Linear algebra for data science

Sridharakumar Narasimhan

IIT Madras

Contents

Contents

Contents

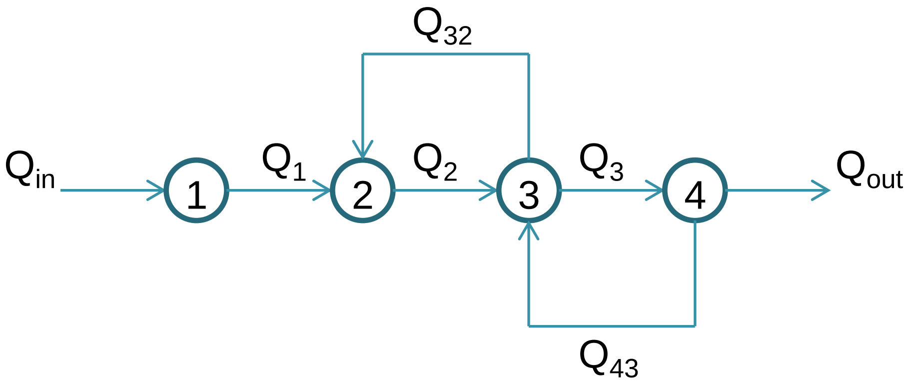
Outcome

Module learning outcomes:

1. Participants will be able to identify relationships between variables in large datasets
2. Participants will be able to identify information sufficiency in terms of both equations and variables
3. Participants will be able to understand basic linear algebra concepts that underlie the complicated data analytics algorithms

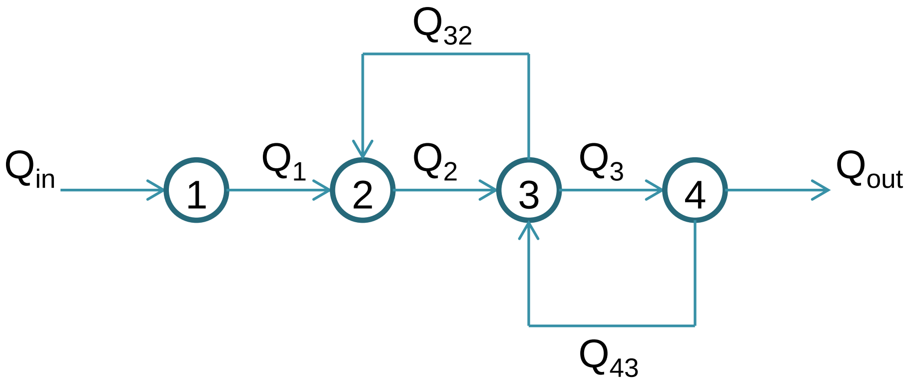
Linear Algebra

System of Equations



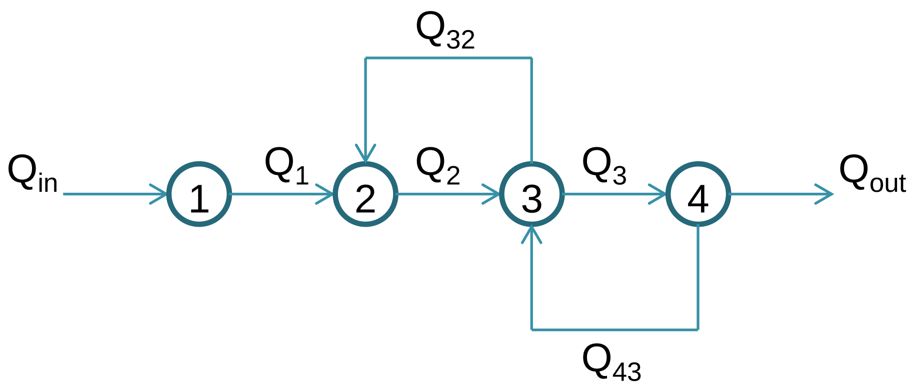
How many equations could be formed from this?

System of Equations



- $Q_{in} = Q_1$
- $Q_1 + Q_{32} = Q_2$
- $Q_2 + Q_{43} = Q_{32} + Q_3$
- $Q_3 = Q_{out} + Q_{43}$

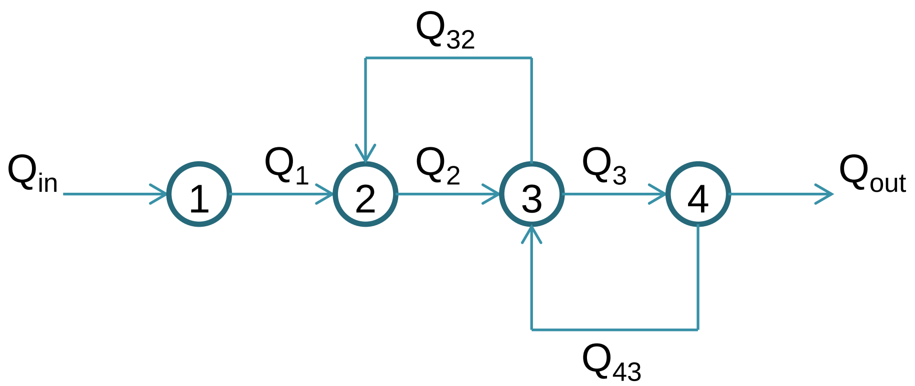
System of Equations



- $Q_{in} = Q_1$
- $Q_1 + Q_{32} = Q_2$
- $Q_2 + Q_{43} = Q_{32} + Q_3$
- $Q_3 = Q_{out} + Q_{43}$

- Can we solve the equations when –
  - $Q_{in} = 10$, $Q_{32} = 5$ and $Q_{43} = 3$
  - $Q_{in} = 10$, $Q_{out} = 10$ and $Q_3 = 7$
  - $Q_{in} = 10$, $Q_{out} = 11$ and $Q_3 = 7$

System of Equations

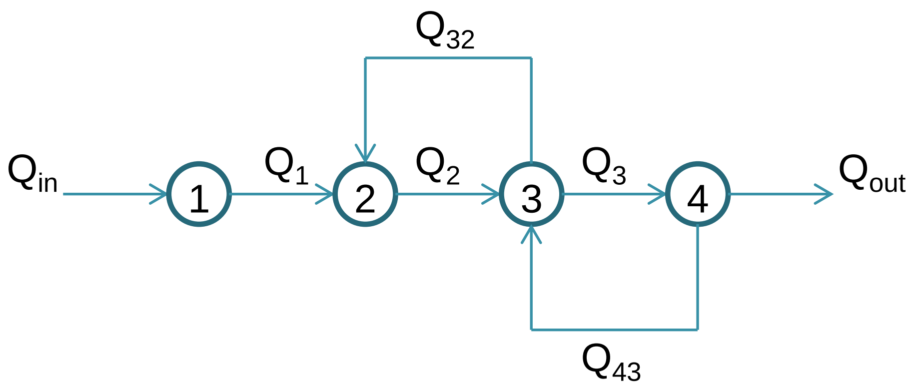


- $Q_{in} = Q_1$
- $Q_1+Q_{32} = Q_2$
- $Q_2+Q_{43} = Q_{32}+Q_3$
- $Q_3 = Q_{out}+Q_{43}$

- Can we solve the equations when $Q_{in} = 10$, $Q_{32} = 5$ and $Q_{43} = 3$?

Unique solution

System of Equations



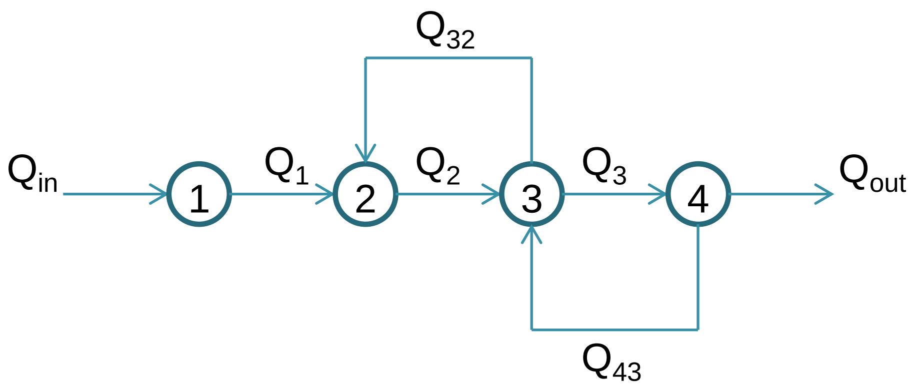
- $Q_{in} = Q_1$
- $Q_1+Q_{32} = Q_2$
- $Q_2+Q_{43} = Q_{32}+Q_3$
- $Q_3 = Q_{out}+Q_{43}$

- Can we solve the equations when $Q_{in} = 10$, $Q_{out} = 10$ and $Q_3 = 17$?

Infinite number of solutions

- Loop involving Q_2 and Q_{32} : both are unmeasured
- Given measurements consistent with equations

System of Equations



- $Q_{in} = Q_1$
- $Q_1+Q_{32} = Q_2$
- $Q_2+Q_{43} = Q_{32}+Q_3$
- $Q_3 = Q_{out}+Q_{43}$

- Can we solve the equations when $Q_{in} = 10$, $Q_{out} = 11$ and $Q_3 = 7$?

No solution

- Given information not consistent with equations

System of Equations: Key Concept

Understanding when do we have:

- Unique solution
- No solution
- Infinite number of solutions

Solving Simultaneous Linear Equations

Solve the two linear equations:

$4x - 2y = 0$; $-2x + 4y = 6$

Solving Simultaneous Linear Equations

Elimination (High–school method)

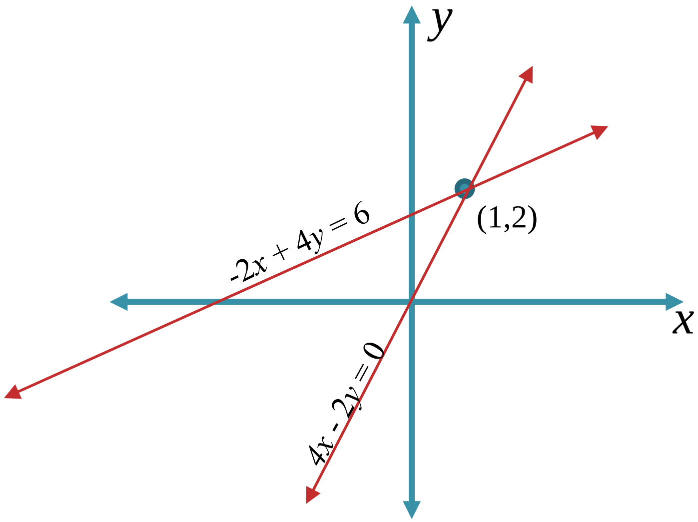
$$
\begin{array}{r}
4\cancel{x} - 2y = 0 \\
2(-\cancel{2}x + 4y = 6) \\
\hline
6y = 12
\end{array}
$$

$\Rightarrow y = 2$

$\Rightarrow x = 1$

(1,2) is the solution

Graphical method



(1,2) is the solution

Geometry of Linear Equations

- We may view a system of linear equations in three different ways –
  - Matrix form – **A***x* **= b** where **A** forms a matrix with the coefficients of the unknowns and ***x*** forms a matrix with the unknowns and **b**, a matrix with the values in the R.H.S
  - Row picture – viewing one equation at a time
  - Column picture – two separate equations as one vector equation

Geometry of Linear Equations

Geometry for a system of 2 equations:

Matrix form

- Consider two linear equations:

$$4x - 2y = 0$$

$$-2x + 4y = 6$$

- A matrix is a rectangular arrangement of numbers in rows and columns
- Rows run horizontally and columns run vertically

- Order of a matrix: $m \times n$ where m is the # of rows and n is the # of columns

Geometry of Linear Equations

<u>Matrix form</u>

$$\begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

- This is of the form **A*x* = b**

  where –

  - **A** – matrix with coefficients of the unknowns
  - ***x*** – unknowns
  - **b** – R.H.S of the equations

- n equations and n unknowns $\Rightarrow$ $n \times n$ matrix (square matrix)

- m equations and n unknowns $\Rightarrow$ $m \times n$ matrix (rectangular matrix)

Geometry of Linear Equations

Row picture

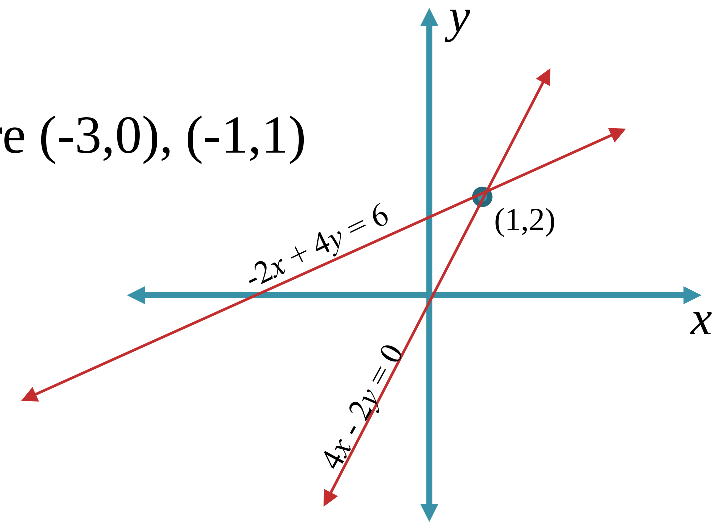
$$4x - 2y = 0 \; ; -2x + 4y = 6$$

- Taking one row at a time and plotting it in the x–y plane

Few points that satisfy $4x - 2y = 0$ are (0,0), (1,2) and (1/2, 1)

Few points that satisfy $-2x + 4y = 6$ are (-3,0), (-1,1) and (1,2)

So the solution of the system is (1,2)



Geometry of Linear Equations

Geometry for a system of 3 equations

- Consider a system of 3 equations:

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

Matrix form

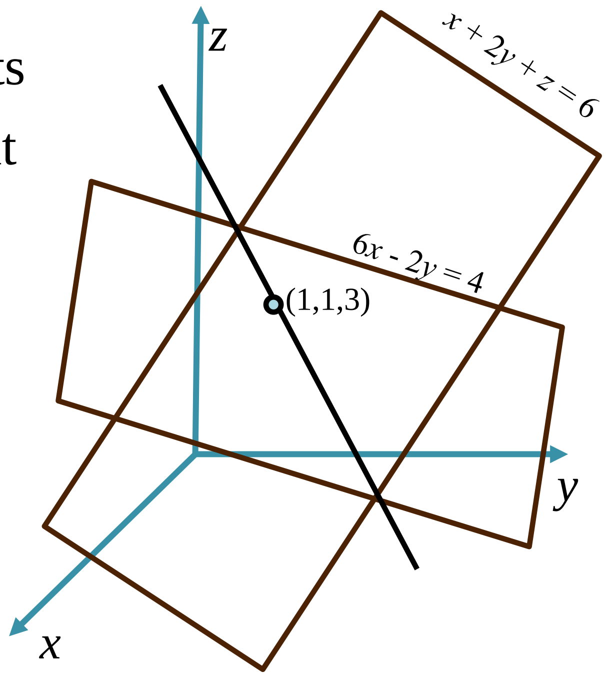
$$
\begin{bmatrix} 1 & 2 & 1 \\ 6 & -2 & 0 \\ -3 & -1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 8 \end{bmatrix}
$$

Geometry of Linear Equations

Row picture

$$x + 2y + z = 6 \;;\; 6x - 2y = 4 \;;\; -3x - y + 4z = 8$$

- Each equation describes a plane in 3 dimensions. The intersection of the first plane with the second plane is a line
- The 3rd plane (not shown in the figure) intersects the line of intersection of the other two planes at a point (1,1,3)
- Solution for the system of equation is (1,1,3)



Types of Solutions

A system of linear equations can have –

- Unique solution

  or

- No solution

  or

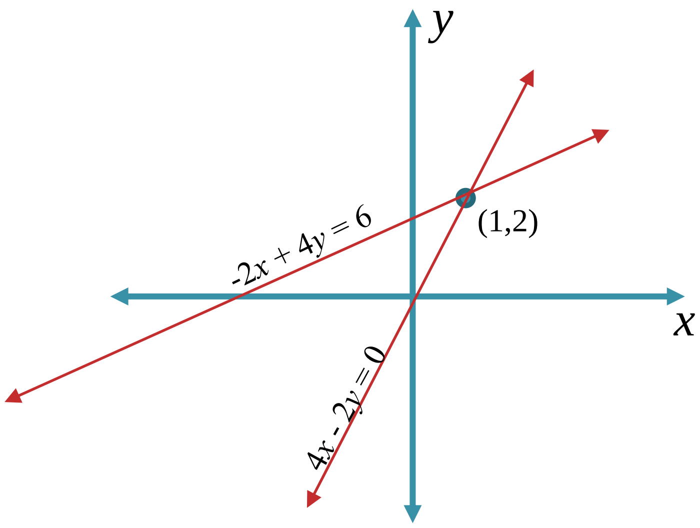
- Infinite number of solutions

Types of Solutions

<u>2 dimensional case</u>

<u>Unique solution</u>

- This is the nice case where the system will have a point of intersection and hence a unique solution. $4x - 2y = 0$ and $-2x + 4y = 6$ has a unique solution (1,2)



One solution (x, y) = (1,2)

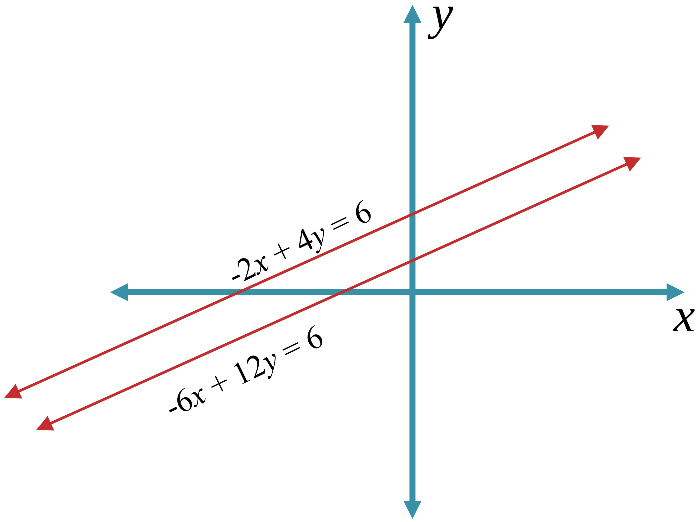
Types of Solutions

<u>No solution</u>

- A system has no solution if the equations are inconsistent. For example, $-2x + 4y = 6$ and $-6x + 12y = 6$ has no solution

$$-6x + 12y = 6 \Rightarrow -2x + 4y = 2$$

which contradicts with the first equation and hence the system has no solution



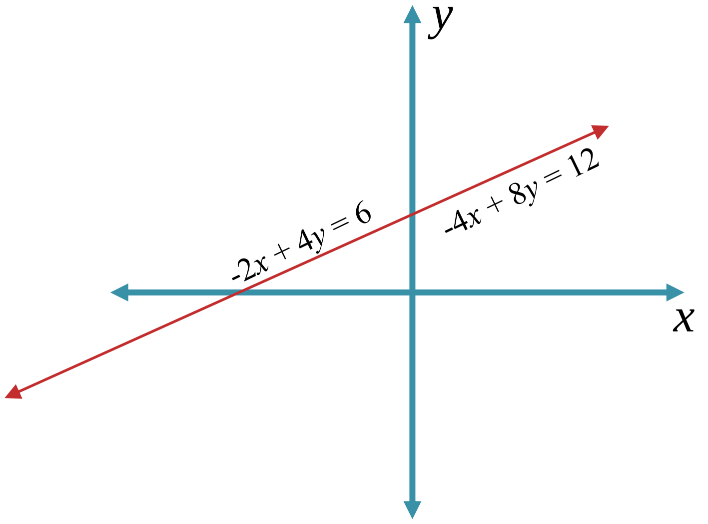
Parallel: No solution

Types of Solutions

<u>Infinite number of solutions</u>

- The other case is when one equation is just some multiple of the other. Then we will get infinite number of solutions

  $-2x + 4y = 6$; $-4x + 8y = 12$

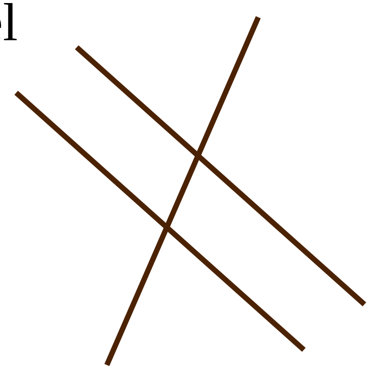


Whole line of solutions

Types of Solutions

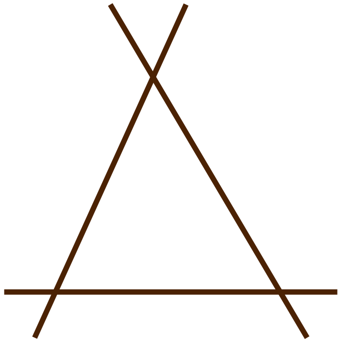
<u>Singular case for three dimensions</u>

- Two planes may be parallel



Two parallel planes: No solution

- One plane is parallel to the line of intersection of the other two planes

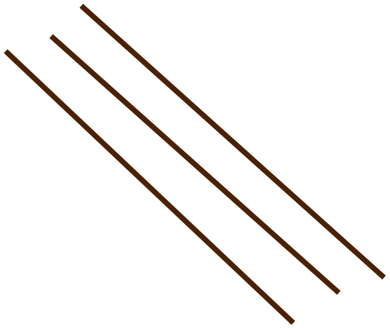


No intersection: No solution

Types of Solutions

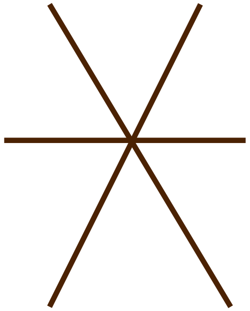
<u>Singular case for three dimensions</u>

- Three parallel planes



All planes parallel: No solution or a whole plane of solutions

- One equation is just the sum of the other two equations, the three planes have a whole line in common

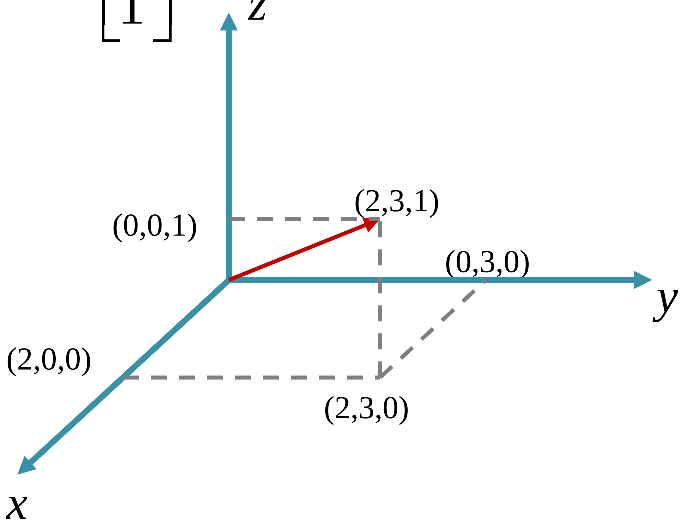


Line of intersection: Infinite # of solutions

Geometry of Linear Equations

Vector

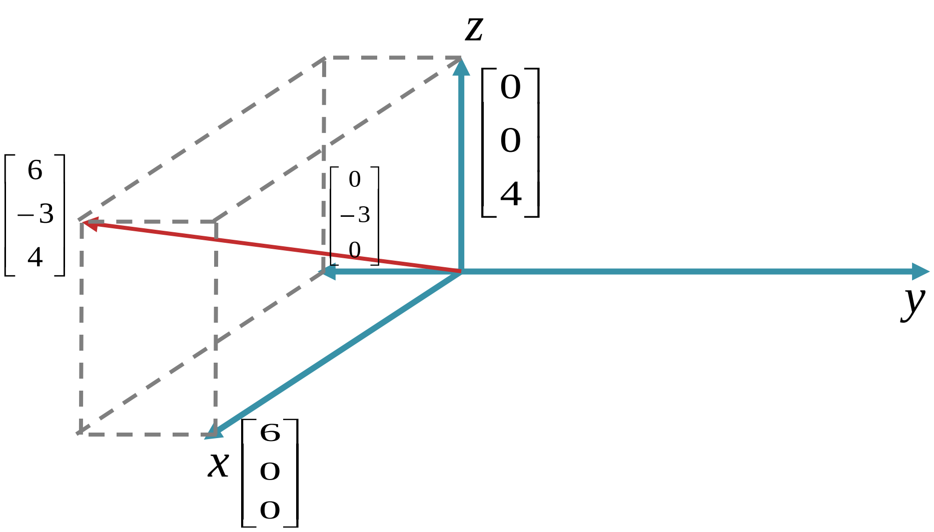
- A vector is defined as an ordered collection of numbers
- Elements of a vector arranged as a column $\longrightarrow$ column vector
- Elements of a vector arranged as a row $\longrightarrow$ row vector
- If a vector v contains three real numbers say, $v = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$, then v belongs to the vector space R^3
- The vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} b \\ a \end{bmatrix}$ are not the same



Vector Addition

Addition of a vector $\begin{bmatrix} 6 \\ -3 \\ 4 \end{bmatrix}$ is done component by component and can be written as –

$$\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \\ 4 \end{bmatrix}$$



Geometry of Linear Equations

<u>Column picture</u>

$$4x - 2y = 0 \;\; ; \;\; -2x + 4y = 6$$

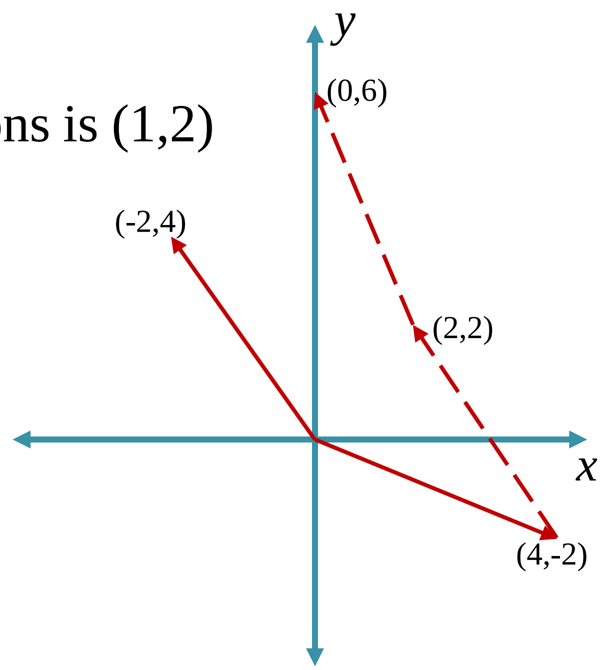
- Column picture: Linear combination of columns

$$x\begin{bmatrix}4\\-2\end{bmatrix} + y\begin{bmatrix}-2\\4\end{bmatrix} = \begin{bmatrix}0\\6\end{bmatrix}$$

We know that the solution for the two equations is (1,2)

