A Review of Concepts Concerning Linear Algebra and the Applications of Similarity Transformations

Advait Pravin Savant
Department of Engineering Maths
Sardar Patel Institute of Technology

Abstract

Linear algebra is a field of mathematics which concerns itself with the properties and operations of linear spaces and linear transformations which are functions mapping from one linear space to another based on a set of axioms they must follow. Linear transformations are represented via matrices. While linear spaces and the elements that constitute a linear space are defined formally in terms of axioms/properties concerning the addition, multiplication and other operations on the elements [5], it is best to start linear algebra with Euclidean spaces and vectors that constitute the Euclidean space. The 3 dimensional Euclidean space and the Cartesian geometry which we do in that space provide good intuitions for understanding linear algebra. These can be generalized to n dimensions. Concepts such as lines, planes, parallelism, projections and orthogonality have neat formulations and generalizations in linear algebra. We start with vectors to represent points in 3 D space, and matrices as mathematical objects which stretch and rotate these vectors. We see that this is related to 3 linear equations with 3 unknowns. We then review and formalize concepts such as linear spaces, basis vectors, independence, projections, orthogonality, diagonalization of a matrix etc. Linear algebra is deeply connected with multivariable calculus and optimization; concepts such as the derivative of a vector field and multivariable Taylor Series require the understanding of linear algebra. Today, linear algebra is getting even more important, from principal component analysis in machine learning to Fourier analysis in digital signal processing to the Lorentz transformation in Einstein's Relativity, all are based on the concepts of linear transformations and change of basis.
Keywords- Linear Transformations, Cauchy Schwartz Inequality, Eigen-Decomposition, Similarity Transformations, Quadratic Forms, Rayleigh Principle

I. INTRODUCTION

A Euclidean vector is a mathematical object which has magnitude and direction. It can be added or subtracted to other vectors based on the laws of vector algebra such as the triangle law of vector addition [4]. In 3 D space, a vector is visualized geometrically as a ray starting from the origin to the point under consideration. This ray has a direction based on the point and the length of the ray is the magnitude of the vector. We can also represent this vector as a tuple whose components are the coordinates of the point based on the axes. The representation of a vector depends on the choice of the axes. We can generalize these concepts in n dimensional space. Consider a system of linear equations.

$$3x + 5y + 4z = 10$$
$$2x - 2y + z = -2$$
$$x + y + 3z = 6$$

We can consider this as 3 planes in the 3 D space with each row representing a plane. We can solve these equations by Gaussian elimination [1] and find the point where the planes intersect. We will discuss the singular cases soon. We can also think about these equations in terms of the columns. We can write the same system of equations as follows,

$$x\begin{bmatrix}3\\2\\1\end{bmatrix} + y\begin{bmatrix}5\\-2\\1\end{bmatrix} + z\begin{bmatrix}4\\1\\3\end{bmatrix} = \begin{bmatrix}10\\-2\\6\end{bmatrix}$$

We can think of this problem as finding the scalars x, y and z to multiply the corresponding vectors with so that the linear combination gives us the vector on the right hand side of the equation. These column vectors give us the column picture of matrix multiplication. We can write the coefficients of the variables in a 3*3 array which is a matrix (here, a mathematical object that transforms vectors). This leads us to our first kind of equation in linear algebra which is,

$$\begin{bmatrix}3 & 5 & 4\\2 & -2 & 1\\1 & 1 & 3\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}10\\-2\\6\end{bmatrix}$$

If we go row wise, we multiply each element in the first row of the matrix with the corresponding element in the vector for the variables to get the value equal to the first element of the vector on the right hand side. We proceed similarly for the next

two rows. This gives us the original equations. We thus understand the matrix notation and the two ways to multiply a matrix with a vector(via columns and via rows). This notation enables us to generalize concepts for n equations in n unknowns.
In general, for a matrix A and vectors x, b we have,

$$Ax = b$$

In the singular case for 3 dimensions, if the three column vectors lie in the same plane, then any linear combination of the vectors will also lie in that plane. If the vector on the right hand side is outside the plane, the system will have no solution. The process of Gaussian elimination fails. If the vector on the right hand side lies in the same plane as the three column vectors, it follows from vector algebra that the system will have infinitely many solutions. In the n dimensional case for a system of equations, we talk about column vectors lying in a subspace of the n dimensional vector space and the column vectors being unable to span the entire space. We will see that this is closely related to the concept of the rank of a matrix and the linear independence of the columns. If the matrix A maps each vector in n space to another vector in n space uniquely, we can define another matrix to be the inverse of A which essentially does the opposite mapping relative to A [1].

