is to reach a system of equations for which the solution is clear. The following operations result in an equivalent system:

1) interchange any two equations
2) multiply an equation by a *nonzero* scalar
3) add a multiple of one equation to another

The key reason why these operations result in equivalent systems is that they are reversible using the same type of operations: if we transform system S into system T via an operation of type 1, 2 or 3, then system T can be transformed into system S via an operation of the same type.

Now, if $x_1, \ldots, x_n$ satisfy a system S, then $x_1, \ldots, x_n$ will satisfy the system T. Because S is obtained from T also via an operation of type 1, 2 or 3, the same will hold: if $x_1, \ldots, x_n$ satisfy the system T, then $x_1, \ldots, x_n$ satisfy the system S. This implies that systems S and T have the same set of solutions.

Proof of equivalence via example. Consider the system below.

$$\begin{cases} 2x_1 & +3x_2 & -x_3 & = & 0 \\ x_1 & +x_2 & +3x_3 & = & 4 \\ 5x_1 & +6x_2 & -10x_3 & = & 12 \end{cases}$$

Now, let's solve this system.

Note that a general system of equations

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

can be written as

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

Therefore, our system can be viewed as a matrix equation $A\mathbf{x} = \mathbf{b}$, and we are to find all vectors $\mathbf{x}$ that satisfy it. The matrix A is called the *coefficient matrix* of the system. By adjoining column $\mathbf{b}$ to the matrix A on the right side, we get the *augmented matrix* of the system, denoted $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$.

We can associate an augmented matrix to every system of equations. It is easily seen that if system T is obtained from system S by one of above operations 1,2 or 3 on the system, the augmented matrix of system T is obtained from the augmented matrix of system S by a similar operation performed on *rows*.

We have *elementary row operations* on a matrix:

1) interchange any two rows
2) multiply a row by a *nonzero* scalar
3) add a multiple of one row to another

By performing row operations we are essentially transforming a system of equations into an equivalent system of equations with less writing.

Example. Solve the previous system using row operations on the augmented matrix.

What form should an augmented matrix have so that "the solution is clear?" The same one as when we solved the example above, which is, for example ($*$ is any number)

$$\begin{bmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & * & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

This form is called the *reduced row echelon* form.

Definition. A *zero row* of a matrix is one where all entries are 0, otherwise, the row is a *nonzero row*. The leftmost nonzero entry in a row is called a *leading entry*. A matrix is in *row echelon* form if

1) each nonzero row lies above every zero row
2) the leading entry of a nonzero row lies to the right of the leading entry of any preceding row
3) if a column contains a leading entry, all the entries of the column below the leading entry are 0

A matrix is in *reduced row echelon* form if it additionally satisfies

4) if a column contains a leading entry, all the other entries of the column are 0
5) the leading entry of each nonzero row is 1

Once the matrix is in reduced row echelon form, the solutions are easy to find.

Example. Solve systems whose augmented matrices are:

$$\begin{bmatrix} 0 & 1 & 2 & 0 & -4 & 3 \\ 0 & 0 & 0 & 1 & 6 & -11 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

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

Theorem 1.4. Every matrix can be transformed into exactly one matrix in reduced row echelon form via a sequence of elementary row operations.

The theorem says that for a given matrix, any sequence of elementary row operations that ends with a matrix in reduced row echelon form gives the same matrix.

Procedure for Solving a System of Linear Equations.

1) Write the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ of the system.
2) Find the reduced row echelon form $\begin{bmatrix} R & \mathbf{c} \end{bmatrix}$ of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$
3) If $\begin{bmatrix} R & \mathbf{c} \end{bmatrix}$ contains a row in which the only nonzero entry is in the last column, the system $A\mathbf{x} = \mathbf{b}$ has no solution.
   Otherwise, the system has at least one solution.
   Write the system corresponding to $\begin{bmatrix} R & \mathbf{c} \end{bmatrix}$, and solve the system for basic variables (corresponding to columns with leading 1's) in terms of the free variables (corresponding to remaining columns other than the rightmost one) to get a general solution of $A\mathbf{x} = \mathbf{b}$.

1.4 Gaussian Elimination

A system of equations is easy to solve when its augmented matrix is in reduced row echelon form. Here we find out how to get any matrix into reduced row echelon form by row operations.

The idea is to use a nonzero entry in a column — called a *pivot position* — to turn all entries above and below it into zeroes via row operations ("eliminate" the entries). It is preferable if the pivot position is a 1 (swap rows to achieve this).

We start in the leftmost nonzero column, choose a pivot position and bring it into the first row if needed. Then eliminate all entries *below it.*

Once this is done, move down and to the right from the pivot position to find the next pivot position and repeat eliminating all entries *below* the pivot position. In this stage, rows above the current pivot position remain as they were — only the rows below are affected. Once the pivot positions are exhausted, the matrix is in row echelon form, and the pivot positions will be the leading entries in each row. This is the end of the *forward pass* of the algorithm.