Substitute them

$$1\begin{bmatrix}4\\-2\end{bmatrix} + 2\begin{bmatrix}-2\\4\end{bmatrix} = \begin{bmatrix}0\\6\end{bmatrix}$$



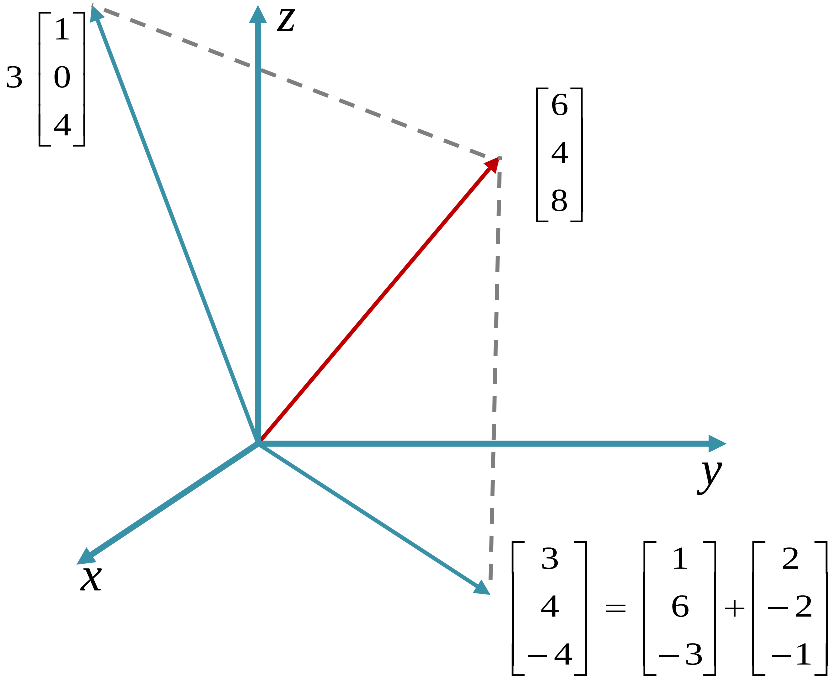
Geometry of Linear Equations

<u>Column picture</u>

$$x + 2y + z = 6 \;;\; 6x - 2y = 4 \;;\; -3x - y + 4z = 8$$

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

Solution for the system of equation is (1,1,3)

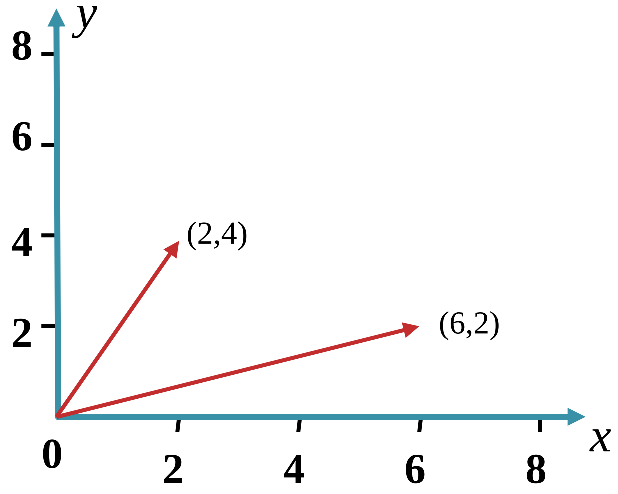


Vector Space

- Let V be a set of all vectors that lie in the first quadrant of R^2 and F be R

- Consider two vectors $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 6 \\ 2 \end{bmatrix} \in V$

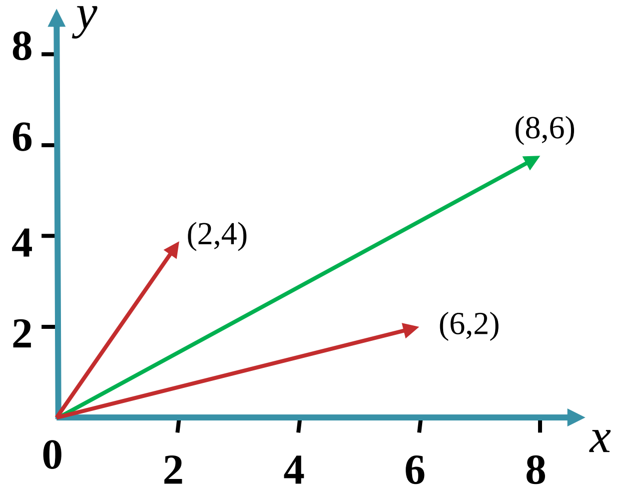
- It can easily be noted that both these vectors lie in the first quadrant of R^2



Vector Space

- Addition: $\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix} \in V$

- The resulting vector also lies in the first quadrant of R^2



Vector Space

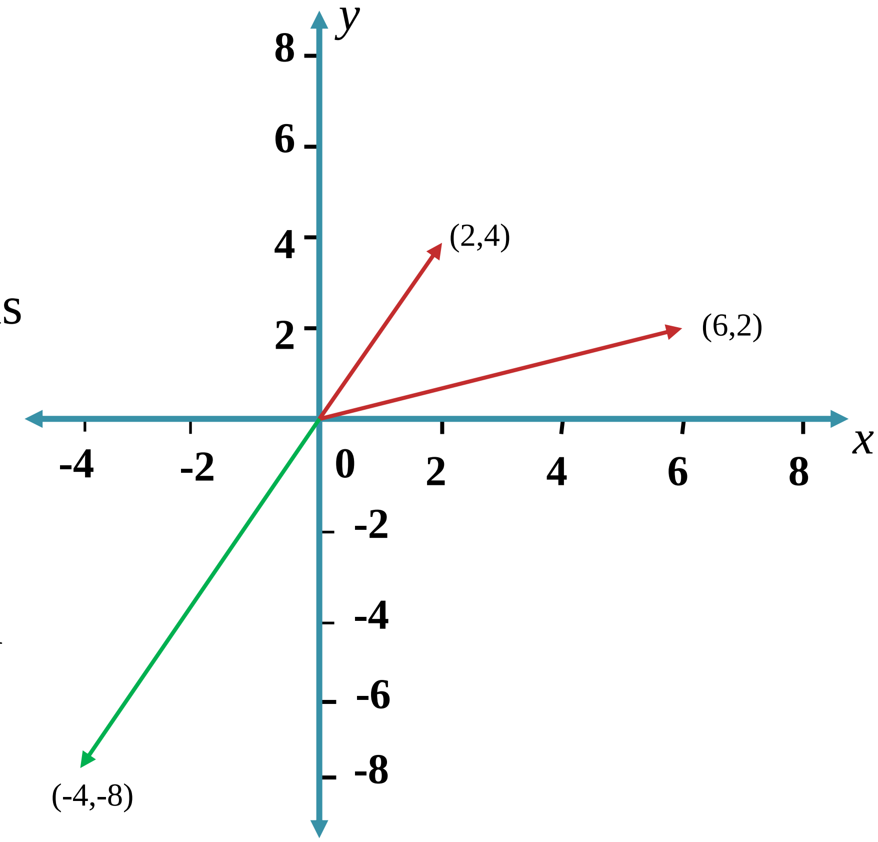
- Scalar multiplication: $a \in F$ where F is a field R.

  Consider a = -2

  Then, -2 × $\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ -8 \end{bmatrix}$

- It is clear that the resulting vector is outside V and hence it is not a vector space

- So, the first quadrant of R^2 is not a vector space whereas, R^2 is a vector space



Vector Space

- Let V be a set of vectors and F be a field. Then V is called a vector space over a field F if the following requirements are met

$$\forall x, y \in V, (x + y) \in V$$

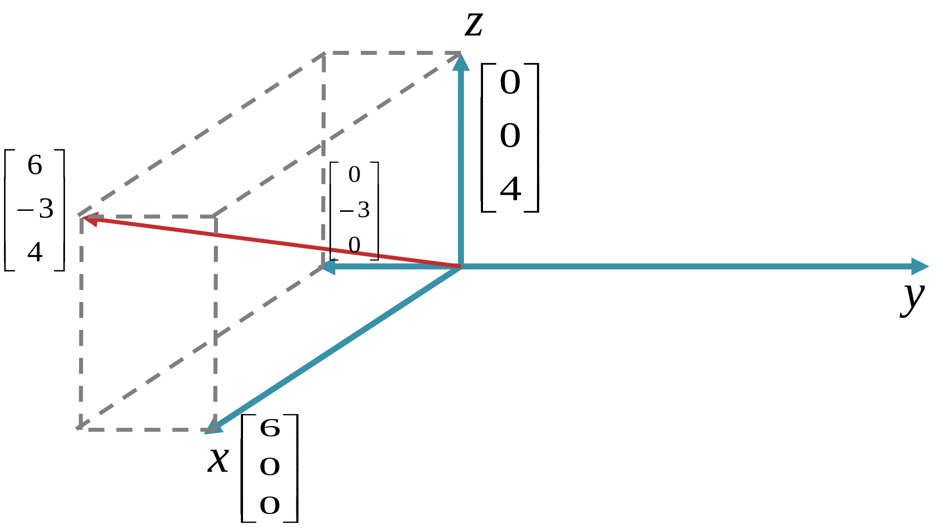
$$\forall x \in V \text{ and } \forall a \in F, a \times x \in V$$

- In other words, a set of vectors is called a vector space if the set is closed under vector addition and scalar multiplication of a vector

Linear Combination

- The vector $\begin{bmatrix} 6 \\ -3 \\ 4 \end{bmatrix}$ can be expressed as a linear combination of vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ as below

$$\begin{bmatrix} 6 \\ -3 \\ 4 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$



Linear Combination

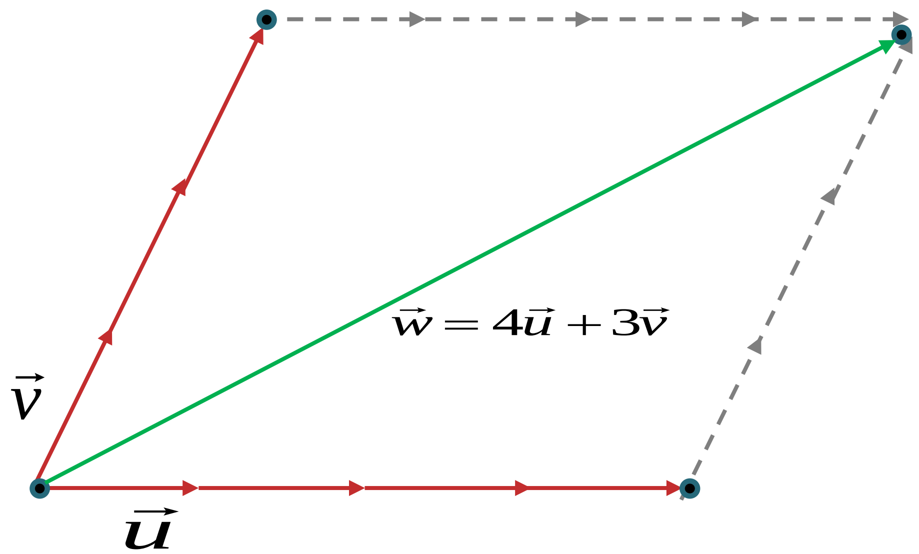
- A vector v can be written as a **linear combination** of vectors $u_1, u_2, u_3, \ldots, u_n$ such that

$$v = c_1 u_1 + c_2 u_2 + \ldots + c_n u_n$$

where $c_1, c_2, \ldots, c_n$ are all scalars

Linear Combination

- Geometrically, we may see the linear combination as –



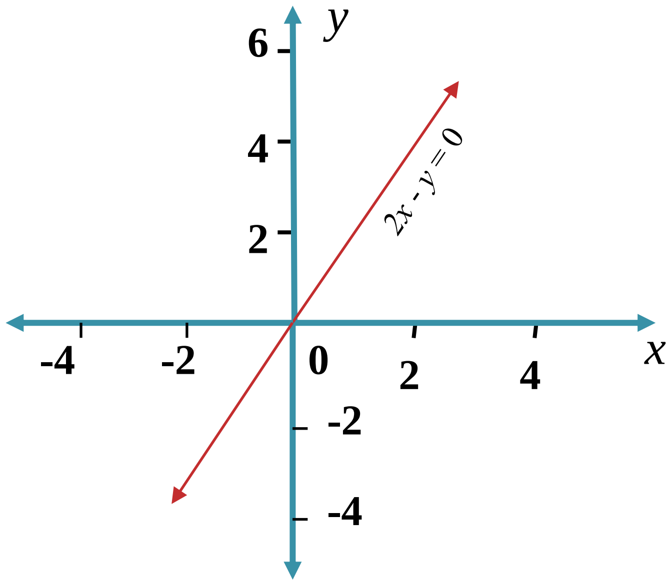
- We have the vector $\vec{w}$ which is expressed as a linear combination of the other two vectors $\vec{u}$ and $\vec{v}$

Subspace

- R^2 satisfies the conditions for a vector space – vector addition and scalar multiplication
- What about the line $2x - y = 0$? Does it satisfy the conditions?

  Answer: Yes

  Any line which passes through the origin in R^2 is a subspace of R^2



Subspace

- A subspace of a vector space is a nonempty subset that satisfies the requirements for a vector space –
  - Vector addition: For any vectors x, y in the subspace, $x + y$ is in the subspace
  - Scalar multiplication: For any scalar c, cx is in the subspace
- Column space: Contains all linear combinations of the columns of **A**
- Row space: Contains all linear combinations of the rows of **A**

Subspace

Is column space a subspace?

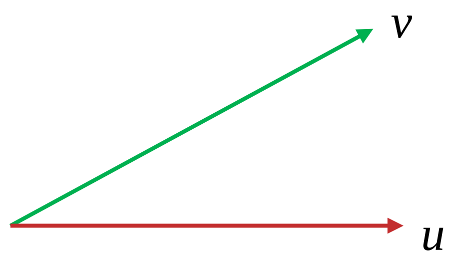
- Column space contains all linear combinations of the columns of **A**
- A = $\begin{bmatrix} 2 & 0 \\ 1 & 4 \\ 0 & 1 \end{bmatrix}$; Column space of A contains all linear combinations of

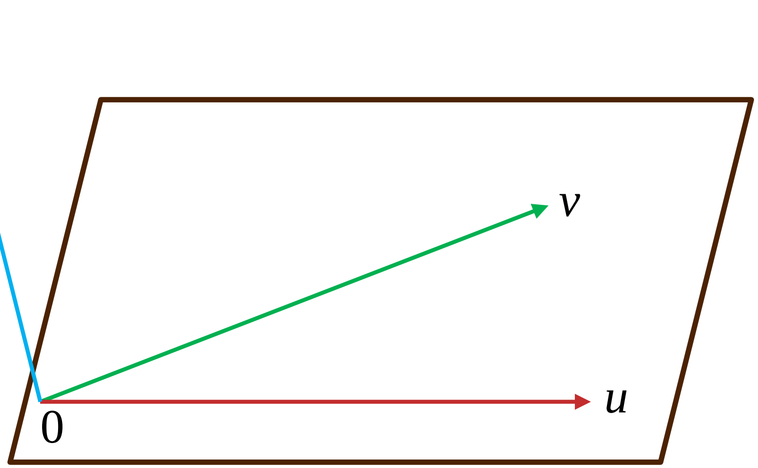
$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}$$

- If we take all the linear combinations of the columns in A, it will fill a plane in $R^3 \Rightarrow$ it is a subspace of R^3

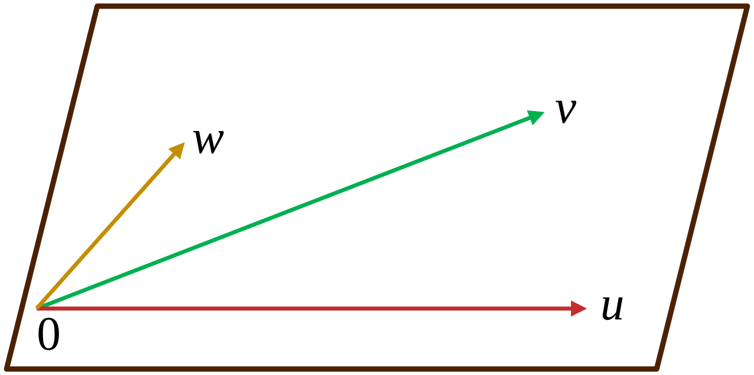
Linear Dependence & Independence

- Linearly independent:





- Linearly dependent:



Linear Independence

- A set of vectors $\{v_1, v_2, v_3, \dots, v_n\}$ is said to be **linearly independent** if $c_1 = c_2 = c_3 = \dots = c_n = 0$ is the only solution to the following equation

$$c_1v_1 + c_2v_2 + c_3v_3 + \dots + c_nv_n = 0$$

- Any one vector in the set cannot be expressed as a linear combinations of the rest of the vectors in the set

- If any other combination gives zero, they are **linearly dependent** and at least one of them is a linear combination of the others

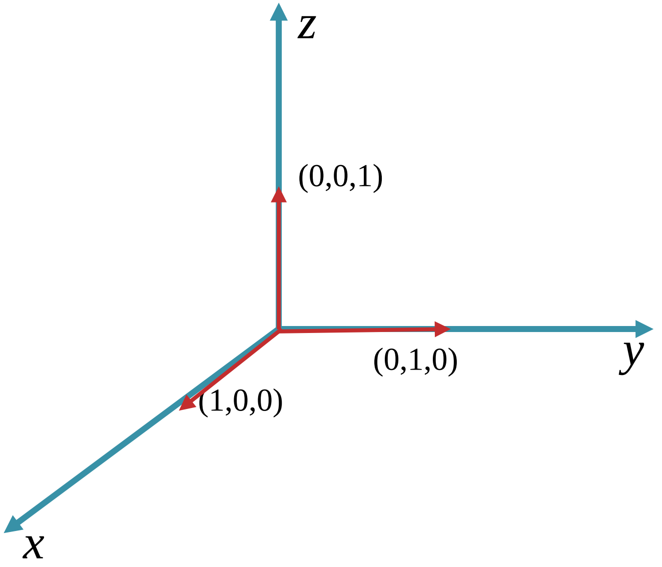
Linear Independence

Example:

- Consider the set $V = \{ [1,0,0]^T , [0,1,0]^T , [0,0,1]^T \}$. Now we represent the zero vector as

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The only values c_1, c_2 and c_3 can take in the above equation is zero. So, the set V is linearly independent



Linear Dependence & Independence

Example:

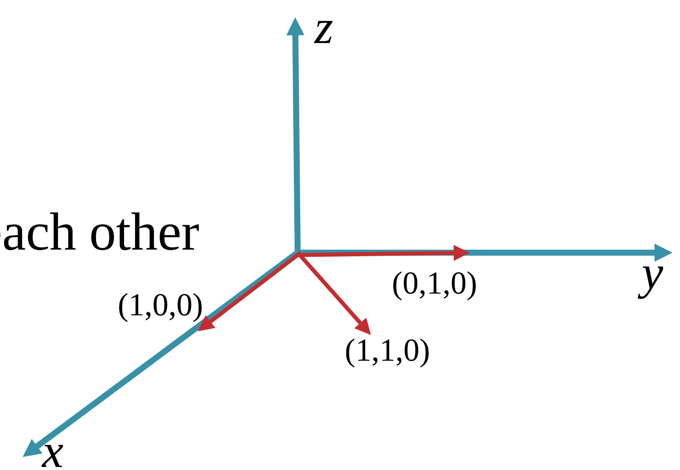
- Consider the set **U** = $\{[1,0,0]^T, [0,1,0]^T, [1,1,0]^T\}$

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Here, in addition to the solution $c_1 = c_2 = c_3 = 0$, there exists other solutions such as $c_1 = c_2 = -1$ and $c_3 = 1$. So they are linearly dependent on each other. Hence one can be expressed in terms of the rest. For instance, here,

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

Hence the vectors are linearly dependent on each other



Linear Dependence & Independence

- So the vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}$, $\begin{bmatrix}1\\0\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\\0\end{bmatrix}$ are linearly dependent
- The vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\end{bmatrix}$ are linearly dependent or independent?
- Answer: Linearly independent
- Zero is the only value that c_1 and c_2 could take in

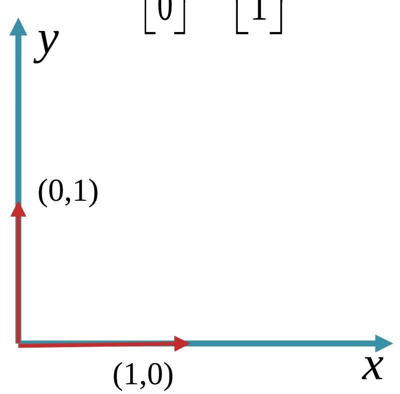
$$c_1\begin{bmatrix}1\\1\\0\end{bmatrix} + c_2\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Basis & Dimension of a Vector Space

- **Basis**: A basis for a vector space is a set of vectors having two properties at once:
  - It is linearly independent
  - It spans the space, i.e. a linear combination of the basis vectors can generate any other vector in the vector space
- In more general terms, a basis is a linearly independent spanning set
- Standard basis:

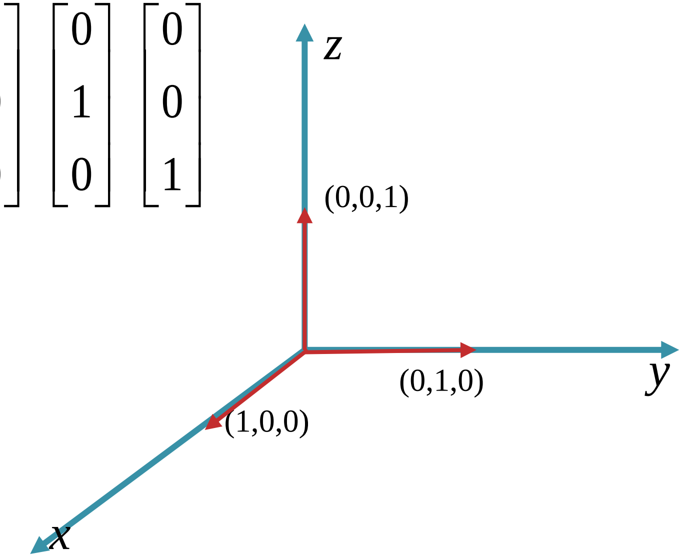
2D:

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



3D:

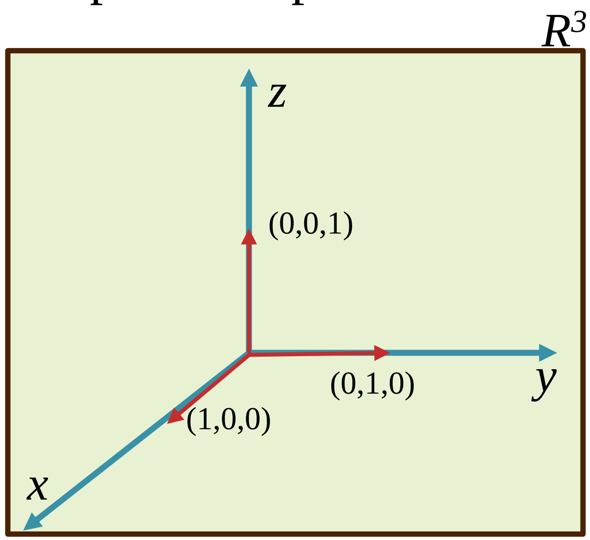
$$\begin{bmatrix}1\\0\\0\end{bmatrix} \quad \begin{bmatrix}0\\1\\0\end{bmatrix} \quad \begin{bmatrix}0\\0\\1\end{bmatrix}$$



Basis & Dimension of a Vector Space

- **Span:** Span of a vector set V refers to the vector space generated by all possible linear combinations of vectors present in V

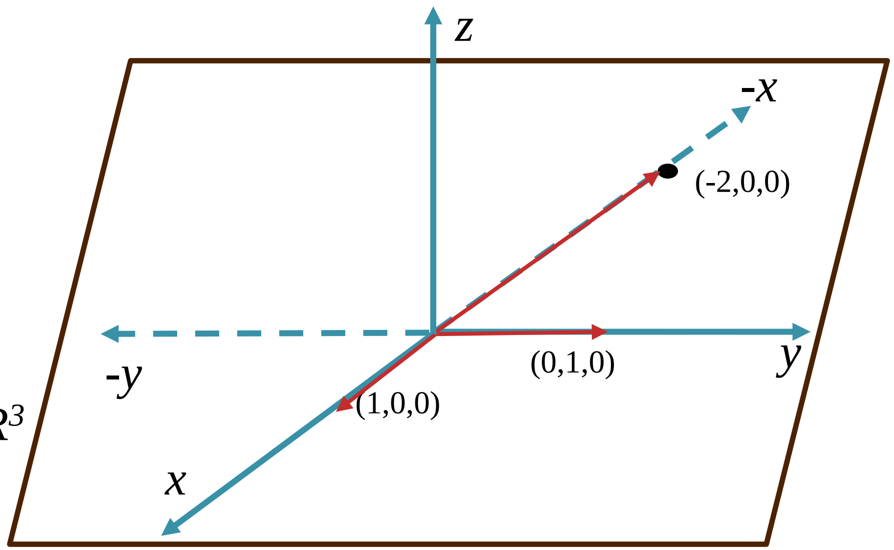
  Example: Consider the vector set $V = \{[1,0,0]^T, [0,1,0]^T, [0,0,1]^T\}$. The span of V is R^3 . This means that the possible linear combinations of the vectors of V fill the complete R^3 space

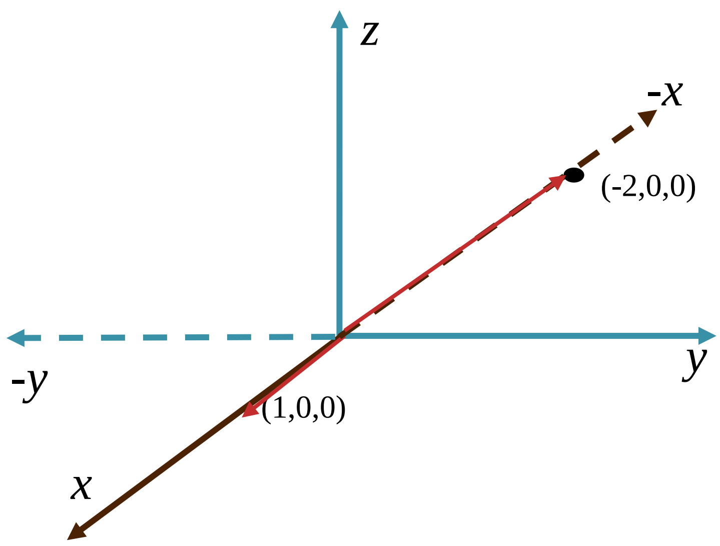


- **Dimension:** The maximum number of linearly independent vectors that can be obtained from a vector space is called the **dimension** of that vector space

Basis & Dimension of a Vector Space

- $V = \{(1,0,0)\ ,\ (0,1,0)\ ,\ (-2,0,0)\}$
- Spans a plane (the x–y plane) in R^3
- The vector (-2,0,0) and (1,0,0) are linearly dependent
- Hence, the vector set V doesn't form a basis in R^3
- ***Dimension = 2***

