E I G E N A X E S

an aid to understanding eigenvalues and eigenvectors

Naveen K . S .

Stephen Vadakkan

R. V. College of Engineering

Bangalore

Karnataka – 560059

INDIA

Beloved: What is in a line ?

Lover:

The early rays through the mist of dawn ,
Drew verse with petals on the lawn.
And when I rose to read a line ,
I saw my beloved thy face divine.

Anonymous

PROLOGUE

Lines and vectors lie at the heart of Linear Algebra. In this module we introduce the basic meaning, method of calculation and purpose of *eigenvalues* and *eigenvectors* at the Higher Secondary (Junior High/ Class XI) level. Our intention is to be intuitive rather than definitive (rigorous and exact). So we shall use a few simple examples to illustrate the concepts (≈ motivation + idea).

In the process we also show how the concepts in Linear Algebra may have evolved and the connection between the different branches of Mathematics like Analytical Geometry, Algebra and Analysis. Through some of the examples we show how some of the operations in Analysis can also be achieved in Linear Algebra. While the main text is meant for Higher Secondary students, the links, references and further reading is meant for undergraduate students and faculty.

TABLE OF CONTENTS

1. ANALYTICAL GEOMETRY versus ALGEBRA

"Pola ninna bela thula."[1]

"There is no royal road to Geometry"

Euclid to King Ptolemy I (323 – 283 BC)
in commentary by Proclus Lycaes (8 February 412 – 17 April 487 AD).

Given the information:

father's age + son's age = 70

father's age – son's age = 30

we have at least three different ways of extracting the information of the ages of the father and son. With a little abstraction (Abstract Algebra) we may represent the given information as:

$$x + y = 70$$

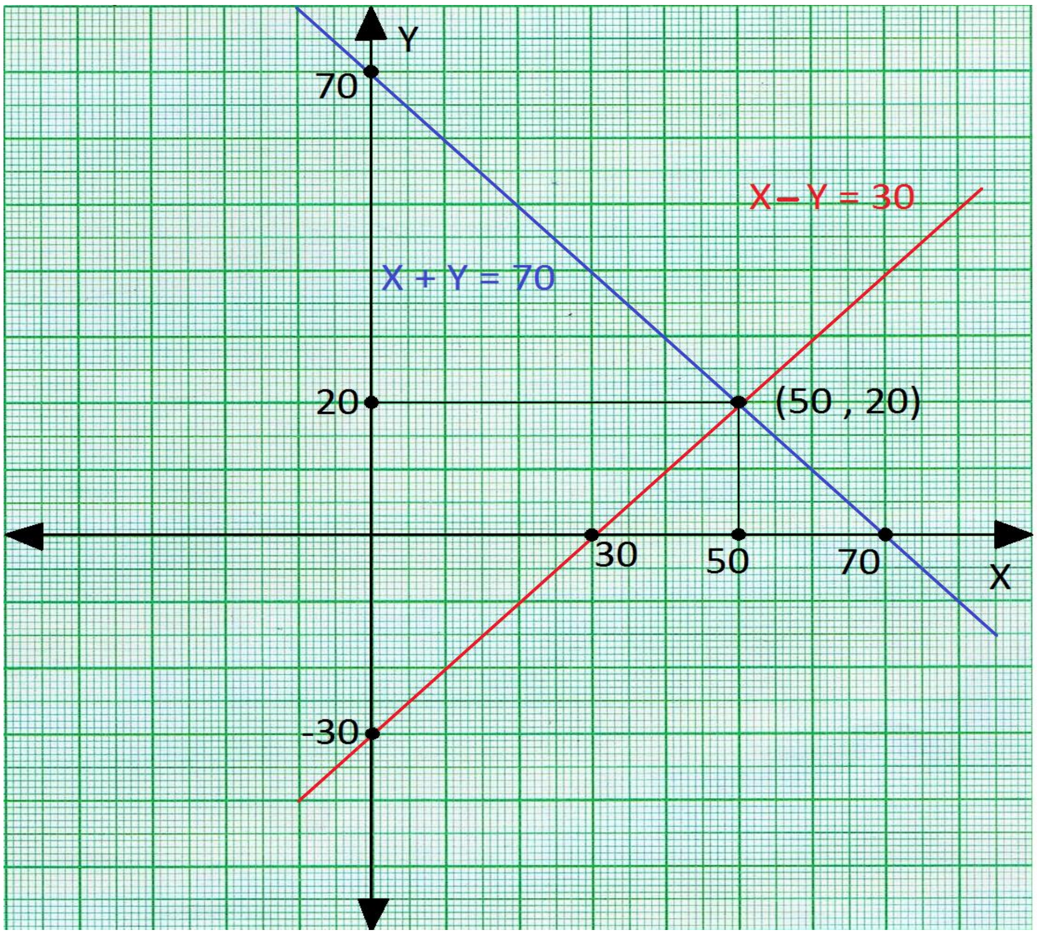
$$x - y = 30$$

where x = father's age and y = son's age.

A fragment of Euclid's elements

[1]What Euclid meant in Tulu, a language without a script in South Canara, western coast of India.

We know enough of Analytical Geometry to see that these are equations of straight lines and so we may draw these lines in the x-y plane and find the point of intersection. We get x = 50 and y = 20.



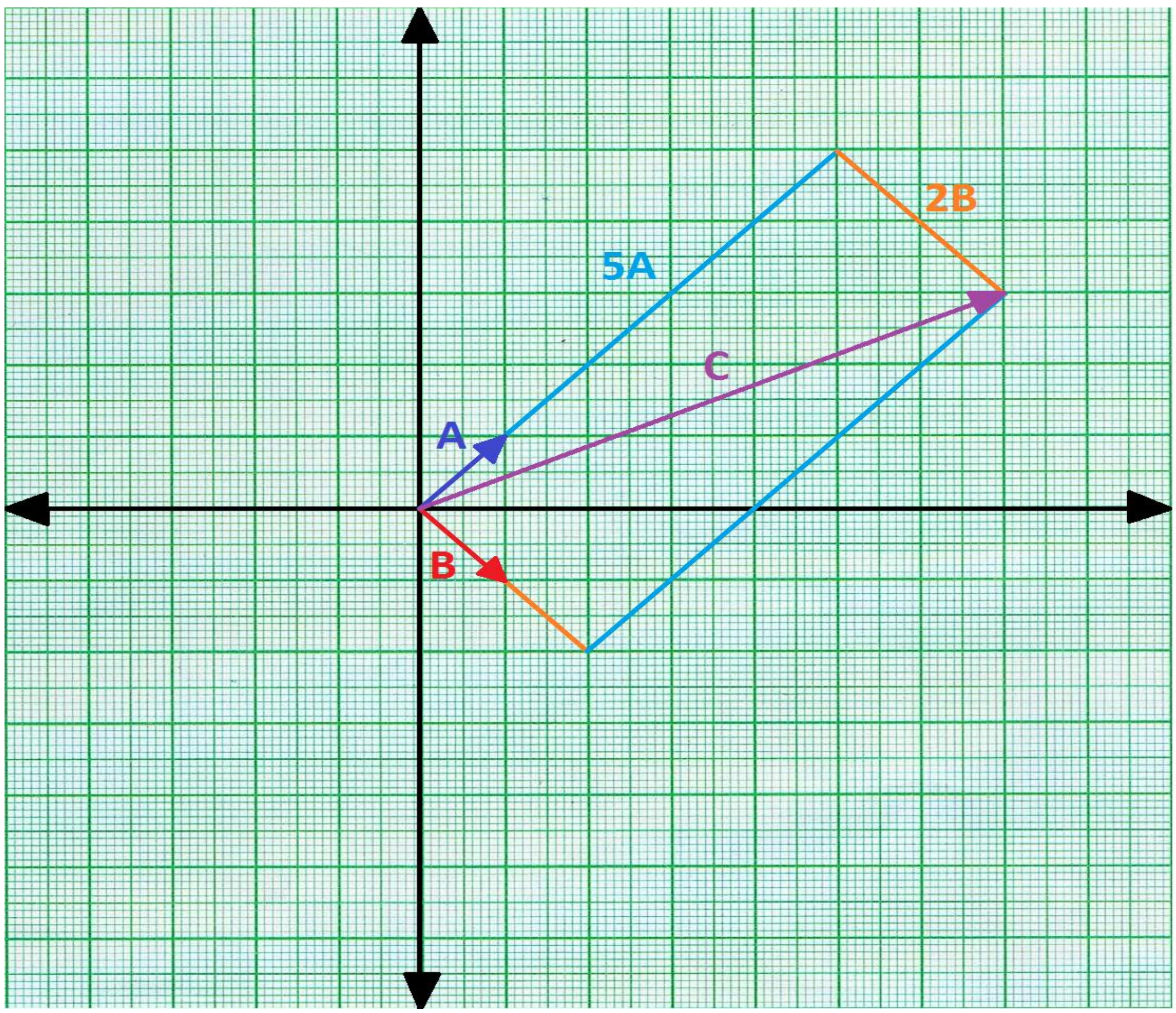
Thanks to the French Mathematician Rene Descartes (1596 - 1650) we have ***Analytical Geometry*** – the link between Analysis and Geometry.

Further reading: Mathematics in Western Culture by Morris Kline, © Oxford University Press, Inc., 1953.

Further reading: Geometry and Algebra in Ancient Civilizations by B. C. Van der Waerden.

Algebra in Ancient and Modern times by V. S. Varadarajan, American Mathematical Society, Providence, Rhode Island, 1988.

Instead of Analytical Geometry, we could have used methods in Linear Algebra to extract the necessary information. In Linear Algebra there are two basic ways of viewing the given abstracted information. There is the ROW picture and the COLUMN picture.



In the figure, A = $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, B = $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and C = $\begin{bmatrix} 7 \\ 3 \end{bmatrix}$ which is a scaled down version of $\begin{bmatrix} 70 \\ 30 \end{bmatrix}$.

We first represent the information in matrix form:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$$

$$A \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$$

More generally we may write this as:

$A \cdot u = b$ where $u = \begin{bmatrix} x \\ y \end{bmatrix}$ and $b = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$

Now depending on which view (ROW or COLUMN) we take we may find the x and y. With the ROW picture in mind we may apply the Gaussian Elimination method which works fine for a small number of variables.

Thus:

$$x + y = 70$$

$$x - y = 30$$

Adding both equations we get:

$$2x = 100$$

$$x = 50 \Longrightarrow y = 20$$

With the COLUMN picture in mind we may apply the methods of Matrix Algebra (Vector Spaces) to find a Linear Combination:

$$x \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$$

we get $x = 50$ and $y = 20$.

2. ROW Picture to COLUMN Picture

The line x + y = 70 represents a ***continuous*** set of information.

Likewise the line x – y = 30 represents a ***continuous*** set of information.

The branch of mathematics that deals with ***continuous*** operands is ***Analysis***. The branch of mathematics that deals with ***discrete*** operands is ***Algebra***. Refer: "A Little Bit of Calculus" by S. Vadakkan, © 1998.

How can we go from ***Analysis*** to ***Algebra***?

2.1 THE NOTION OF VECTORS

Lines are ***continuous*** and ***infinite*** – spread over from –∞ to +∞.

How can we convert this to something ***finite*** and ***discrete***?

There are three concepts we employ to achieve this – ***scalars, vectors*** and ***linear combinations***.

Recall how we defined the number line. The ***continuous*** number line from –∞ to +∞ is the geometric representation of the ***complete*** algebraic set of Real numbers.

Each and every point on the number line can be uniquely identified with a real number and vice versa. Lines and planes do not have ***direction***. Let us introduce this sense of ***direction*** to a line. Let us define something with ***direction*** called ***vector*** anchored at a point on this line. Usually the point we select as the anchor point we label it zero and call it the origin. For simplicity let this vector be of ***unit length***. Let us call this vector v.

Now any point on the number line can be uniquely identified by **a** · v, where **a** is a real number, $-\infty < a < +\infty$, called a ***scalar***.

For ease and consistency in our calculations we associate the + sign with the direction in which the vector points.

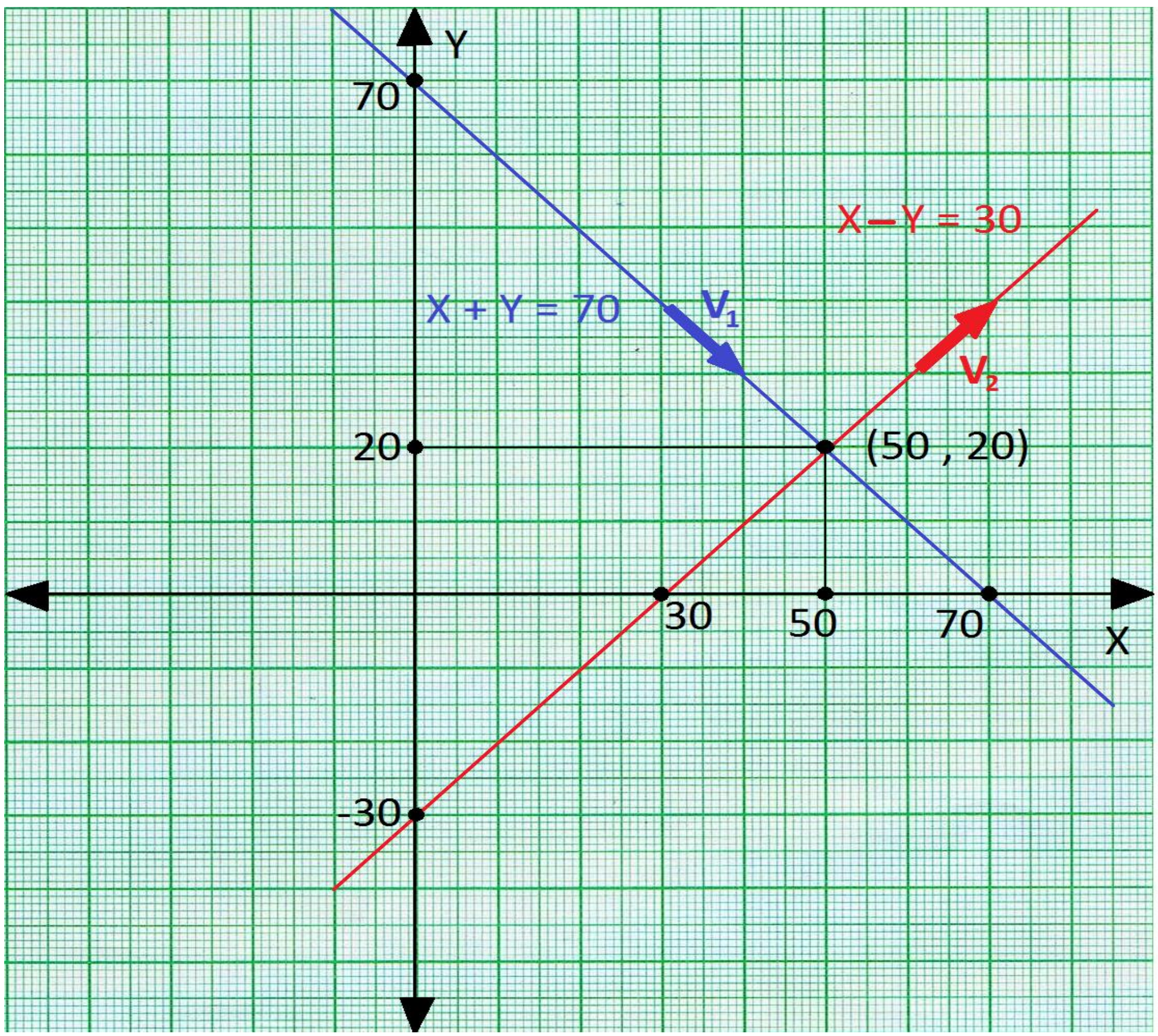
Unlike lines which are continuous entities, ***vectors*** are ***discrete*** and ***finite*** entities.

Rene Descartes (1596 – 1650)

"Cogitate, Ergo Sum"

"I Think, Therefore I Am."

Let us now apply this notion of ***vectors*** to the x + y = 70 and x – y = 30 lines and see how we can get from the ROW picture to the COLUMN picture.



Let us give the line x + y = 70 a ***direction***. Now we can define a ***vector*** anchored somewhere on this line. For simplicity let this vector be of ***unit length***. Let us call this vector v_1.

Now any point on this line x + y = 70 can be uniquely identified by **a** $\cdot$ v_1, where **a** is a real number, $-\infty <$ **a** $< +\infty$, called a ***scalar***.

Likewise, let us define the vector v_2 of unit length along the line x – y = 30. Again any point on this line can be uniquely identified by b $\cdot$ v_2 where b is a ***scalar***.

So far we have not said anything about where exactly along these lines the vectors v_1 and v_2 are defined. For the moment let us assume that each vector is fixed at (anchored at) some point on its line. It does not have to be a common point. Later we shall define a special point and call it the ***origin***.

We are almost done with the use of the concepts of ***scalars*** and ***vectors.*** Let us now employ the concept of ***Linear Combinations***.


DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX.
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.
AVCTORE CLAVDIO GASPARE BACHETO
LVTETIAE PARISIORVM,
Sumptibus SEBASTIANI CRAMOISY, via
Iacobæa, sub Ciconiis.
M. DC. XXI.
CVM PRIVILEGIO REGIS.

Work by Diophantus (3rd century AD)

2.2 LINEAR COMBINATIONS

We know that the solution must be a unique point common to both lines. So we can say

$$\mathbf{a} \cdot v_1 = \mathbf{b} \cdot v_2$$

which is $$\mathbf{a} \cdot v_1 - \mathbf{b} \cdot v_2 = 0$$

This can be rewritten as $$\mathbf{a} \cdot v_1 + (-\mathbf{b}) \cdot v_2 = 0$$

The zero on the R.H.S. is of no use. It means exactly what it says: zero information.

So we ask: What **a** and **b** can we find such that

$$\mathbf{a} \cdot v_1 + \mathbf{b} \cdot v_2 = \text{the solution point } (50\ ,\ 20)\ ?$$

We cannot find such **a** and **b** . The solution point (50 , 20) is defined with respect to the ***origin*** (0 , 0). The vectors v_1 and v_2 are not defined so.

The only information we have that we can put on the R.H.S. of the equation is the vector $\begin{bmatrix} 70 \\ 30 \end{bmatrix}$ which is again defined with respect to the ***origin*** (0 , 0).

So we cannot say

$$\mathbf{a} \cdot v_1 + \mathbf{b} \cdot v_2 = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$$

and try to find **a** and **b** .

While the above way of thinking is intuitive and good, it is not useful from the solution point of view. So let us refine it.

2.3 CHOOSING THE RIGHT VECTORS

In 2-D space a line is ***continuous*** information. In 2-D space a line (eg: x + y = 70) is actually a 1-D "***plane***". Lines and planes do not have direction.