A. Linear Spaces

A linear space V is a set of elements which satisfy 10 axioms as listed in [2]. These axioms include closure under addition, closure under multiplication, commutativity, associativity, distributive property etc. Such a formalism follows from the laws of algebra, can be understood intuitively for 2 and 3 dimensional Euclidean spaces and enables us to create a more general and elegant mathematical theory. The set of real numbers is a linear space, so is the set of complex numbers. The set of vectors which are n-tuples is also a linear space with addition and multiplication by scalars defined as usual for vectors [4]. Linear spaces can also include functions [3]. Each linear space has a zero element which when added to any other element gives the same element and when multiplied to any other element gives back the zero elements. There is also a unique negative element for each element both of which add to give the zero elements. A subset of the linear space which satisfies the closure axioms is known as a subspace of the linear space. A set of elements in the linear space is said to be dependent if we can find a non-trivial set of scalars such that the linear combination of those elements adds up to the zero element. In such a set, we can express an element as a linear combination of the other elements. Otherwise, the set of elements is independent. For example, a set of 3 or more vectors lying in the same plane in 3 D space is dependent. A set of elements is said to span the linear space V if any element in V can be expressed as a linear combination of elements in that space. A finite set S is called the basis of a linear space if the set of elements is linearly independent and spans the entire space. For a finite n dimensional vector space, any set of n independent vectors can act as a basis. Any vector in the space can be expressed as a linear combination of these vectors. A basis essentially consists of the minimum number of independent vectors required to span the entire space. This is akin to choosing the orientation of the coordinate axes. For basis vectors e_1 up-to e_n, if the vector x can be written as,

$$x = \sum_{i=1}^{n} c_i e_i$$

Then $c_{1,}c_2$ up-to c_n are the components of x relative to the choice of the basis.
An important metric property is the dot product between two vectors x and y. It is defined as,

$$x.y = (x, y) = \sum_{i=1}^{n} x_i\, y_i = x^T y$$

We can see that the root of the dot product of a vector with itself is the norm [3] of the vector. A broader concept is that of the inner product (a,b) of two elements a and b in a linear space. It is defined axiomatically in terms of commutativity, linearity, distributivity, associativity with respect to scalar multiplication and positivity. The dot product is an inner product. Consider two vectors a and b in n dimensional space. Let α be the angle between the two vectors. We have from the cosine law,

$$|a - b|^2 = |a|^2 + |b|^2 - 2(|a| * |b|)\cos\alpha$$
$$(a - b).(a - b) = a.a + b.b - 2 * |a| * |b| * \cos\alpha$$
$$a.a - a.b - b.a + b.b = a.a + b.b - 2 * |a| * |b| * \cos\alpha$$
$$a.b = |a| * |b| * \cos\alpha$$

Since the value of the cosine function lies between -1 and 1, we get the Cauchy Schwartz inequality by squaring and simple substitution in the previous equation. For any two vectors a and b,

$$\left(\sum_{i=1}^{n} a_i^2\right) * \left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i\, b_i\right)^2$$

Two vectors in a vector space are orthogonal if their dot product is zero. This is in line with our geometric intuitions coming from the cosine function. Extending this, two elements in a linear space are orthogonal if their inner product is zero. We can easily see that a set of mutually orthogonal elements are linearly independent [3]. Let us say we express a vector x as a linear combination of vectors which constitute an orthogonal basis. An orthonormal basis is one where the magnitude of each basis vector is unity. For an orthogonal basis, if we simply take the dot product of x with a particular basis vector, all other terms in the summation will be zero except the term with that particular basis vector. We thus get the component of x along that particular basis vector to the dot product of x and that basis vector divided by the dot product of that basis vector with itself. As we extend from our 2D and 3D intuitions, the projection of vector a along vector b will be given by (a,b)/(b,b) times the vector b.