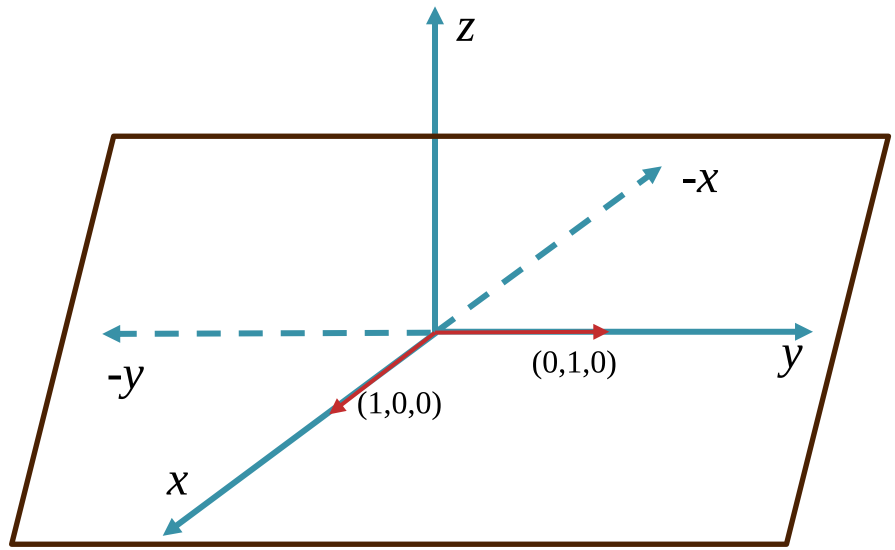


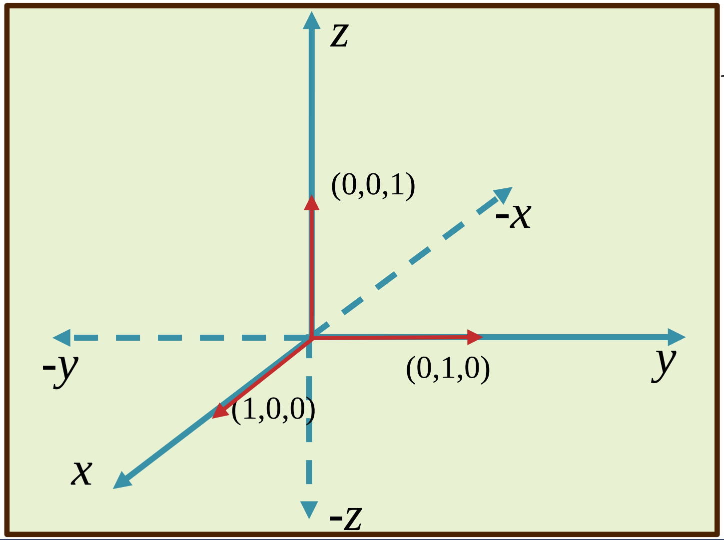


- $V = \{(1,0,0)\ ,\ (-2,0,0)\}$
- Spans only a line
- Linearly dependent
- Hence V doesn't form a basis in R^3
- ***Dimension = 1***

Basis & Dimension of a Vector Space

- $V = \{(1,0,0)\ ,\ (0,1,0)\}$
- Spans a plane (the x–y plane) in R^3
- Linearly independent
- Hence, V is a basis for the x–y plane
- ***Dimension = 2***





R^3

- $V = \{(1,0,0)\ ,\ (0,1,0)\ ,\ (0,0,1)\}$
- Spans the whole space R^3
- Linearly independent
- So, V is a basis for the space R^3
- ***Dimension = 3***

Rank of a Matrix

- Rank of a matrix refers to the number of linearly independent rows or columns of the matrix: Dimension of column space= Dimension of row space

- It can also be viewed as the number of pivots in Gaussian elimination process

Example:

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

Number of linearly independent rows = 2

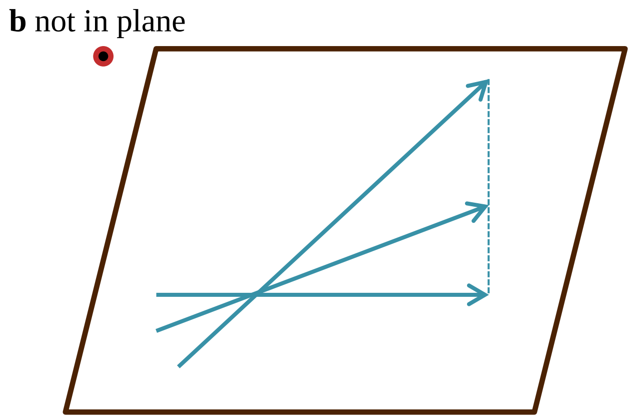
$\Rightarrow$ Rank of the matrix = 2

Existence and Uniqueness of a solution

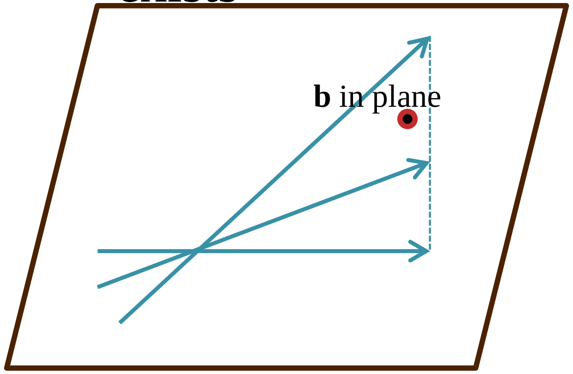
- **Ax=b**

- When does solution exist?

No solution



At least one solution exists

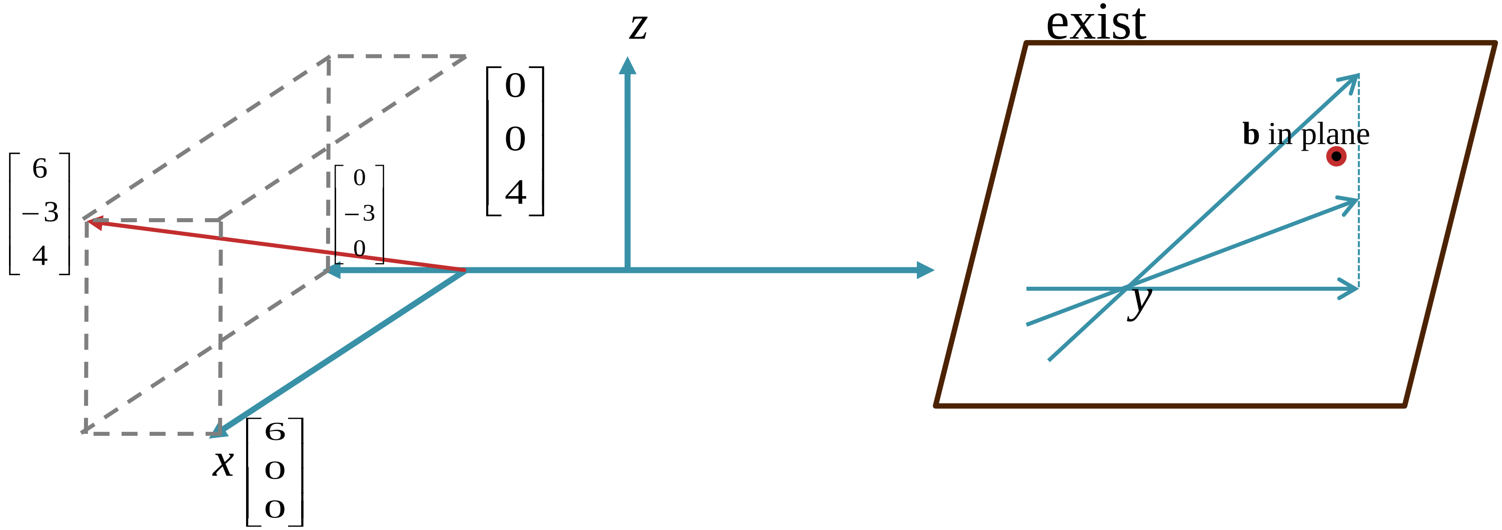


Existence and Uniqueness of a solution

- Uniqueness?

Unique solution

Infinite # solution exist



Null Space

- The null space of a matrix **A** consists of all vectors $\boldsymbol{x}$ such that $\mathbf{A}\boldsymbol{x} = \mathbf{0}$

- The set of solutions ($\boldsymbol{x}$'s) in $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ is itself a vector space which is called as a null space of **A**

$$N(\mathbf{A}) = \{\, x \in R^n \,/\, \mathbf{A}\boldsymbol{x} = \mathbf{0} \,\}$$

- If the column vectors are linearly independent, the null space contains only the zero vector
- The null space of an invertible matrix contains only zero vector

Null Space

To find null space for the matrix $A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 3 \end{bmatrix}$

- Null space: $N(\mathbf{A}) = \{x \in R^n \,/\, \mathbf{A}\boldsymbol{x} = \mathbf{0}\}$

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 2 & 2 & 3 & 0 \end{array}\right] \xrightarrow{R_3 = R_3 - R_1} \left[\begin{array}{cccc|c} 1 & 1 & 2 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

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

Null Space

- We have the equations:

$$x_1 + 2x_3 + x_4 = 0 \Rightarrow x_1 = -2x_3 - x_4$$

$$x_2 + x_4 = 0 \Rightarrow x_2 = -x_4$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

- $N(A)$ = scalar multiples of the vectors $\begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$

Null Space

- **If the vectors are linearly independent, the null space contains only the zero vector**
- The vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\end{bmatrix}$ are linearly independent
- What is the null space of $\begin{bmatrix}1 & 1\\1 & 0\\0 & 0\end{bmatrix}$?
- $\mathbf{A}\boldsymbol{x} = \mathbf{0} \Rightarrow \begin{bmatrix}1 & 1\\1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$ .What are the values of x_1 and x_2?

Null Space

- We have –

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- $x_1 = 0$ and $x_2 = 0$ is the only solution that the system of equations could take

$\Rightarrow$ If the vectors are linearly independent, the null space contains only the zero vector

Null Space

- **If the vectors are linearly independent, the null space contains only the zero vector. What about the dependent vectors?**

- Consider two dependent vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$

- What is the null space of $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$?

- $\mathbf{A}\boldsymbol{x} = \mathbf{0} \Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- x_1= ? and x_2= ?

Null Space

- We have –

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

- We have the equations –

$$x_1 + 2x_2 = 0 \Rightarrow x_1 = -2x_2$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

- We have $N(\mathrm{A})$ = scalar multiples of the vector $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$
- **If the vectors are linearly dependent, the null space also contains non-zero vector(s)**

Null Space: Cause of Non-Uniqueness

- Solutions of $\mathbf{A}x = \mathbf{b}$ with **A** having n columns.

- Solution does not exist if **b** does not belong to the column space of **A**

- At least one solution exists if **b** belongs to the column space of **A**

  - Solution unique if null space of **A** has only the zero vector

  - Infinite solutions if null space of **A** has non-zero vectors

Existence and Uniqueness of a solution

- **A***x* = **b** has no solution or is inconsistent iff
rank(**A**) < rank(augmented matrix [**A b**])
[**b** does not belong to column space of **A**]

- **A***x* = **b** has a unique solution iff
rank(**A**) = rank(augmented matrix [**A b**]) = *n*
[**b** belongs to column space of **A** and null space of **A** has only the zero vector]

- **A***x* = **b** has infinitely many solutions iff
rank(**A**) = rank(augmented matrix [**A b**]) < *n*
[**b** belongs to column space of **A** and null space of **A** has non-zero vectors as well]

Existence and Uniqueness of a solution

- Consider the system,

$$\begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \\ 3 \end{bmatrix}$$

- We will include the right hand side as an extra column to **A**. That matrix is called as an 'Augmented matrix' –

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 8 \\ 3 & 1 & 2 & 12 \\ 1 & 0 & 1 & 3 \end{array}\right]$$

- The system is inconsistent as the rank(**A**) < rank(augmented matrix)

Existence and Uniqueness of a solution

- Consider the system,

$$\begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 11 \\ 6 \end{bmatrix}$$

- The system has a unique solution as the rank(**A**) = rank(augmented matrix) = n

Existence and Uniqueness of a solution

- Consider the system,

$$\begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 11 \\ 3 \end{bmatrix}$$

- The system has infinitely many solutions as the rank(**A**) = rank(augmented matrix) < n

Determinants

- Determinant is the value associated with the *square matrix* (matrix with same number of row and columns)

- Used to determine whether a nxn matrix has rank n or not

- Equivalent to finding if inverse of a square matrix exists or not
- Used for computing eigenvalues as well

Determinants

How to calculate the determinant of a square matrix ?

Expand using the first row

i+j is even assign a positive sign

$$A = \begin{array}{cc} + & - \\ \end{array}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_{2\times 2}$$

i+j is odd assign a negative sign

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

A matrix is said to be <u>singular</u> if the determinant value is zero

Determinants

Expand using the first row

$$
\begin{array}{cc}
 & \begin{array}{ccc} + & - & + \end{array} \\
A = & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}_{3\times 3}
\end{array}
$$

$$
|A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}
$$

Determinants

- If the determinant of a matrix is zero, the matrix is singular
  - Rows or columns are dependent
  - Rank < n
- Singular matrices are not invertible
- If the determinant of a matrix is non-zero, the matrix is non-singular or invertible
  - Rows and columns are independent
  - Rank =n

Solving a System of Linear Equations

- For solving a system of n equations with n unknowns, there are two ways
  - Gaussian elimination –We eliminate unknowns by performing operations on it until there is only one equation and one unknown, which can be solved
  - Cramer's rule – We use determinants to solve the equations
- To solve large systems of equations, we prefer Gaussian elimination over Cramer's rule. This is because it is hard to find determinant for an $n \times n$ matrix with a larger n

Gaussian Elimination

- Consider a system of 3 equations

$$3x + y + 2z = 3 \;;\; 2y + z = 0 \;;\; 6x + y + 9z = -5$$

This can be written in a matrix form as –

$$\begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 1 \\ 6 & 1 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -5 \end{bmatrix}$$

- This is of the form **A*x* = b**. Now our goal is to eliminate x from the last two equations and y from the last equation

$$\begin{aligned} 3x + \quad y + 2z &= 3 \\ \cancel{0x} + 2y + 1z &= 0 \\ \cancel{6x} + \cancel{1y} + 9z &= -5 \end{aligned}$$

Gaussian Elimination

Augmented matrix

- We will include the right hand side as an extra column to **A**. That matrix is called as an 'Augmented matrix' –

$$\begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 1 \\ 6 & 1 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -5 \end{bmatrix} \Rightarrow \left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 6 & 1 & 9 & -5 \end{array}\right]$$

- For convenience, the augmented matrix is used to solve the equations since we don't have to carry over the x, y, z and = for every step

Gaussian Elimination

Manipulating rules

- There are three manipulating rules which may be applied for transforming an augmented matrix and also leaves the value of the solution set unchanged
  - Any two rows can be exchanged
  - Any row may be multiplied (or divided) by a nonzero constant
  - A multiple of any row can be added to any other row

Gaussian Elimination

Augmented matrix and pivots

- Augmented matrix: $\left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 6 & 1 & 9 & -5 \end{array}\right]$

- The coefficient of *x* in the 1st equation is known as the first pivot. (First nonzero number from left in a row is called a pivot)

- **Pivots should never be zero**
- What if they happen to be zero?
  - Exchange rows or columns so that there is no zero in the pivot position

- We have 1 as pivot (at the $\mathbf{A}_{11}$ position). We need to make all the values below the pivot as zero

Gaussian Elimination

Step by step process of elimination $3x + y + 2z = 3 \,;\, 2y + z = 0 \,;\, 6x + y + 9z = -5$

Step 1

Pivot $\left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 6 & 1 & 9 & -5 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & -1 & 5 & -11 \end{array}\right]$ $[R_3 = R_3 - 2R_1]$

Original system:	$3x + y + 2z = 3$ $2y + z = 0$ $6x + y + 9z = -5$	On 1st step of elimination $\Rightarrow$	$3x + y + 2z = 3$ $2y + z = 0$ $-y + 5z = -11$

- The value at the position $\mathbf{A_{21}}$ is already zero. Hence we may leave the second row unchanged
- Subtract two times the first row from the last row to get $\mathbf{A_{31}}$ as 0

Gaussian Elimination

Step 2

- Now, we have the second pivot to be 2 at the $\mathbf{A_{22}}$ position.
- At the second step, we need to make the -1 at $\mathbf{A_{32}}$ to be zero. So, add two times the third row with the second row

$$\left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & -1 & 5 & -11 \end{array}\right] \Longrightarrow \left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & 0 & \boxed{11} & -22 \end{array}\right] \quad [R_3 = 2R_3 + R_2]$$

$$\begin{aligned} 3x + y + 2z &= 3 \\ 2y + \ \ z &= 0 \\ -y + 5z &= -11 \end{aligned} \quad \xrightarrow{\text{On 2nd step of elimination}} \quad \begin{aligned} 3x + y + 2z &= 3 \\ 2y + \ \ z &= 0 \\ 11z &= -22 \end{aligned}$$

- Now, we have the third pivot to be 11 at the $\mathbf{A_{33}}$ position

Gaussian Elimination

<u>Step by step process of elimination</u> $3x + y + 2z = 3$; $2y + z = 0$; $6x + y + 9z = -5$

$$\left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 6 & 1 & 9 & -5 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & -1 & 5 & -11 \end{array}\right] \quad [R_3 = R_3 - 2R_1]$$

$$\Rightarrow \left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & 0 & \boxed{11} & -22 \end{array}\right] \quad [R_3 = 2R_3 + R_2]$$

- The forward elimination of the system of equations could be seen as –

Original system:

$$\begin{aligned} 3x + y + 2z &= 3 \\ 2y + z &= 0 \\ 6x + y + 9z &= -5 \end{aligned} \Rightarrow \begin{aligned} 3x + y + 2z &= 3 \\ 2y + z &= 0 \\ -y + 5z &= -11 \end{aligned} \Rightarrow \begin{aligned} 3x + y + 2z &= 3 \\ 2y + z &= 0 \\ 11z &= -22 \end{aligned}$$

Gaussian Elimination

<u>Forward elimination and back–substitution</u>

- On **Forward elimination**, we have three equations,

$$3x + y + 2z = 3$$
$$2y + z = 0$$
$$11z = -22 \Rightarrow z = -2$$

- Now, we perform **Back–substitution** to obtain the solution
- Substitute $z = -2$ in the second equation

$$2y + z = 0 \Rightarrow y = 1$$

- Substitute $y = 1$ and $z = -2$ in the first equation

$$3x + y + 2z = 3 \Rightarrow x = 2$$

- Hence the solution for the given set of linear equations is $x = 2$, $y = 1$ and $z = -2$

Reduced Row Echelon Form (rref)

- We obtained the below matrix by Gaussian elimination –

$$\left[\begin{array}{ccc|c} \boxed{3} & 1 & 2 & 3 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & 0 & \boxed{11} & -22 \end{array}\right]$$

- The reduced row echelon form further reduces the matrix by making the pivots 1 and making the elements even above the pivots to be zero

- It gives the solution (***x***) directly without performing forward elimination or back–**s**ubstitution

Reduced Row Echelon Form (rref)

- So, the rref of the matrix $\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 11 & -22 \end{array}\right]$ can be calculated as follows:

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 11 & -22 \end{array}\right] \xrightarrow[\substack{R_1 = R_1 / 3 \\ R_2 = R_2 / 2 \\ R_3 = R_3 / 11}]{} \left[\begin{array}{ccc|c} 1 & 1/3 & 2/3 & 1 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right] \xrightarrow[R_2 = R_2 - (1/2)R_3]{} \left[\begin{array}{ccc|c} 1 & 1/3 & 2/3 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

$$\xrightarrow[R_1 = R_1 - (2/3)R_3]{} \left[\begin{array}{ccc|c} 1 & 1/3 & 0 & 7/3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right] \xrightarrow[R_1 = R_1 - (1/3)R_2]{} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

- Thus the solution for the given set of linear equations is $x = 2$, $y = 1$ and $z = -2$

Gaussian Elimination

When could the process break down?

- ***Zero in the pivot position***
- Pivots should not be zero. Note that a zero can appear in a pivot position, even if the original coefficient in that place was not zero
- This could be cured by row exchanges

Example: Non singular case:

$$\begin{aligned} 3x + y + 2z &= \_ \\ 6x + 2y + 9z &= \_ \\ 2y + z &= \_ \end{aligned} \xrightarrow{(R_2 = R_2 - 2R_1)} \begin{aligned} 3x + y + 2z &= \_ \\ 5z &= \_ \\ 2y + z &= \_ \end{aligned} \xrightarrow{\text{(On exchanging rows 2 and 3)}} \begin{aligned} 3x + y + 2z &= \_ \\ 2y + z &= \_ \\ 5z &= \_ \end{aligned}$$

- Now the system could be solved by **back–substitution**

Gaussian Elimination

Example: Singular case:

$$\begin{aligned} 3x + y + 2z &= \_ \\ 6x + 2y + 9z &= \_ \\ 3x + y + z &= \_ \end{aligned} \xrightarrow[\substack{(R_2 = R_2 - 2R_1 \\ \text{and} \\ R_3 = R_3 - R_1)}]{} \begin{aligned} 3x + y + 2z &= \_ \\ 5z &= \_ \\ -z &= \_ \end{aligned}$$

- No exchange of equations could be done here to avoid zero in the pivot position
- If we have the last two equations to be $5z = 10$ and $-z = 3$, there is no solution
- If we have $5z = 10$ and $-z = -2$, then we have infinity number of solutions. (Since we have $z = 2$ but the first equation cannot decide both x and y)

Infinite Solutions: Null space use

- Suppose we solve:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

- How many solutions?
- Infinite solutions : Rank(A) = Rank(A|b) < 2
- How do we characterize these infinite solutions?

Infinite Solutions: Null space use (2)

- Gaussian elimination leads to:

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

- Infinite solutions characterized as:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Inverses – Gauss–Jordan Method

<u>To find inverse of A</u>

- Consider a matrix $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 2 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] (R_2 \leftrightarrow R_1)$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 2 & 1 & 0 & 1 & 0 \\ 0 & -3 & -2 & 1 & -2 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] (R_2 = R_2 - 2R_1)$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 2 & 1 & 0 & 1 & 0 \\ 0 & \boxed{1} & 2 & 0 & 0 & 1 \\ 0 & -3 & -2 & 1 & -2 & 0 \end{array}\right] (R_2 \leftrightarrow R_3)$$

Inverses – Gauss–Jordan Method

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 0 & -3 & 0 & 1 & -2 \\ 0 & \boxed{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 4 & 1 & -2 & 3 \end{array}\right] \quad \begin{array}{l}(R_1 = R_1 - 2R_2\ ; \\ R_3 = R_3 + 3R_2)\end{array}$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 0 & -3 & 0 & 1 & -2 \\ 0 & \boxed{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & \boxed{1} & 1/4 & -1/2 & 3/4 \end{array}\right] \quad (R_3 = R_3/4)$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} \boxed{1} & 0 & 0 & 3/4 & -1/2 & 1/4 \\ 0 & \boxed{1} & 0 & -1/2 & 1 & -1/2 \\ 0 & 0 & \boxed{1} & 1/4 & -1/2 & 3/4 \end{array}\right] \quad \begin{array}{l}(R_2 = R_2 - 2R_3\ ; \\ R_1 = R_1 + 3R_3)\end{array}$$

Inverse of $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$ is $\begin{bmatrix} 3/4 & -1/2 & 1/4 \\ -1/2 & 1 & -1/2 \\ 1/4 & -1/2 & 3/4 \end{bmatrix}$

Inverses – Gauss–Jordan Method

- Inverse of $\mathbf{A}$ is denoted as $\mathbf{A}^{-1}$ and when $\mathbf{A}$ is multiplied with its inverse, it produces an identity matrix : $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

- We have $\mathbf{A}\boldsymbol{x} = \mathbf{b}$,

  multiply by $\mathbf{A}^{-1} \Rightarrow \boldsymbol{x} = \mathbf{A}^{-1}\mathbf{b}$

  $\Rightarrow$ we may get the solution $\boldsymbol{x}$ directly by multiplying $\mathbf{b}$ with $\mathbf{A}^{-1}$

- Not all matrices can have inverses

  - Singular matrices cannot have an inverse i.e. if its determinant is zero, the matrix cannot have an inverse
  - If $\mathbf{A}\boldsymbol{x} = 0$ and $\boldsymbol{x} \neq 0$, $\mathbf{A}$ cannot have an inverse

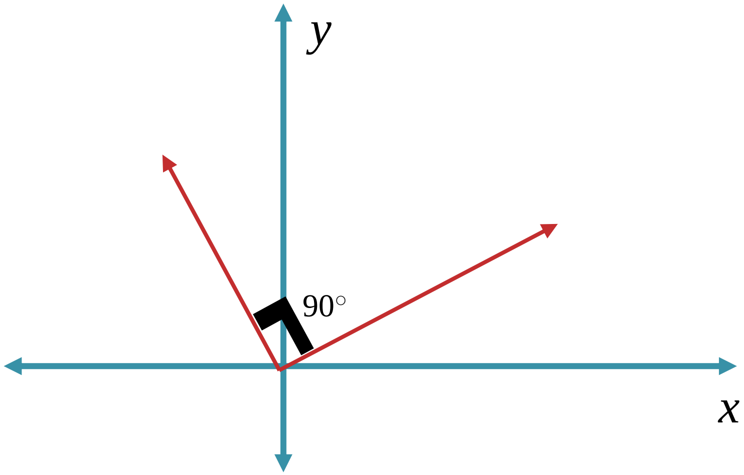
Orthogonality

- A basis is a linearly independent spanning set. Geometrically, it is a set of coordinate axes
- We need a basis to convert geometric constructions into algebraic calculations and we need an orthonormal basis to make those calculations simple
- To achieve that, we need to know–
  - length of a vector
  - test for orthogonality
  - how to create perpendicular vectors from linearly independent vectors

Orthogonality

- Two vectors are orthogonal if their inner product $u^T v$ is zero. It means that the angle between the two vectors is 90°

- If $u^T v > 0$, their angle is less than 90° and if $u^T v < 0$, their angle is greater than 90°



Orthogonality

- Consider two vectors u = [u_1,u_2,…,u_n] and v = [v_1,v_2,…,v_n]

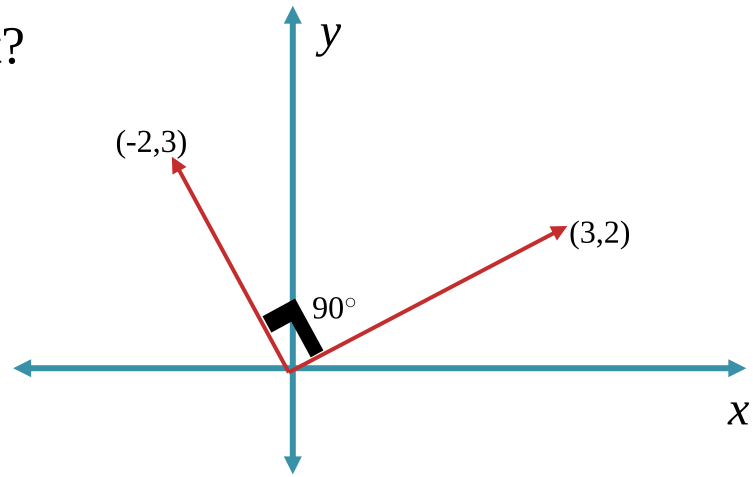
$$u^T v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

If $u^T v = 0$, then the two vectors are said to be orthogonal to each other

Orthogonality

Example:

The vectors $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ are orthogonal or not?