We can define one 1-D reference vector along this line in 2-D space by choosing an arbitrary point as ***origin*** and choosing a finite length as the value of its component. For convenience we choose unit length. Since this is a 1-D vector, a special case, the sign (+ or –) of the ***component*** will give it direction along this line (this 1-D "***plane***"). Let us call the vector v_1. Now **any point** on this line can be **uniquely identified** by $\mathbf{a} \cdot v_1$ where **a** is a real number, which is again ***continuous*** information. So We have not lost any information.

In 3-D space a 2-D plane is **continuous** information. We can select an arbitrary point as ***origin*** and define two 2-D reference vectors, say v_1 and v_2. The vectors should not lie on the same line. They should be ***independent*** – one should not be a linear combination of other, in this case $v_1 \neq \mathbf{a} \cdot v_2$, where **a** is real. Again, **any point** in this 2-D plane can be **uniquely identified** as a linear combination of v_1 and v_2 .

Note that in both cases we have a ***finite*** number of vectors, they are ***discrete*** and ***finite*** in their dimensions and each component of each vector is ***finite*** in value.

Also, in both cases, the vectors ***span*** the whole space: each and every point in the space can be uniquely identified as a linear combination of the vectors.

Clubbing the ROWS in ***matrix*** form we can say:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$$

Our x-axis and y-axis intersect at the ***origin*** (0 , 0) .The ***columns*** become our ***vectors*** anchored at the ***origin*** and we re-write this as a ***linear combination*** : $x \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 70 \\ 30 \end{bmatrix}$

where x and y are ***scalars*** and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are the right vectors.

Notice how the operands have changed. Instead of two ***continuous*** lines of ***infinite*** length we have two ***vectors*** that are ***discrete*** and ***finite*** in dimensions. Also each of the components in each vector is ***finite*** in value. Now we can search for ***scalars*** x and y that satisfy this equation.

We get x = 50 and y = 20 , the same solution as we got with the ROW picture.

So instead of searching for a solution that is the point of intersection of lines (ROWS), we search for a solution that is the linear combination of vectors (COLUMNS).

Here we took the column vectors out of the matrix and then searched for ***linear combinations***. Properly speaking the matrix is got by clubbing the columns. We shall see this again, the formation of the matrix from column vectors, when we define a space in chapter 3. We can extend this kind of thinking to n-dimensional spaces.

Note that vectors are ***discrete*** and ***finite*** entities. Vectors are ***discrete*** in their dimension. The dimensions can be 1, 2, 3, · · · , n, · · · and each component is ***finite*** in value.

Thus we can go from Analytical Geometry to Algebra - the Algebra of ***vector spaces*** and ***matrices***. More generally, from ROW picture to COLUMN picture, even though there is no one-to-one correspondence between a ROW and a COLUMN.

2.4 LINEAR COMBINATIONS REVISITED

The Algebra that deals with ***vectors*** and ***linear combinations*** of ***vectors*** is called ***Linear Algebra,*** in particular, the Algebra of ***Vector Spaces***.

Notice how with ***linear combinations*** only two operations on ***vectors*** are allowed.

1. We can multiply a ***vector*** by a ***scalar***.
2. We can add any number of ***vectors***.

The vectors are ***discrete*** and ***finite*** in dimensions. Also each component of each vector is finite in value. Finiteness is one of the requirements to work with Algebra. The other requirement is ***discrete***.

The first time we encounter the concept of ***linear combinations*** is in 4th grade. We are taught to find the GCD without being aware that we are toying with the notion of ***linear combinations*** – a very powerful and fundamental concept.

Euclid (c. 300 BC – Greek Geometrician and educator of Alexandria) was one of the first to recognize this concept. He used it to find the GCD of two positive integers without having to factor them – Euclid's algorithm! See book 7 of the 13 books by Euclid.

http://aleph0.clarku.edu/~djoyce/java/elements/toc.html

These ideas were further developed by the 3rd century AD Greek mathematician Diophantus.

Further reading: Chapter 2, Section 1, Diophantine Equations "***Exploring the Real Numbers***" by Frederick W. Stevenson, © Prentice Hall, 2000, University of Arizona, Tucon, Arizona, USA.

2.5 SIMULTANEOUS LINEAR COMBINATIONS

Each equation is a ***linear combination*** of x and y,

$$x + y = 70 \quad \text{as in} \quad a_1 \cdot x + b_1 \cdot y = c_1$$

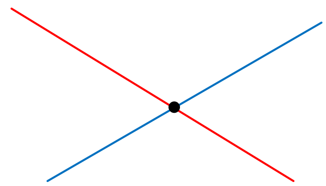
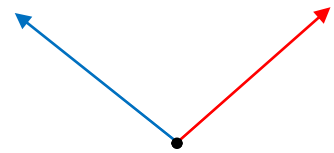
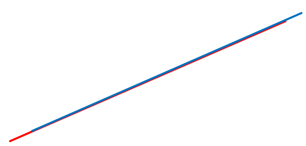
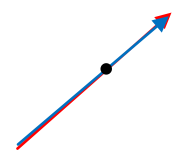
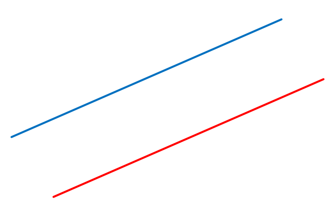
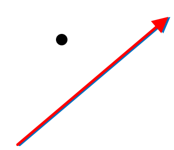
and

$$x - y = 30 \quad \text{as in} \quad a_2 \cdot x + b_2 \cdot y = c_2$$

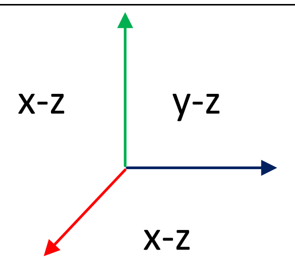
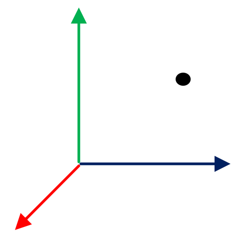
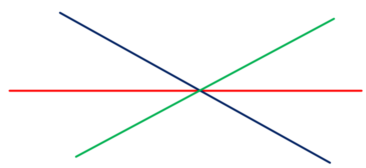
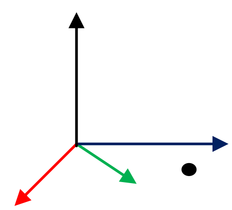
when taken separately has "continuously" many solutions. They form the ***continuous*** or the ***complete*** set of real numbers **R**. But when we take them ***simultaneously*** the possible solutions form a discrete set. There are certain exceptions known as ***singular*** cases as we shall see.

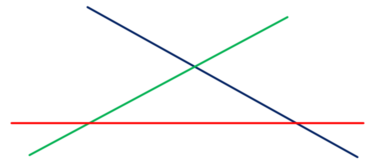
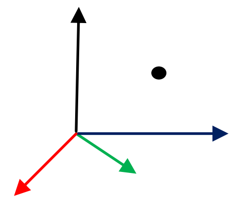
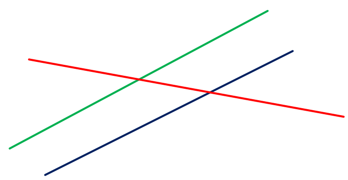
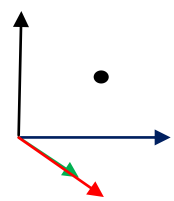
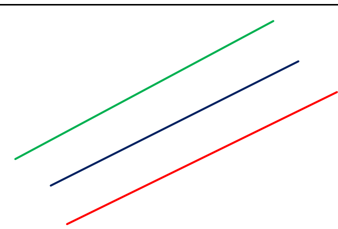
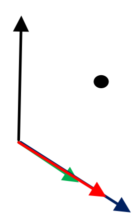
Linear Combinations of the vectors guarantee that whatever happens in the ROW picture has a one-to-one correspondence with the same thing happening in the COLUMN picture even though there is no one-to-one correspondence between a ROW and a COLUMN.

In two dimensions there are 3 possibilities:

possibility	ROW picture	COLUMN picture
unique solution Equations are independent and consistent	 lines intersect at a point	 the vectors are independent
infinitely many solutions Singular case Equations are dependent and consistent	 lines are coincident	 vectors are dependent and solution lies in line with them
no solution Singular case Equations are inconsistent	 lines are parallel	 vectors are dependent and solution lies off the lines

In 3-D, this can be represented as follows, assuming in the ROW picture that the third dimension passes through the plane of paper and perpendicular to it.

possibility	ROW picture	COLUMN picture
unique solution	 three planes intersect at a point, like the x-y, y-z and x-z planes intersect at the origin.	 vectors are independent
infinitely many solutions singular case	 planes intersect at a line. The intersecting planes rise out of the page.	 vectors are dependent and the solution lies in the plane of those vectors.

no solution singular case 1	 two planes intersect at a line. The intersecting planes rise out of the page.	 vectors are dependent but the solution lies off the plane of those vectors.
no solution singular case 2	 two planes are parallel and the third plane intersects them.	 two vectors lie on same line and the solution is off the plane of three vectors.
no solution singular case 3	 all three planes are parallel.	 all three vectors lie on same line and the solution is off that line.

Taking a real life problem, abstracting the information and representing it as vectors in a Vector Space will allow us to use the tools of Linear Algebra to find the solution. Typically the mathematical modeling is in the ROW picture and the solution is found using the COLUMN picture.

Usually when we use the word ***space***, the notion of ***points***, ***lines*** and ***planes*** in 2-D and 3-D space comes to mind. But the word ***space*** can be used in a more general or generic sense.

As an example we may define a 3-dimensional ***name space*** where the three dimensions are ***first name***, ***middle name*** and ***last name***. Each full name can be represented as a point in this 3-D ***name space***.

We may extend this way of thinking and apply it to indexed information to form a multidimensional ***Information Space***. Each index is an axis from a set of keywords. We may search for some particular information in this ***Information Space*** using some keywords. This is the basic idea behind search engines.

Ref. Fadeeva, V. N., Computational methods in Linear Algebra, Dover, New York, 1959.

Main ref.: Lecture notes on Linear Algebra by S. Vadakkan.

Further reading: Hermann Grassmann and the Creation of Linear Algebra by Desmond Fearnley-Sander (an expanded version of a paper presented at the 1979 Summer Research Institute of the Australian Mathematical Society).

Further reading: Linear Algebra by Jim Hefferon, 13-Aug-2008, Saint Michael's College, Colchester, Vermont, USA 05439.

2.6 INSIGHT ≈ FUNDAMENTALS + INTUITION

Sets and relations are the most fundamental concepts in mathematics.

The ROW picture is a ***relation*** as in ordered pairs or ordered n-tuples. The COLUMN picture is also a ***relation***. We mentioned earlier that there is no one-to-one correspondence between a ROW and a COLUMN.

Linear Combinations of vectors help us to preserve or carry forward the ***relations*** in the ROW picture to ***relations*** in the COLUMN picture even though there is no one-to-one correspondence between a ROW and a COLUMN.

We may think of a row in the ROW picture as horizontal information defining relationships across a generation between siblings and first cousins, like **a** is the brother of **b** , **b** is the sister of **c**, … . We may then try to infer or extract from the ROW picture: is **x** the cousin twice removed of **y**?

A column in the COLUMN picture maybe viewed as vertical information defining relationships across a hierarchy of generations as in grandfather, father, son. Given all the columns that we can form from the ROW picture we can try to extract the same information.

Later you may see this vertical relationship of hierarchy of generations in the form: $\begin{bmatrix} position \\ speed \\ acceleration \end{bmatrix}$ or other forms as columns in matrices.

Further reading: Data Structures and Algorithms by Niha N. Sheikh *et al*, 2009

So far we have only developed a different view of the mathematical model of the ROW picture. We call this view the COLUMN picture.

We have not defined any operations or algorithms to extract information from this new view. In the ROW picture we have the Gaussian elimination algorithm to extract information. We may put the COLUMNS together and get a matrix. Now we may think of operations or algorithms on matrices like factoring a matrix into ***lower triangular*** and ***upper triangular*** matrices **(LU)** or into ***lower triangular***, ***diagonal*** and ***upper triangular*** matrices **(LDU)** to extract information. This is the realm of Linear Algebra.

Euclid (325 BC – 265 BC)

3. SPACE and BASIS VECTORS

To define a Space all we need is a *system of axes* (frame of reference) – origin and direction are taken care of, and a *unit length* – magnitude is taken care of. Now any point (x, y) in the Space can be uniquely identified as a vector – magnitude and direction.

The point (x, y) is uniquely identified with respect to the *system of axes*. We may also refer to or uniquely identify the point (x, y) as a Linear Combination of a set of *basis vectors*. Thus

$$(x, y) = x \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

in the x-y *system of axes*.

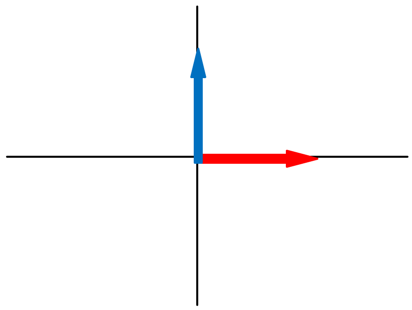
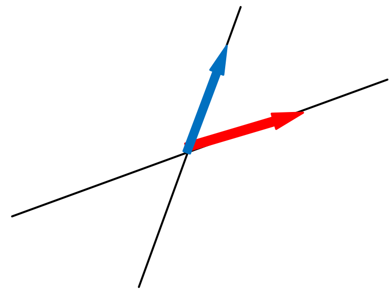
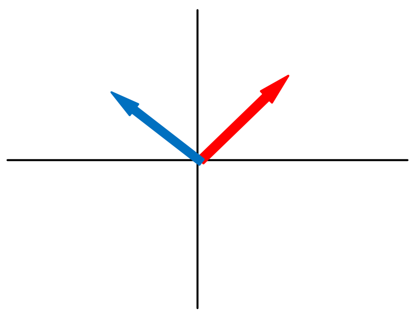
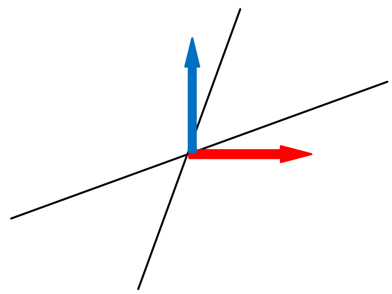
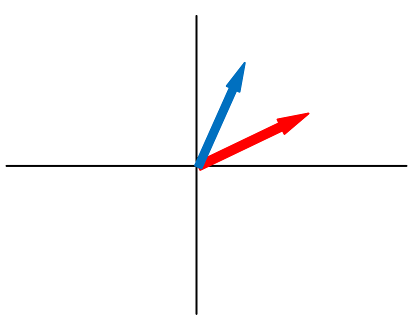
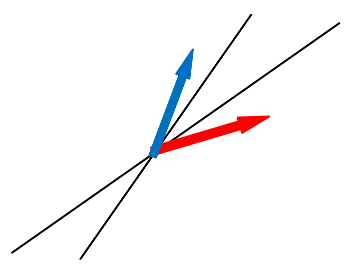
The system of axes in 2-D space is the standard orthogonal (or rectangular co-ordinate) system of the x-axis and y-axis.

The *basis vectors* are:

$$e_x = (1, 0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad e_y = (0, 1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Notice how the basis vectors are defined with respect to the system of axes. Usually we choose the basis vectors along the system of axes. But this is not necessary. One or more or even all the basis vectors need not be along an axis in the frame of reference. To keep the calculations simple we usually choose basis vectors of unit length.

The system of axes does not have to be orthogonal. The axes could be *skewed*.

 Both Frame of Reference and Basis Vectors are orthogonal and in line	 Both Frame of Reference and Basis Vectors are skewed and in line
 Both Frame of Reference and Basis Vectors are orthogonal, but not in line	 Basis Vectors are orthogonal, but Frame of Reference is skewed
 Basis Vectors are skewed, but Frame of Reference is orthogonal	 Both Frame of Reference and Basis Vectors are skewed, but not in line

There are two fundamental aspects of *basis vectors*:

1. Basis vectors are *independent*. A basis vector cannot be expressed as a *linear combination* of the other basis vectors.
2. Any point in 2-D space can be uniquely identified by a *linear combination* of the basis vectors e_x and e_y.

These notions can be easily generalized and applied to n-Dimensional space.

More formally, the space with points uniquely identified as *linear combinations* of basis vectors is a Vector Space. Thus point and vector are synonyms.

A few more points concerning *basis vectors*:

1. The set of basis vectors is not necessarily unique.
2. The basis vectors need not be orthogonal. They may be *skewed*.
3. A basis vector need not be of unit length.

Thus the point s = (2, 2) in the 2-D plane (defined by an orthogonal system of axes) may be represented as a linear combination of different sets of basis vectors:

$$e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \; : \quad s = 2 \cdot e_x + 2 \cdot e_y$$

$$e_{2x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \; e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \; : \quad s = 1 \cdot e_{2x} + 2 \cdot e_y$$

$$e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \; : \quad s = 2 \cdot e_1 + 0 \cdot e_y$$

$$e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \; : \quad s = 0 \cdot e_x + 2 \cdot e_1$$

4. SPACE and MATRIX

A Space may be represented by an Identity Matrix and a set of basis vectors.

For 2-D space the Identity Matrix $\mathbf{I}_2$ and basis vectors are:

$$\mathbf{I}_2 = \begin{array}{l} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array}\begin{array}{l} \text{"plane" } x = 0 \\ \text{"plane" } y = 0 \end{array} \qquad e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

For 3-D space the Identity Matrix $\mathbf{I}_3$ and basis vectors are:

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{array}{l} \text{y-z plane} \\ \text{x-z plane} \\ \text{x-y plane} \end{array} \qquad e_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ e_y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } e_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Each row represents a ***plane*** in the ***space*** and each column is an axis in the ***space***. Looking back at the father's age – son's age example: column 1 is the father's age axis and column 2 is the son's age axis.

The particular values as in father's age = 50 and son's age = 20 is known as the ***state***.

Together this is known as the ***State-Space*** view

Notice how the basis vectors $e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ appear as columns in the matrix $\mathbf{I}_2$ with the magnitudes of the basis vectors along the ***diagonal***.

Likewise, in matrix $\mathcal{I}_3$ we note the columns are the basis vectors $e_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $e_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, with the magnitudes of the basis vectors along the ***diagonal***.

We know that in both 2-D space and 3-D space there is the standard orthogonal system of axes with the basis vectors lying along those axes.