For any two vectors a and b, if we subtract the component of b along a from b what we are left with is a vector perpendicular to a. Extending this concept, from any given basis, we can construct an new orthogonal/orthonormal basis using the Gram-Schmidt orthogonalization process [3]. What we do is we start with one of the original basis vectors, scale it to be of unit norm and make it the first vector in our new set of basis vectors. We then take the next vector in our original basis, take the component of that vector along the first vector in our new basis, subtract that component from the concerned vector and scale it to be of unit norm, this will be the second vector in our new basis. We then take the third vector in the original basis, take its components along the two vectors in our new basis and subtract them from it in order to get the third vector in our new basis. We continue this process to cover all the vectors in our original basis.

B. Linear Transformations

If V and W are linear spaces, a linear transformation T:V->W is a function that maps each element x in V to an element T(x) in W. V is the domain and W is the range of T. A linear transformation must satisfy the following properties related to addition and scalar multiplication. For all x and y in V and all scalars c,

$$T(x+y) = T(x) + T(y)$$
$$T(cx) = cT(x)$$

The set of elements in V which are mapped onto the zero element constitute the null space of T. This null space is a subspace of V. The nullity plus rank theorem states that the dimension of the null space plus the dimension of the transformation will give us the dimension of the domain space V. Let n be the dimension of V. Let k be the dimension of the null space of T. We have elements e_1 upto e_k which constitute the basis of the null space of T and further if we consider elements e_{k+1} to e_{k+r} we get a complete basis for V. Hence k plus r equals n. If we show that elements $T(e_{k+1})$ to $T(e_{k+r})$ constitute a basis for T(V), we have r as the dimension of T(V) and hence the nullity plus rank theorem follows.
Let x be an element in V, let y=T(x).

$$x = \sum_{i=1}^{k+r} c_i\, e_i$$
$$y = T(x) = c_i \sum_{i=1}^{k+r} T(e_i) = \sum_{i=1}^{k} c_i\, T(e_i) + \sum_{i=k+1}^{k+r} c_i\, T(e_i)$$

Since the first k basis elements are in the null space,

$$y = \sum_{i=k+1}^{k+r} c_i\, T(e_i)$$

Suppose elements e_{k+1} to e_{k+r} are dependent, then there exist scalars such that,

$$\sum_{i=k+1}^{k+r} c_i\, T(e_i) = O$$

Therefore from the property of linear transformations,

$$T\left(\sum_{i=k+1}^{k+r} c_i\, e_i\right) = O$$

Therefore the element $z=c_{k+1}e_{k+1} + \ldots . c_{k+r}e_{k+r}$ lies in the null space of T. This means it can be expressed in terms of e_1 upto e_k which are the basis of the null space. We let $z= s_1e_1 + \ldots . s_ke_k$.

Subtracting the two values of z we get a linear combination of all the basis elements of V (e_{k+1} upto e_{k+r}) as equal to the zero element(z-z). But since the basis elements of V are independent, each of the scalars must be zero. Therefore, the set of scalars c_{k+1} upto c_{k+r} are zero and $T(e_{k+1})$ to $T(e_{k+r})$ are independent. For any element x in V which we can express in terms of the basis elements as a summation, T(x) can be expressed in terms of $T(e_{k+1})$ to $T(e_{k+r})$ as the basis elements from the null space of T give us the zero element in the transformation of the summation of x.

Every linear transformation can be expressed in terms of a matrix. Consider a linear transformation T which maps from linear space V of dimension n to linear space W of dimension m. Let e_1 upto e_n be the basis for V and w_1 upto w_m be the basis for W. We can express every element x in V as a linear combination of the basis vectors, and due to the property of linear transformations over addition, we can know everything about a linear transformation by looking at what happens to the basis elements in V. We express the transformations of basis elements e_k in terms of the basis elements of W.

$$T(e_k) = \sum_{i=1}^{m} t_{ik} w_i$$

For each basis element e_k of V, we have m elements which are the scalars for expressing the transformation of e_k in terms of the m basis elements of W. We put these m elements in the k^{th} column of a matrix M. Doing this for all n basis elements of V, we have a m * n matrix M. For an element x in V with compenents $x_1,..x_n$ relative to the basis $e_1,\ldots e_n,$ the product Mx (of matrix M and vector representation of x) will give us a vector of size m whose values are the components of the transformation of x relative to the basis of W. We get this as follows,

$$x = \sum_{k=1}^{n} x_k e_k$$

$$T(x) = \sum_{k=1}^{n} x_k T(e_k) = \sum_{k=1}^{n} x_k \sum_{i=1}^{m} t_{ik} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} t_{ik} x_k \right) w_i$$