- Test for orthogonality:

$u^T v = 0$

$u^T v = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

$= -6 + 6 = 0$

and hence the vectors are orthogonal to each other

Orthogonality

- Length of a vector: $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
- Length of a vector (1,2) is $\sqrt{(1^2 + 2^2)} = \sqrt{5}$
- Two vectors which are orthogonal and of length 1 are said to be **orthonormal**
- Orthogonality always implies linear independence but not vice–versa
- Zero vector is orthogonal to all vectors

Orthogonality

- **Orthogonality always implies linear independence but not vice–versa**

- Orthogonality implies linear independence?

- The vectors $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ are orthogonal. Are they linearly independent?

$$c_1 \begin{bmatrix} 3 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- The only values that c_1 and c_2 could take is 0 and hence they are linearly independent

Orthogonality

- Linear independence implies orthogonality always?
- The vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\end{bmatrix}$ are linearly independent. Are they orthogonal?
- Test for orthogonality: $x^T y = 0$

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

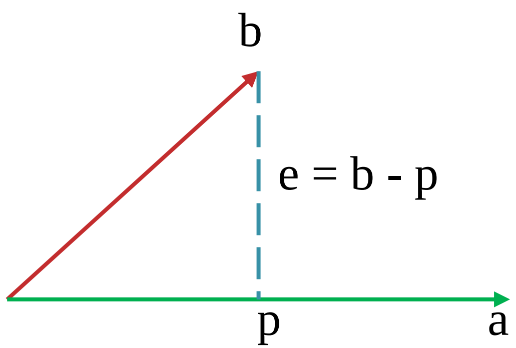
- The vectors are linearly independent but not orthogonal

$\Rightarrow$ Orthogonality always implies linear independence but not vice–versa

Projection

- Consider 2 vectors *a* and *b*. Project *b* on *a* and *p* is the point on *a* which is closest to *b*. The point *p* is called as the projection of *b* onto line through *a*

- Geometrical fact: The line from b to the closest point $p = \hat{x}a$ is perpendicular to the vector *a*



$$p = \hat{x}a \quad ; \quad \hat{x} = ? \; ; \; e = b - \hat{x}a$$

$$(b - \hat{x}a) \perp a$$

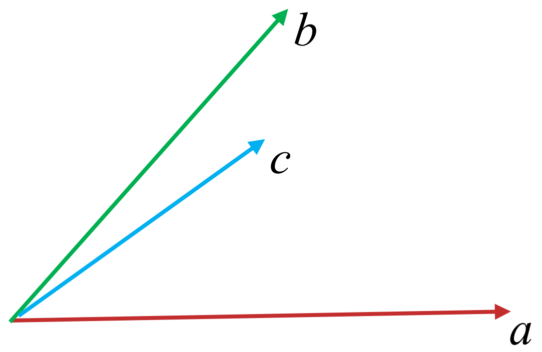
$$\Rightarrow a^T(b - \hat{x}a) = 0$$

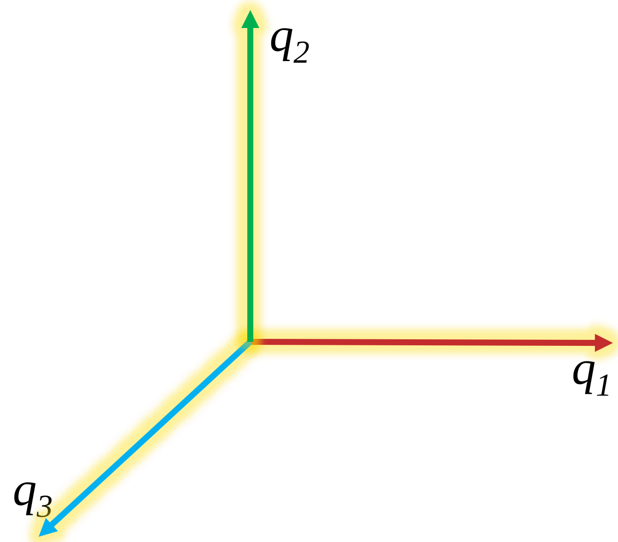
$$\hat{x} = \frac{a^T b}{a^T a}$$

$$p = \hat{x}a = \frac{a^T b}{a^T a} a$$

Gram–Schmidt Orthogonalization

Goal: To make the vectors a, b, c into new glowing vectors q_1, q_2, q_3





Gram–Schmidt Orthogonalization

- In an orthogonal basis, every vector is perpendicular to every other vector. The coordinate axes are mutually orthogonal

- The vectors $q_1, \ldots, q_n$ are orthonormal if:

$$q_i^T q_j = \begin{cases} 0 \text{ whenever } i \neq j \text{, giving the orthogonality;} \\ 1 \text{ whenever } i = j \text{, giving the normalization} \end{cases}$$

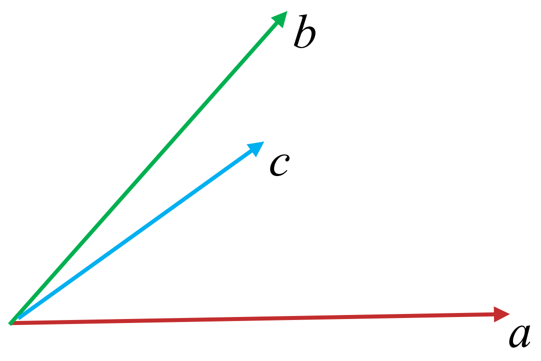
- To change an orthogonal basis into orthonormal basis, we may simply divide each vector by its length which will make it a unit vector

Gram–Schmidt Orthogonalization

The Gram–Schmidt process:

- Consider 3 independent vectors a, b, c and we seek orthonormal vectors q_1, q_2, q_3

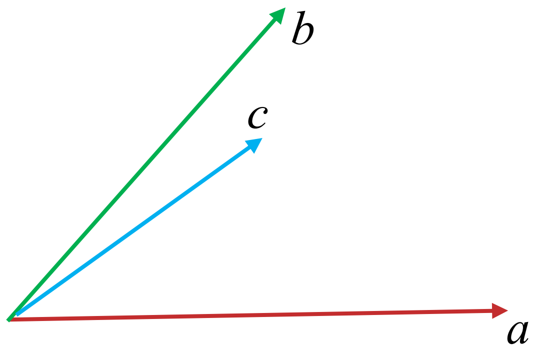
  i.e. to make the vectors a, b, c perpendicular to each other and also their length has to be 1