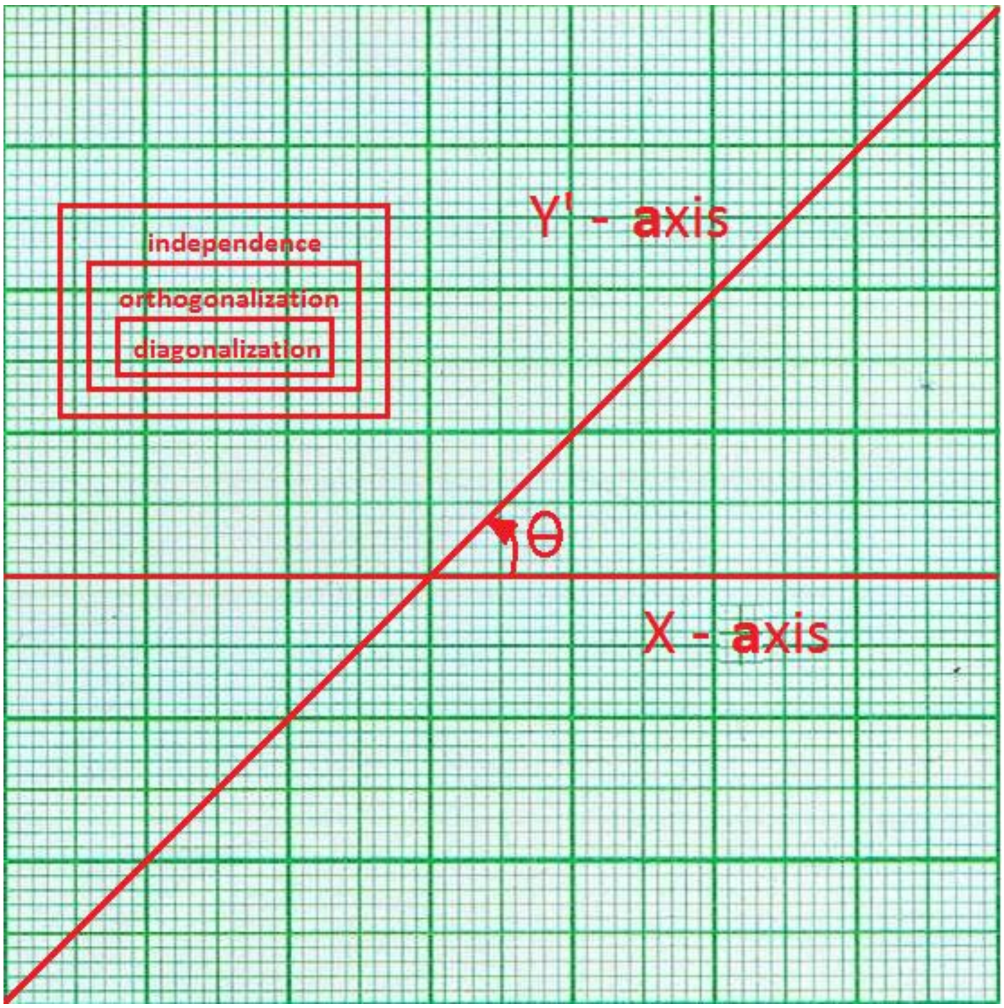
In other words, the matrix that represents the space and the set of basis vectors (more fundamentally – the system of axes on which they lie or defined with respect to) go hand-in-hand.

Once we have selected a *frame of reference* (system of axes) we would like it to remain fixed. There are two things that can affect it.

1. ***Translation:*** In this case we lose our origin. So we disallow this.
2. ***Rotation*** of the axes***:*** If one or more of the axes rotate we lose our sense of direction.

This we can identify and then try to choose a system of axes (frame of reference) that does not rotate or shift direction. Here we used the word ***rotation*** in a general way to represent any change or shift in direction. Later we shall see a very specific action known as ***rotation***.

Rotation implies ***skew,*** and ***skew*** implies a kind of dependence or ***coupling***. The axes are not entirely independent of each other – not orthogonal as in:



So y’ has direction component cosθ along the x axis. A change along one axis causes a change along the other.

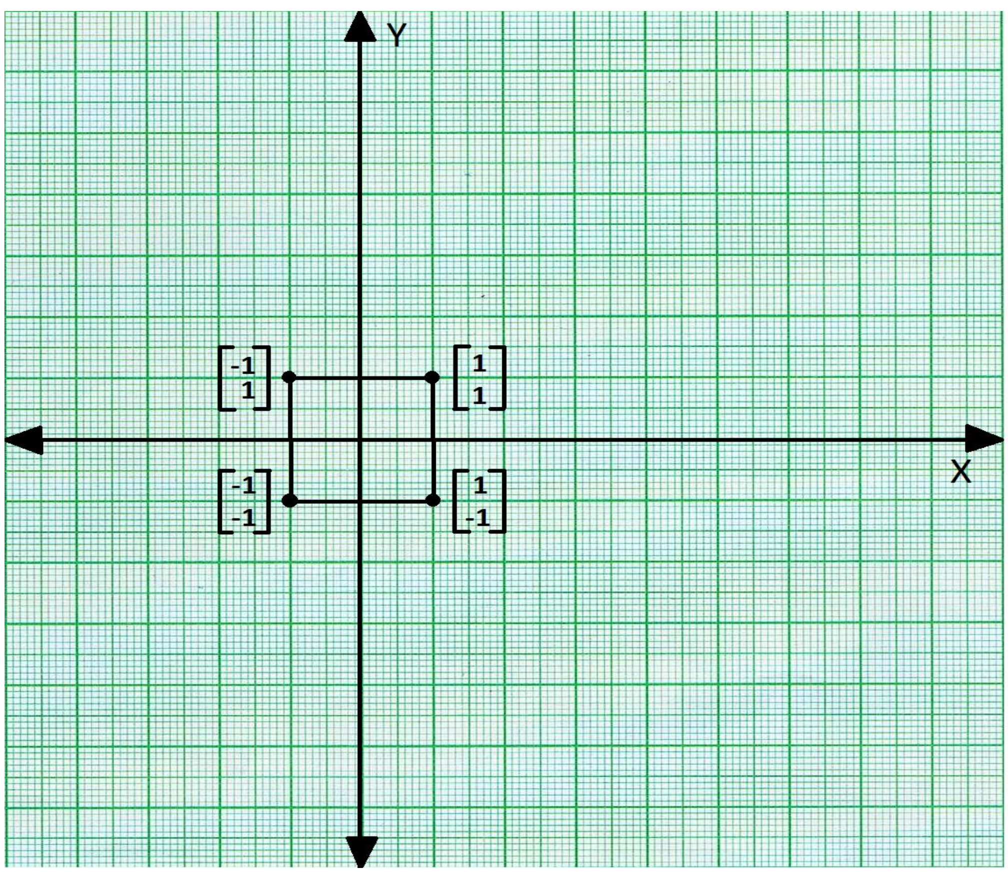
Separating the axes, getting rid of this dependence, is known as ***decoupling*** or ***orthogonalization***. ***Orthogonalization*** is a special case of independence. ***Diagonalization*** is a very special case of ***orthogonalization.*** Once the axes are decoupled we may play around with changes along one axis at a time and study the behavior of a system. The application of numerical methods becomes easy.

5. The ACTION of a Matrix on a Space

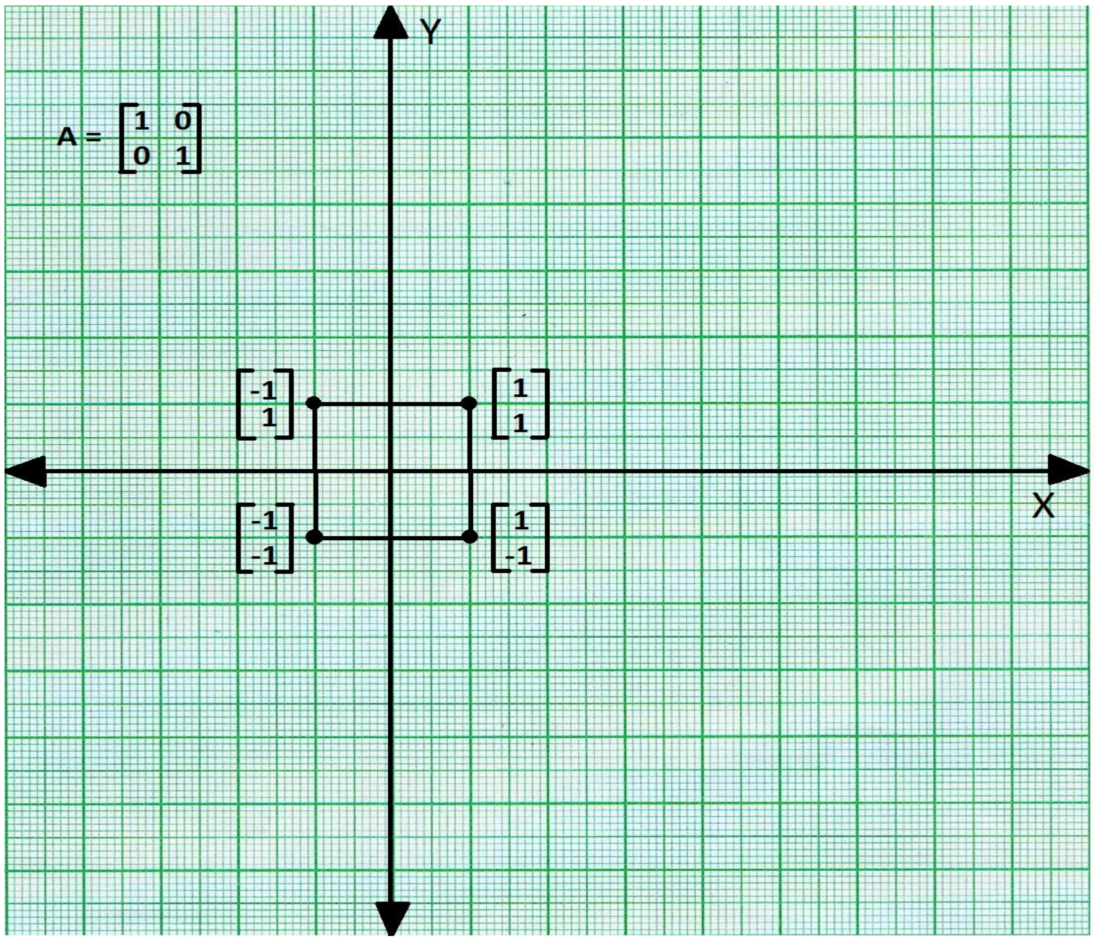
In addition to a matrix representing a space, a matrix also ***acts*** on a space. Some of the ***actions*** are: stretch (scaling), shear, project, reflect and rotate. These ***actions*** deform or ***transform*** the space and hence are called *transformations*.

One action we disallow is *translation*. Translation implies/requires a shift in the origin. But once the origin is shifted we lose control over the space – lost in space/spaced out. However, in computer graphics there are techniques (fancy word for tricks) to achieve this effect.

Let us see a few examples of actions on 2-D space by seeing what happens to a small sample of the space - the square with corners (1, 1), (-1, 1), (-1, -1) and (1, -1) located around the origin.



Example 1: apply the action matrix to the corners.

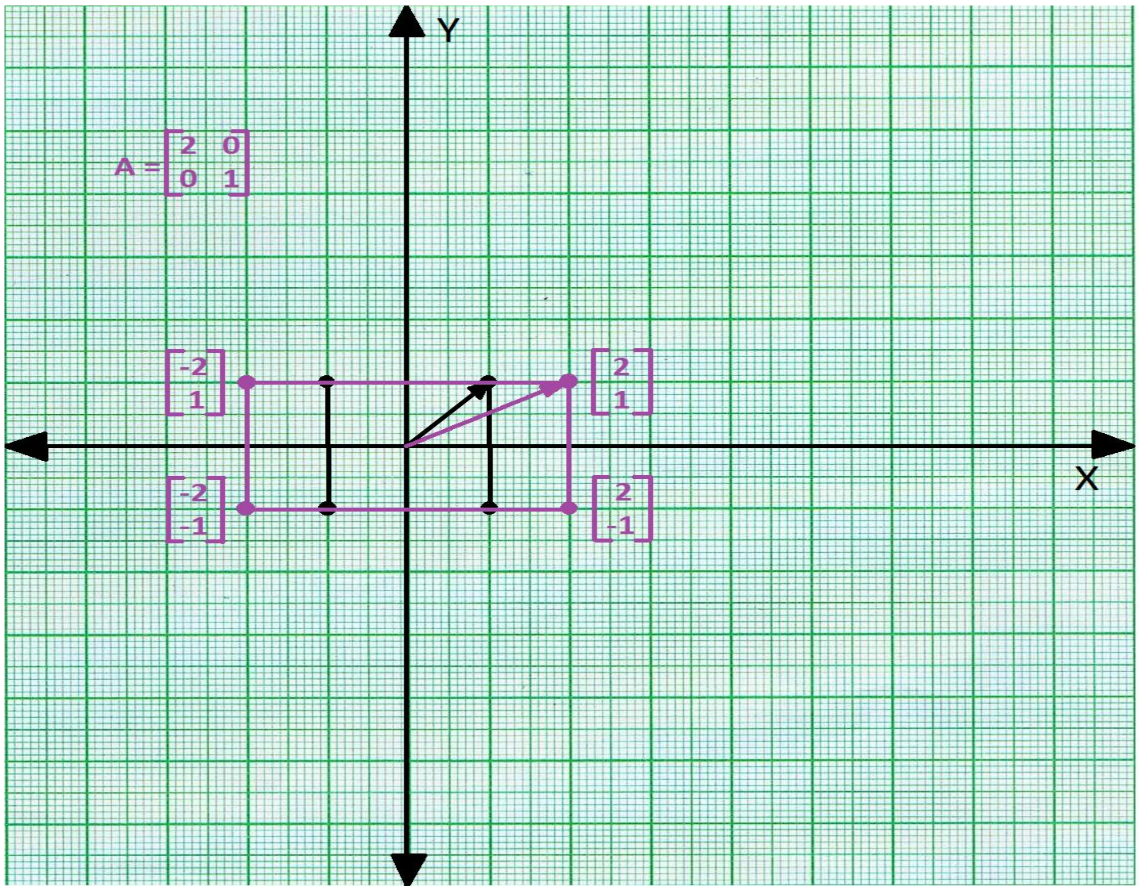


$$A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$A \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $A \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$; $A \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$; $A \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

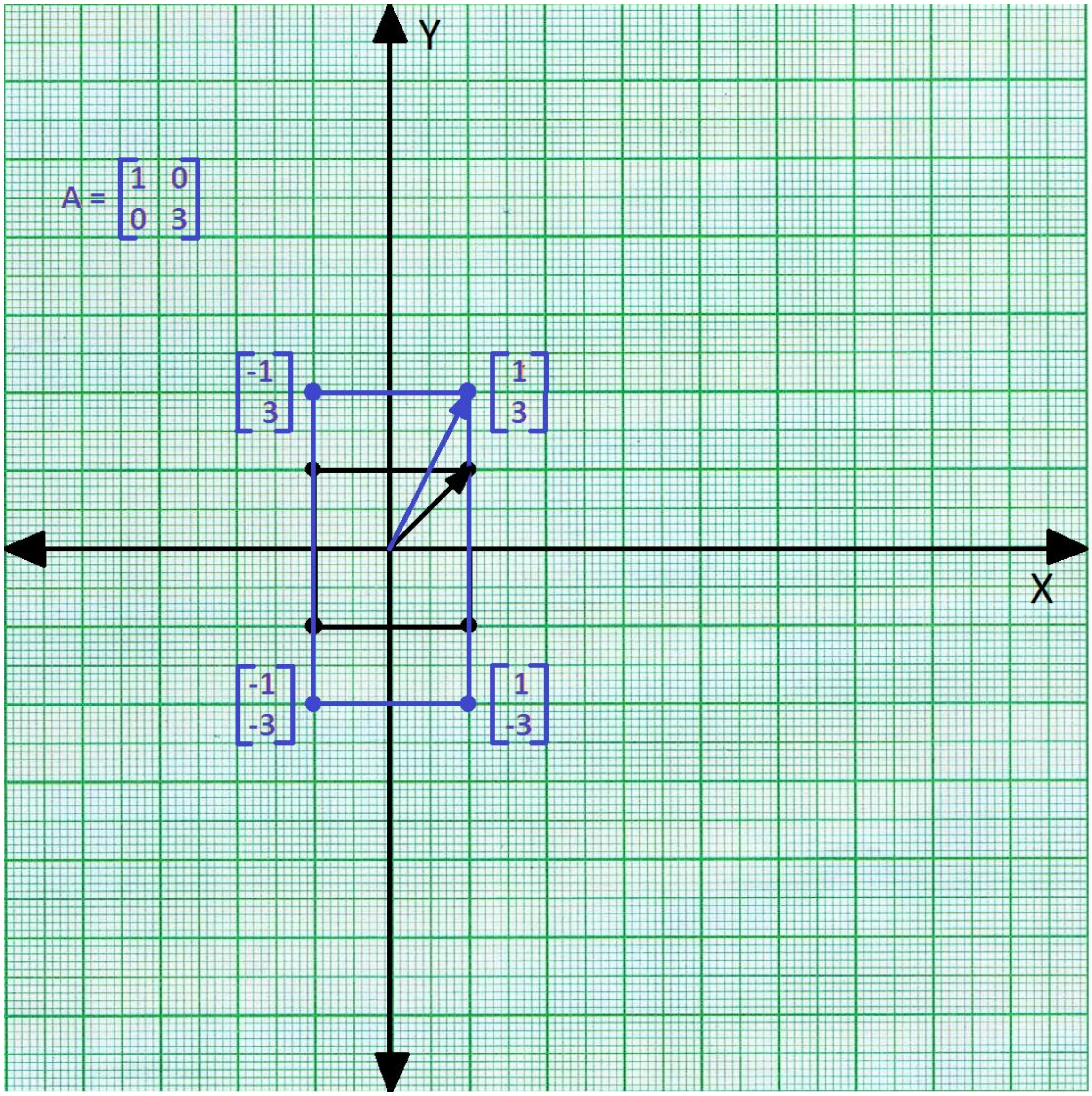
The square remains unchanged. This is the property of the ***identity matrix.***

Example 2: Stretch along x-axis.



Note how the square got stretched along the x-axis. Notice how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became $\begin{bmatrix}2\\1\end{bmatrix}$ due to the action of A. But the points along the x-axis still lie along the x-axis even though they have been stretched and the points on the y-axis still lie on the y-axis. In other words neither of the axes in the frame of reference has ***shifted*** or ***rotated*** despite the action of the matrix.