Taking a look at the last term above, the elements of the inner summation which gives us scalars in order to represent the transformation of x in terms of the basis of W are essentially coming from the multiplication of matrix M and vector x. This lets us represent the transformation T:V->W in terms of matrix M for a given choice of basis elements for spaces V and W. The matrix representation of a linear transformation is important. This can be extrapolated further in order to define the multiplication of two matrices for a composition of linear transformations [3]. We get the column and row pictures of matrix multiplication. Consider two n*n matrices A and B and the resulting matrix multiplication A*B. We can look at the columns of B, multiply each column by matrix A to obtain another column, and the matrix of the resulting columns placed one after the other in order is our result. We can look at the rows of A, multiply each row of A to matrix B to obtain another row vector, and the matrix of the resulting rows placed one below the other in order is our result. The properties of the matrix such as its rank, row space and column space tell us a lot about the linear transformation [1].

C. Eigenvalues and Eigenvectors

Consider a linear transformation T:V->V where we are using the same basis for the input space and output space. For each linear transformation, we want to find vectors which change by only a scalar factor when the linear transformation is applied to them. Geometrically speaking, the vectors do not undergo any rotation, they are only stretched. The eigenvalues denote the extent to the vectors are stretched, a negative eigenvalue denotes the flipping of the vector where the direction is reversed. For the linear transformation T represented by matrix A, we have a scalar λ as the eigenvalue and x as the eigenvector such that,

$$T(x) = \lambda x$$
$$Ax = \lambda x$$
$$(A - \lambda I)x = O$$

Since x must lie in the null space of (A – λI), the matrix (A – λI) must be singular. We find the determinant [1] of the matrix and equate it to zero in order to obtain an equation for the eigenvalues. We substitute back the eigenvalues in order to obtain the components of the eigenvectors by solving simultaneous equations.

We now talk about the eigen-decomposition of a matrix A(n*n). Consider A to have n linearly independent eigenvectors x_1 upto x_n. Let Q be the matrix whose columns are $x_1 \ldots . x_n$.

$$AQ = A[x_1 \quad \cdots \quad x_n] = [Ax_1 \quad \ldots \quad Ax_n] = [\lambda_1 x_1 \quad \ldots \quad \lambda_n x_n]$$

Now consider the following matrix multiplication,

Let, $\wedge = \begin{bmatrix} \lambda_1 & 0 & 0 \\ .. & .. & .. \\ 0 & 0 & \lambda_n \end{bmatrix}$ be a diagonal matrix.

$$Q \wedge = \begin{bmatrix} | & | & | \\ x_1 & \ldots & x_n \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ .. & .. & .. \\ 0 & 0 & \lambda_n \end{bmatrix} = \left[\begin{bmatrix} | & | & | \\ x_1 & \cdots & x_n \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 \\ .. \\ 0 \end{bmatrix} \ldots \ldots \ldots . \begin{bmatrix} | & | & | \\ x_1 & \cdots & x_n \\ | & | & | \end{bmatrix} \begin{bmatrix} 0 \\ .. \\ \lambda_n \end{bmatrix} \right]$$

For each matrix vector multiplication in the last term, for the multiplication of Q and the column vector with λ_k at the kth position, using the column picture, we get the vector x_k multiplied by λ_k added to other zero vectors. Hence,

$$Q \wedge = [\lambda_1 x_1 \quad \ldots \quad \lambda_n x_n]$$

Comparing this with our equation for AQ, we get,

$$Q^{-1} A Q = \wedge$$
$$A = Q \wedge Q^{-1}$$

This is also known as the diagonalization of a matrix. The diagonalization of a matrix depends on the matrix having enough eigenvectors while the invertibility of a matrix depends on having non zero eigenvalues [1]. If we have distinct eigenvalues, we can show by mathematical induction that the eigenvectors are linearly independent [1]. If we have enough linearly independent eigenvectors, they can form a basis for the space V in which we are doing the linear transformation T.

D. Similarity Transformations

Consider a linear transformation T:V->V where V is a linear space of dimensionality n. We have a set of basis vectors e_1 upto e_n. We represent what happens to each basis vector in the transformation in terms of other basis vectors.

