Using Matrix Exponential to Solve System of Differential Equations

1st Chandrakishor Singh
Department of Information Technology
Indian Institute of Information Technology, Allahabad
mit2021117@iiita.ac.in

2nd Dablu Chauhan
Department of Information Technology
Indian Institute of Information Technology, Allahabad
mit2021115@iiita.ac.in

3rd Madhu Donipati
Department of Information Technology
Indian Institute of Information Technology, Allahabad
mit2021116@iiita.ac.in

***Abstract*—This paper presents a fundamental introduction to the matrix exponential problem. We begin with the discussion on Taylor's series to calculate the value of e^x where x is some real number. We then define the e^A where A is some square matrix in a similar way as the Taylor's series but with matrix multiplication. Then we present some mathematical preliminaries which are required to understand the matrix exponential. This includes the understanding of eigenvalues and eigenvectors, diagonalization of a matrix in terms of eigenvalue matrix and eigenvector matrix and power matrix. After that we introduce the matrix exponential problem. Finally we demonstrate the application of matrix exponential to solve a problem of the system of differential equations and give some other important applications of eigenvalues and eigenvectors.**

***Keywords*—Matrix Exponential, Eigenvalue, Eigenvector, System of Differential Equations**

I. Introduction

The matrix exponential e^A is a function of a square matrix A which is similar to the regular exponential function e^x in mathematics. It's used to solve system of linear differential equations.

We know from the Taylor's series that the value of e^x can be represented as a sum of following infinite series.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + ...$$

In a similar fashion, the matrix exponential for a square matrix A(which could be a real or complex matrix) is defined as follows.

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \frac{A^4}{4!} + ...$$

Just like how the sum of the series of e^x always converges, the sum of the matrices obtained in individual terms also converges.[4]

For example, if A is a $1 * 1$ matrix then the matrix exponential of A will also be a $1*1$ matrix whose only element will be equal to the ordinary exponential of the only element present in A.

So the straightforward way to calculate the matrix exponential would be to calculate powers of A and then find their sum according to this series. It is only practical for those matrices which become zero matrix after some power. We will see later in this paper how to obtain the value of matrix exponential easily using its eigenvalue and eigenvector.

II. Prerequisites

Before trying to understand matrix exponential and its application of solving systems of linear differential equations, one must first understand a few related concepts that are used to define the properties of matrix exponential.

A. Eigenvalue and Eigenvector

Eigenvalue and eigenvector are some of the most important properties of a matrix. They contain a lot of information regarding the nature of the matrix.

Mathematically, for a matrix A, the vector x is said to be its eigenvector if for some scalar λ the following holds true.[4]

$$Ax = \lambda x$$

λ is called the corresponding eigenvalue of A associated with the eigenvector x.

There may be multiple eigenvector and eigenvalue of a matrix A. If the order of matrix A is n then, at maximum, it will have n different eigenvalues and eigenvectors.[4]

An eigenvector, which corresponds to a real nonzero eigenvalue, points in the direction that the transformation(the matrix multiplication Ax) stretches it, and the eigenvalue is the factor by which it stretches it. The orientation is reversed if the eigenvalue is negative.[4]

The eigenvalues of a matrix A can be found out from the solution of its characteristic equation $|A - \lambda I| = 0$

Below are some properties of eigenvalues and eigenvectors which will be essential in understanding further content of this paper.

- If $Ax = \lambda x$ then $A^n x = \lambda^n x$ i.e., the eigenvectors of the powers of the same matrix remain unchanged while the eigenvalues change in an exponential manner.[4]
- If $Ax = \lambda x$ then $(A + cI)x = (\lambda + c)x$[4]

B. Diagonalization of a Matrix

In many practical applications of linear algebra, we want to convert a matrix into a diagonal matrix(if possible). This is because these types of matrices are very easy to process for a computer once their eigenvalue and eigenvectors are known. For example, to find the power of a diagonal matrix, one just has to raise the diagonal elements to the same power as all other elements are zero. Similarly, the determinant of a diagonal matrix is simply the product of the diagonal entries which is, again, easy to compute rather than calculating the determinant of the matrix from the definition of determinant.