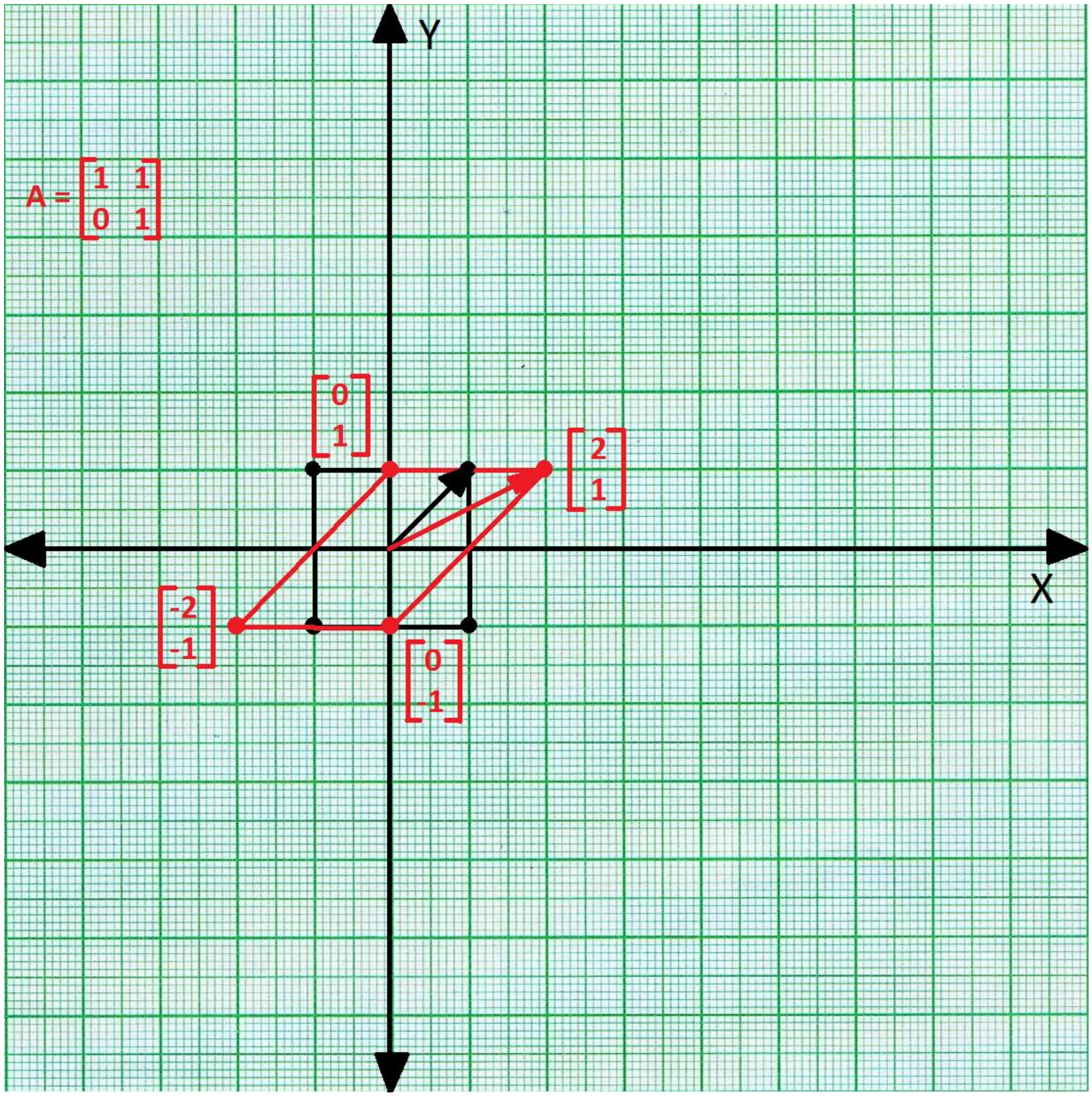
Example 3: Stretch along y-axis.



Note how the square got stretched along the y-axis. Notice how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became $\begin{bmatrix}1\\3\end{bmatrix}$ due to the action of A.

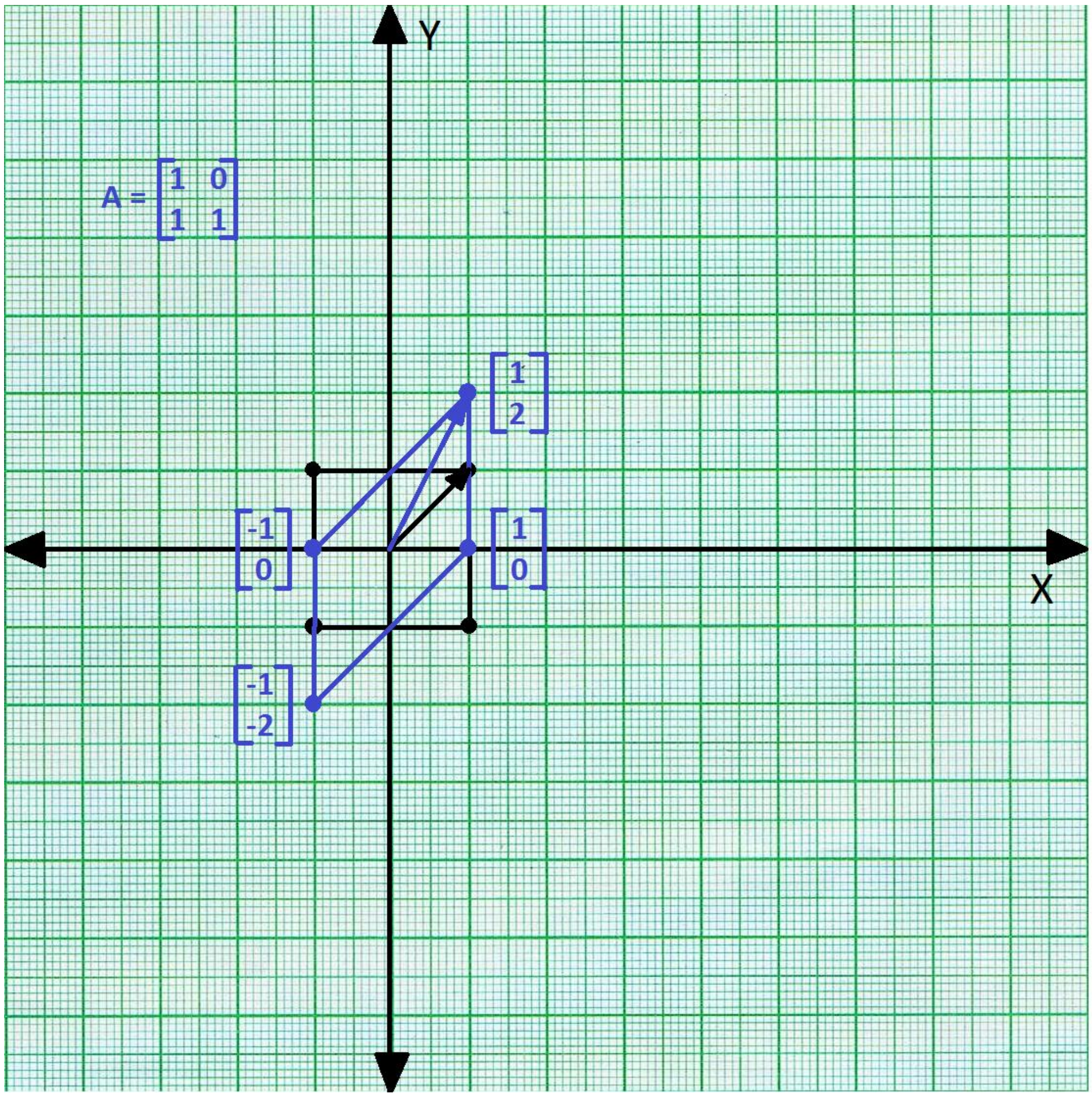
Again the x-axis and y-axis have not ***shifted*** or ***rotated***.

Example 4: Shear along y-axis.



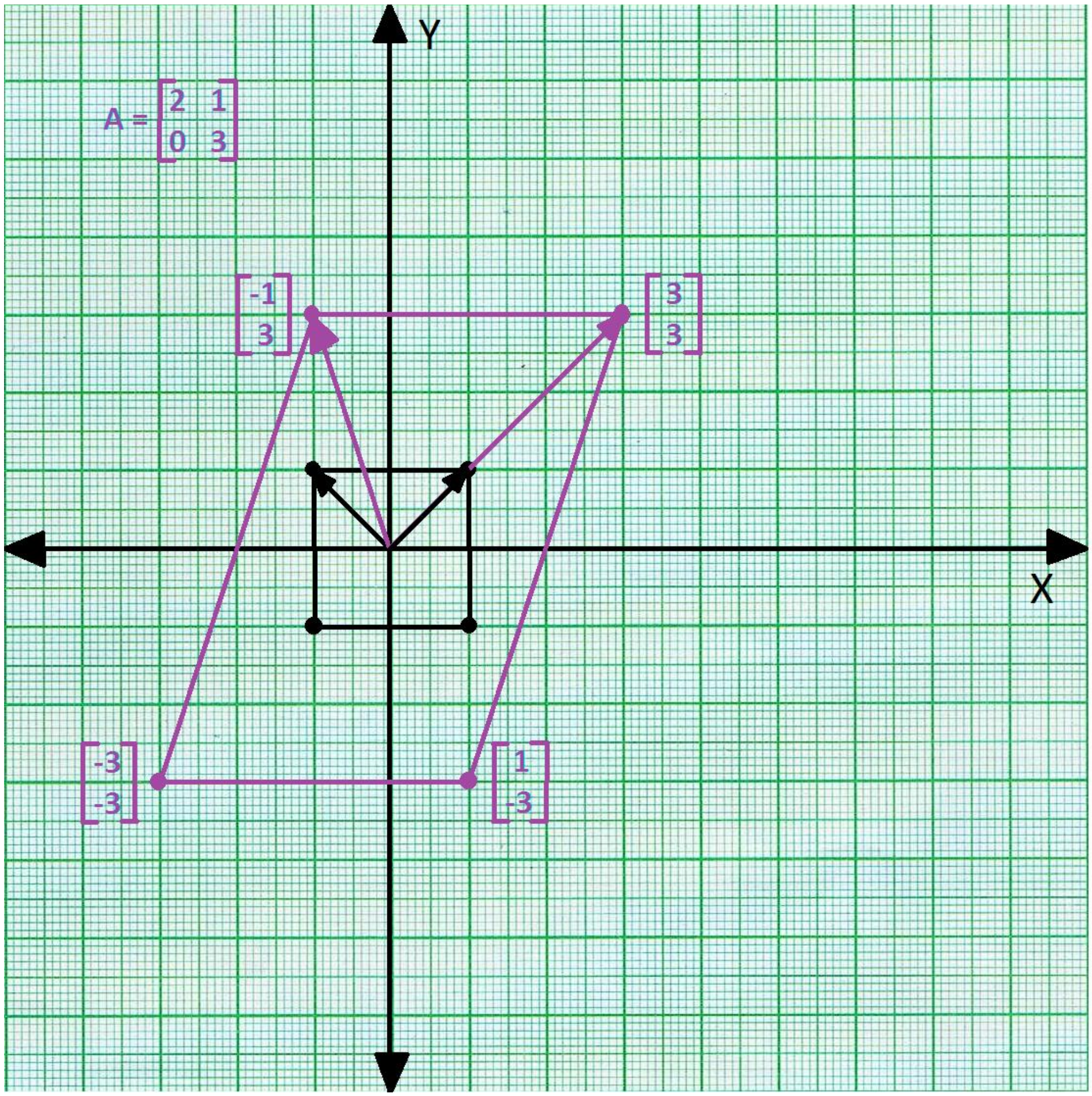
Notice how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became $\begin{bmatrix}2\\1\end{bmatrix}$ due to the action of the matrix A. Also note how the point $\begin{bmatrix}0\\1\end{bmatrix}$ lying on the y-axis shifted to the point $\begin{bmatrix}1\\1\end{bmatrix}$. The whole y-axis has ***shifted*** or ***rotated***.

Example 5: Shear along the x-axis.



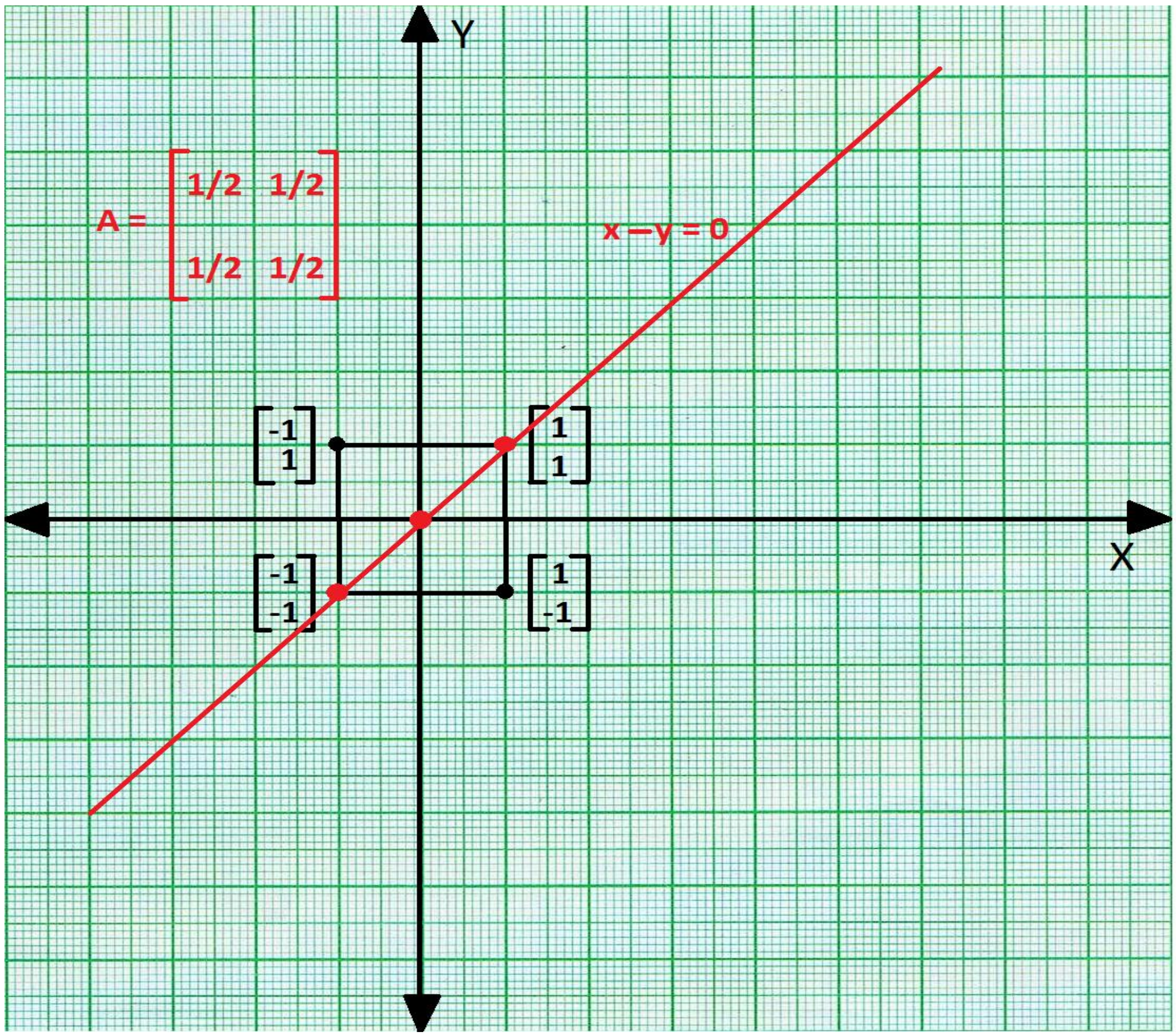
Notice how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became $\begin{bmatrix}1\\2\end{bmatrix}$ due to the action of the matrix A. Here, the whole x-axis has ***shifted*** or **rotated**. Also notice how the change in position of a single value (a_{12} became a_{21}) caused a different action.

Example 6: We may combine two stretching actions and a shearing action into one matrix.



Note how the square got stretched along the both the x-axis and y-axis. Notice how the vector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ rotated and became $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ due to the action of A. What can you say about the points lying on the x-axis and y-axis?

Example 7: Projection onto the line y = x.

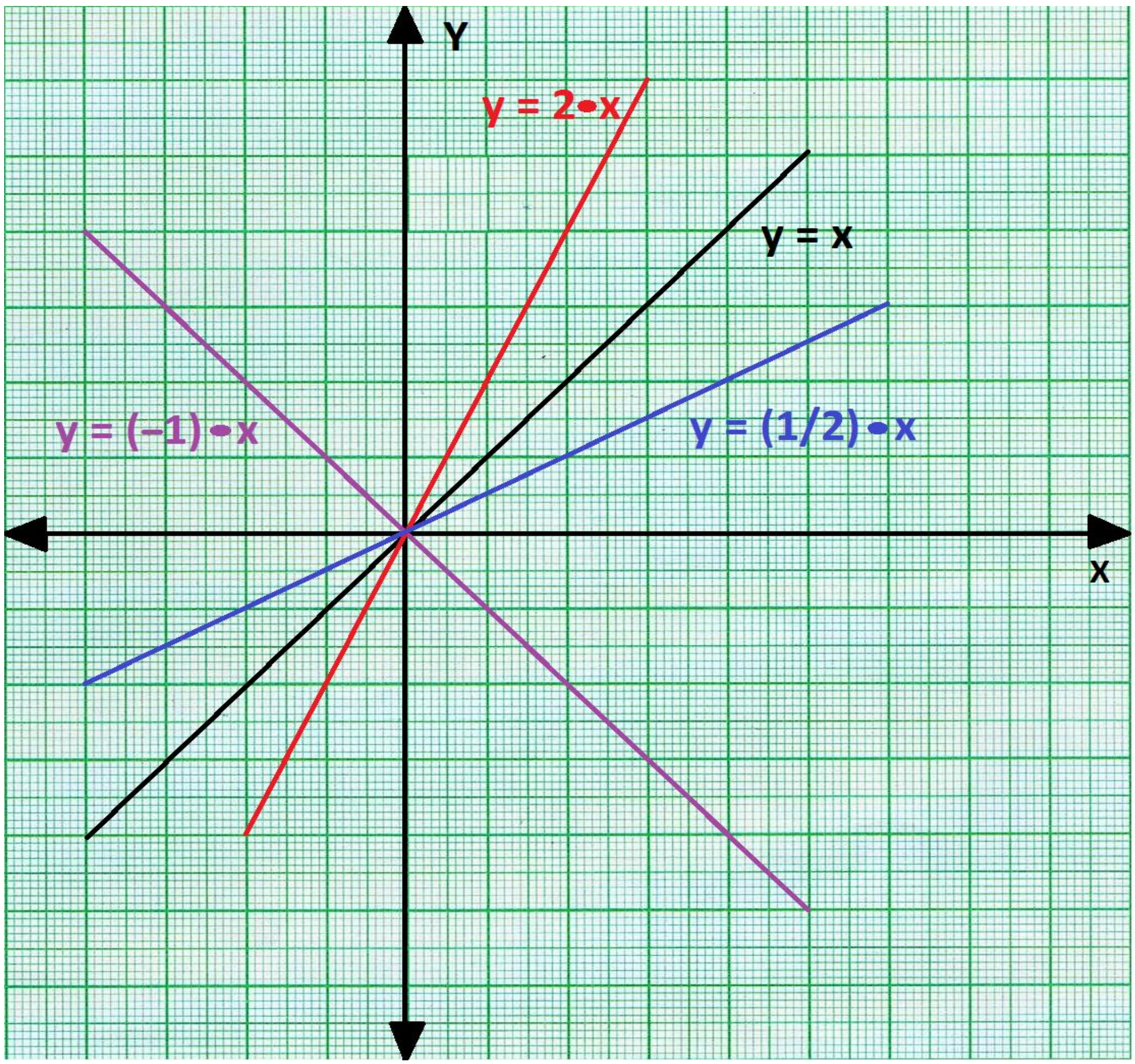


The entire 2-D space (the x-y plane) is projected onto the line x – y = 0. Obviously both the x-axis and y-axis rotated. Which are the points (more precisely line of points – axis) that remained ***invariant*** due to the action of the matrix?