The *backward pass* of the algorithm is to turn the leading entries in the rows into 1's and eliminate all the entries *above them*, as the ones below are already all zero. We start with the rightmost (and lowest) leading entry and work upwards. When done, the matrix will be in reduced row echelon form.

Example. Find the reduced row echelon form of the matrix. Write the solution of the system, assuming the matrix was the augmented matrix of the system.

$$\begin{bmatrix} -2 & -2 & 10 & -7 & 3 & 9 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ 1 & 3 & -1 & 2 & 2 & -1 \\ -2 & 0 & 14 & -9 & 3 & 11 \end{bmatrix}$$

Example. Solve the system.

$$\begin{cases} 2u & -8v & +14w & +12z & = & 6 \\ 3u & -10v & +26w & +7z & = & 22 \\ & -3v & -8w & +17z & = & -13 \\ u & -3v & +10w & +z & = & 7 \end{cases}$$

Definition. The *rank* of an $m \times n$ matrix A, denoted $\operatorname{rank} A$, is the number of nonzero rows in the reduced row echelon form of A, which is same as the number of pivot columns. The *nullity* of A, denoted $\operatorname{nullity} A$, is the number of columns of A in the reduced row echelon form that do not contain leading entries, same as the number of nonpivot columns. Note that

$$\operatorname{rank} A + \operatorname{nullity} A = n$$

Example. Go back to the previous two examples and find the rank and nullity of the augmented matrices.

Note. An $n \times n$ matrix A has rank n if and only if its reduced row echelon form is I_n.

Definition. A system of equations is called *consistent* if it has at least one solution, otherwise it is *inconsistent*.

Consider the system $A\mathbf{x} = \mathbf{b}$ and note that the reduced row echelon form of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ is the reduced row echelon form of A with an additional column. When the additional column produces an additional nonzero row, the system is inconsistent, otherwise it is consistent.

In other words, we see that the system is inconsistent if and only if

$$\operatorname{rank} \begin{bmatrix} A & \mathbf{b} \end{bmatrix} > \operatorname{rank} A.$$

Now assume the system $A\mathbf{x} = \mathbf{b}$ is consistent. Going back to how we were able to solve the system once it was in reduced row echelon form, *considering just the reduced row echelon form of* A, we see:

1) $\operatorname{rank} A$ is equal to the number of basic variables in the general solution
2) $\operatorname{nullity} A$ is equal to the number of free variables in the general solution, so the system has infinitely many solutions if and only if $\operatorname{nullity} A > 0$.

Theorem 1.5. The following conditions are equivalent.

a) The system $A\mathbf{x} = \mathbf{b}$ is consistent.
b) The vector $\mathbf{b}$ is a linear combination of the columns of A.
c) The reduced row echelon form of A has no row of form $\begin{bmatrix} 0 & \dots & 0 & d \end{bmatrix}$, where $d \neq 0$.

1.6 The Span of a Set of Vectors

Definition. For a nonempty set of vectors $\mathcal{S} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of vectors in $\mathbf{R}^n$ we define the span of $\mathcal{S}$ to be the set of all linear combinations of $\mathbf{u}_1, \ldots, \mathbf{u}_k$. It is denoted by $\operatorname{Span}\mathcal{S}$ or $\operatorname{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$.

Example. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be three nonzero vectors in $\mathbf{R}^3$. Describe the possibilities for $\operatorname{Span}\{\mathbf{a}\}$, $\operatorname{Span}\{\mathbf{a}, \mathbf{b}\}$, $\operatorname{Span}\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

Example. Do the vectors $\begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 11 \\ 7 \\ -3 \end{bmatrix}$, and $\begin{bmatrix} 2 \\ -2 \\ 9 \end{bmatrix}$ span $\mathbf{R}^3$?

The problem illustrates the following theorem.

Theorem 1.6. The following statements about an $m \times n$ matrix A are equivalent:

(a) The span of columns of A is $\mathbf{R}^m$.

(b) The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbf{R}^m$.

(c) The rank of A is m, the number of rows of A.

(d) The reduced row echelon form of A has no zero rows.

(e) There is a leading 1 in each row of the reduced row echelon form.

Proof. $a \iff b, c \iff d \iff e$ clear, show $b \iff c$

For the example above, we found out that the four vectors span $\mathbf{R}^3$. We know that, in general, three vectors are enough to span $\mathbf{R}^3$. Can we drop one of the four in the example?

Theorem 1.6. Let $\mathcal{S} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a set of vectors in $\mathbf{R}^n$, and $\mathbf{v}$ a vector in $\mathbf{R}^n$. Then $\text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}\} = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ if and only if $\mathbf{v}$ is in $\text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$.

This means that if one of the vectors is a linear combination of the others, it can be dropped from the spanning set.

Proof.