Mathematically speaking, a matrix A is called a diagonalizable matrix if there exists an invertible matrix P and a diagonal matrix D such that $A = PD^{-1}P$. We will now see how to diagonalize a matrix in terms of its eigenvectors and eigenvalues.[5]

Let A be a square matrix of order n. Let its eigenvectors be $x_1, x_2, ..., x_n$ and the corresponding eigenvalues be $\lambda_1, \lambda_2, ..., \lambda_n$. We define the eigenvector matrix of A to be $V = [x_1, x_2, ..., x_n]$ and the eigenvalue matrix of A to be $\Lambda = [\lambda_1, \lambda_2, ..., \lambda_n]$.

$$A\Lambda = A[x_1, x_2, ..., x_n]$$

$$= [Ax_1, Ax_2, ..., Ax_n]$$

$$= [\lambda_1 x_1, \lambda_2 x_1, ..., \lambda_n x_n]$$

$$= \Lambda V$$

Hence, $A\Lambda = \Lambda V$.
Also, $A = \Lambda V \Lambda^{-1}$.
Hence,

$$A^2 = (\Lambda V \Lambda^{-1}) * (\Lambda V \Lambda^{-1}) = \Lambda V^2 \Lambda^{-1}$$

This can be generalized as $A^k = \Lambda V^k \Lambda^{-1}$.

So, we can see that the eigenvector matrix of the powers of A remains unchanged while the eigenvalues changes in an exponential manner.

Note that formula $A^k = \Lambda V^k \Lambda^{-1}$ is only applicable when A has n independent eigenvectors because otherwise we can't find the inverse of the eigenvector matrix Λ. Hence, if some eigenvalues of A are repeated then there might not be n independent eigenvectors.[5]

C. Power of Matrix

Consider vector u_k which are related to the matrix A by the equation $u_{k+1} = Au_k$. Then,

$$u_k = Au_{k-1} = A * (Au_{k-2}) = ... = A^k u_0$$

This type of relationship shows up very frequently when solving a system of linear differential equations with some initial value. These types of problems are also referred to as initial value problems or IVPs.

Here, u_0 is some initial vector and A is some matrix. We are supposed to calculate the value of u_k. It is easy to see from the equation that to calculate the value of u_i, we need to multiply the matrix A with itself i times. In other words, we have to calculate the matrix A^i in order to calculate the value of u_k.

Note that the initial vector u_0 could be any vector. If A has n independent eigenvectors $x_1, \ldots, x_n$ then they can be used as a basis. So, now u_0 can be written as some linear combination of the eigenvectors $x_1, \ldots, x_n$ as follows.

$$u_0 = c_1 x_1 + c_2 x_2 + ... + c_n x_n$$

Then we have,

$$Au_0 = A(c_1 x_1 + c_2 x_2 + ... + c_n x_n) = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + ... + c_n \lambda_n x_n$$

where $\lambda_1, \lambda_2, ..., \lambda_n$ are eigenvalues.
In the same manner,

$$A^2 u_0 = A(c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + ... + c_n \lambda_n x_n) = c_1 \lambda_1^2 x_1 + c_2 \lambda_2^2 x_2 + ... + c_n \lambda_n^2 x$$

Hence, it is easy to follow that

$$u_k = A^k u_0 = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2 + ... + c_n \lambda_n^k x_n$$

In matrix notation we can write this as $u_k = A^k u_0 = V \Lambda^k V^{-1} u_0$.

Notice that now, in order to calculate the vector u_k, we only need to know the eigenvalues of matrix A and don't have to actually multiply the entire matrix A with itself repeatedly. This is a much more convenient way of calculating the vector u_k.

III. Matrix Exponential

At first glance, the matrix exponential seems a vague concept but it is actually presented as a definition. Mathematically, it is well defined as follows.[4]

$$e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \frac{(At)^4}{4!} + ...$$

Just like how an exponential function e^{at} is a solution to the differential equation $y' = ay$.

We can extend this concept into a system of linear differential equations using matrix exponential. For example, if we have n linear differential equations of type $y' = Ay$, where y and y' are vectors, then its solution can be given as $y(t) = e^{At} y(0)$.[4]

We can verify that the above given solution is correct by substituting this definition into the original differential equation as follows.

$$\frac{d(e^{At} * y(0))}{dt} = y(0) \frac{d(e^{At})}{dt}$$

$$= y(0)(\frac{d(I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \frac{(At)^4}{4!} + ...)}{dt})$$

$$= y(0)(\frac{dI}{dt} + At + \frac{d(\frac{(At)^2}{2!})}{dt} + \frac{d(\frac{(At)^3}{3!})}{dt} + \frac{d(\frac{(At)^4}{4!})}{dt} + ...)$$

$$= y(0)(A + A^2t + \frac{A^3t^2}{2!} + ...)$$

$$= y(0)A(I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \frac{(At)^4}{4!} + ...)$$

$$= y(0)Ae^{At} = Ay(t)$$

Hence, it is easy to see that the matrix exponential gives a solution to the system of linear differential equations in a similar manner as done by the exponential function e^{at} for an ordinary differential equation.