It is possible to demonstrate some of the actions (like shear, projection and rotation) in 3-D space using a ***cuboid carton*** (with the top and bottom removed) as a sample of the 3-D space centered around the origin.

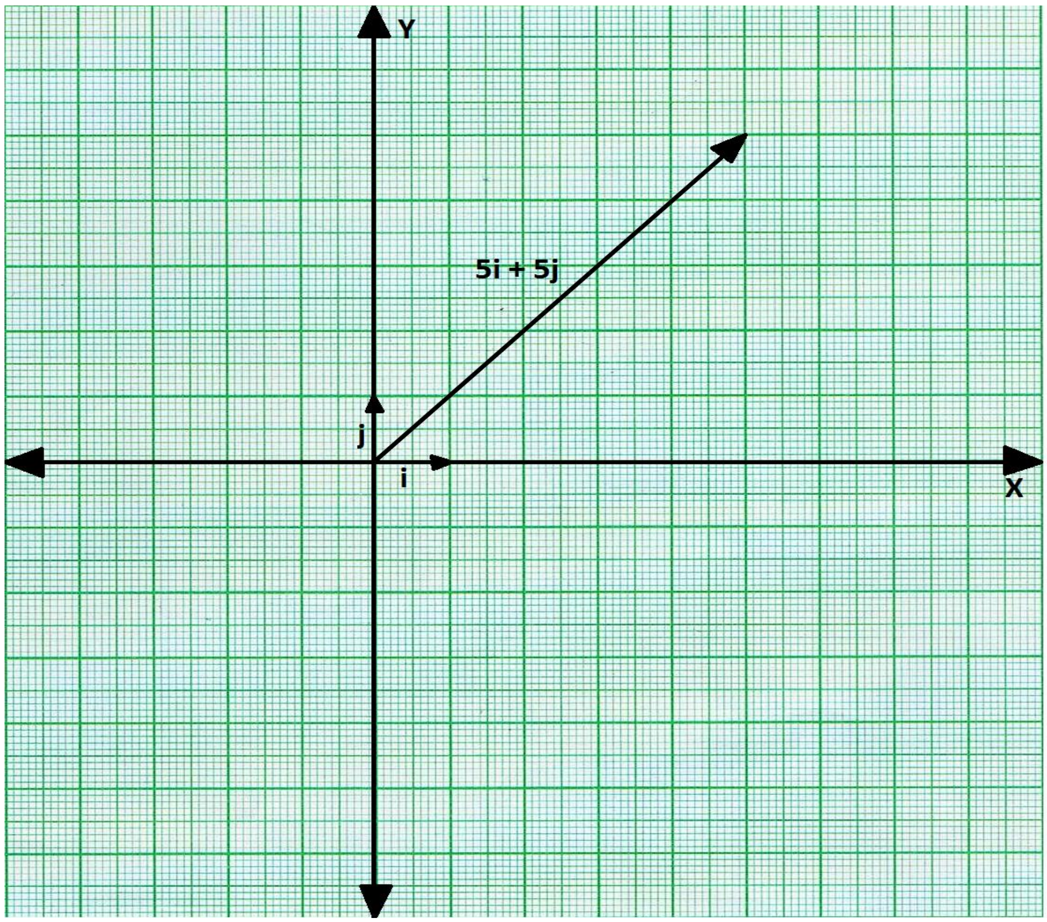
6. **Rotate a Line** versus **Rotate a Vector**

It is easy to rotate the line y = x. Just multiply it by a suitable scalar.



How can we rotate a vector?

Consider the vector v = 5i + 5j

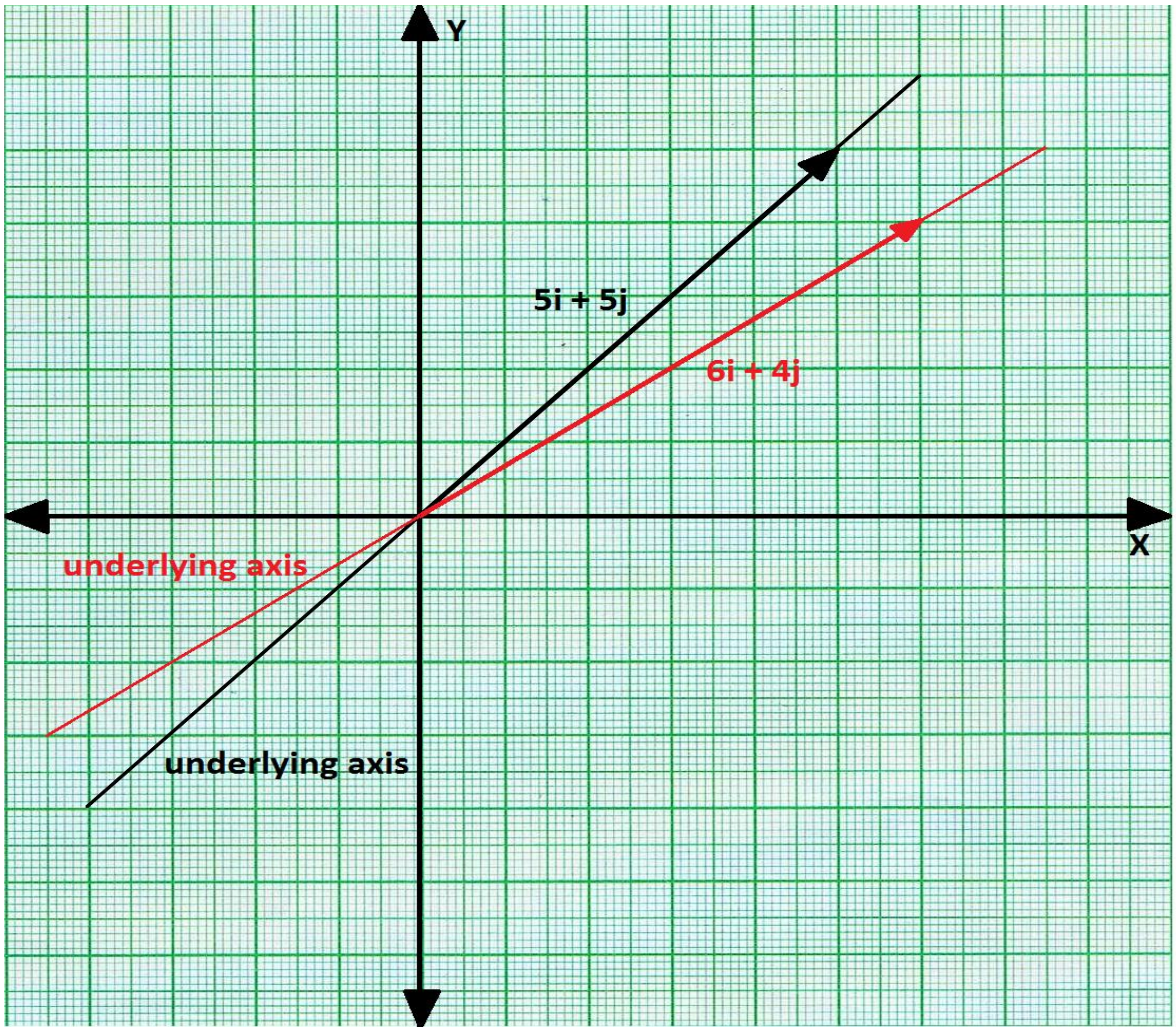


Multiplying ***v*** by a scalar will only change its magnitude. It will not rotate or change direction or shift from the line (axis) on which it lies. Even if we multiply ***v*** by – 1 to get ***–v*** it will still lie on the same line (axis) but point in the opposite direction.

To rotate a vector we have to add or subtract something from the components. At this point let us not worry about the magnitude. Thus:

$$\begin{aligned} \boldsymbol{v} &= 5i + 5j \\ &+ 1i - 1j \\ \text{becomes } & 6i + 4j \end{aligned}$$

The vector has rotated, changed direction or shifted from the line (axis) on which it was lying.



If we now review Example 2: stretch along the x-axis, we observe how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became vector $\begin{bmatrix}2\\1\end{bmatrix}$

Likewise if we look at Example 3: Stretch along the y-axis, we observe how the vector $\begin{bmatrix}1\\1\end{bmatrix}$ rotated and became $\begin{bmatrix}1\\3\end{bmatrix}$

We note three main points:

1. An action matrix transforms (changes the shape of) a space. A change in just a single element in the matrix, be it in value or just position, causes a different action.
2. An action matrix can cause a vector to rotate, including maybe the basis vectors.
3. Since the basis vectors are represented along the diagonal, we intuit that we have to add or subtract something along the diagonal to neutralize or cancel out the rotation caused by the action of the matrix.

How do we compute how much to add or subtract from each component?

What is the calculation that takes into account each and every value in the matrix and its position as well?

Linear Algebra offers us the calculation of the ***determinant*** of a matrix. The method of calculation of the ***determinant*** of a matrix takes into account each and every value of the matrix and its position as well.

7. Determinant of a Matrix

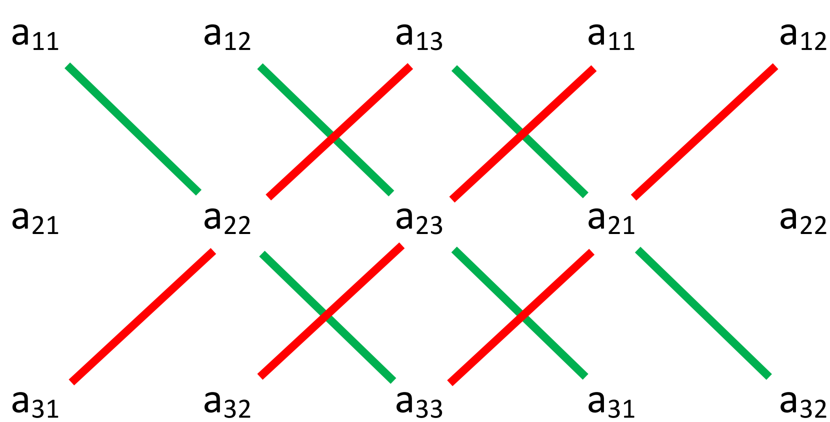
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad \text{det. } A = +a_{11}a_{22} - a_{12}a_{21}$$

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

$$\text{det. } A = +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

$$- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

This is got by application of Sarrus' rule



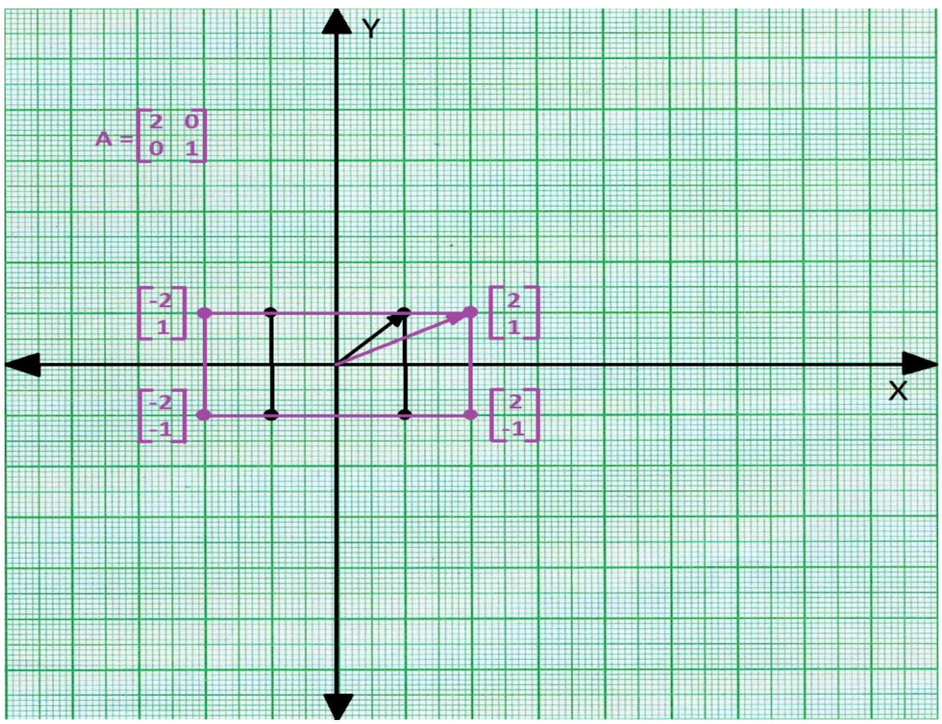
The products of the numbers connected by green lines are added and the products of the numbers connected by the red line are subtracted.

The determinant is a linear function of the entries of each row. The determinant can be computed from the elements of any fixed row and their *cofactors*. The *cofactors* can be computed from *minors*.

8. Eigenvalues and Eigenvectors

Let us now use the method of the calculation of the ***determinant*** to compute the values denoted by λ and called *eigenvalues*, how much to add or subtract from each component to neutralize the rotation caused by an action matrix, and to find an axis along which there is no rotation.

Let us consider ***Example 2*** again: Stretch along the x-axis.



$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{becomes} \quad \begin{bmatrix} 2-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix}.$$

The calculation of the ***determinant*** gives us a polynomial in λ.

$$\lambda^2 - 3\lambda + 2$$

This is known as the ***characteristic polynomial***.

The roots of this polynomial are:

$$\lambda_1 = 2 \quad \text{and} \quad \lambda_2 = 1.$$

What is the meaning of these values? How do we use them?

In Calculus the first derivative = 0 means there is an ***extremum*** or a point of inflection.

In Geometry, when we draw the graph of a polynomial, the real roots are the intercepts on the x-axis.

In Linear Algebra the roots of the ***characteristic polynomial***, the *eigenvalues*, are used to find a new set of basis vectors known as the *eigenvectors*.

What is the advantage of the eigenvectors?

We saw how a matrix acts on a space and rotates vectors. So a matrix may rotate even the basis vectors – shift them off the system of axes, cause them to change direction.

More fundamentally, since the system of axes is our frame of reference (origin and direction) we would like our system of axes to remain unchanged (invariant) by the action of a matrix.

We mentioned earlier that we disallow a ***translation*** operation because we will lose the origin of the frame of reference. Now, if an axis in the frame of reference (system of axes) rotates we will lose its direction.

So had we used the eigenvectors as our basis vectors with the underlying system of axes (i.e. the system of axes on which the eigenvectors lie, which we call the eigenaxes) hopefully the action matrix will not cause them to shift axis or change direction. They will lie on the same axes as before. *Eigen =* ***same*** *or* ***characteristic*** *in German.*

With this in mind we have matrices

$A - \lambda_1 I = \begin{bmatrix} 2-2 & 0 \\ 0 & 1-2 \end{bmatrix}$ and $A - \lambda_2 I = \begin{bmatrix} 2-1 & 0 \\ 0 & 1-1 \end{bmatrix}$

The solutions to $(A - \lambda_i I) \cdot x = 0$ are eigenvectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$.

Note that $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ is on the same axis as $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

So now we have a new set of basis vectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

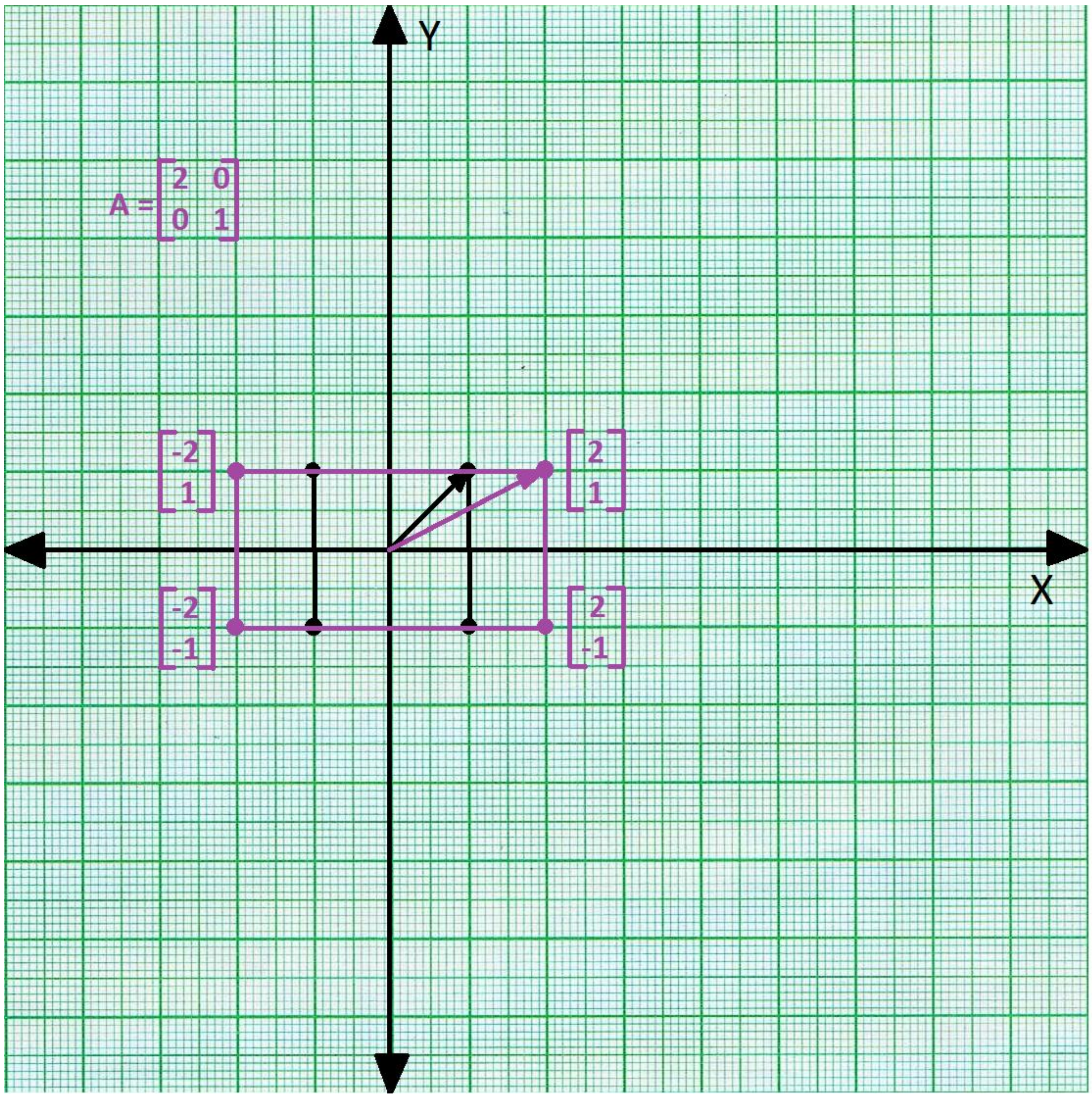
What have we gained? What is the difference between this set of basis vectors and the set of basis vectors $e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$?