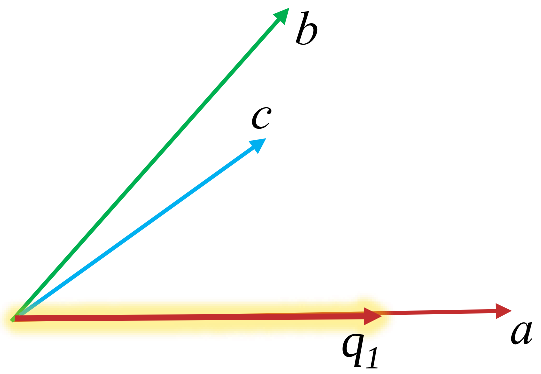
The vector c is in a different plane

Gram–Schmidt Orthogonalization

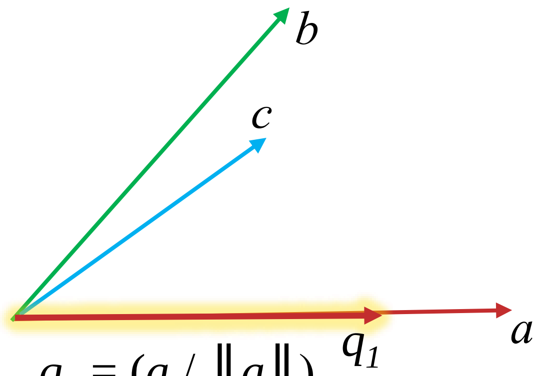
Gram–Schmidt Orthogonalization



Gram–Schmidt Orthogonalization

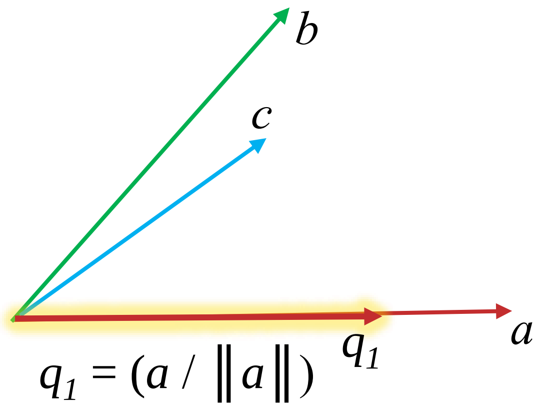


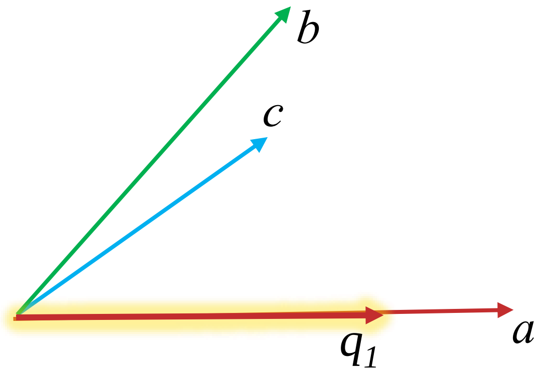
Gram–Schmidt Orthogonalization



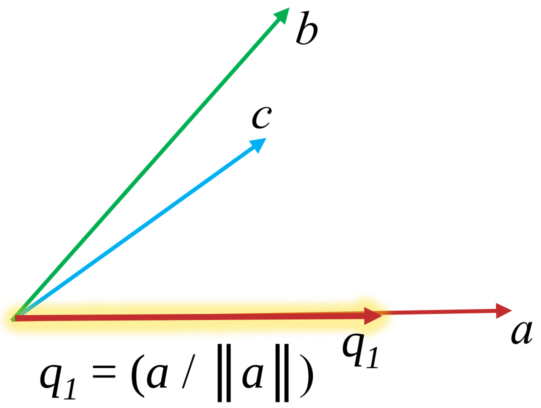
$$q_1 = (a / \|a\|)$$

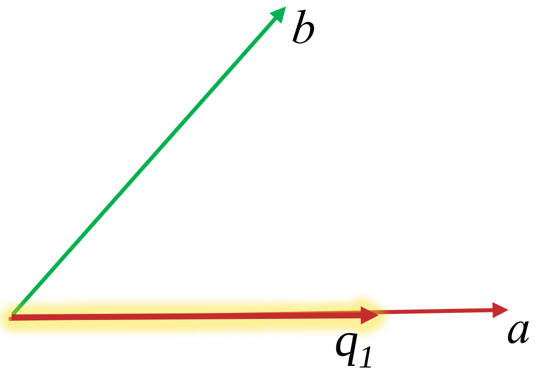
Gram–Schmidt Orthogonalization



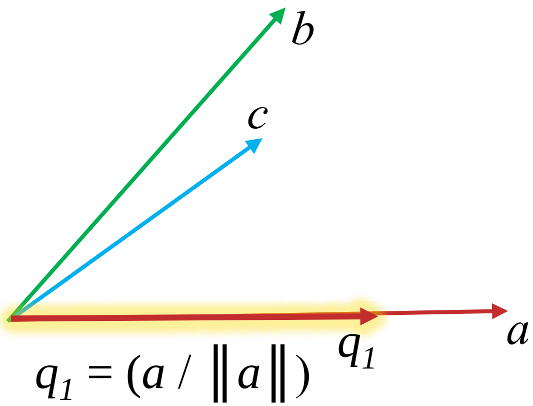


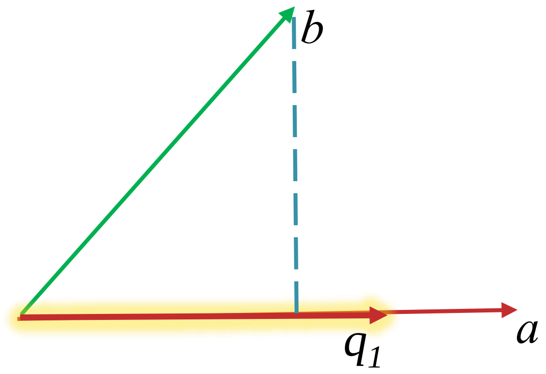
Gram–Schmidt Orthogonalization



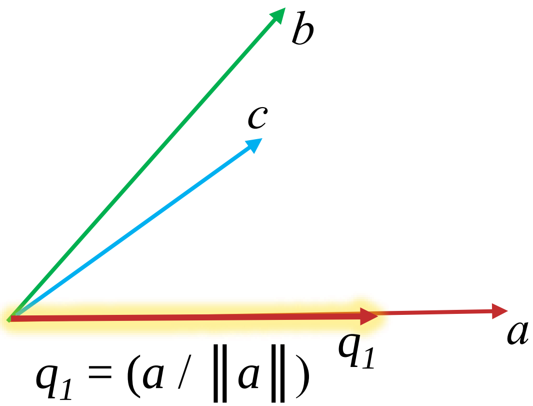


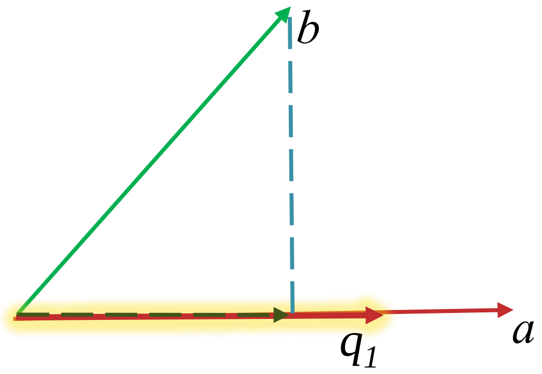
Gram–Schmidt Orthogonalization



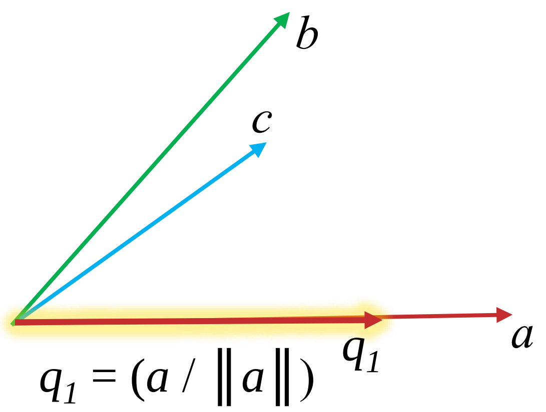


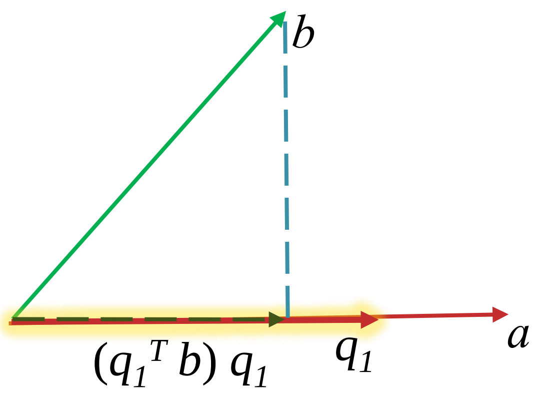
Gram–Schmidt Orthogonalization



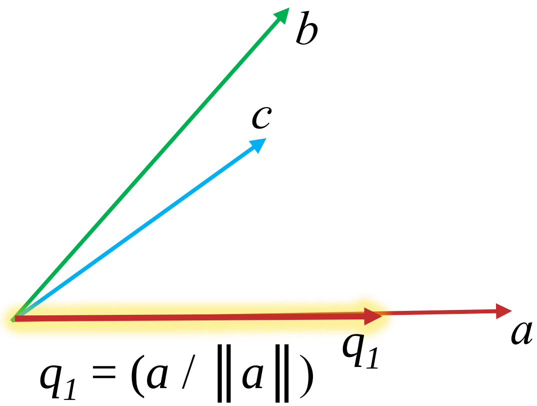


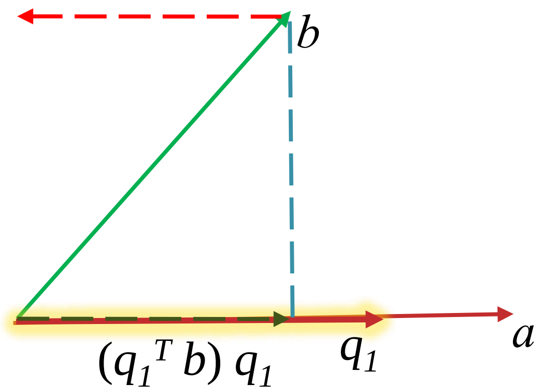
Gram–Schmidt Orthogonalization



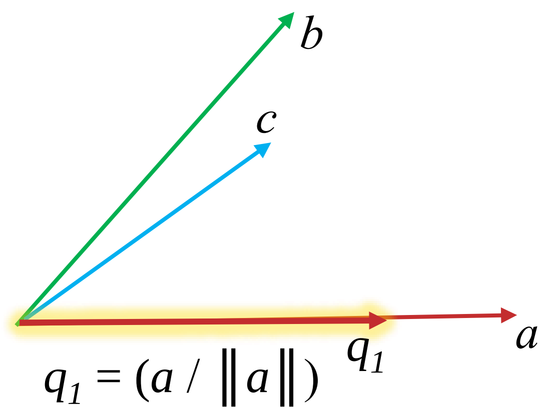


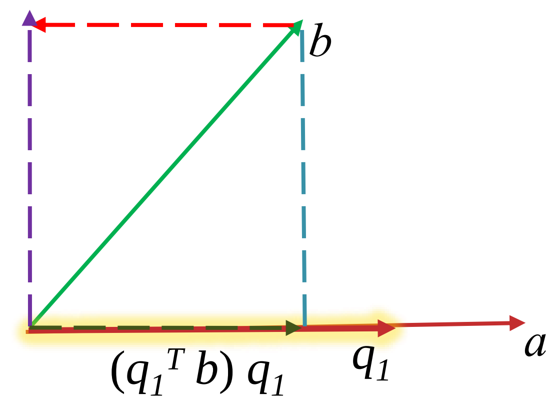
Gram–Schmidt Orthogonalization



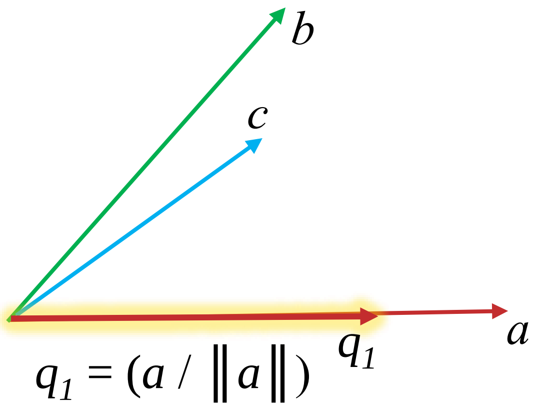


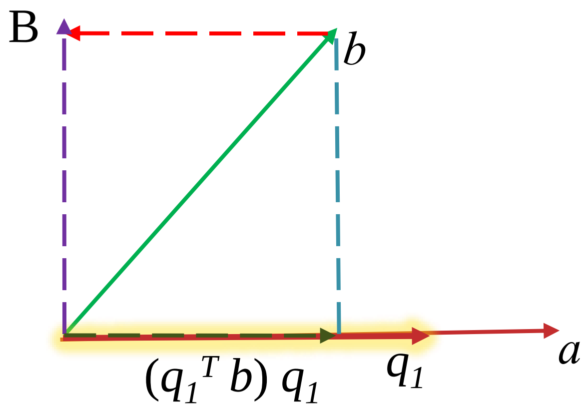
Gram–Schmidt Orthogonalization



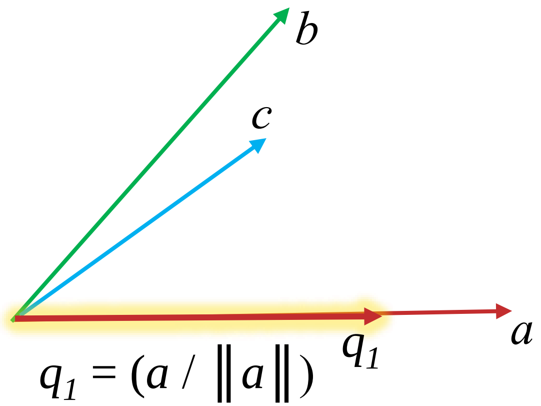


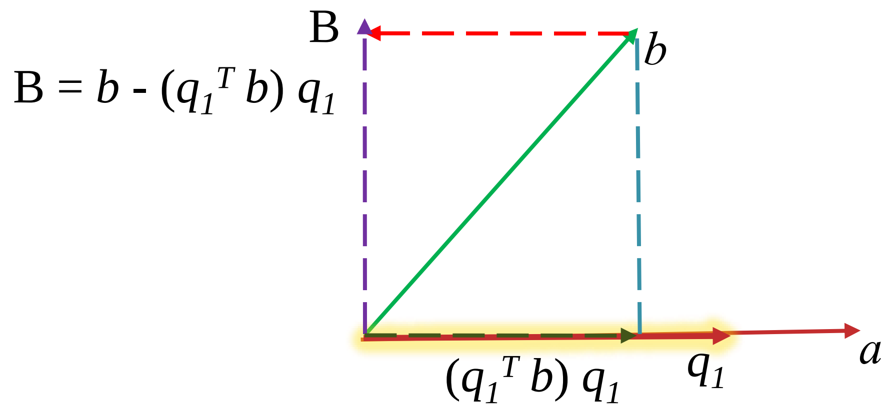
Gram–Schmidt Orthogonalization



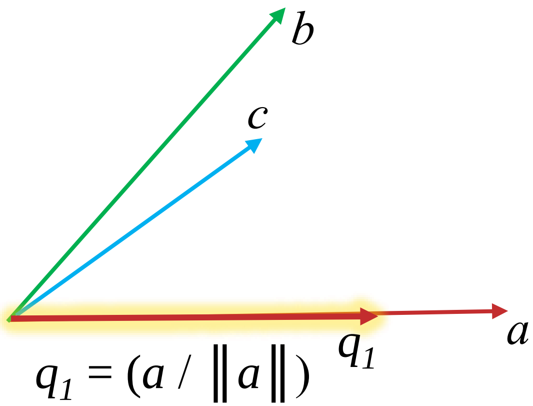


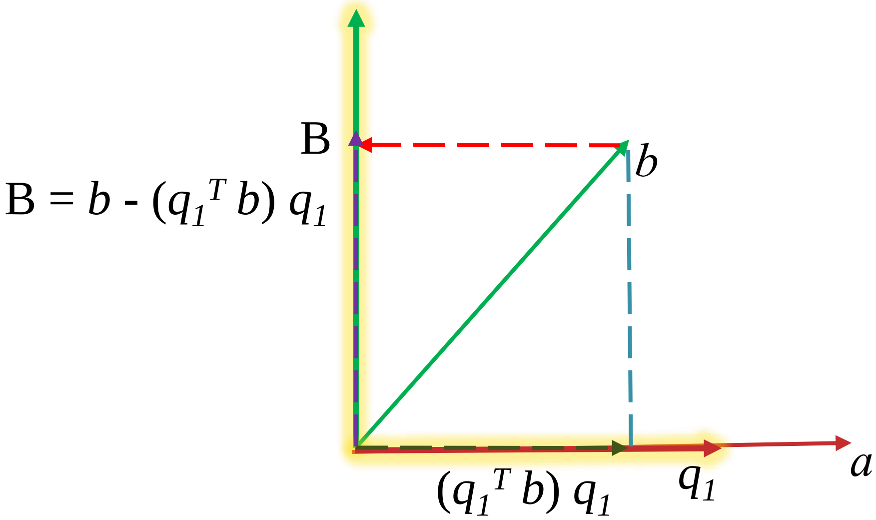
Gram–Schmidt Orthogonalization



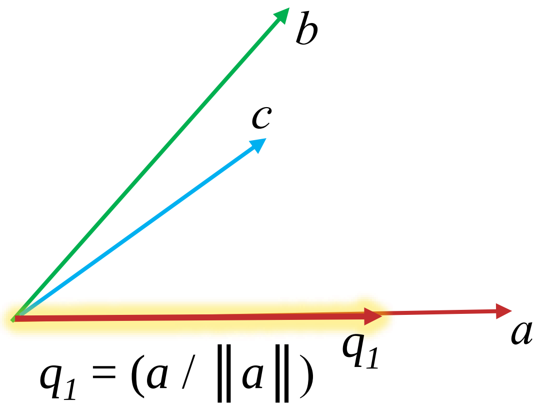


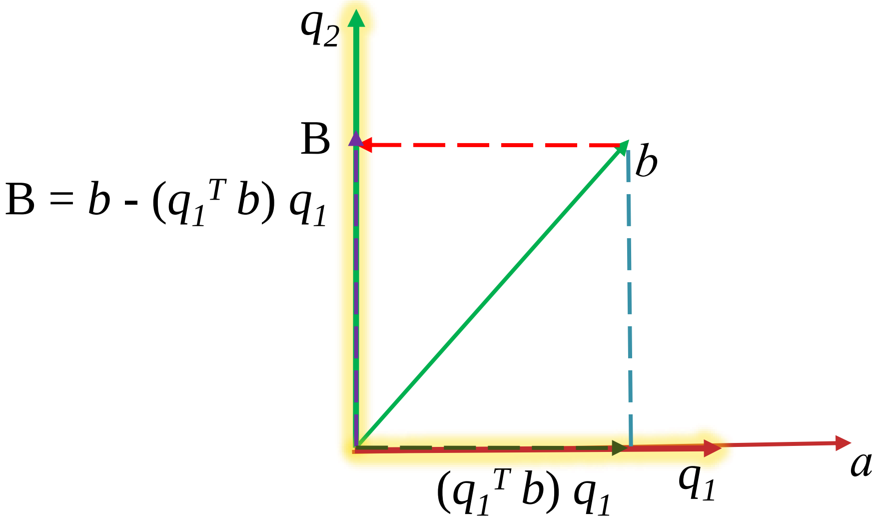
Gram–Schmidt Orthogonalization



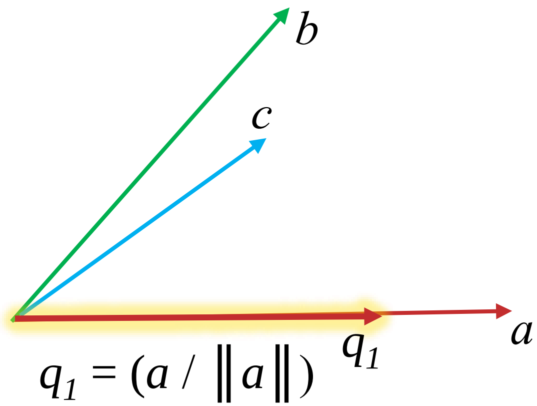


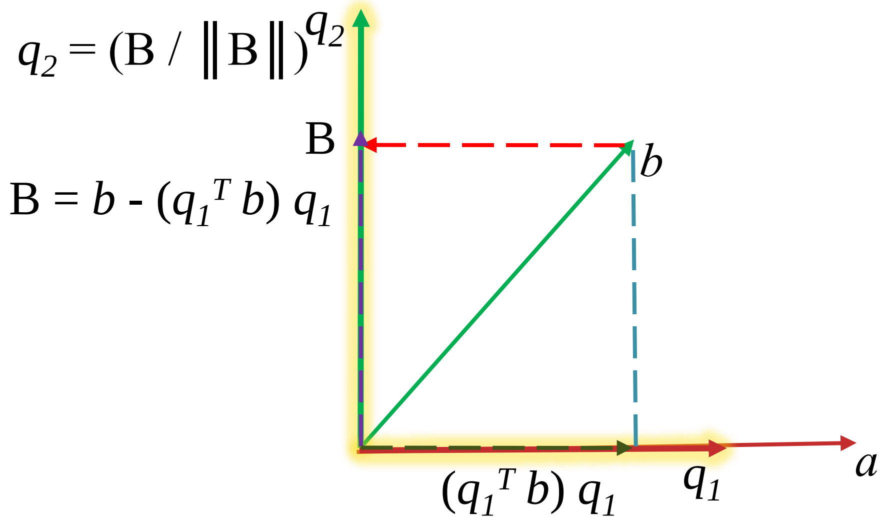
Gram–Schmidt Orthogonalization