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i$$

The scalars a_{ik} corresponding to $T(e_k)$ which express it in terms of other basis vectors go into the kth column of matrix A. As we have discussed before, for any vector x relative to this basis, multiplication with A gives us the transformation of x, again relative to this basis. A is the matrix of T relative to the basis vectors e_1 upto e_n.

We have another set of basis vectors u_1 upto u_n. We again repeat the same procedure and let B be the matrix of T relative to this basis.

$$T(u_k) = \sum_{i=1}^{n} b_{ik} u_i$$

Since the two sets of basis vectors span the same space V, we can express each vector in set u in terms of the basis vectors in set e.

$$u_j = \sum_{k=1}^{n} c_{kj} e_k$$

Let E=[e_1,….e_n], U=[u_1,…u_n] and C as e the matrix represented by the scalars c_{kj} for all k and for all j.
We can the above equation in matrix form as,

$$U = EC$$

This makes sense as when we apply matrix E to each column of C based on the column picture of matrix multiplication [], we get the corresponding column of U in terms of the columns of E. This is exactly what the summation denotes.
Applying transformation T to the previous equation based on the summation,

$$T(u_j) = \sum_{k=1}^{n} c_{kj} T(e_k)$$

Let E'=[T(e_1),….T(e_n)], U'=[T(u_1),…T(u_n)]
For each of the two sets of basis vectors, we have equations showing summations which express the transformation of each basis vector in terms of the vectors in that basis. Just as we did before, the can be represented in matrix form as,

$$E' = EA$$
$$U' = UB$$

We also have an equation representing the transformation of basis vectors in set u in terms of the transformations of basis vectors in set e as a linear combination using the scalars c. This can be represented in matrix form as,

$$U' = E'C$$

Substituting E'=EA and U'=UB in the previous equation we get,

$$UB = EAC$$

We have an earlier equation which represents U in terms of E and C,

$$U = EC$$
$$E = UC^{-1}$$

Substituting E=UC^{-1} in UB=EAC,

$$UB = UC^{-1}AC$$
$$B = C^{-1}AC$$

Matrices B and A are known as similar matrices. They represent the same linear transformation relative to two different sets of basis vectors. Going from one to another is known as a similarity transformation [1][3].

II. APPLICATIONS

We remember the equation B=C^{-1}AC in the context of the previous section.

1) Consider the problem of rotating a vector by 45 degrees but in another basis P in a 2 D plane. Let our basis be the standard basis which is [1,0],[0,1]. We have the other basis P as [a,b],[c,d]. Let this be matrix Z.

We have,

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

This is analogous to U=EC from the previous section where Z is analogous to C. For 45 degree rotation relative to our basis we have the matrix R to represent the linear transformation,

$$R = 1/\sqrt{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

This is analogous to matrix A from the previous section, we find matrix Y analogous to matrix B of the previous section using the equation B=C^{-1}AC from the previous section. This becomes,

$$Y = Z^{-1}RZ$$

For any vector x represented with respect to P, we multiply it by Y in order to obtain the vector which denotes the 45 degree rotation of x relative to P.

2) Consider another problem of reflecting a vector about an arbitrary plane IN 3 D space passing through the origin. We have the basis vectors for the plane which correspond to the X,Y and Z axes in that plane. Let M be the matrix which represents the basis vectors of the plane relative to the standard axes. This is analogous to C. In the basis of the plane, the matrix for reflection about the plane is given by R,

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

R can be explained using the column picture of matrix vector multiplication [1], we are essentially keeping the x and y coordinates the same and reversing the sign of the z coordinate. This amounts to reflection about the plane.
Here B=R and we have to find A.

$$R = M^{-1}AM$$
$$A = MRM^{-1}$$

For any vector x with respect to the standard basis, we multiply it by A to reflect it about the concerned plane.
3. Consider the Quadratic form for a symmetric matrix A which is a scalar function of a vector x.