If we look more closely we see that $e_1 = e_x$ and $e_2 = e_y$.

In other words, the ***underlying system of axes*** is the same (***invariant***). But we half expected this because our intuition tells that a stretch action on one axis alone does not introduce any kind of rotation of a vector along the other axis. Here the basis vectors are *orthogonal* or *decoupled*.

The system of axes on page 29 is not orthogonal. Assume the basis vectors lie on the axes. Will the action of A cause axis y’ to rotate or change direction?

In practical terms, for a stretch action like A = $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$



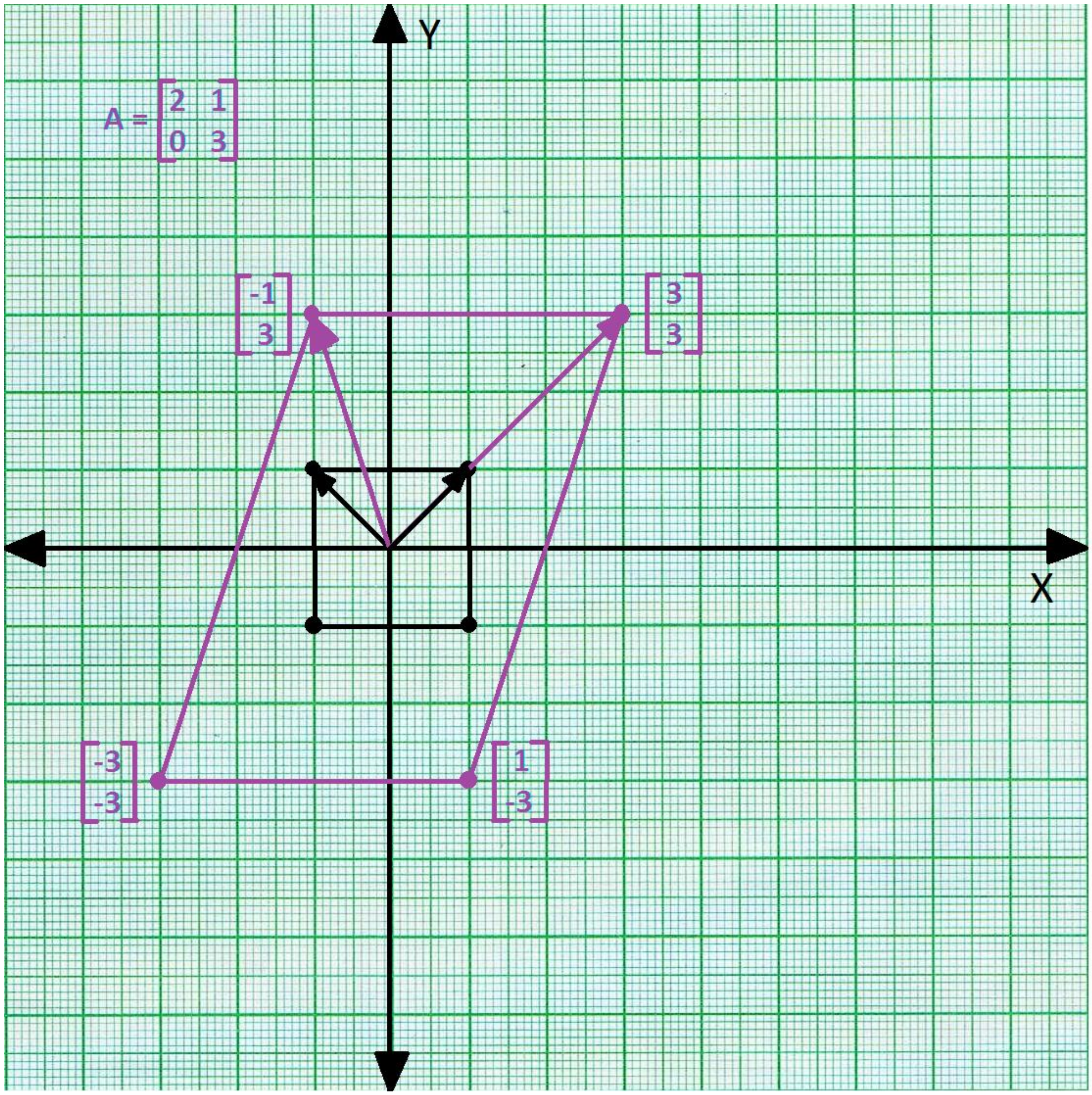
acting in the x-direction, only the points on the x-axis and the y-axis will not rotate or shift axis. Likewise, only the points along the x-axis and y-axis will not rotate if stretched only in the y-direction.

The ***eigenvalues*** help us to find the ***eigenvectors***. The eigenvectors help us to identify the underlying ***eigenaxes***.

eigenvalues $\longrightarrow$ ***eigenvectors*** $\longrightarrow$ ***eigenaxes***

Note that for large matrices trying to find the eigenvalues (roots of the characteristic polynomial) by first computing det $(A - \lambda_i I)$ is not the best way. Galois and Abel proved that there can be no algebraic formula for the roots of a polynomial of degree five and higher in terms of its co-efficients (Abel-Ruffini Theorem).

Let us consider ***Example 6*** again.



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

The det. of $\begin{bmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{bmatrix}$ is the characteristic polynomial

$$(2 - \lambda) \cdot (3 - \lambda) - 0 \cdot 1 = \lambda^2 - 5\lambda + 6$$

The roots are $\lambda_1 = 2$ and $\lambda_2 = 3$.

$$A - \lambda_1 I = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } A - \lambda_2 I = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}.$$

The solutions to $(A - \lambda_i I) \cdot x = 0$ yield eigenvectors

$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, a set of *skewed* basis vectors.

The underlying system of axes is the x-y' axes where y' is the line passing through (-3, -3), (0, 0) and (3, 3).

Note that

$\begin{bmatrix} 2 & 1 \\ 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ basis vector e_x does not rotate due to A.

$\begin{bmatrix} 2 & 1 \\ 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ basis vector e_y does rotate due to A.

$\begin{bmatrix} 2 & 1 \\ 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ basis vector e_2 does not rotate due to A.

We may say: a basis vector that does not change direction (i.e. shift axis) due to the action of action matrix A is an eigenvector and its underlying axis is an eigenaxis.

From the practical point of view we expect the points on the x-axis to remain on the x-axis and the points on the y'-axis to remain on the y'-axis even after the action of the matrix, which agrees with our intuition.

Looking at the diagram the question may arise: why is it we did not get an eigenvector lying on the axis x' passing through (-1, 3), (0, 0) and (1, -3)?

Looking more closely at the diagram, we can see the axis passing through (1, -1) and (-1, 1) is different from the axis passing through (1, -3) and (-1, 3). Intuitively we expect this rotation to happen due to the action of the matrix. Hence we did not get an eigenvector on either of these axes.

9. A Word of Caution

Text books say the eigenvector does not shift or change direction. To bring out this point more clearly we used the notion of *eigenaxes*.

The vectors $e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $-1 \cdot e_x = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $0 \cdot e_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

all lie on the same x-axis.

In fact the zero vector is common to all the axes, and hence implicitly does not change direction.

We mentioned earlier, that in order to make a vector rotate we need to add or subtract something from its components. We used the example

$$\begin{aligned} \boldsymbol{v} &= 5i + 5j \\ &\;\; + 1i - 1j \\ \text{becomes } &\; 6i + 4j \end{aligned}$$

Here we both added and subtracted something. Why is it then, in computing the eigenvalues, we only subtract λ along the diagonal?

Note that $6i + 4j = 2\,(3i + 2j)$.

They both lie on the same axis.

$3i + 2j$ can be got from $5i + 5j$ by just subtracting from the components. All we need and can do is identify the underlying axis.

$$\begin{aligned} \boldsymbol{v} &= 5i + 5j \\ &\;\; - 2i - 3j \\ \text{becomes } &\; 3i + 2j \end{aligned}$$

10. A Closer Look at the Determinant

Let us look more deeply into the calculation of the ***determinant*** and ask ourselves:

What is so special about the ***determinant*** calculation?

We can think of so many other calculations that take into account each and every value in the matrix and its position.

Let us do the calculation from a vector cross product point of view.

$$a = a_1 \cdot i + a_2 \cdot j$$

$$b = b_1 \cdot i + b_2 \cdot j$$

$$a \times b = (+ a_1 b_2 - a_2 b_1) \cdot k$$

When we drop the unit vector k (in the new third dimension) we are left with $+ a_1 b_2 - a_2 b_1$ which is $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$

With three vectors:

$$a = a_1 \cdot i + a_2 \cdot j + a_3 \cdot k$$

$$b = b_1 \cdot i + b_2 \cdot j + b_3 \cdot k$$

$$c = c_1 \cdot i + c_2 \cdot j + c_3 \cdot k$$

$$a \cdot (b \times c) = a_1 \cdot (b_2 c_3 - b_3 c_2) + a_2 \cdot (b_3 c_1 - b_1 c_3) + a_3 \cdot (b_1 c_2 - b_2 c_1)$$

which is again $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = b \cdot (c \times a) = c \cdot (a \times b)$

What is the significance of this?

Vector cross product is a kind of ***combined effect of each on each*** and analogous to the ***convolution operation***.

Refer p. 76, **Convolution & the LTI Systems** by S. Suresh *et al*, © May 2008

We can see this happen in nature when we look at $\vec{F} = \vec{E} \times \vec{B}$ in motors and in the Poynting vector of electromagnetic waves as in $\vec{E} \times \vec{H}$.

So in computing the ***determinant***, taking into account the position and value of each element in the matrix, we are actually computing a kind of ***combined effect*** – as it happens in nature!

This has deep implications in understanding the role of the ***Jacobian***.

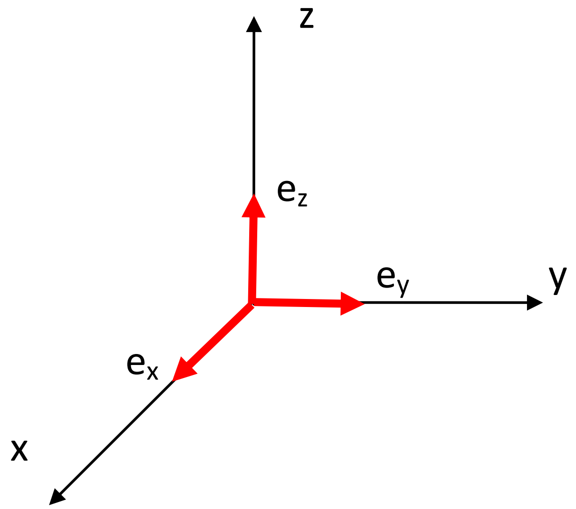
From the ***State-Space*** view: each column in the matrix is a dimension – an axis of a variable. Each column in the ***Jacobian*** matrix has the ***partial derivatives*** (rates of change) of each of the old variables with respect to a new (due to the change in) variable.

The secret of the ***Jacobian*** is: ***Between the two systems of variables, to get the change (as in integration) to match, you have to get the rates of change (as in differentiation) to match***.

Vector cross product, ***determinant*** and ***convolution*** operations are closely related in the sense that they model the ***combined effect*** behavior observed in nature.

11. The Case of Repeated Roots

Let us project the 3-D space onto the x-y plane.



The action matrix is A = $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Given any point in 3-D space, the x and y co-ordinates remain the same. The z co-ordinate goes to zero – projected onto the x-y plane.

1. axes: x-y-z, the usual orthogonal system of axes.
2. basis vectors:

   $e_x = (1, 0, 0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_y = (0, 1, 0) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $e_z = (0, 0, 1) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
3. action matrix: A as above
4. action on the basis vectors:

   $A \cdot e_x = e_x$, $A \cdot e_y = e_y$ and $A \cdot e_z = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Since the basis vectors e_x , e_y and e_z lie on the same system of axes despite the action of A, we call them the *eigenvectors* and the x-y-z system of axes is the *eigenaxes*.

Since the action of A on the basis vectors is correct, it must be correct for all points in the 3-D space. Also since the basis vectors are *eigenvectors*, we really cannot make the action matrix more simple. The *eigenvectors* e_x , e_y and e_z are a *best basis* or *optimal basis*.

Let us use the method to find the *eigenvalues* of A and then find the *eigenvectors* and *eigenaxes*.

1. Form the matrix $A - \lambda I$

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

2. Form the characteristic polynomial from det. $(A - \lambda I)$.

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

   $- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$

   We get: $+ (1 - \lambda) \cdot (1 - \lambda) \cdot (- \lambda) \quad + 0 \cdot 0 \cdot 0 \quad + 0 \cdot 0 \cdot 0$

   $- (1 - \lambda) \cdot 0 \cdot 0 \qquad - 0 \cdot 0 \cdot (- \lambda) - 0 \cdot (- \lambda) \cdot 0$

   $= \lambda^3 - 2\lambda^2 + \lambda$

3. Find the roots of the characteristic polynomial

$$\lambda_1 = 1,\ \lambda_2 = 1 \text{ and } \lambda_3 = 0 \text{ are the eigenvalues.}$$

4. Form the matrices $A - \lambda_1 I$, $A - \lambda_2 I$ and $A - \lambda_3 I$

$$A - \lambda_1 I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},\ A - \lambda_2 I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ and } A - \lambda_3 I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

5. Find the eigenvectors ev_i as solutions to $(A - \lambda_i I) \cdot x = 0$.
   We can spot them in the columns

$$ev_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = -1 \cdot e_z,\ ev_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = e_x \text{ and } ev_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = e_y.$$

Intuitively we chose e_x , e_y and e_z as the eigenvectors and the x-y-z system of axes as the eigenaxes.

Had we not known how to find the *best basis* for action matrix A, this method of eigenvalues and eigenvectors would have found it for us.

Notice the repeated eigenvalues. This led to

$$(A - \lambda_1 I) = (A - \lambda_2 I)$$

We still got two independent eigenvectors ev_1 and ev_2 by looking at the columns of $(A - \lambda_3 I)$.

One may argue that

$$(A - \lambda_1 I) \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so vector $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ should be a basis vector.

However, $(A - \lambda_2 I) \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ also $= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

We can take another vector for example, $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, which satisfies

$$(A - \lambda_i I) \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ for } i = 1,\ 2$$

So where is the catch?

Since the roots are repeated any two independent (orthogonal or skewed) vectors in the x-y plane can be eigenvectors.

The x-y-z *eigenaxes* is not unique. We could have chosen any other x’-y’-z as the system of *eigenaxes* where x’ and y’ are ***skewed*** (not orthogonal) or x’ and y’ are orthogonal with different orientation.

For example $x' = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $y' = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Moreover, we said that the basis vectors must cover or ***span*** the whole space – uniquely identify each and every point in the space by linear combinations.

Now we need a third vector that is not in the x-y plane to be able to define any point in the 3-D space. We would like this third vector to be an eigenvector. The only choice we have is to choose a vector perpendicular to the x-y plane.

We choose $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, which, after the action of the matrix becomes $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ implying it still lies on the same axis.

We can generalize this situation to higher dimensions. For example, if we get three repeated eigenvalues for a 4 x 4 matrix we can choose any three vectors in the 3-D ***eigenspace*** which is a 3-D "plane" in 4-D space and find the fourth vector which is orthogonal to this 3-D "plane". We used the word ***eigenspace*** because the points in the projected plane, (here the x-y plane), are ***invariant*** - not affected by the action of projection matrix.

12. Applications

There are two main types of applications. One type of application looks for invariance, modes and solutions to differential equations. Another type of application involves repeated action. We shall see a few examples of both types.