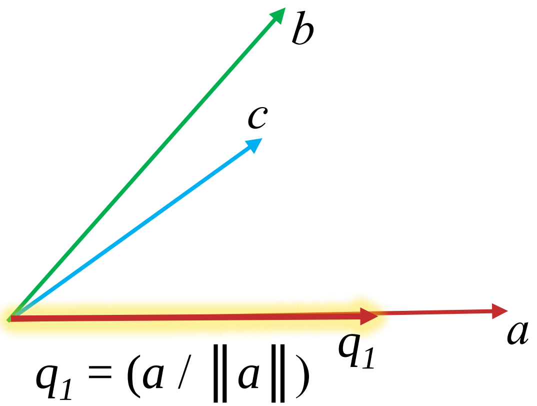


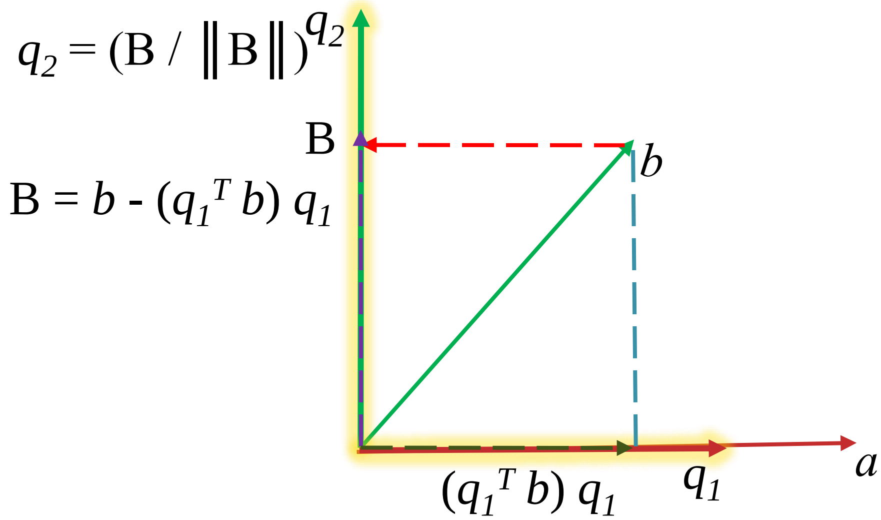
Gram–Schmidt Orthogonalization

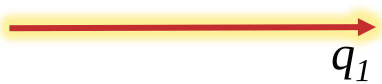




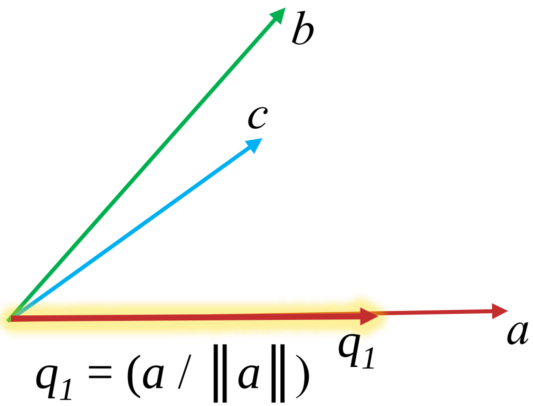
Gram–Schmidt Orthogonalization

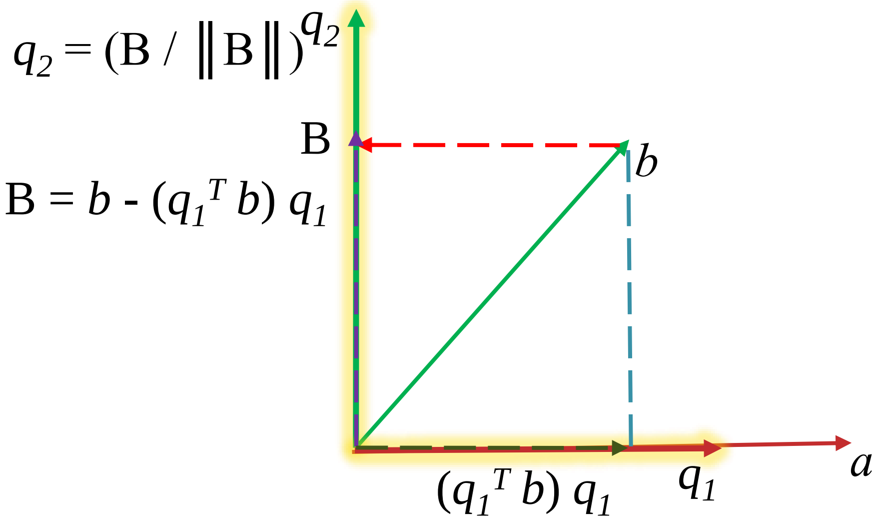


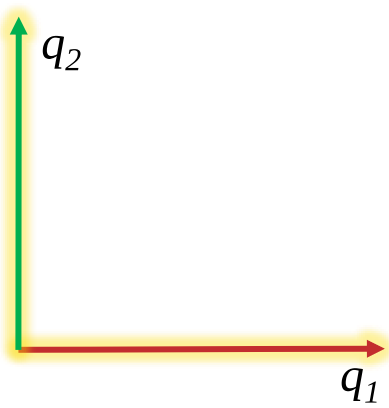




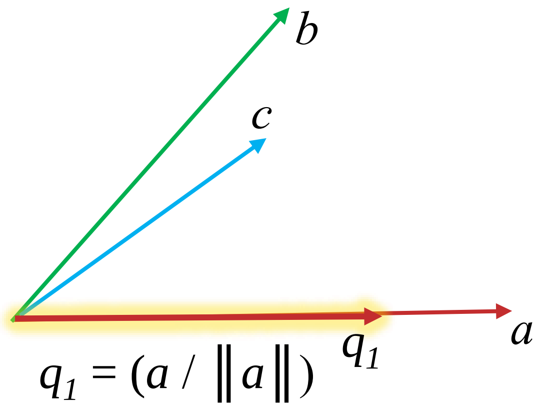
Gram–Schmidt Orthogonalization

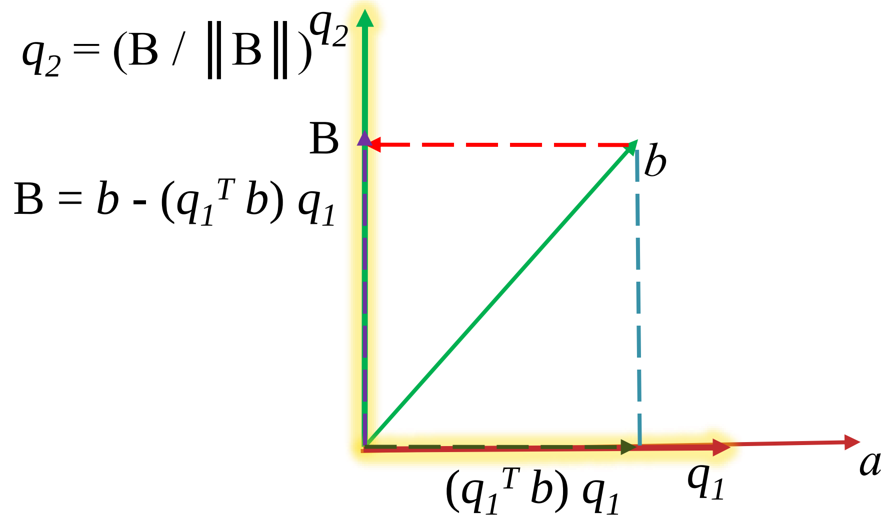


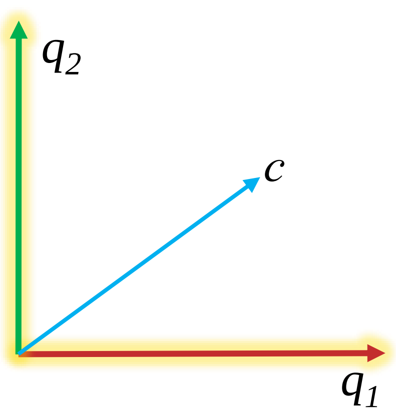




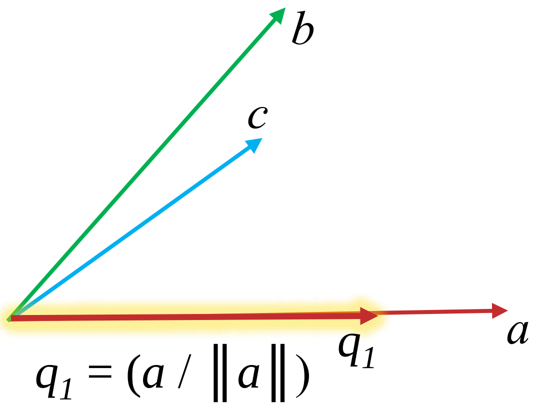
Gram–Schmidt Orthogonalization

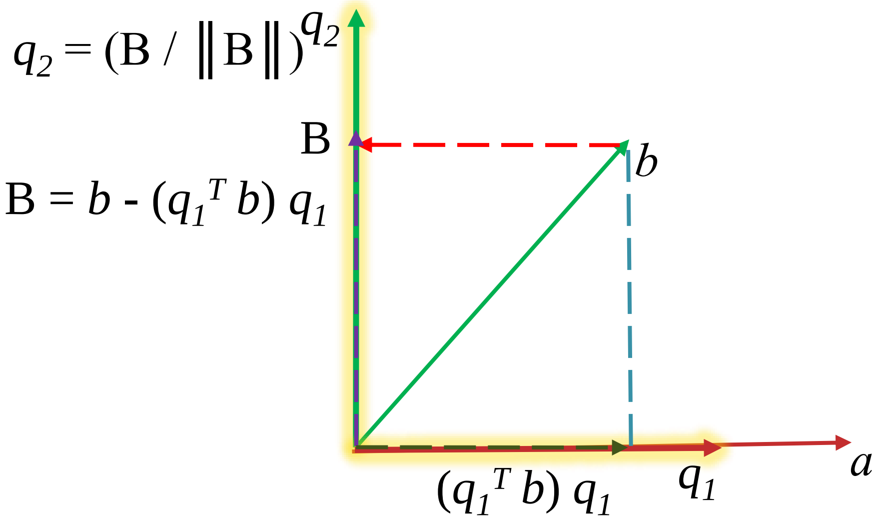


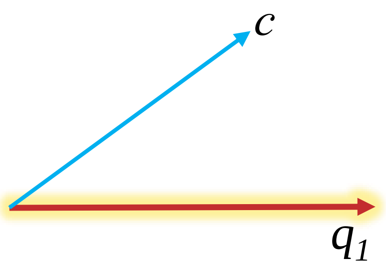




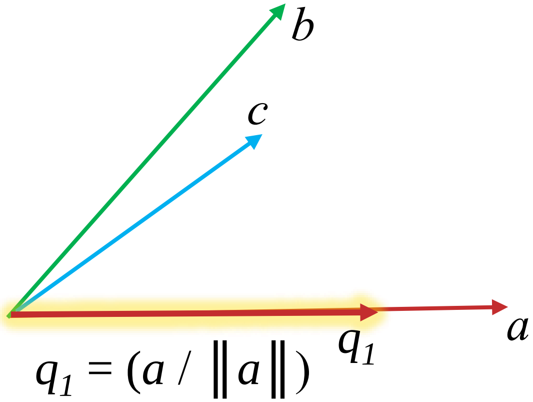
Gram–Schmidt Orthogonalization

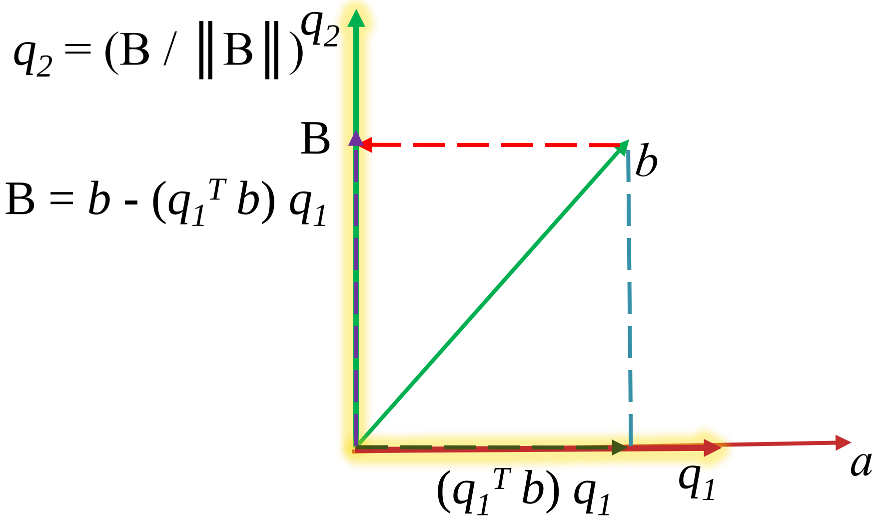


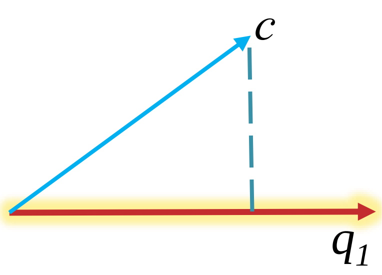




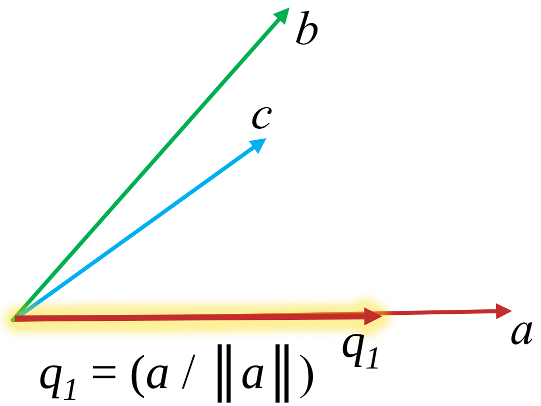
Gram–Schmidt Orthogonalization

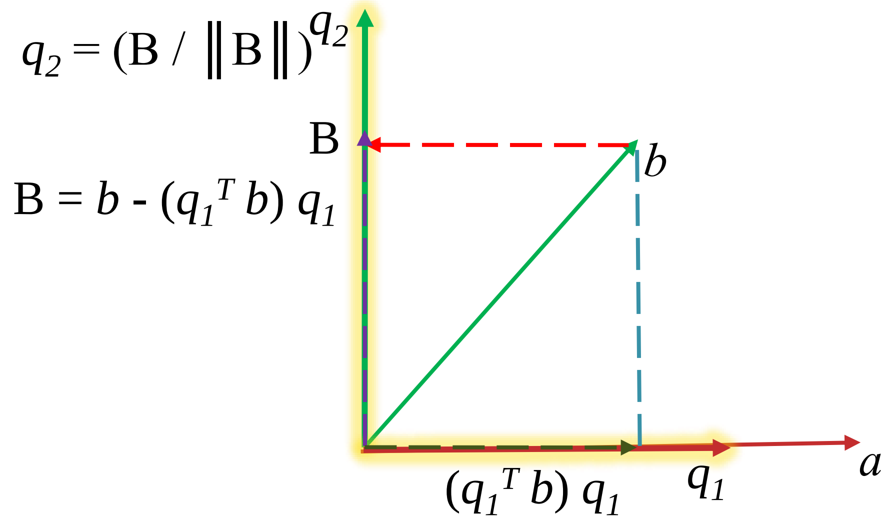


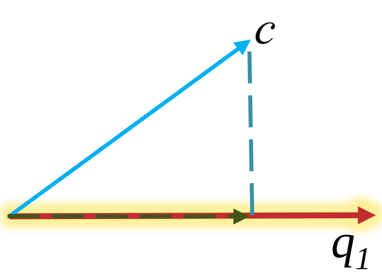




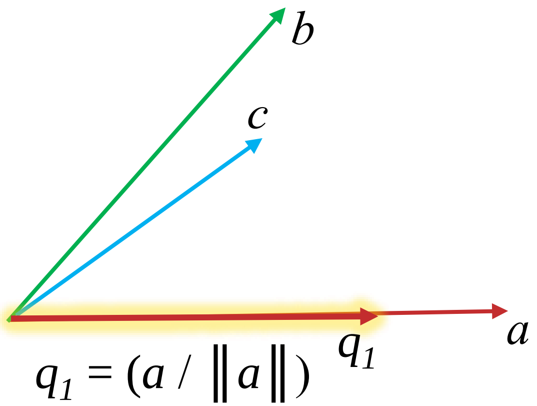
Gram–Schmidt Orthogonalization

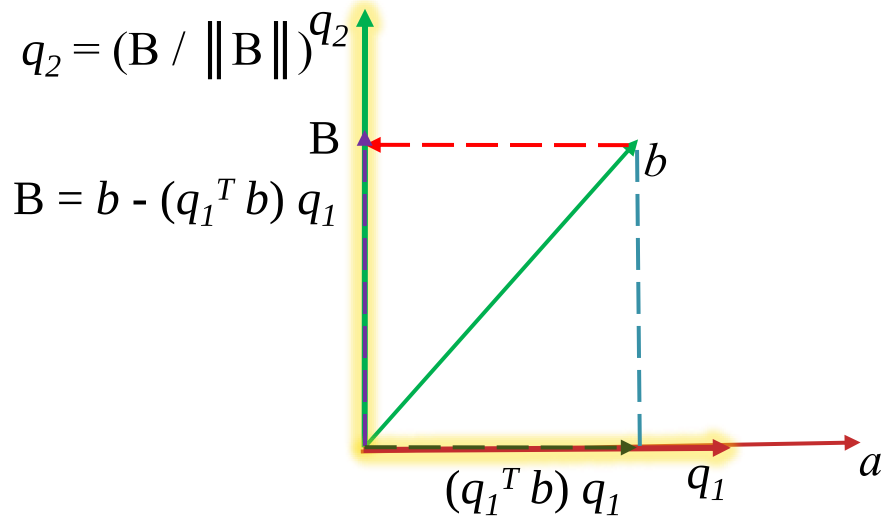


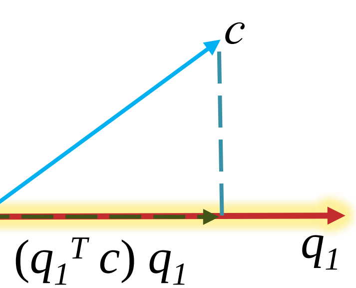




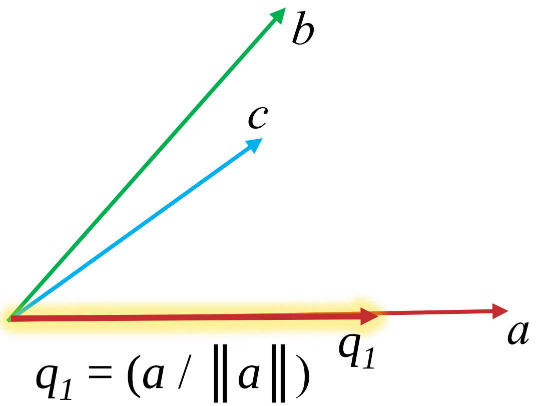
Gram–Schmidt Orthogonalization

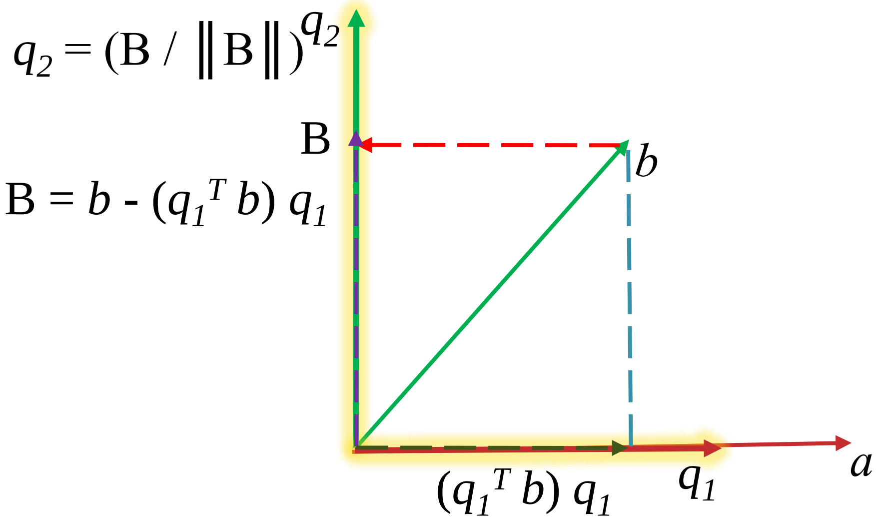


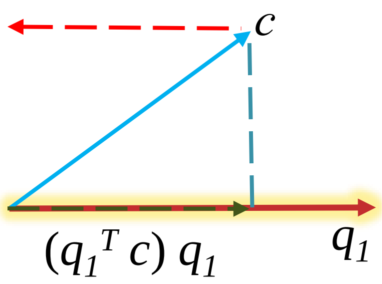




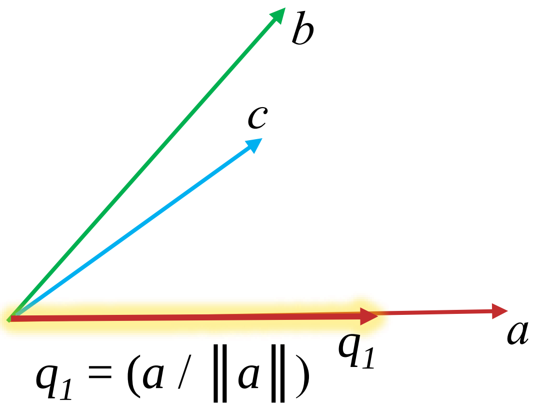
Gram–Schmidt Orthogonalization

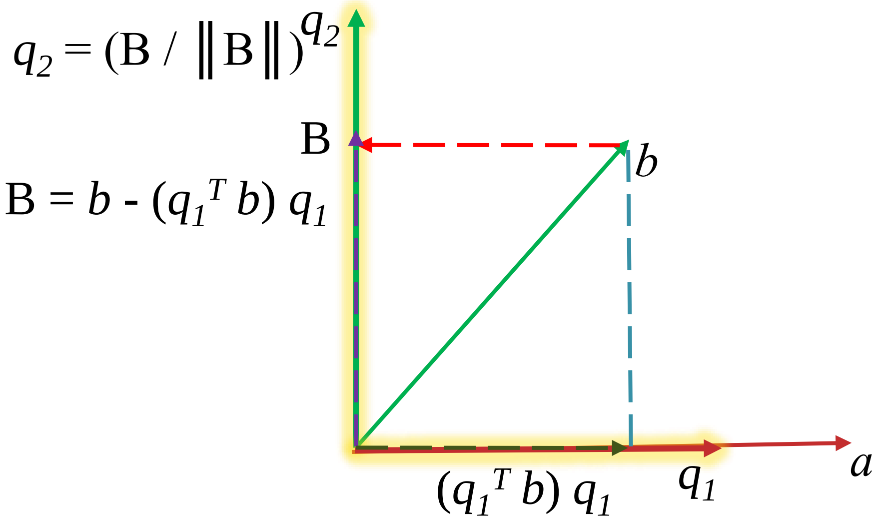


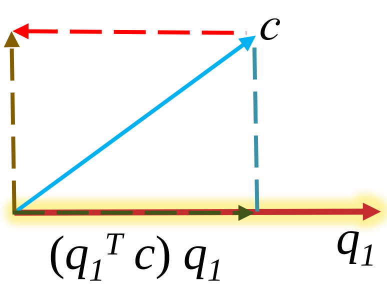




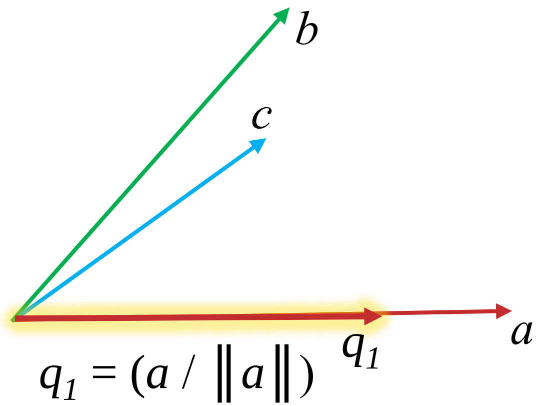
Gram–Schmidt Orthogonalization

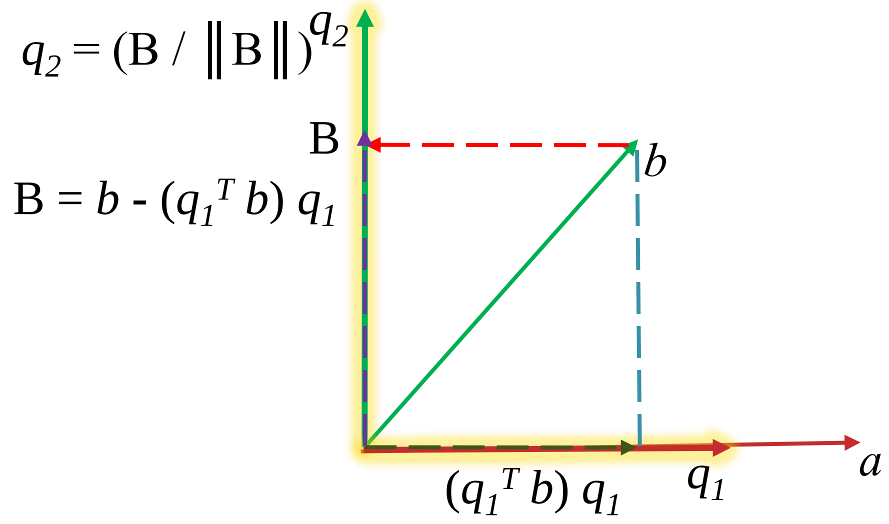


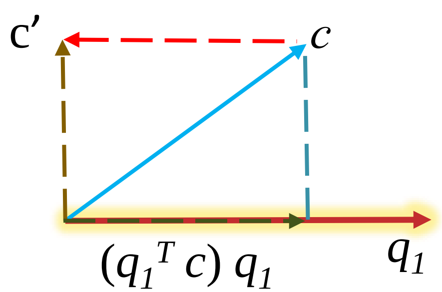




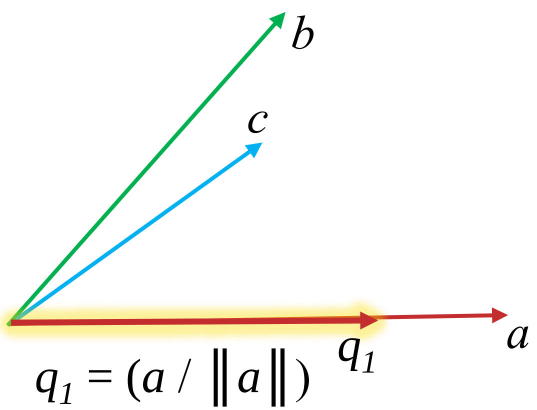
Gram–Schmidt Orthogonalization