$$f(x) = x^T Ax$$

Symmetric matrices arise in many contexts. The Hessian which is a matrix of the second order partial derivatives is symmetric. It arises in the context of the Taylor series expansion of a scalar field in is important in the optimization of functions. The covariance matrix for a random vector is symmetric. An easy to prove property of symmetric matrices is that the eigenvectors are orthogonal [3]. We know the transpose of an orthogonal matrix is it's inverse. This implies that the diagonalization of A if Q is the corresponding eigenvector matrix will happen as follows,

$$A = Q \wedge Q^T$$

We can represent x in terms of the eigenvectors y assuming that the eigenvectors form a basis for the space. We normalize to have an orthonormal basis of y's.

$$x = \sum_{i=1}^{n} c_i\, y_i$$

Multiplying by A,

$$Ax = \sum_{i=1}^{n} c_i\, \lambda_i y_i$$

Also,

$$x^T = \sum_{i=1}^{n} c_i\, {y_i}^T$$

Therefore,

$$f(x) = x^T Ax = \sum_{i=1}^{n} c_i^2\, \lambda_i$$

Since y transpose multiplied by y gives us unity for the same y's and zero for different y's as we have an orthonormal basis. Here we get an important result that the quadratic form for a square symmetric matrix is always positive if all the eigenvalues of that matrix are positive. Such a matrix is said to be positive definite.
4. Consider the following optimization problem for a symmetric matrix A,

$$R(x) = \frac{x^T Ax}{x^T x}$$

R(x) is known as the Rayleigh Quotient [1].
We have the diagonalization of A,

$$A = Q \wedge Q^T$$

Let,

$$z = Q^T x$$

Therefore,

$$R(x) = \frac{(Qz)^T A(Qz)}{(Qz)^T (Qz)}$$
$$R(x) = \frac{z^T (Q^T AQ)z}{z^T z} = \frac{z^T \wedge z}{z^T z}$$
$$R(x) = \frac{\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2}{z_1^2 + \cdots + z_n^2}$$

If λ_1 is the smallest eigenvalue,

$$\lambda_1 (z_1^2 + \cdots + z_n^2) \leq \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$$
$$\lambda_1 \leq \frac{\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2}{z_1^2 + \cdots + z_n^2}$$

We have shown that the Rayleigh Quotient is always greater than λ_1. Similarly we can show that it is always less than the largest eigenvalue λ_n. It takes the minimum and maximum values when x is the corresponding eigenvector. For $x=x_1$,

$$R(x) = \frac{{x_1}^T A x_1}{{x_1}^T x_1} = \frac{\lambda_1 {x_1}^T x_1}{{x_1}^T x_1} = \lambda_1$$

Thus is the minimum value of the Rayleigh quotient R(x) is the minimum eigenvalue and the and it is obtained when x is the corresponding eigenvector.

III. CONCLUSION

We started with the intuition for linear algebra with simultaneous linear equations, we introduced vectors and matrices as mathematical objects which stretch and rotate vectors. We understood linear spaces and the operations and properties of elements in a linear space such as orthogonality and projections. We understood that the representation of a vector depends on the choice of basis vectors for a linear space. We understood linear transformations and how they are represented via matrices. We reviewed various perspectives on matrix vector multiplication. We saw eigenvalues and eigenvectors and how we can create a diagonal representation of a matrix using them. We understood similarity transformations and reviewed a few applications of them. Linear algebra is a vast and beautiful field of mathematics which is very useful to engineers and scientists in a variety of domains. In multivariable calculus, we approximate non-linear functions with linear functions in the locality and linear algebra comes into the picture. The very successful backpropagation algorithm with gradient descent for iterative optimization in artificial neural networks relies heavily on linear algebra for a computationally efficient implementation. Deep neural networks and their variants are used in state of the art AI technologies. Having an elegant theoretical basis and use as a mathematical discipline, linear algebra has applications from machine learning to game theory to economics and is a beautiful thinking tool by the side of an engineer, a scientist and a mathematician for solving and conceptualizing a vast variety of problems.

REFERENCES

[1] Gilbert Strang, Linear Algebra and its applications, MIT press,2006.
[2] Tom Apostol, One Variable Calculus with an introduction to linear algebra, Wiley India, 2017.
[3] Tom Apostol, Multi Variable Calculus and linear algebra with applications to differential equations and probability, Wiley India, 2017.
[4] Spiegel, Lipschutz and Spellman, Schaum’s Outline Series - Vector Analysis, McGraw Hill Education India, 2009.
[5] R.D.Sharma and Ritu Jain, Theory and Problems of linear algebra, I.K. International Publishing House, 2012.