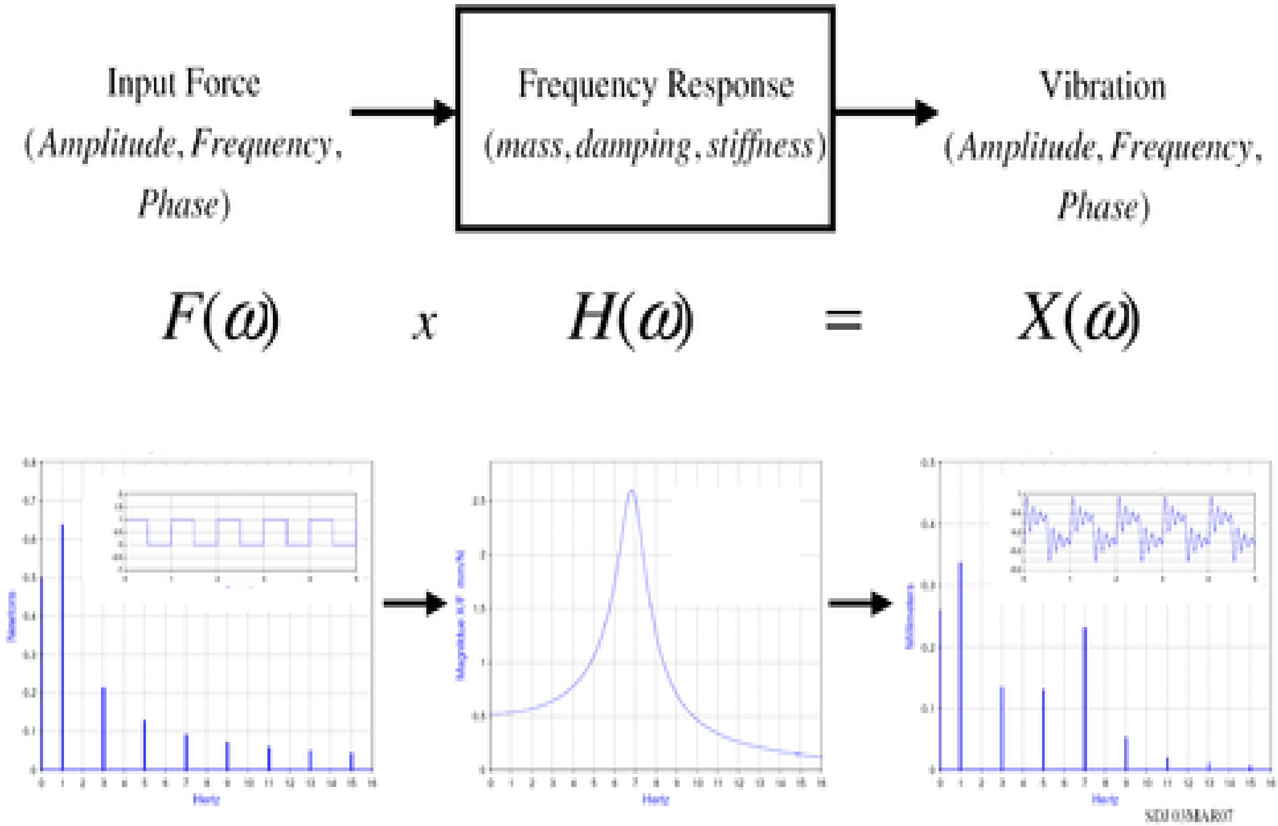
We have an interesting problem which illustrates the importance of eigenvalues and eigenvectors. In the year 1831 a bridge collapsed in Manchester when soldiers were marching on it. Refer: http://en.wikipedia.org/wiki/Broughton_Suspension_Bridge

In this, the eigenvectors for the vibration of the bridge (***modes of vibration***) fell in the range of frequency of soldiers' march, thus causing the bridge to resonate to it and leading to its collapse. This is why soldiers break cadence when crossing a bridge.

Refer this link to find out more about how eigenvalues help in solving problems concerning vibrations:

http://en.wikipedia.org/wiki/vibrations

eigenvalues and eigenvectors help us to find the dominant modes(frequencies) of vibration of a given system. By knowing the dominant modes, we can design the system as to avoid resonance due to vibrations caused in its surroundings.



The analogy here for going from ROW picture to COLUMN picture is:

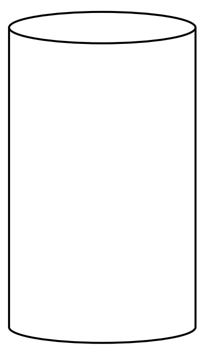
Analysis	Algebra
line $\longrightarrow$	(scalar, vector)
signal $\longrightarrow$	(amplitude, frequency)

In analysis we use Fourier methods to extract the same information.

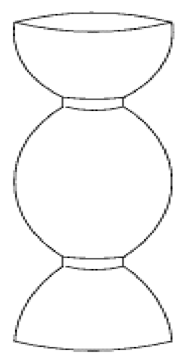
Refer: The Continuous Time Fourier Series by S. Vadakkan *et al*, © April 2009

The matter below was found on the internet. We apologize to the author(s) for failing to record the link.

The typical design of a load-bearing column is cylindrical. However, this is not necessarily the strongest. Steve Cox of Rice University and Michael Overton of New York University proved, based on the work of J. Keller and I. Tadjbakhsh, that the column would be stronger if it was largest at the top, middle, and bottom. At the points $1/4^{th}$ of the way from either end, the column could be smaller because the column would not naturally buckle there anyway. A cross-section of this column would look like this:



cylindrical column



"optimal" column

This new design was discovered through the study of the eigenvalues of the system involving the column and the weight from above. Note that this column would not be the strongest design if any significant pressure came from the side, but when a column supports a roof, the vast majority of the pressure comes directly from above.

Refer Modern Digital Signal Processing by Roberto Cristi © 2003 p.208 Example 5.5. (Caution: if the matrix multiplication function on the calculator is used an error message may appear. So work out the matrix multiplication calculation step-by-step to avoid the round off.)

13. Repeated Action and Diagonalization

So far we have looked at only simple spaces to illustrate the role of ***eigenvalues*** and ***eigenvectors***. There are many spaces varying from Engineering to Games to Economics that may be modeled using Vector Spaces.

Take for example an Information Space – data on websites. If the information is stored as vectors in a Vector Space, the *action of searching* – as in a search engine – for some particular information (= vector) may be represented by an action matrix.

Now searching for this particular information (vector) may be by the *repeated action* of the matrix on some key vector. We would like the ***search engine*** to proceed more or less in the direction of the desired data (vector) rather than go round in circles.

So if we can find the ***eigenvalues*** and the ***eigenvectors*** of the action matrix, then we can come up with a more simple action matrix where there is less computational effort involved due to more direct nature of the search.

If the eigenvalues are: $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_n$

and the eigenvectors are: $e_1, e_2, e_3, \cdots, e_n$

and the key of the desired particular information is a vector

$$v = a_1e_1 + a_2e_2 + a_3e_3 + \cdots + a_ne_n ,$$

searching for the desired information (vector) may be via an action matrix B = a more simple action matrix (from computation procedure point of view) with the eigenvectors as the set of basis vectors.

Now if the search involves applying action matrix A repeated m times on search key vector then the computation is simple.

$A^m (v) = \lambda_1^m a_1 e_1 + \lambda_2^m a_2 e_2 + \lambda_3^m a_3 e_3 + \cdot \cdot \cdot + \lambda_n^m a_n e_n$.

$$A = P \cdot D \cdot P^{-1} \Longrightarrow A^m = P \cdot D^m \cdot P^{-1} \text{ where } D = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_n \end{bmatrix}$$

is the diagonal matrix.

Refer p.30 and p.876, ***The Princeton Companion to Mathematics*** edited by Timothy Gowers *et al*, © 2008 by Princeton University Press

Further reading: THE $25,000,000,000 EIGENVECTOR THE LINEAR ALGEBRA BEHIND GOOGLE by KURT BRYAN and TANYA LEISE

Further reading: The Linear Algebra Behind Search Engines by Amy Langville.

There are many applications in genetics and other areas which can be modeled by using ***Markov chains***.

Let us now see how we make the action matrix more simple. By this we mean a more simple procedure to compute A^m.

Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$

We found out that the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 3$.

The eigenvectors are $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, a set of *skewed* basis vectors.

Now to construct a more simple action matrix we have to put the eigenvectors as the columns of the matrix **P** and put the eigenvalues along the diagonal of the matrix **D**. Then find $\mathbf{P}^{-1}$.

Thus:

$$P = \begin{array}{c} \begin{matrix} e_1 & e_2 \end{matrix} \\ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \text{ and } D = \begin{array}{c} \begin{matrix} \lambda_1 & \lambda_2 \end{matrix} \\ \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \end{array} \text{ and } P^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Now you can verify that $A = P \cdot D \cdot P^{-1}$.

Notice that we have kept the order of eigenvalues and eigenvectors the same. This is important to get a correct action matrix. Still we can change the order of both eigenvectors and eigenvalues to get different **P** and **D**.

$$P = \begin{array}{c} \begin{matrix} e_2 & e_1 \end{matrix} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array} \text{ and } D = \begin{array}{c} \begin{matrix} \lambda_2 & \lambda_1 \end{matrix} \\ \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \end{array} \text{ and } P^{-1} = \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$$ is also a valid simple action matrix.

For another example you may refer Linear Algebra by Gilbert Strang, 4th edition, p.262, Problem Set 5.3, Problem 1.

While this method of calculation may seem new and complex the thought process behind it is analogous to finding the solution of a ***recurrence relation***. The Fibonacci sequence 1, 1, 2, 3, 5, 8, 13, **. . .** is modeled by the recurrence relation $F_{n+2} - F_{n+1} - F_n = 0$. We may find the 1000th term by finding sequentially the first 999 terms or we may find the solution to the ***recurrence relation*** and directly find the 1000th term.

Refer The Linear Algebra View of the Fibonacci Sequence, www.the-idea-shop.com/article/218/the-linear-algebra-view-of-the-fibonacci-sequence
Refer Linear Algebra by Gilbert Strang, 4th edition, p.255.
Refer Digital Signal Processing – Principles, Algorithms and Applications by Proakis *et al* © 2006, p.436. Example 7.4.2.

Recurrence relations are also known as difference equations. Difference equations are known to be the discrete versions of differential equations. While ***Analysis*** deals with continuous operands ***Algebra*** deals with discrete operands. So eigenvalues help us to solve differential equations especially for a coupled system of differential equations.

Ref. http://tutorial.math.lamar.edu/Classes/DE/DE.aspx**.**

Lecture notes of Paul Dawkins on Differential Equations.

Ref. p.30, ***The Princeton Companion to Mathematics*** edited by Timothy Gowers *et al*, © 2008 by Princeton University Press.

Ref. Linear Algebra by Gilbert Strang, 4th edition, p.233 – p.237.

14. Non Diagonalizable Matrices

Some action matrices ***rotate*** or ***shift*** an axis no matter which system of axes we choose. We can detect this because the basis vector (along that axis) shifts or changes direction due to the action. In this case we cannot hope for a *best basis* or *optimal basis* that will give us a more simple version of the action matrix.

Consider the example $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

1. Form the matrix $A - \lambda I$

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

2. Form the characteristic polynomial from det. $(A - \lambda I)$.

$$\text{We get: } (1 - \lambda) \cdot (1 - \lambda) - 0 \cdot 1$$

$$= \lambda^2 - 2\lambda + 1$$

3. Find the roots of the characteristic polynomial

$\lambda_1 = 1, \ \lambda_2 = 1$ are the eigenvalues.

4. Form the matrices $A - \lambda_1 I$ and $A - \lambda_2 I$

$$A - \lambda_1 I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } A - \lambda_2 I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

5. Find the eigenvectors ev_i as solutions to $(A - \lambda_i I) \cdot x = 0$.
   Try to spot them in columns. We get:

$$ev_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ which is } e_x.$$

We cannot find another eigenvector except $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, which is a trivial eigenvector. This means that the 2-D plane cannot be ***spanned*** by the eigenvectors of the given action matrix since the two eigenvectors are not independent. So we cannot find a simpler action matrix for given action matrix. Such matrices are called ***non-diagonalizable*** matrices.

15. SVD and PCA

Decomposition is a fundamental concept in mathematics. We decompose a natural number into its prime factors.

$$360 = 2^3 \cdot 3^2 \cdot 5^1$$

We decompose a polynomial into its factors.

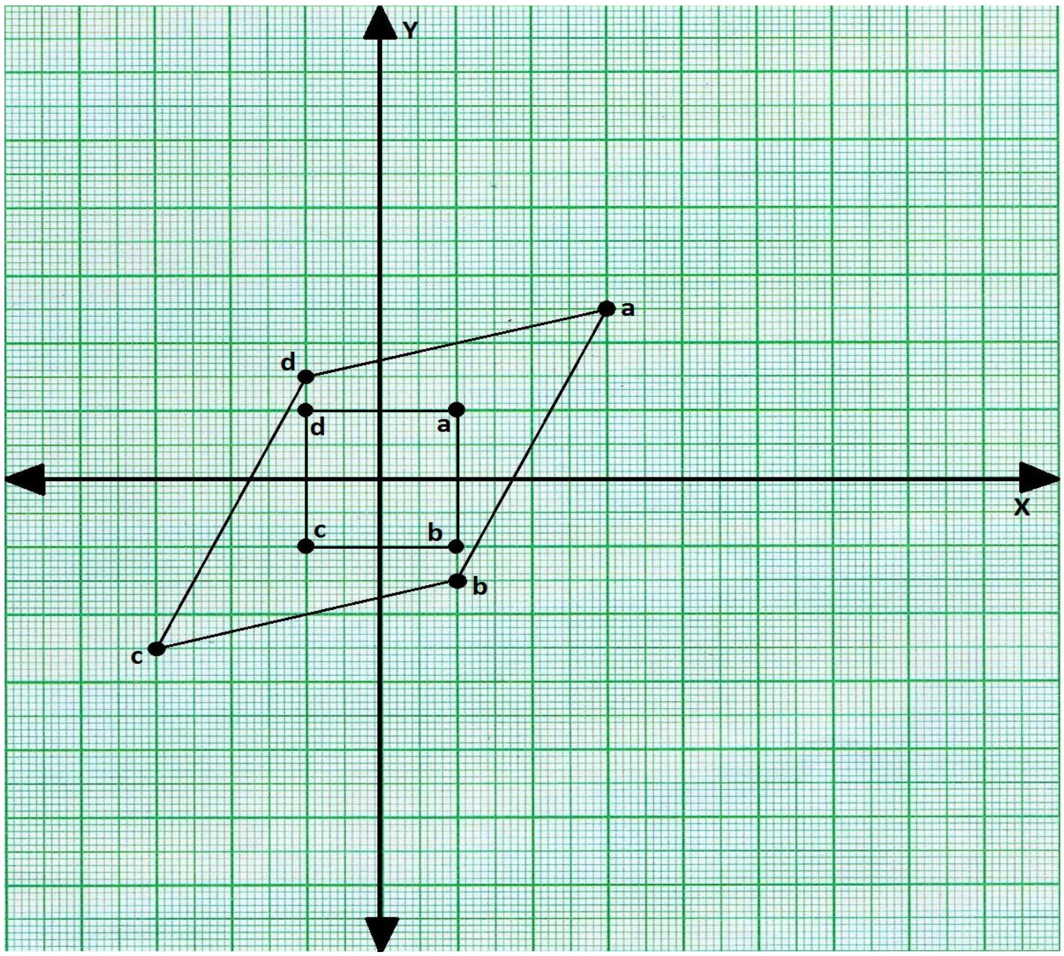
$$x^2 - 3x + 2 = (x - 1) \cdot (x - 2)$$

We also decompose a signal into its frequency components as we saw in the mechanical vibrations example.

So we ask: can we decompose an action matrix that has multiple actions like stretch, shear and rotate into a composition of individual actions? Suppose we are given only two transformations – stretch and rotate, how can we achieve the shear transformation?

There are methods in Linear Algebra known as SVD (Singular Value Decomposition) and PCA (Principal Components Analysis) to achieve this.

Consider the example: A = $\begin{bmatrix} 1 & 2 \\ 1/2 & 2 \end{bmatrix}$. The action of this matrix on the square abcd is as shown below

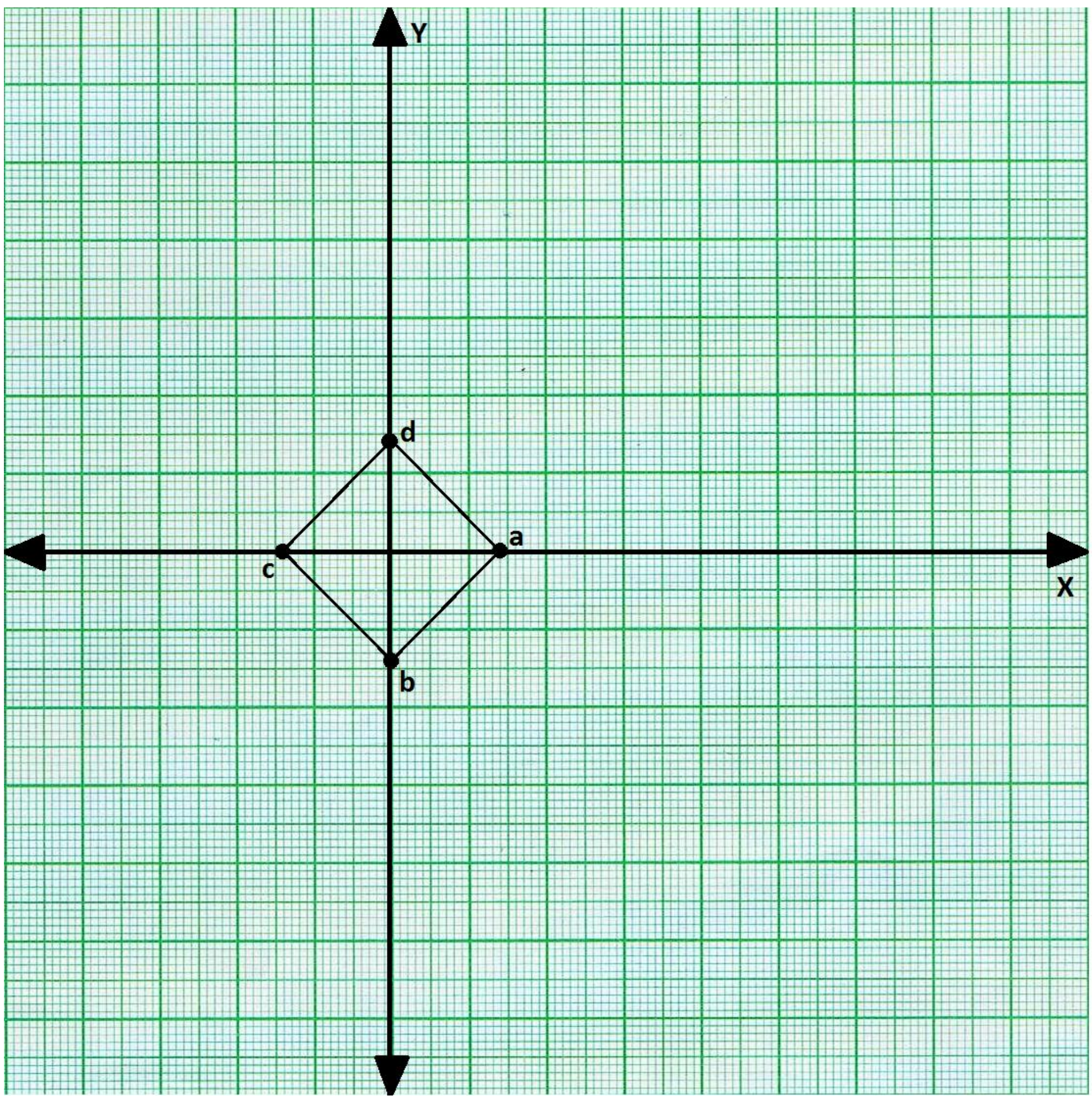


Instead of doing this in one step, let us see if we can decompose it into three transformations: two rotations and a stretching action.