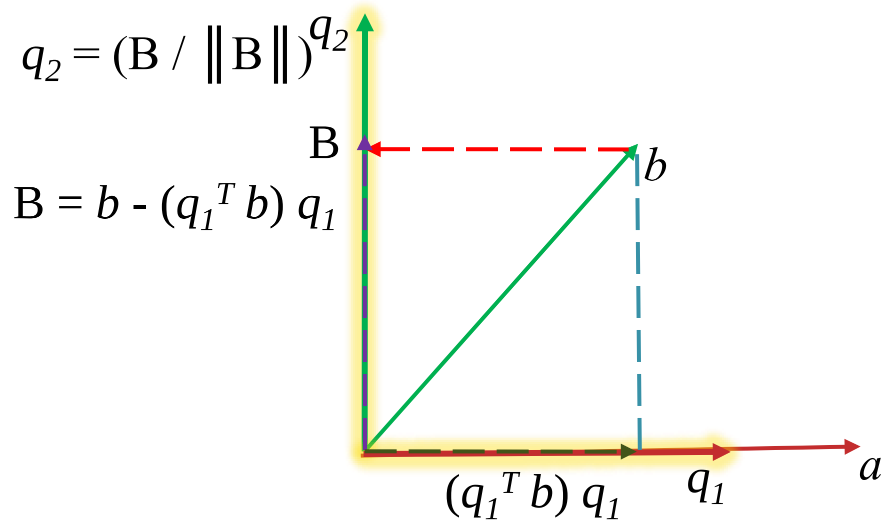


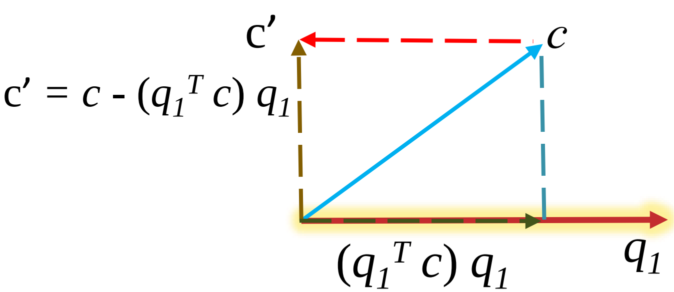




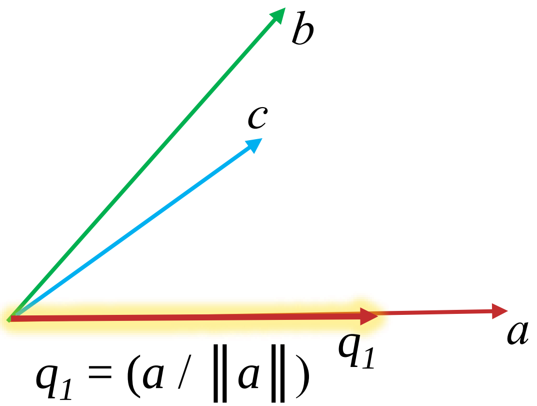
Gram–Schmidt Orthogonalization

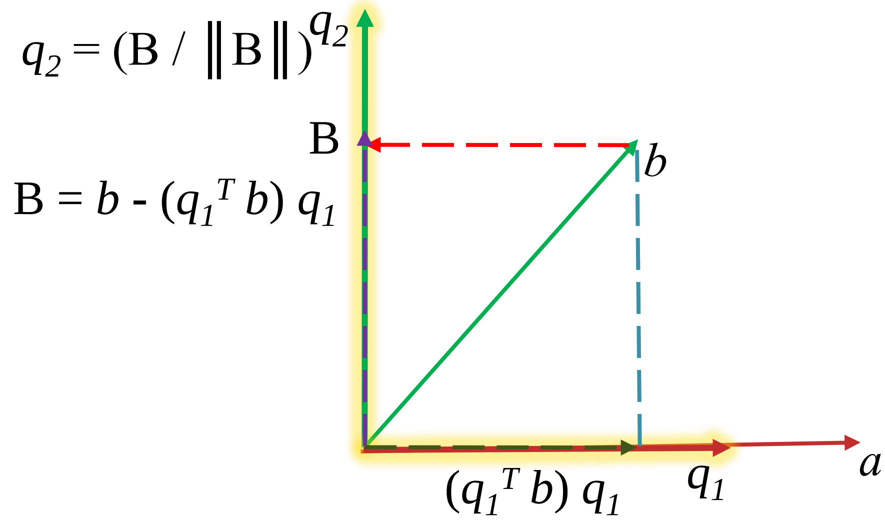


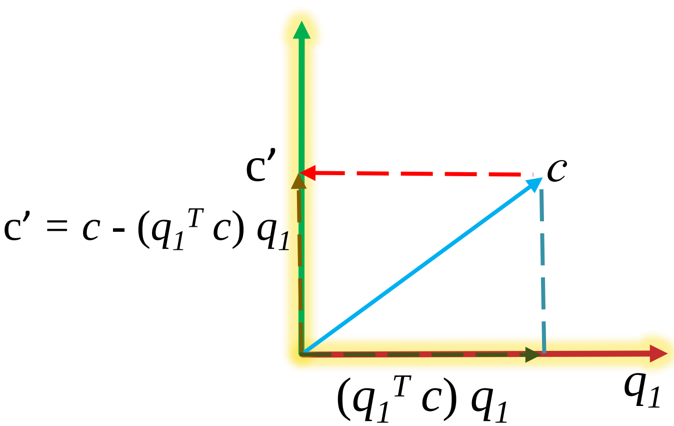




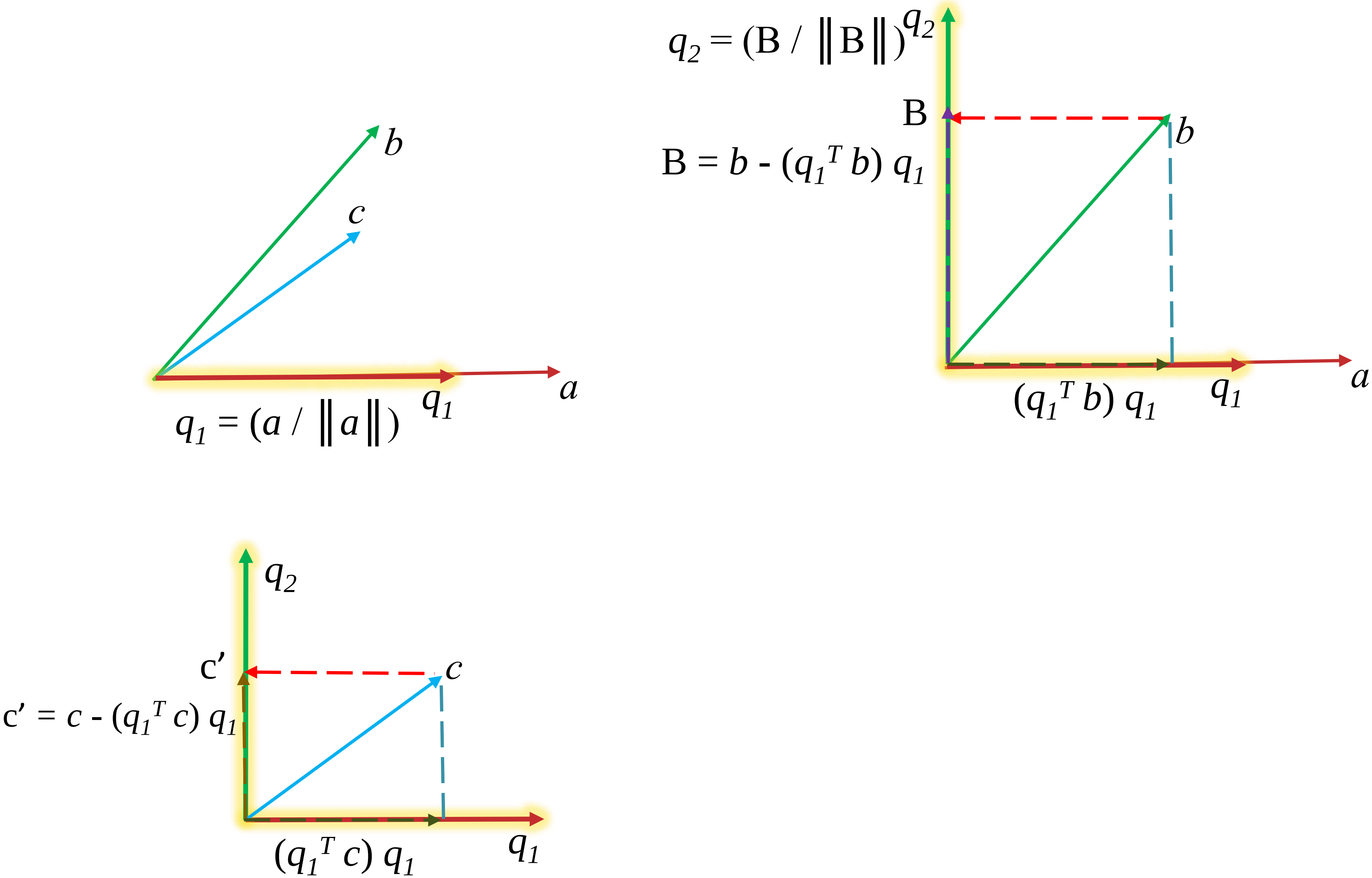
Gram–Schmidt Orthogonalization



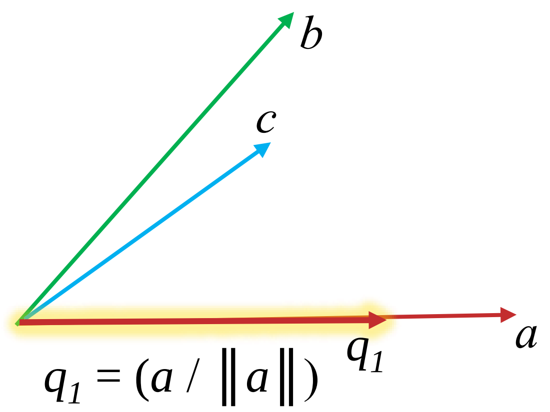


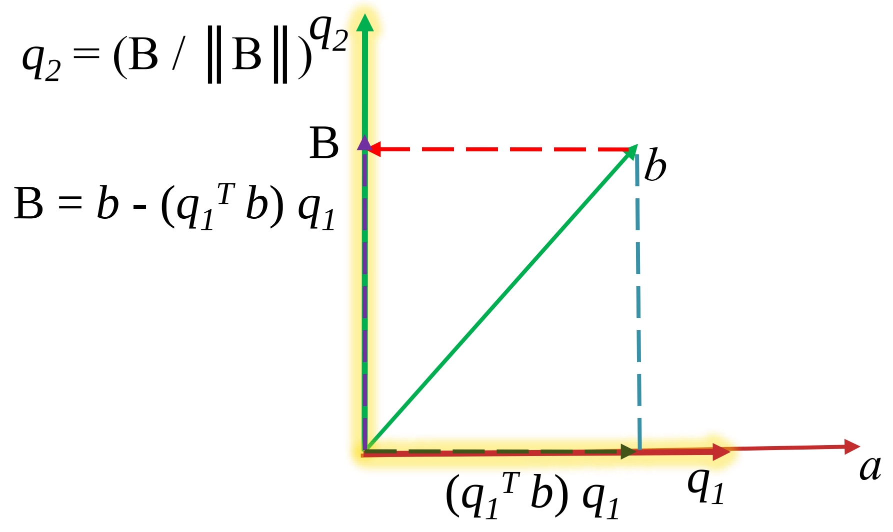


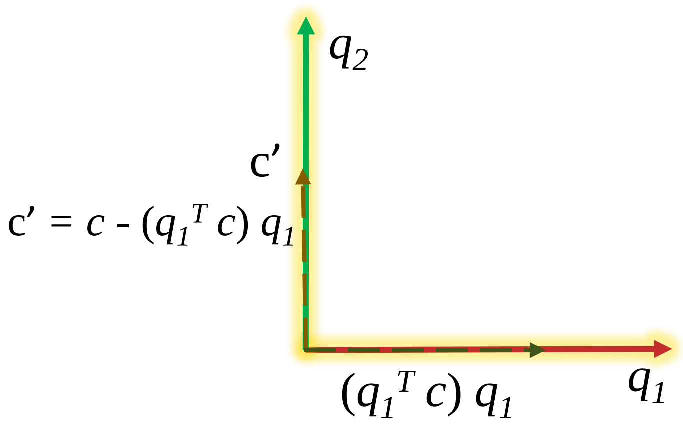
Gram–Schmidt Orthogonalization



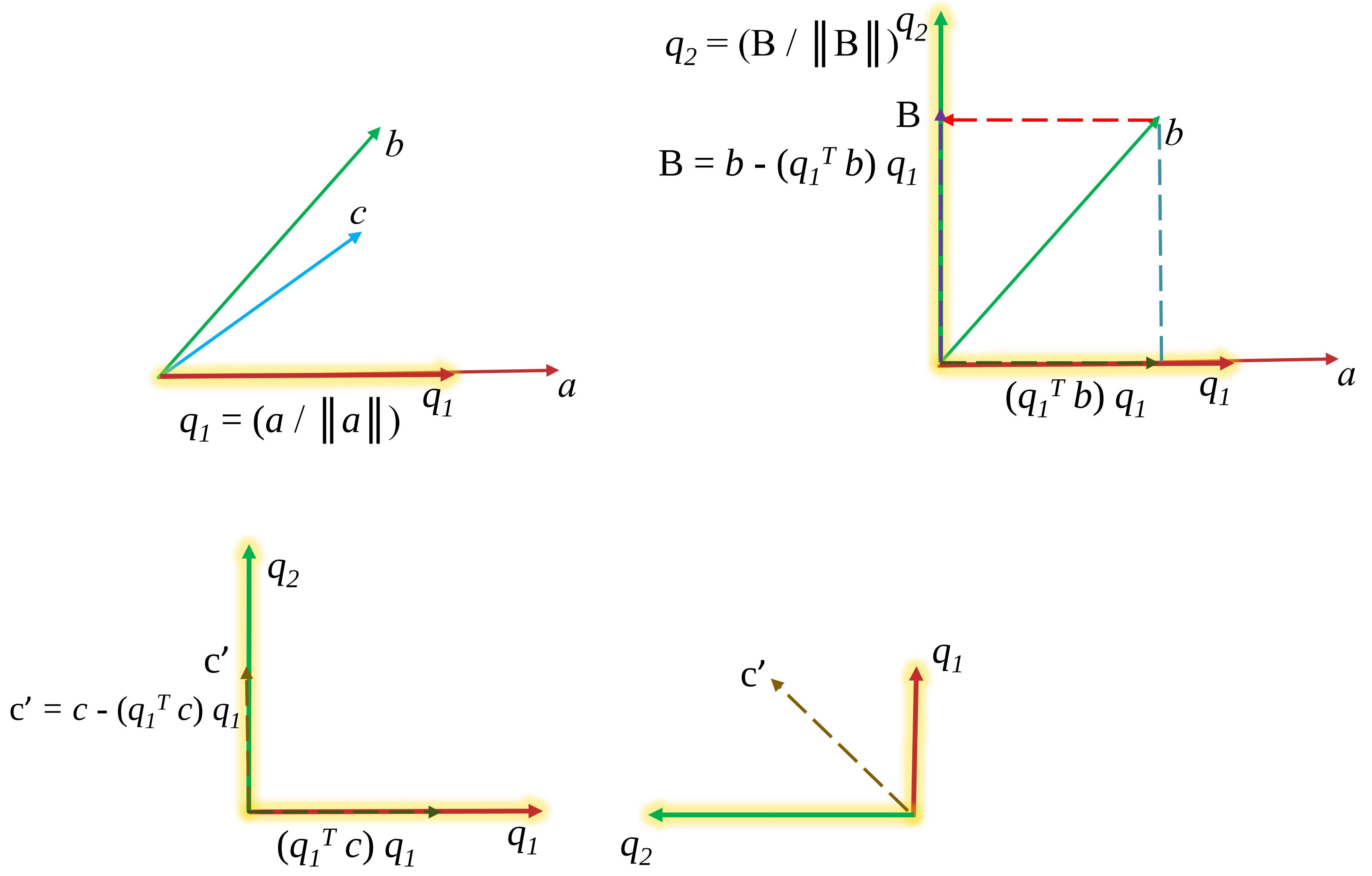
Gram–Schmidt Orthogonalization



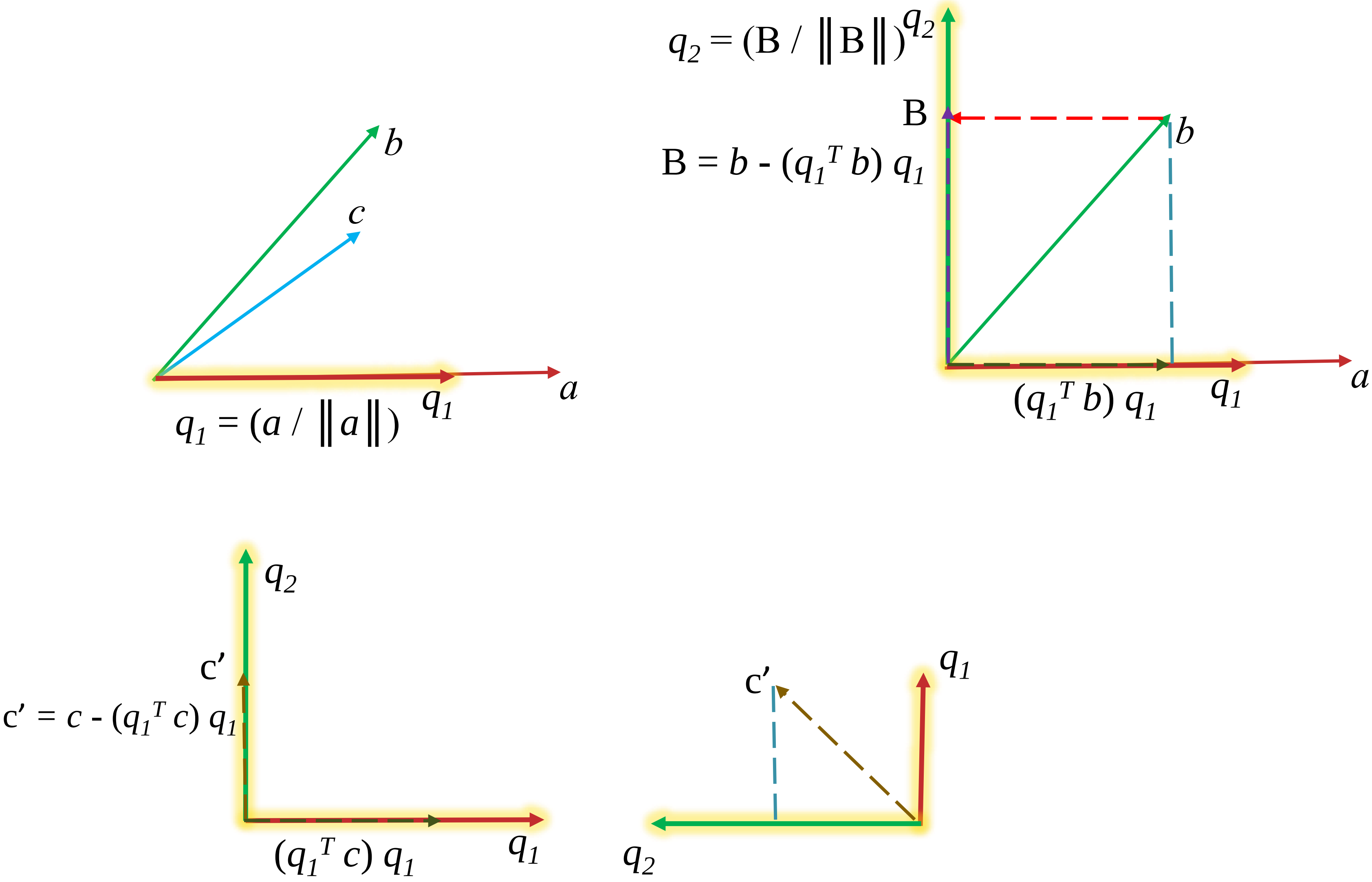




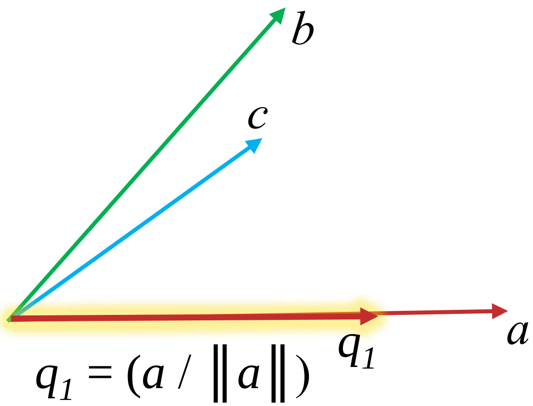
Gram–Schmidt Orthogonalization

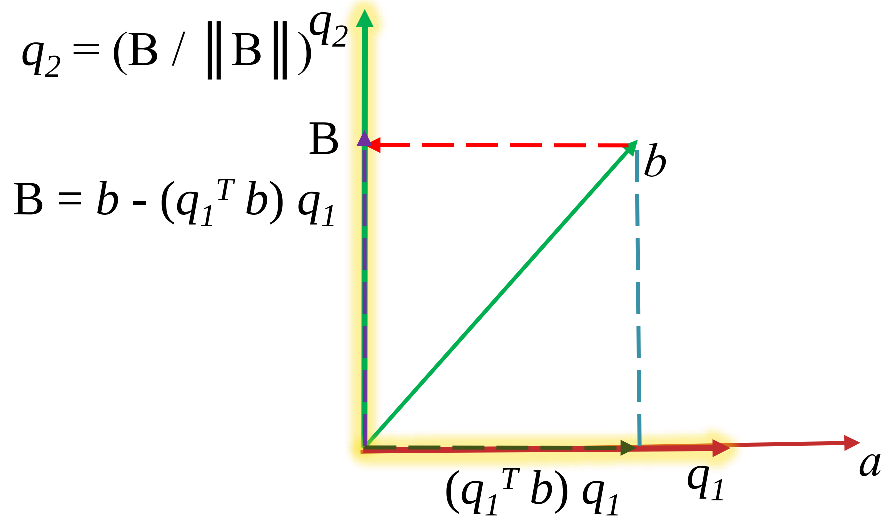


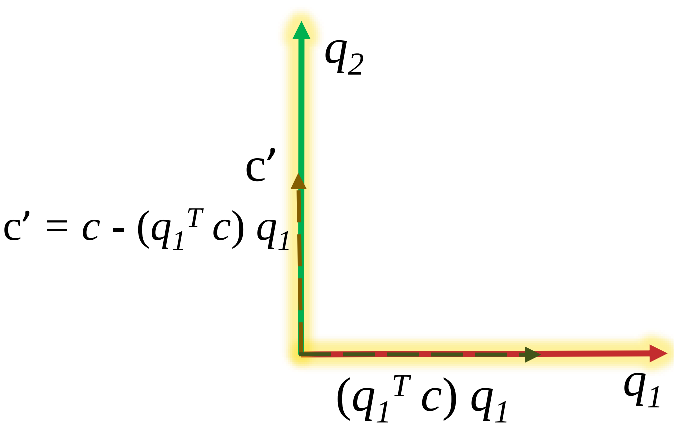
Gram–Schmidt Orthogonalization

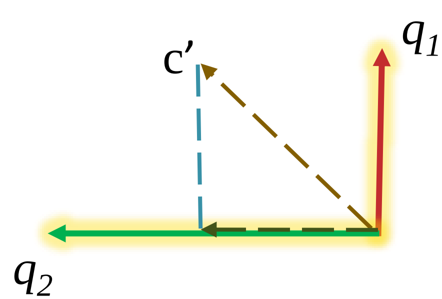


Gram–Schmidt Orthogonalization

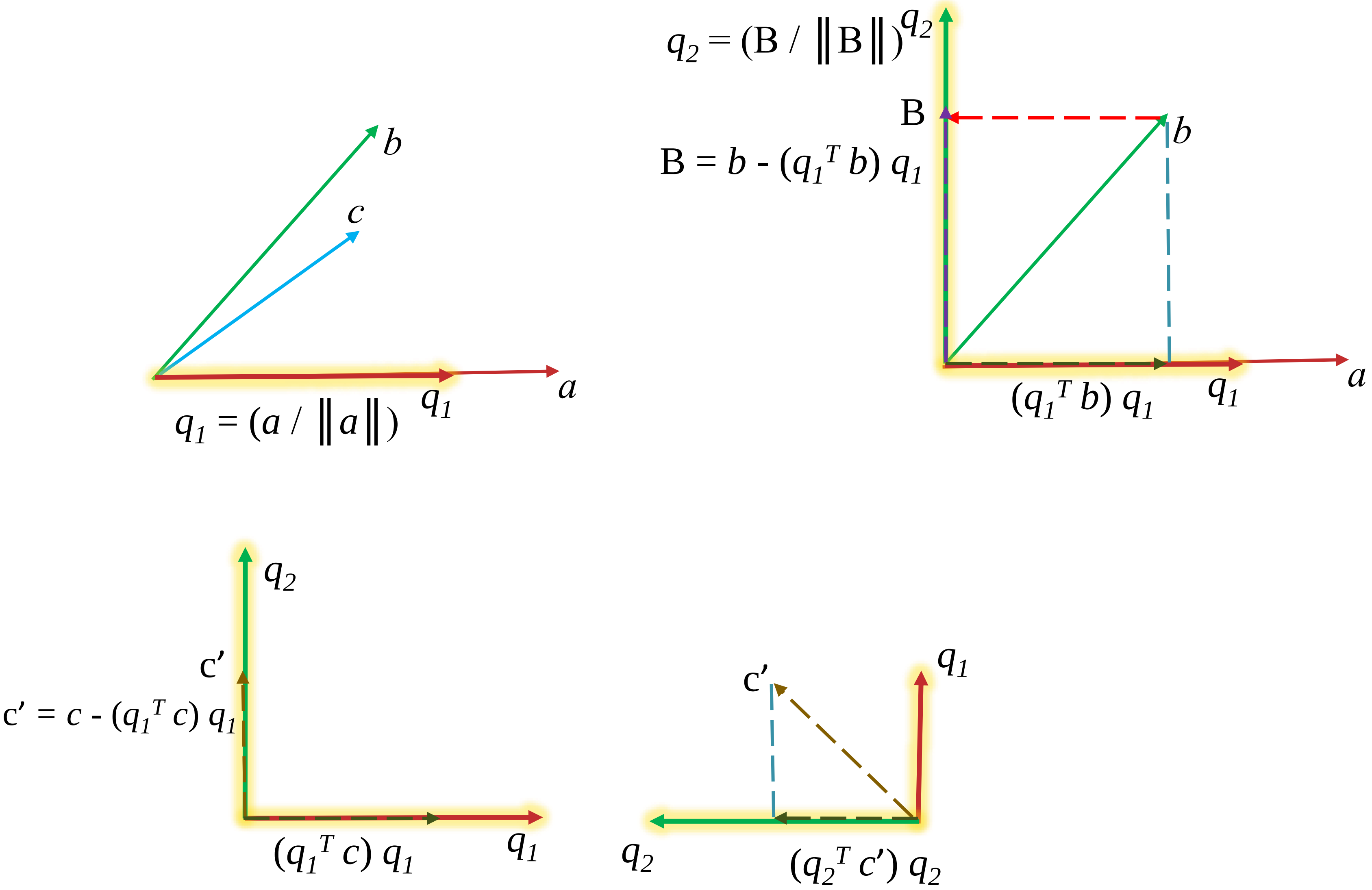




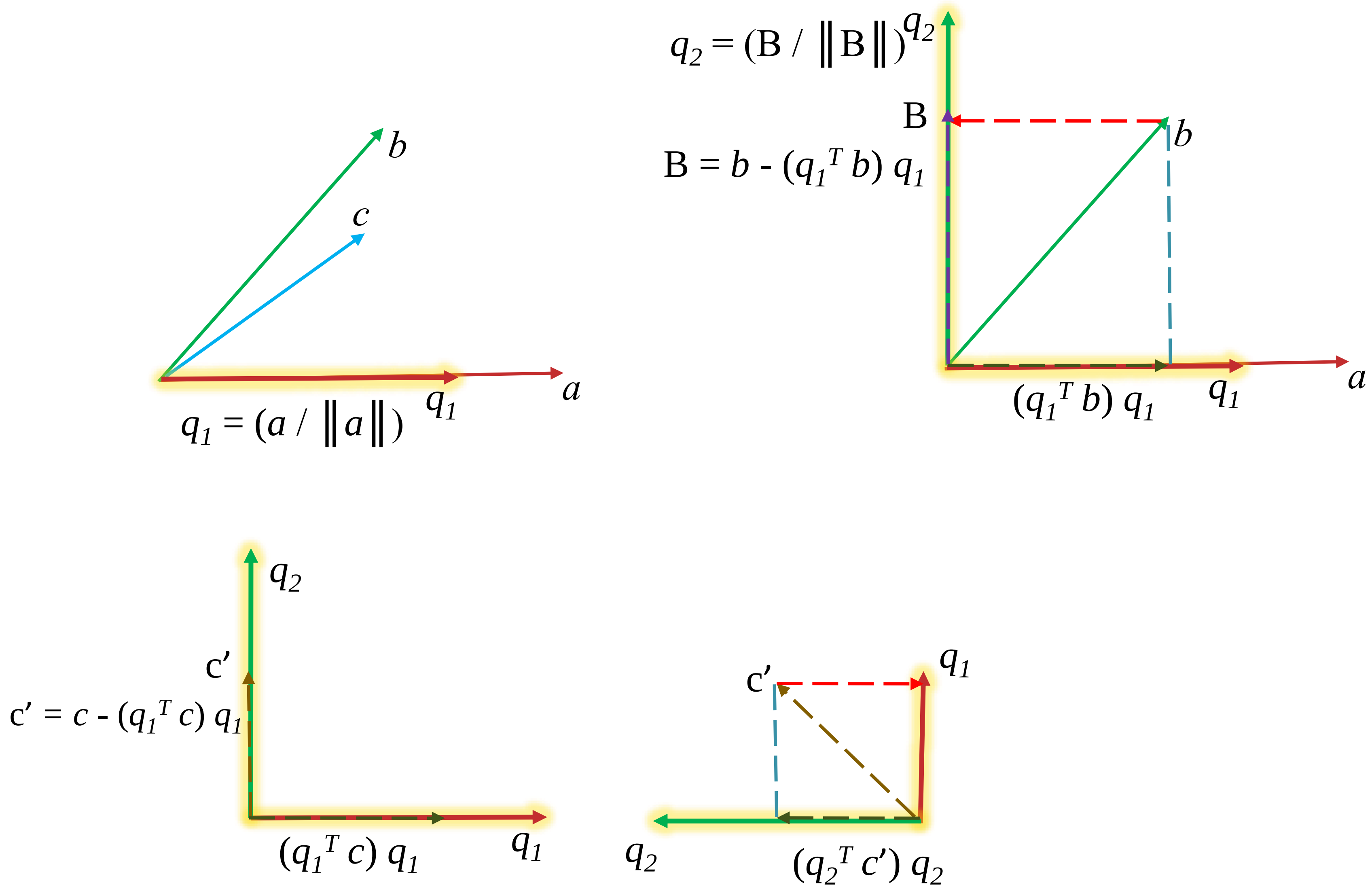




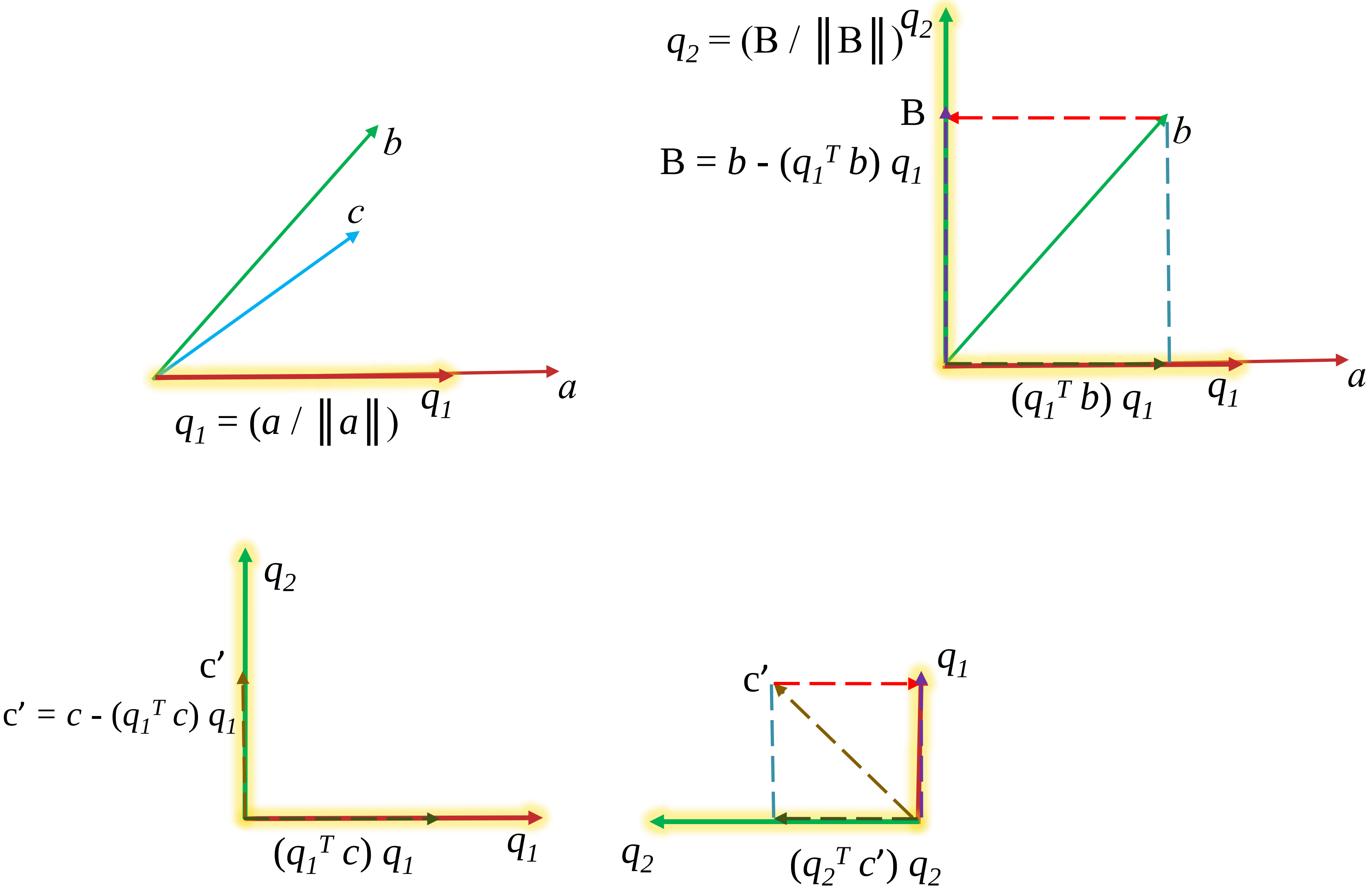
Gram–Schmidt Orthogonalization



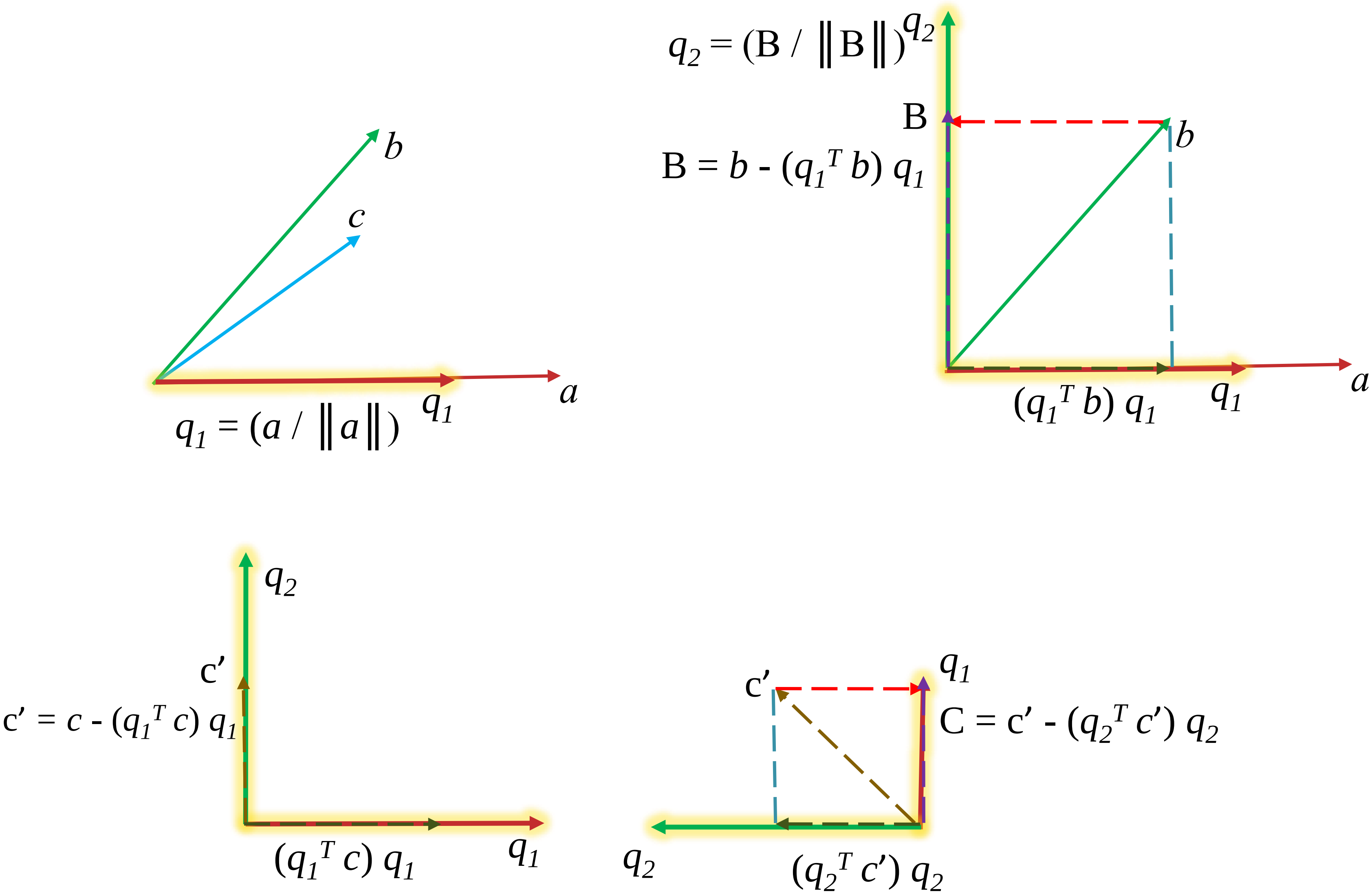
Gram–Schmidt Orthogonalization



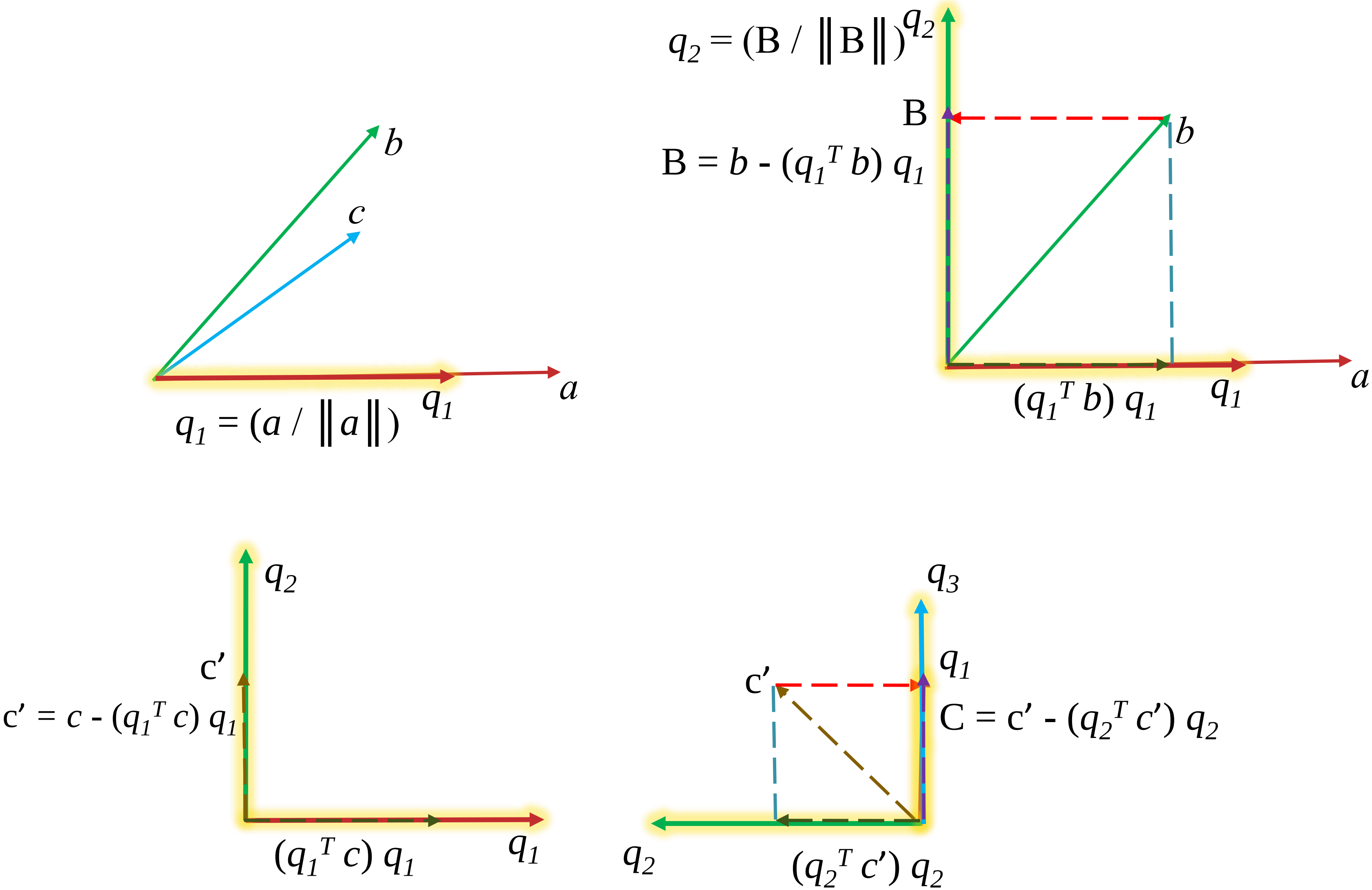
Gram–Schmidt Orthogonalization



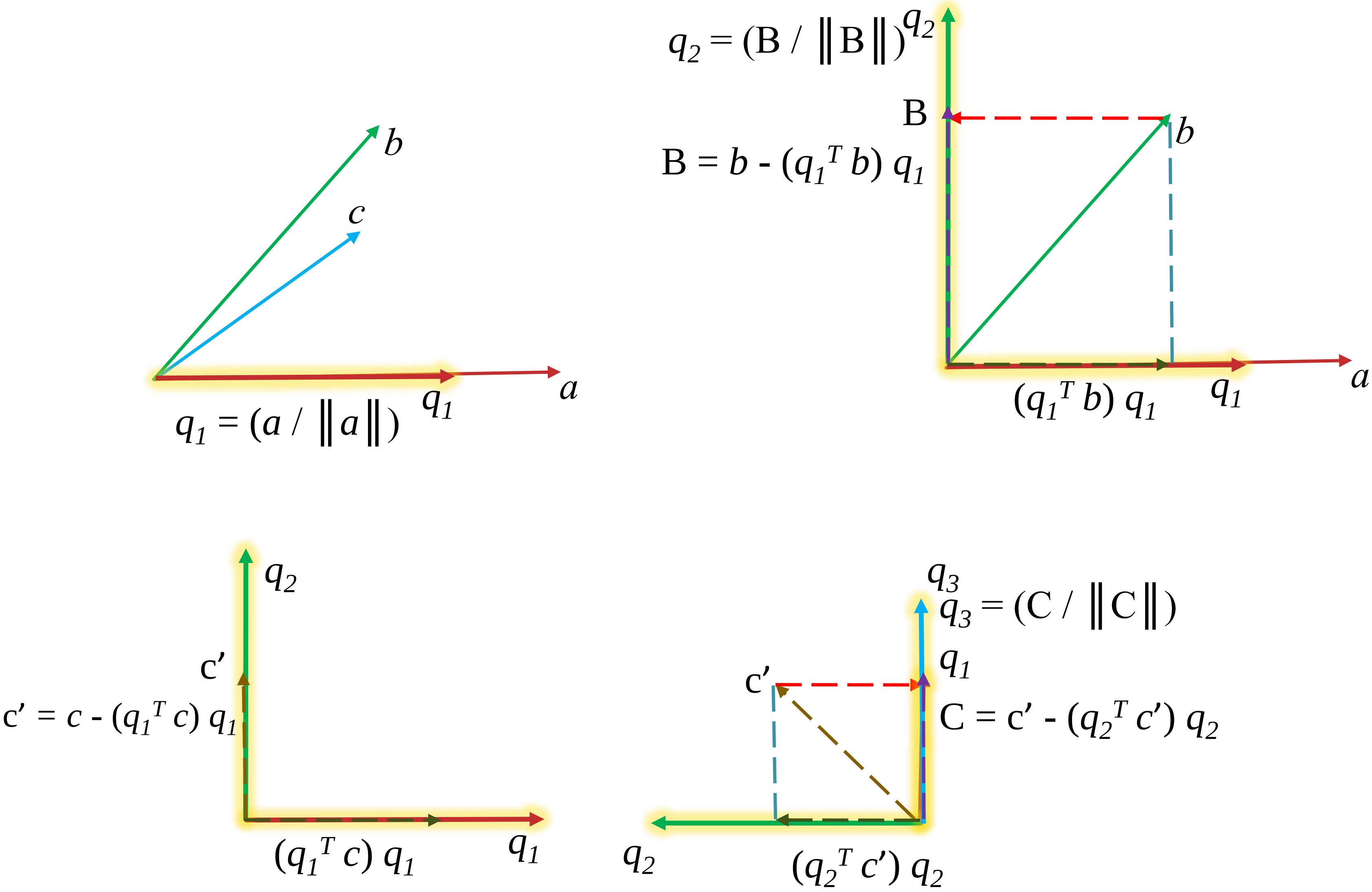
Gram–Schmidt Orthogonalization



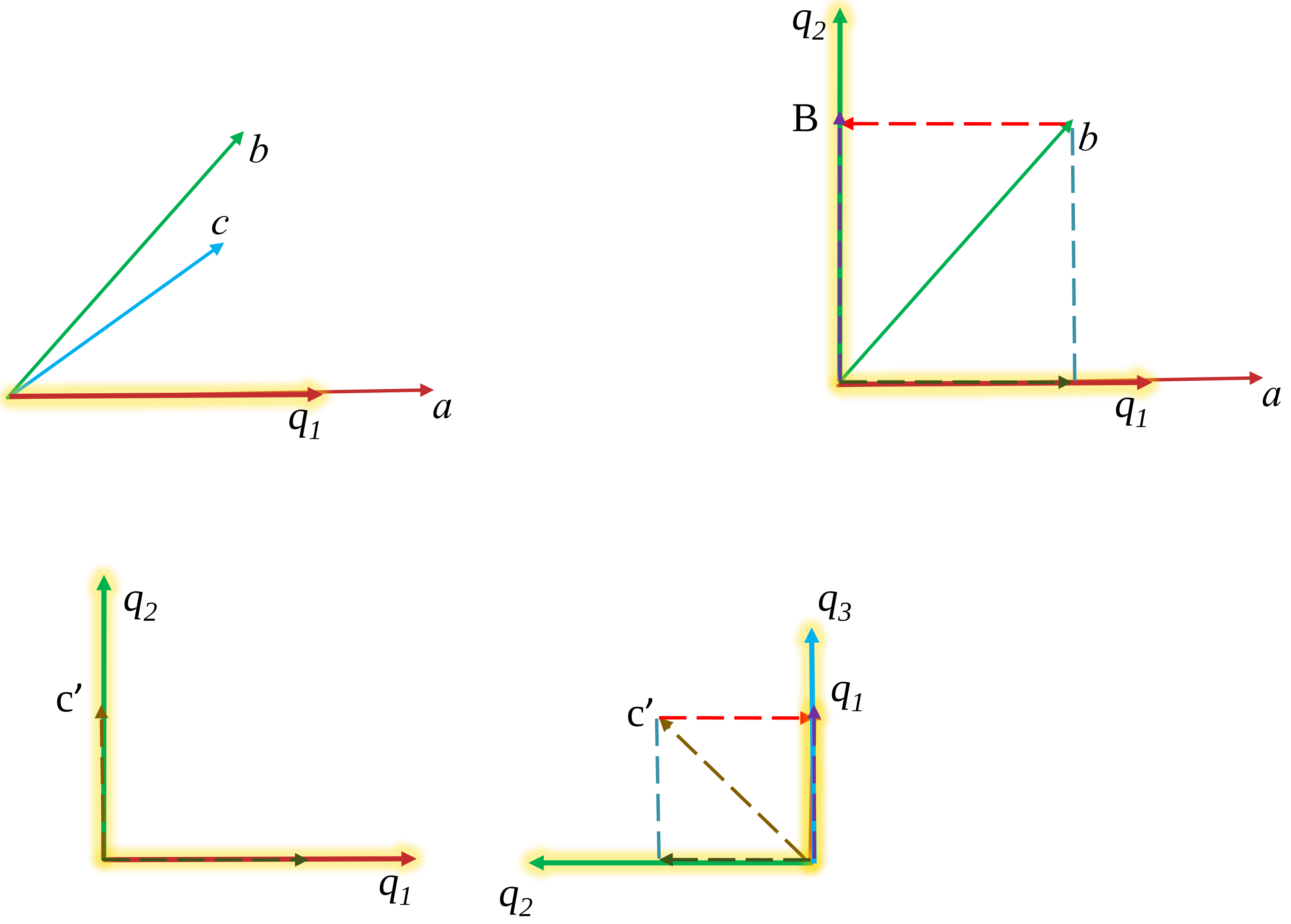
Gram–Schmidt Orthogonalization



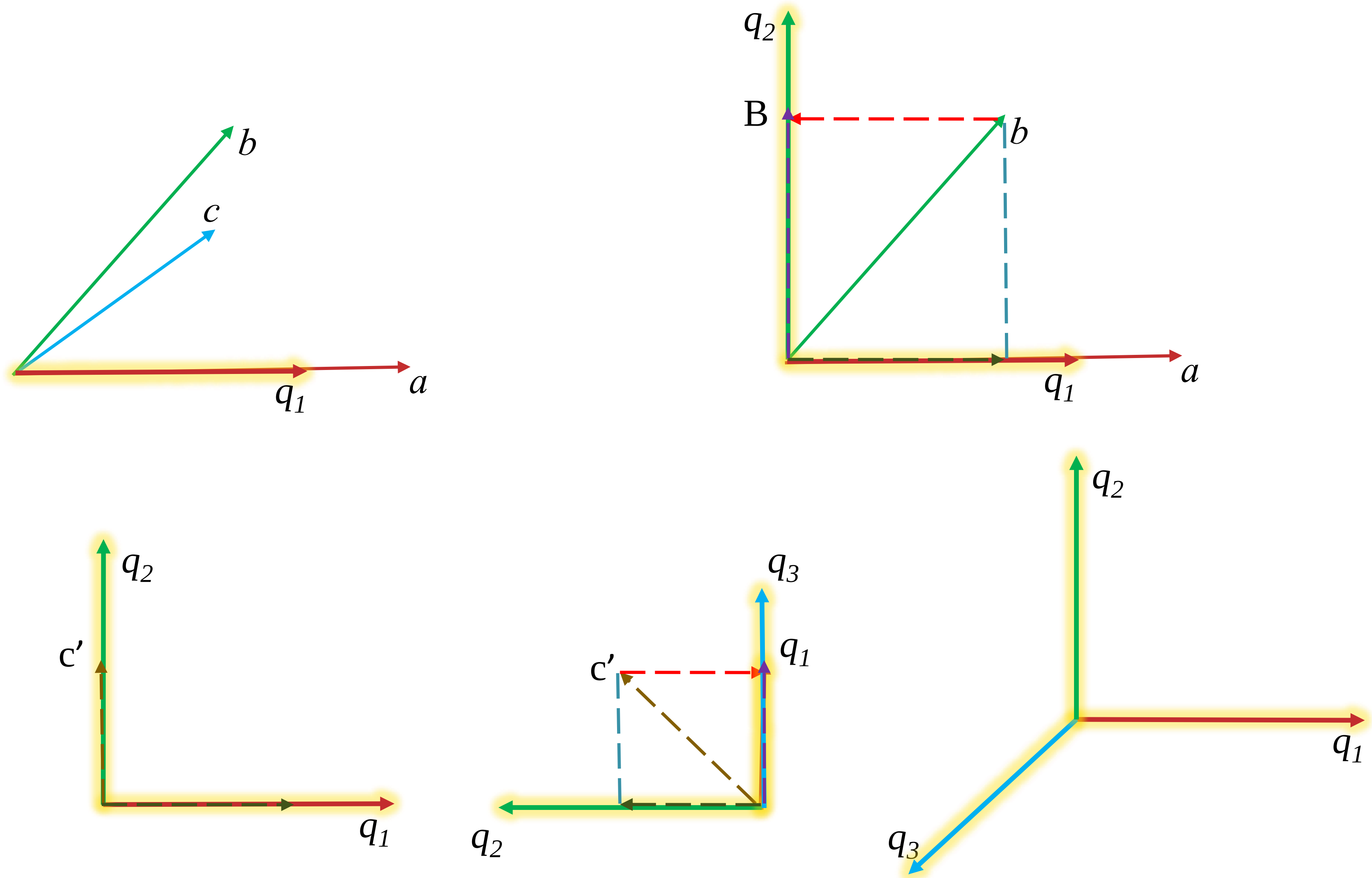
Gram–Schmidt Orthogonalization



Gram–Schmidt Orthogonalization



Gram–Schmidt Orthogonalization



Gram–Schmidt Orthogonalization

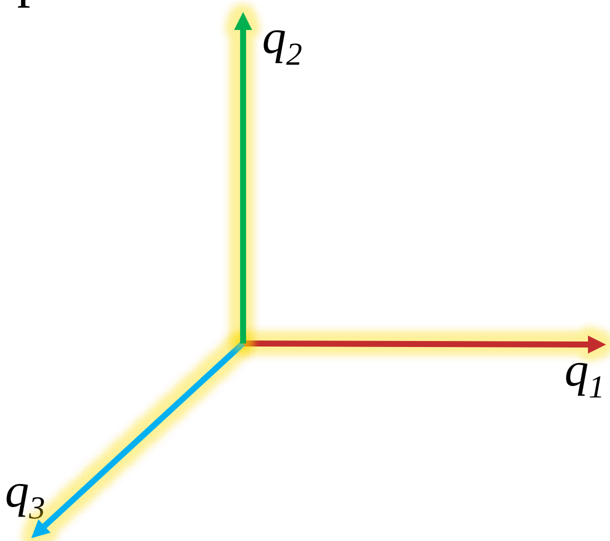
- Thus the three independent vectors *a*, *b*, *c* have been converted into three orthonormal vectors q_1, q_2, q_3 by Gram–Schmidt process