Example. In the example above, which vector can be dropped from the spanning set, because it is a linear combination of the other vectors?

1.7 Linear Independence and Linear Dependence

The example at the end of section 1.6 is instructive for determining when one vector in a set is a linear combination of the others. We tried to solve the system

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3 + x_4\mathbf{u}_4 = \mathbf{0}$$

and the existence of a solution which was not all zeroes allowed us to express some vectors in terms of the others.

Definition. A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is called *linearly independent* if the only scalars $c_1, \ldots, c_k$ that satisfy

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$$

are all zero (that is, $c_1 = c_2 = \cdots = c_k = 0$)

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is *linearly dependent* if it is not linearly independent, which means that there are scalars $c_1, \ldots, c_k$, of which at least one is not zero, so that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}.$$

Example. Any set of vectors that contains a zero vector is linearly dependent.

Note. A set of vectors is linearly dependent if and only if one of them is a linear combination of the others. Therefore, in a linearly independent set, no vector is a linear combination of the others.

Note. A set of two vectors is linearly dependent if and only if one is a multiple of the other.

Note the language: One does **not** say that:
— a vector is linearly independent of some other vectors, or
— a vector is linearly dependent on some other vectors,

because linear (in)dependence is not a property of a single vector, but of a set of vectors.

To capture what may have been attempted in the above two statements, one would say
— a vector is not a linear combination of some other vectors, or
— a vector is a linear combination of some other vectors.

Example. Are the following sets of vectors linearly independent?

a) $\left\{ \begin{bmatrix} -1 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ -1 \\ -1 \\ 2 \end{bmatrix} \right\}$

b) $\left\{ \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}$

Figuring whether vectors are linearly independent boils down to solving a system of type $A\mathbf{x} = \mathbf{0}$, which comes up often.

Definition. The system of equations $A\mathbf{x} = \mathbf{b}$ is called *homogeneous* when $\mathbf{b} = \mathbf{0}$.

Note. A homogeneous system of equations always has at least one solution — all zeroes. The question usually is whether it has only one or infinitely many solutions.

The solution set of a homogeneous system is the span of a finite set of vectors. If the solution was obtained using Gauss elimination, the vectors in the spanning set are linearly independent.

Theorem 1.8. The following statements about an $m \times n$ matrix A are equivalent:

(a) The columns of A are linearly independent.

(b) The equation $A\mathbf{x} = \mathbf{b}$ has at most one solution for each $\mathbf{b}$ in $\mathbf{R}^m$.

(c) The nullity of A is zero.

(d) The rank of A is n, the number of columns of A.

(e) The reduced row echelon form of A has the first n standard basis vectors or $\mathbf{R}^m$ as its columns.

(f) The only solution of $A\mathbf{x} = \mathbf{0}$ is $\mathbf{0}$.

(g) There is a leading 1 in each column of the reduced row echelon form.

Proof. $a \iff f \iff g$ clear, show $b \implies c \implies d \implies e \implies f \implies b$

Note. A subset of $\mathbf{R}^m$ containing more than m vectors is linearly dependent.

Theorem 1.9. Vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ in $\mathbf{R}^n$ are linearly dependent if and only if $\mathbf{u}_1 = \mathbf{0}$ or there is an $i \geq 2$ such that u_i is a linear combination of the preceding vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}$ (Thus, some vector in the set can be written as a linear combination of its *predecessors*, and not just other vectors in the set.)

Proof.

Table summarizing essential takeaways for an $m \times n$ matrix A.

rank of A	Solutions of $A\mathbf{x} = \mathbf{b}$	Columns of A	Reduced row echelon form of A
rank $A = m$	at least one for every $\mathbf{b}$ in $\mathbf{R}^m$	span $\mathbf{R}^m$	has a leading 1 in every row
rank $A = n$ nullity $A = 0$	at most one for every $\mathbf{b}$ in $\mathbf{R}^m$	are linearly independent	has a leading 1 in every column

Geometric interpretation of the solution of a linear system. For a consistent system $A\mathbf{x} = \mathbf{b}$, the set of solutions can be described as

(general solution of $A\mathbf{x} = \mathbf{b}$) = (particular solution of $A\mathbf{x} = \mathbf{b}$)+(general solution of $A\mathbf{x} = \mathbf{0}$)

Geometrically, it is a shift of the span of some vectors.

Example. Write the solution of the system whose augmented matrix is below.

$$\begin{bmatrix} -2 & -2 & 10 & -7 & 3 & 9 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ 1 & 3 & -1 & 2 & 2 & -1 \\ -2 & 0 & 14 & -9 & 3 & 11 \end{bmatrix}$$

has reduced row echelon form

$$\begin{bmatrix} 1 & 0 & -7 & 0 & -33 & -19 \\ 0 & 1 & 2 & 0 & 7 & 4 \\ 0 & 0 & 0 & 1 & 7 & 3 \end{bmatrix}$$