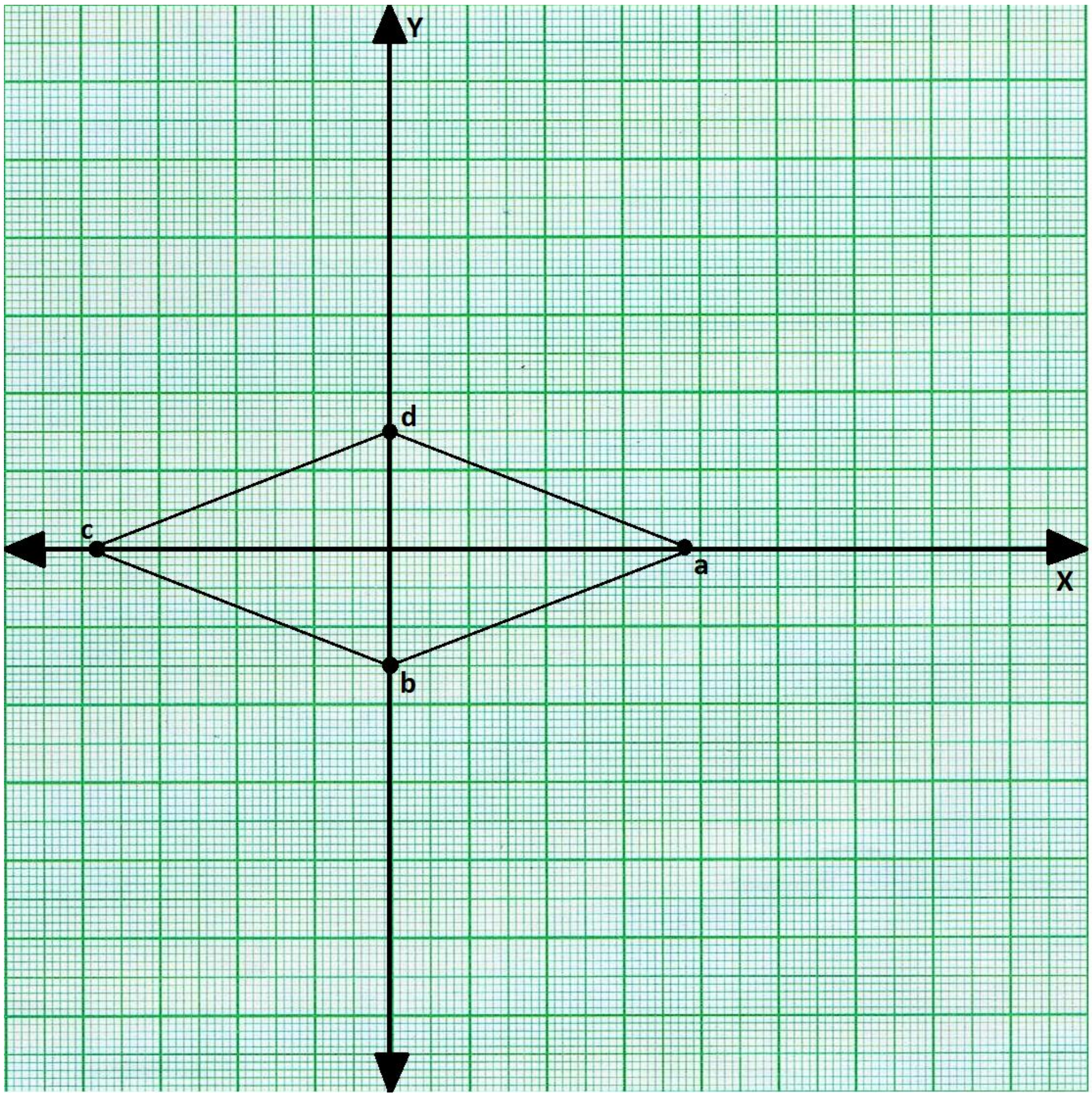
We can write A = U Σ V^T where

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \frac{\sqrt{2.25}}{\sqrt{2}} & 0 \\ 0 & \frac{\sqrt{15.25}}{\sqrt{2}} \end{bmatrix} \text{ and } V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

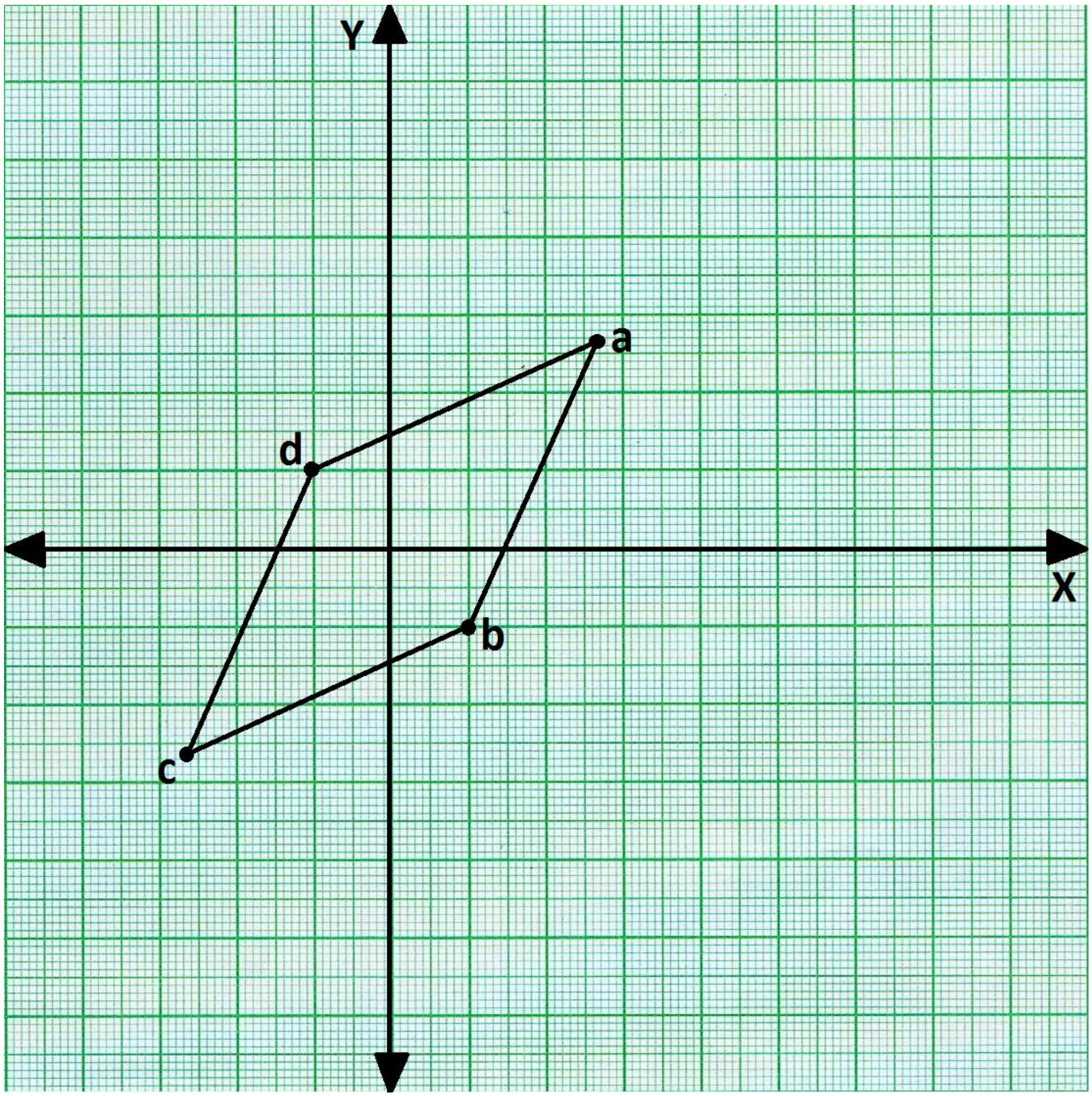
V^T will first rotate the square abcd by 45 degrees.



Σ will stretch it along both dimensions



and finally U will rotate it back by the same angle of 45 degrees.



While this way of thinking looks right, it doesn't match the original action of A. Compare this diagram with the diagram on page 69. Let us see how we can correct it.

Let us try and find the determinant of A using an indirect method.

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

$$|A|^2 = (ad)^2 + (bc)^2 - 2abcd$$

$$\text{Now, } A \cdot A^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}.$$

$$|A \cdot A^T| = (ad)^2 + (bc)^2 - 2abcd = |A|^2$$

$$|A| = \sqrt{|A \cdot A^T|}.$$

Determinants can be found only for n x n matrices. But life is not always fair and square. It is not always possible to represent reality in n x n matrix. We often have to deal with m x n matrices.

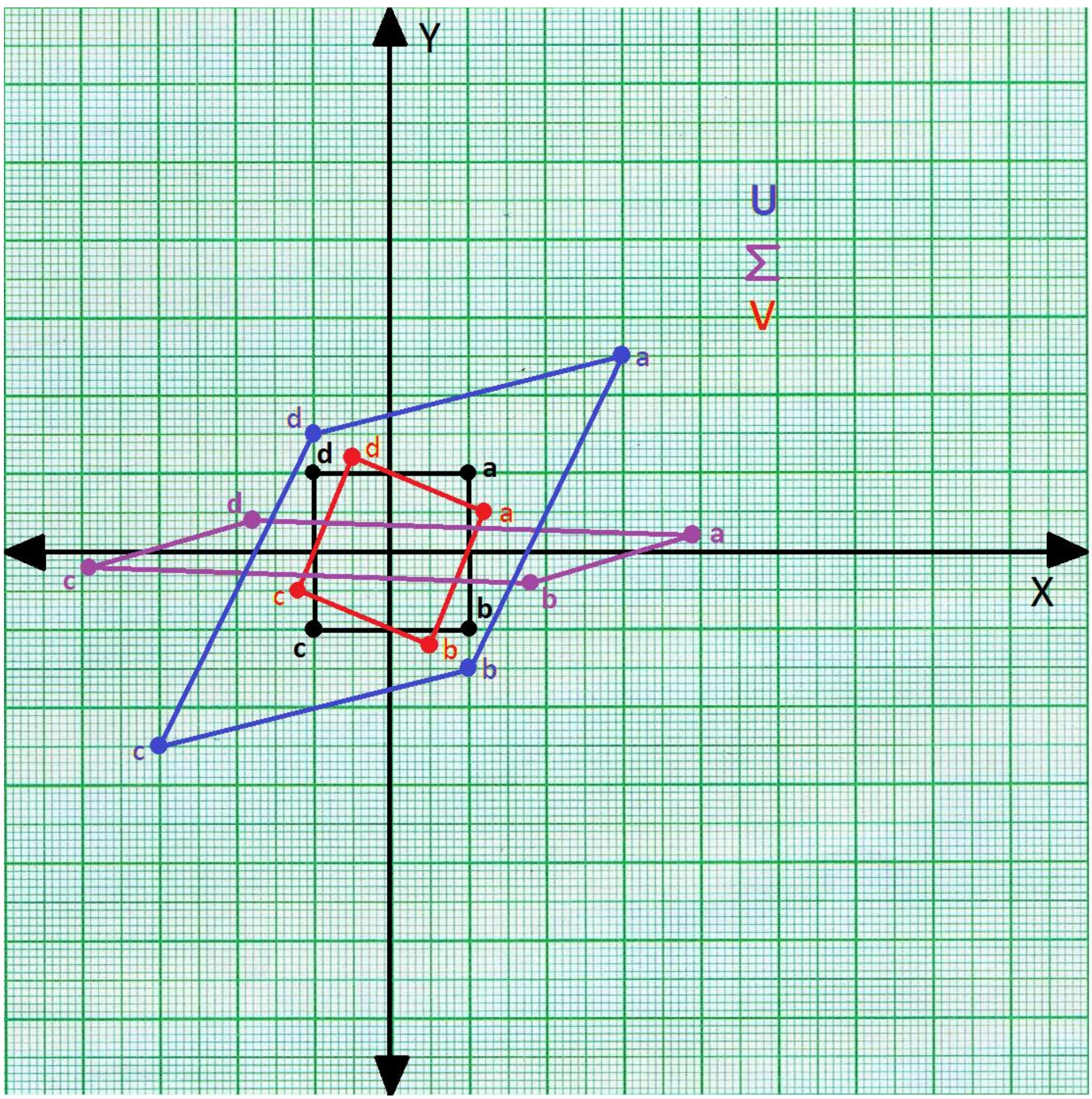
So how can we find the “eigenvalues” of a m x n matrix?

Well, maybe we should not call them eigenvalues except when m = n.

But they are special values. So we call them ***singular values***.

Let us now apply the method we just saw for A. Let us find the ***eigenvectors*** of $A \cdot A^T$ and use them as columns to construct U. Then we shall use the ***eigenvectors*** of $A^T \cdot A$ to find V. For Σ we shall find ***square roots*** of the ***eigenvalues*** of $A^T \cdot A$. Using MATLAB we get the matrices:

$U = \begin{bmatrix} -0.7359 & -0.6771 \\ -0.6771 & 0.7359 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 3.0233 & 0 \\ 0 & 0.3308 \end{bmatrix}$ and $V = \begin{bmatrix} -0.3554 & -0.9347 \\ -0.9347 & 0.3554 \end{bmatrix}$



Note that the angles in the rotations are not the same. But the composition of the three matrices U Σ V^T on square abcd has the same as the action of A.

Compare the transformation here with the transformations due to the change of variables in the Jacobian.

Refer:
http://www.math.oregonstate.edu/home/programs/undergrad/CalculusQuestStudyGuides/vcalc/change_of_variables_multiple_integrals.html

Is the angles of rotation not being the same in any way related to the problem of being unable to trisect a 60 degree angle using only a compass(rotation) and straight edge(stretch or shear) or the problem of being unable to square a circle?

Refer this link to see how an action matrix transforms a circle into an ellipse and how the action matrix is decomposed into two rotations and a stretching:

http://en.wikipedia.org/wiki/Singular_value_decomposition

Now if we look more closely, in the ***Jacobian*** we are doing the same thing when we use it for changing the variables. This once again assures us that whatever can be done in ***Analysis*** can also be done in ***Algebra.*** The only exception is translation, which can be done in ***Analysis***, but not in ***Algebra*** unless you use vector graphics.

From a Mathematical point of view, when we apply SVD to an m x n matrix, we decompose it into three matrices:

U which is m x m

Σ which is m x n and

V which is n x n.

This means that to ***map*** a n-D vector onto an m-D space, we first rotate it in the n-D space with V^T, then take a ***simpler mapping*** Σ and again rotate the resultant m-D vector with U.

Also, U and V are ***orthonormal*** matrices, which means that applying them to a vector does not change its magnitude. So, if we want to multiply the action matrix A by a ***scalar***, we can multiply just Σ and can have the same result.

In diagonalization, we do not change the space, but represent the same vector in terms of different and convenient basis vectors using the matrices P and P^{-1} . In SVD, We do not have a set of linearly independent basis vectors to change. So to make the computation (many-to-many mapping) simpler we change the space itself using U and V^T .

For more on SVD, refer the work of Stanford Prof. Gene Golub.
http://www.mathworks.in/company/newsletters/news_notes/oct06/clevescorner.html

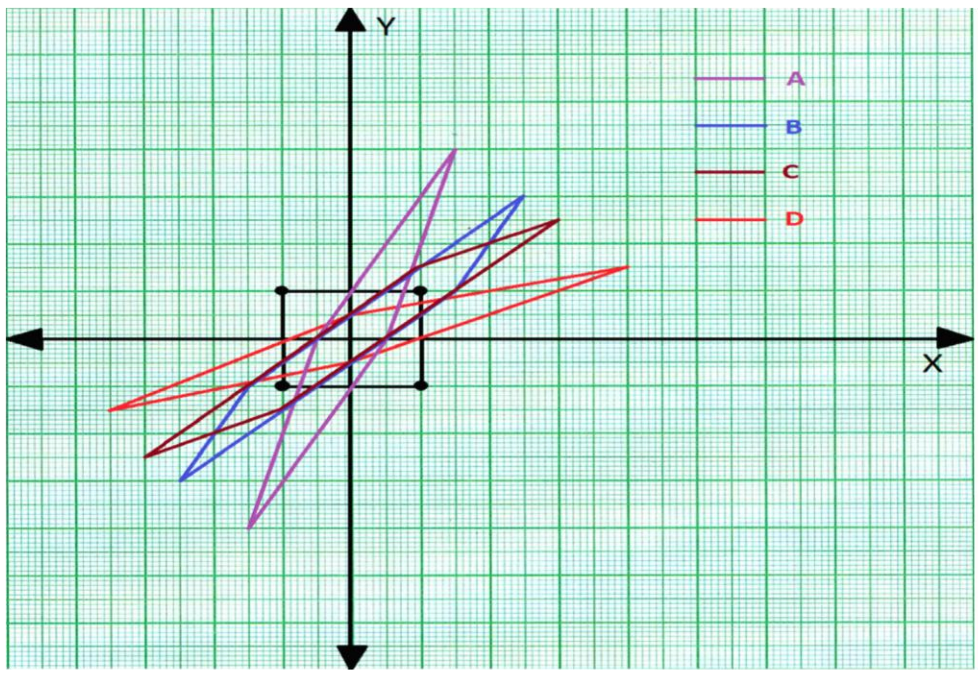
16. Some more applications

The matrices

$$A = \begin{bmatrix} 1 & 1/2 \\ 2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1/2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 1/2 & 2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 2 & 2 \\ 1/2 & 1 \end{bmatrix}$$

all have the same eigenvalues.

Find the eigenvectors and eigenaxes in each case. Are any two sets the same? Apply each matrix to the square and sketch the trapezium.



So now we have four different views. The information of these views are stored as action matrices on an average or normalized view.

Now play around with the minus sign. Make some of the values negative and see how the figures get reflected

They are like "***shadows cast in different directions by some central solid essence***", p.28, GÖDEL, ESCHER, BACH: *an eternal golden braid* by Douglas R. Hofstadter, © Basic Books Inc., 1979.

Refer: Linear Algebra for Computer Vision, Introduction, CMSC 828 D

In Higher Secondary we learn about the dynamics of a system. We learn linear motion, uniform circular motion, moment of inertia, harmonic motion (as in oscillating pendulum), waves and damped waves.

Refer: Solving the Harmonic Oscillator Equation by Morgan Root. http://www.ncsu.edu/crsc/events/ugw05/slides/root_harmonic.pdf.

The motion of a child on a swing may be described by damped oscillations. What if there is a slight *twist* (*rotation*) in the motion due to an uneven push – right hand pushes a little bit harder than the left hand?

Differentiation and Integration are the tools in Analysis to help us extract the ***instantaneous*** information. If we are willing to sacrifice a little bit of that ***instantaneous*** information and settle for ***average*** information (over a very small interval of time Δt) we can then use the tools of Algebra which are more suited to run on digital computers.

"There is a beautiful footpath on the ascent to Linear Algebra."

Stephen Vadakkan