  $q_1 = (a \,/\, \|a\|)$

  $q_2 = (\mathrm{B} \,/\, \|\mathrm{B}\|)$ where $\mathrm{B} = b - (q_1^T b)\, q_1$

  $q_3 = (\mathrm{C} \,/\, \|\mathrm{C}\|)$ where $\mathrm{C} = c - (q_1^T c)\, q_1 - (q_2^T c)\, q_2$

- The whole idea of the Gram–Schmidt process is to subtract from every new vector its components in the directions that are already settled



Gram–Schmidt Orthogonalization

Example: Suppose the independent vectors are a, b, c:

$$a = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$$

- To find q_1 divide the first vector a by its length:

$$q_1 = (a / \|a\|) \quad ; \qquad \|a\| = \sqrt{5}$$

$$\Rightarrow q_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

Gram–Schmidt Orthogonalization

- To find q_2, subtract from the second vector its component in the first direction: $B = b - (q_1^T b)q_1$

$$B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left(2/\sqrt{5}\right) \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -2/5 \\ 1/5 \\ 0 \end{bmatrix}$$

$$\|B\| = 1/\sqrt{5} \;\Rightarrow\; q_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}$$

Gram–Schmidt Orthogonalization

- To find q_3 , subtract from c its components along q_1 and q_2 :

  $C = c - (q_1^T c)\, q_1 - (q_2^T c)\, q_2$

$$C = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} - \left(\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} \right) \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix} - \left(\begin{bmatrix} -2/\sqrt{5} & 1/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} \right) \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} - \left(2/\sqrt{5}\right) \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix} - \left(1/\sqrt{5}\right) \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

$$\| C \| = 4 \quad \Rightarrow \quad q_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Gram–Schmidt Orthogonalization

- Orthonormal basis:

$$Q = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- So, we have constructed an orthonormal basis Q from a set of linearly independent vectors $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$

Matrix Multiplication

- How to multiply two matrices?
- Consider $A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 & -2 \\ 1 & -6 & 7 \\ 1 & 0 & 2 \end{bmatrix}$
- AB = ?
- $AB = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 1 & -6 & 7 \\ 1 & 0 & 2 \end{bmatrix}$ (High–school method)

$$= \begin{bmatrix} 0 & 16 & -16 \\ 1 & -6 & 7 \\ 1 & 0 & 2 \end{bmatrix}$$

Matrix Multiplication

- We may also view the matrix multiplication by rows and columns

Multiplication by columns

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

$$= \left[\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ -6 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 7 \\ 2 \end{pmatrix} \right]$$

$$= \begin{bmatrix} 0 & 16 & -16 \\ 1 & -6 & 7 \\ 1 & 0 & 2 \end{bmatrix}$$

Matrix Multiplication

Multiplication by rows

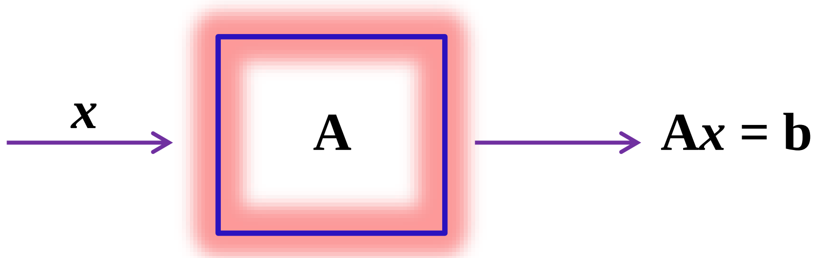
$$AB=\begin{bmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 4 & -2\\ 1 & -6 & 7\\ 1 & 0 & 2\end{bmatrix}$$

$$=\begin{bmatrix}
\begin{pmatrix}1 & -2 & 0\end{pmatrix}\begin{pmatrix}2\\1\\1\end{pmatrix} & \begin{pmatrix}1 & -2 & 0\end{pmatrix}\begin{pmatrix}4\\-6\\0\end{pmatrix} & \begin{pmatrix}1 & -2 & 0\end{pmatrix}\begin{pmatrix}-2\\7\\2\end{pmatrix}\\
\begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix}2\\1\\1\end{pmatrix} & \begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix}4\\-6\\0\end{pmatrix} & \begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix}-2\\7\\2\end{pmatrix}\\
\begin{pmatrix}0 & 0 & 1\end{pmatrix}\begin{pmatrix}2\\1\\1\end{pmatrix} & \begin{pmatrix}0 & 0 & 1\end{pmatrix}\begin{pmatrix}4\\-6\\0\end{pmatrix} & \begin{pmatrix}0 & 0 & 1\end{pmatrix}\begin{pmatrix}-2\\7\\2\end{pmatrix}
\end{bmatrix}=\begin{bmatrix}0 & 16 & -16\\ 1 & -6 & 7\\ 1 & 0 & 2\end{bmatrix}$$

Eigenvalues and Eigenvectors

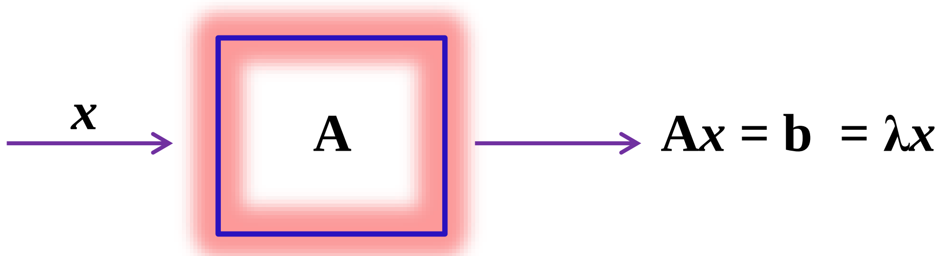
- Defined only for square matrices.
- A system of equations can be expressed as $\mathbf{A}x = \mathbf{b}$

  When a vector x is multiplied by $\mathbf{A}$, it gives a new vector $\mathbf{b}$



- This new vector $\mathbf{b}$, could be in any direction and it depends upon the vector x

- Which x will give a $\mathbf{b}$ that will be in the same direction as x ?

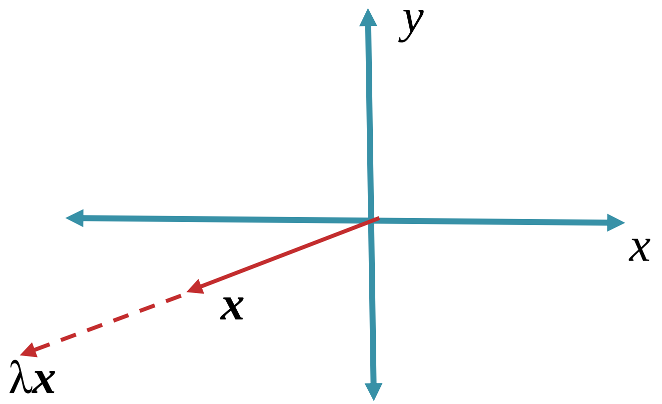


Eigenvalues and Eigenvectors

- Certain exceptional vectors ***x*** are in the same direction as **A*x*** and those are called as the 'eigenvectors'

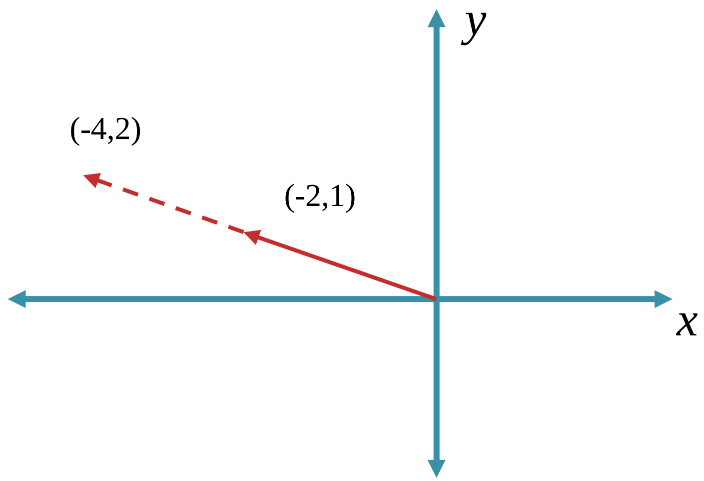
  i.e. it may be written as **A*x* =** λ***x***. The number λ is an eigenvalue of **A**

- The eigenvalue λ tells whether the special vector ***x*** is stretched or shrunk or reversed or left unchanged when it is multiplied by **A**



Eigenvalues and Eigenvectors

- $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ where $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}$ and $\boldsymbol{x}$ is a multiple of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$
- Take $\boldsymbol{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$; when **A** is multiplied with ***x***, we get $\mathbf{A}\boldsymbol{x} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$

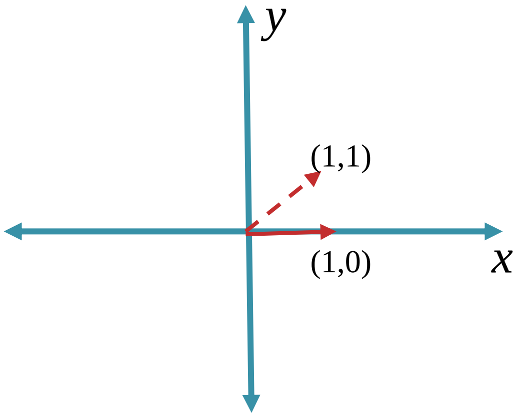


- The direction of the vector $\boldsymbol{x}$ hasn't changed even on multiplication with the matrix **A**. Such a vector $\boldsymbol{x}$ is called as the eigenvector of the matrix **A**

Eigenvalues and Eigenvectors

- Consider some vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Let us see what happens when this vector is multiplied by **A**

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



- Clearly, the vectors $\boldsymbol{x}$ and $\mathbf{A}\boldsymbol{x}$ are not in the same direction. Hence, it is not an eigenvector

Eigenvalues and Eigenvectors

- The eigenvalue λ could be zero. Then $\mathbf{A}\boldsymbol{x} = 0\boldsymbol{x}$ means that this eigenvector $\boldsymbol{x}$ is in the null space

- If **A** is the identity matrix,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \lambda = 1$$

all vectors are eigenvectors of I and all eigenvalues are $\lambda = 1$

Eigenvalues and Eigenvectors

Example–1:

- Consider a matrix A = $\begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}$

- To find eigenvalue: det(A- λI) = 0

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & -2 \\ 1 & 4-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)(4-\lambda) + 2 = 0$$

On solving, we get,

$$\lambda = 3, 2$$

Eigenvalues and Eigenvectors

- Eigenvalues: $\lambda = 3, 2$

- When $\lambda = 3$,

$$A\text{-}3I = \begin{bmatrix} -2 & -2 \\ 1 & 1 \end{bmatrix} \xrightarrow[\substack{R_1 \leftrightarrow R_2 \\ R_2 = R_2 + 2R_1}]{} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

- So, we have the equation –

$$x_1 + x_2 = 0 \Rightarrow x_1 = -x_2$$

- Eigenvector x is any multiple of $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Eigenvalues and Eigenvectors

- when λ=2,

$$A\text{-}2I = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix} \xrightarrow[R_2 = R_2 + R_1]{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

So, we have the equation –

$$x_1 + 2x_2 = 0 \Rightarrow x_1 = -2x_2$$

- Eigenvector x is any multiple of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$

- So the eigenvalues are 3, 2 and corresponding eigenvectors are

$$a\begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ and } b\begin{bmatrix} -2 \\ 1 \end{bmatrix},\ a,b \neq 0$$

Example–2:

- To find the eigenvalues and eigenvectors of the matrix

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

- We first find the eigenvalue λ which satisfies the characteristic equation of the matrix A,

$$\det(A - \lambda I) = 0$$

$$A - \lambda I = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{bmatrix}$$

Eigenvalues and Eigenvectors

- Let us calculate det (A- λI) now

$$\det(A-\lambda I) = (1-\lambda)\begin{vmatrix} -5-\lambda & 3 \\ -6 & 4-\lambda \end{vmatrix} - (-3)\begin{vmatrix} 3 & 3 \\ 6 & 4-\lambda \end{vmatrix} + 3\begin{vmatrix} 3 & -5-\lambda \\ 6 & -6 \end{vmatrix}$$

$$\det(A-\lambda I) = [\,(1-\lambda)(-2+\lambda+\lambda^2)\,] + [\,3(-6-3\lambda)] + [\,3(12+6\lambda)\,]$$

$$= 16 + 12\lambda - \lambda^3$$

- Therefore, det (A- λI) = $-\lambda^3+12\lambda+16 = 0$

Eigenvalues and Eigenvectors

- To solve: $-\lambda^3+12\lambda+16 = 0$

  $-\lambda^3 + 12\lambda + 16 = (\lambda - 4)(\lambda^2 + 4\lambda + 4)$

- On solving $\lambda^2 + 4\lambda + 4$, we get $\lambda = -2$ (repeated root)
- Hence, the eigenvalues of A are $\boxed{\lambda = 4 , -2 , -2}$
- Once the eigenvalues of a matrix have been found, we can find the eigenvectors by Gaussian Elimination

Eigenvalues and Eigenvectors

- For each eigenvalue λ, we have

$$(A - \lambda I)\, x = 0$$

where x is the eigenvector associated with eigenvalue λ

- Case1: when λ=4,

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

$$A - 4I = \begin{bmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix}$$

- Augmented matrix: $\left[\begin{array}{ccc|c} -3 & -3 & 3 & 0 \\ 3 & -9 & 3 & 0 \\ 6 & -6 & 0 & 0 \end{array}\right]$

Eigenvalues and Eigenvectors

$$\begin{bmatrix} -3 & -3 & 3 & 0 \\ 3 & -9 & 3 & 0 \\ 6 & -6 & 0 & 0 \end{bmatrix} \xrightarrow[R_1 = R_1 \times (-1/3)]{} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 3 & -9 & 3 & 0 \\ 6 & -6 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[\substack{R_2 = R_2 - (3R_1) \\ R_3 = R_3 - (6R_1)}]{} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -12 & 6 & 0 \\ 0 & -12 & 6 & 0 \end{array}\right]$$

$$\xrightarrow[R_2 = R_2\,(-1/12)]{} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & -12 & 6 & 0 \end{array}\right]$$

$$\xrightarrow[R_3 = R_3 + 12R_2]{} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[R_1 = R_1 - R_2]{} \left[\begin{array}{ccc|c} 1 & 0 & -1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Eigenvalues and Eigenvectors

- So, we get the equations –

$$x_1 - ½\, x_3 = 0$$

$$x_2 - ½\, x_3 = 0$$

- So the eigenvector x is given by:

$$x = \begin{bmatrix} x_1 = x_3/2 \\ x_2 = x_3/2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- Case2: when $\lambda = -2$

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

$$A + 2I = \begin{bmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix}$$

- Augmented matrix:

$$\left[\begin{array}{ccc|c} 3 & -3 & 3 & 0 \\ 3 & -3 & 3 & 0 \\ 6 & -6 & 6 & 0 \end{array}\right]$$

Eigenvalues and Eigenvectors

$$\left[\begin{array}{ccc|c} 3 & -3 & 3 & 0 \\ 3 & -3 & 3 & 0 \\ 6 & -6 & 6 & 0 \end{array}\right] \xrightarrow{R_1 = R_1 / 3} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 3 & -3 & 3 & 0 \\ 6 & -6 & 6 & 0 \end{array}\right]$$

$$\xrightarrow[R_3 = R_3 - 6R_1]{R_2 = R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So, we get the equations –

$$x_1 - x_2 + x_3 = 0$$

So the eigenvectors x associated with the eigenvalue $\lambda = -2$ are given by:

$$x = \begin{bmatrix} x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- So the eigenvalues are $\lambda = 4, -2, -2$ and the eigenvectors are any multiple of

$$\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Eigenvalue Decomposition

- Let $\mathbf{A}$ be an $n \times n$ matrix with n linearly independent eigenvectors then $\mathbf{A}$ can be factored into

$$\boxed{\mathbf{A} = \mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}}$$

- The columns of $\mathbf{S}$ are eigenvectors of $\mathbf{A}$
- The diagonal matrix $\boldsymbol{\Lambda}$ has eigenvalues of $\mathbf{A}$

- With $\mathbf{A}$, $\mathbf{S}$ and $\mathbf{S}^{-1}$ available, we may find $\boldsymbol{\Lambda}$

  $\boldsymbol{\Lambda} = \mathbf{S}^{-1}\mathbf{A}\,\mathbf{S}$

Eigenvalue Decomposition

Example:

Consider a matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}$

- To find eigenvalue: det(A - λI) = 0

$$\begin{vmatrix} 1-\lambda & -1 \\ -2 & -\lambda \end{vmatrix} = 0 \quad \Rightarrow \lambda^2 - \lambda - 2$$
$$\Rightarrow \lambda = 2, -1$$

- Eigenvalues: $\lambda = 2, -1$

Eigenvalue Decomposition

- To find eigenvector: (A - λI) x = 0

- when λ = 2,

$$A-\lambda I=\begin{bmatrix}-1 & -1\\ -2 & -2\end{bmatrix}$$

$$(A-\lambda I)x=0 \Rightarrow \begin{bmatrix}-1 & -1\\ -2 & -2\end{bmatrix}x=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$x=\begin{bmatrix}-1\\ 1\end{bmatrix}$$

- when λ = -1,

$$A-\lambda I=\begin{bmatrix}2 & -1\\ -2 & 1\end{bmatrix}$$

$$(A-\lambda I)x=0 \Rightarrow \begin{bmatrix}2 & -1\\ -2 & 1\end{bmatrix}x=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$x=\begin{bmatrix}1\\ 2\end{bmatrix}$$

- Eigenvectors: $\begin{bmatrix}-1\\ 1\end{bmatrix}$ and $\begin{bmatrix}1\\ 2\end{bmatrix}$

Eigenvalue Decomposition

- λ = 2, -1 and corresponding eigenvectors are $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

- The matrix S contains eigenvectors of A in its columns –

$$S = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } S^{-1} = \begin{bmatrix} -2/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix}$$

- The diagonal matrix Λ contains eigenvalues of A –

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

- $A = \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{bmatrix} -2/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix}}_{\mathbf{S^{-1}}} = \begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}$

Symmetric matrix

- Symmetric matrix is a square matrix which is equal to its transpose

Example:

$$A = \begin{bmatrix} 4 & 8 & 3 \\ 8 & 5 & 7 \\ 3 & 7 & 1 \end{bmatrix} \qquad A^T = \begin{bmatrix} 4 & 8 & 3 \\ 8 & 5 & 7 \\ 3 & 7 & 1 \end{bmatrix}$$

- Symmetric matrices need not be invertible. A matrix with zero entries is symmetric but not invertible

Eigenvalue Decomposition

- A real symmetric matrix can be factored into

$$\boxed{\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{T}} = \text{(orthogonal) (diagonal) (orthogonal)}}$$

- A symmetric matrix has real eigenvalues
- The eigenvectors of a real symmetric matrix are orthogonal to one another provided that they have distinct eigenvalues

Eigenvalue Decomposition

Example:

For a matrix A = $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ To factorize A as $\mathbf{A} = \mathbf{Q \Lambda Q^T}$

- A is symmetric

- To find eigenvalue: det(A - λI) = 0

$$\begin{vmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = 0 \quad \Rightarrow \lambda^2 - 6\lambda + 5 = 0 \quad \Rightarrow \lambda = 5, 1$$

- Eigenvalues: λ = 5, 1

Eigenvalue Decomposition

- To find eigenvector: (A - λI) x = 0

- when λ = 5,

$$A-\lambda I=\begin{bmatrix}-2 & 2\\ 2 & -2\end{bmatrix}$$

$$(A-\lambda I)x=0\Rightarrow\begin{bmatrix}-2 & 2\\ 2 & -2\end{bmatrix}x=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$x=\begin{bmatrix}1\\ 1\end{bmatrix}$$

- when λ = 1,

$$A-\lambda I=\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$$

$$(A-\lambda I)x=0\Rightarrow\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}x=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$x=\begin{bmatrix}1\\ -1\end{bmatrix}$$

- Eigenvectors: $\begin{bmatrix}1\\ 1\end{bmatrix}$ and $\begin{bmatrix}1\\ -1\end{bmatrix}$

Eigenvalue Decomposition

- For the matrix $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$, we have eigenvalues ($\lambda$) = 5, 1 and eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

- We may factorize the matrix A into $\mathbf{Q\Lambda Q^T}$ as –

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{\mathbf{Q^T}}$$

- **Q** has orthonormal vectors in its columns and **Λ** has eigenvalues in its columns

Singular Value Decomposition

- SVD – factorization of 1 matrix (any matrix) into 3 matrices

- Any $m \times n$ matrix **A** can be factored into

$$\boxed{\mathbf{A} = \mathbf{U\Sigma V}^{\mathbf{T}} = \text{(orthogonal) (diagonal) (orthogonal)}}$$

- The columns of **U** ($m \times m$) are eigenvectors of $\mathbf{AA}^{\mathbf{T}}$
- The columns of **V** ($n \times n$) are eigenvectors of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$

- The diagonal matrix $\mathbf{\Sigma}$ has square roots of eigenvalues from $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ and not from **A**
- Those positive entries will be $\sigma_1, \ldots, \sigma_r$. They are the **singular values** of **A** and they fill the first r places in the main diagonal of $\mathbf{\Sigma}$ – when **A** has rank r. The rest of $\mathbf{\Sigma}$ is zero

Singular Value Decomposition

Example:

For a matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$ To factorize A as $\mathbf{A} = \mathbf{U\Sigma V^T}$

- To find **U**:

$$AA^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \quad ; \text{ eigenvalues} = 4, 2$$

Eigenvectors of $AA^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ for the eigenvalue 4 and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for the eigenvalue 2

- To make the vectors orthonormal, divide them by its length, which will

give $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $\mathbf{U} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Singular Value Decomposition

- To find **V**:

$$A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad ; \text{ eigenvalues} = 4, 2, 0$$

Eigenvectors of $A^TA = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ for the eigenvalue 4 , $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ for the eigenvalue 2

and $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ for the eigenvalue 0

- To make the vectors orthonormal, divide them by its length, which will

give $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$; $\mathbf{V} = \begin{bmatrix} 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Singular Value Decomposition

- The diagonal matrix $\mathbf{\Sigma}$ takes the square root of the nonzero eigenvalues

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{4} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}$$

- So, we have,

$$A = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \sqrt{4} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}}_{\mathbf{V^T}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$$

Singular Value Decomposition

Example:

For a matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ To factorize A as $\mathbf{A} = \mathbf{U\Sigma V^T}$

- To find **U**:

$$AA^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix};\ \text{eigenvalues} = 3, 1, 0$$

Eigenvectors of $AA^T = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ for the eigenvalue 3 , $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ for the eigenvalue 1

and $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ for the eigenvalue 0

Singular Value Decomposition

- To make the vectors orthonormal, divide them by its length, which will give $\begin{bmatrix} \sqrt{2/3} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

$$\Rightarrow \mathbf{U} = \begin{bmatrix} \sqrt{2/3} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$

Singular Value Decomposition

- To find **V**:

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad ; \text{ eigenvalues} = 3, 1$$

Eigenvectors of $A^TA = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for the eigenvalue 3 and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for the eigenvalue 1

- To make the vectors orthonormal, divide them by its length, which will give $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$; $\mathbf{V} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Singular Value Decomposition

- The diagonal matrix $\mathbf{\Sigma}$ takes the square root of the nonzero eigenvalues

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

- So, we have,

$$A = \underbrace{\begin{bmatrix} \sqrt{2/3} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{\mathbf{V}^{\mathbf{T}}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Complex Numbers?

- Till now we have dealt only with real numbers. What about the complex numbers?

- In real life problems, we mostly deal only with real values and hence we have ignored the complex numbers

Thank you