IV. Application in Solving System of Differential Equations

The fact that the matrix exponential can be used to solve systems of linear ordinary differential equations is one of the reasons for its relevance. The solution of $\frac{d}{dt}y(t) = Ay(t)$ where A is a constant matrix, is given by $y(t) = e^{At}y_0$.

Now, we will solve one problem of a system of linear differential equations that has some initial value condition to demonstrate how matrix exponentials are helpful in solving such problems.

Let $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$. Compute e^{At} and use it to solve initial value problem $u^{'}(t) = Au(t)$ with $u_0 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

First, we need to find the eigenvalues and eigenvectors of this matrix.

$$\begin{vmatrix} 6-\lambda & 5 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

$$(6-\lambda)(2-\lambda) - 5 = 0$$

$$\lambda_1 = 1, \lambda_2 = 7$$

Finding eigenvectors $v_1 and v_2$,

$$\begin{bmatrix} 6-1 & 5 \\ 1 & 2-1 \end{bmatrix} v_1 = 0$$

Hence, $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Similarly,

$$\begin{bmatrix} 6-7 & 5 \\ 1 & 2-7 \end{bmatrix} v_2 = 0$$

Hence, $v_2 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$

Now, from the definition of matrix exponential, we can see that, $u = c_1e^tv_1 + c_2e^{7t}v_2$.

$$= \begin{bmatrix} e^t & 5e^{7t} \\ -e^t & e^{7t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

Let,

$$\phi(t) = \begin{bmatrix} e^t & 5e^{7t} \\ -e^t & e^{7t} \end{bmatrix}$$

From the intial condition, we know that, $u = e^{A*0}u(0) = u(0)$.

So,

$$\phi(0) = \begin{bmatrix} e^0 & 5e^{7*0} \\ -e^0 & e^{7*0} \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -1 & 1 \end{bmatrix}$$

Therefore, its inverse will be,

$$\phi^{-1}(0) = \frac{1}{(1)(1)-(-1)(5)} \begin{bmatrix} 1 & 5 \\ -1 & 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 1 & 5 \\ -1 & 1 \end{bmatrix}$$

Hence, $e^{At} = \phi(t)\phi^{-1}(0)$

$$= \begin{bmatrix} e^t & 5e^{7t} \\ -e^t & e^{7t} \end{bmatrix} \frac{1}{6} \begin{bmatrix} 1 & -5 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} e^t + 5e^{7t} & -5e^t + 5e^{7t} \\ -e^t + e^{7t} & 5e^t + e^{7t} \end{bmatrix}$$

Hence, $u(t) = e^{At}u(0)$

$$= \frac{1}{6} \begin{bmatrix} e^t + 5e^{7t} & -5e^t + 5e^{7t} \\ -e^t + e^{7t} & 5e^t + e^{7t} \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} -e^t + 25e^{7t} \\ e^t + 5e^{7t} \end{bmatrix}$$

V. Other Important Applications

Eigenvalues and eigenvectors have many other applications as well. Some of them are listed below.

- **Machine Learning:** Eigenvalues are mostly used in machine learning for dimensionality reduction e.g. face recognition. Computational exhaustive tasks can be reduced to lower-dimensional calculations by removing maximally correlated data to avoid overfitting of data. The principal component analysis is a tool used to reduce dimensions of the multivariate set to a smaller set which can still hold most of the information of a large set with loss of very little information. Eigenvectors corresponding to eigenvalues work as the base for Principal Component Analysis (PCA). PCA is like the fitting of the dimensional ellipsoid to data where each axis represents the principal axis. A smaller axis with minimum variance can be omitted; another axis of maximum variance can coincide with the direction of maximum variation of original observations. Therefore, PCA is the procedure to form a basis of a new set of vectors to represent directions, such that dimension vectors are orthogonal which makes them linearly independent and ranked according to the variance of original data. In mathematical terms, it is an orthogonal linear transformation used to convert a matrix into a diagonalizable matrix. Mathematical operations reduce the number of correlated variables into a small number of uncorrelated variables called principal components. For this purpose, eigenvalues are evaluated on the variables' correlation matrix. [1]

- **Vibrations:**
  Vibrations are widely used in mechanical engineering; civil engineering, communication engineering, and automobile engineering. The axial motion of the bar, beam vibrations, buckling of a beam bar, the stereo system of the car, aerodynamics, etc. The continuous cyclic motion of an equilibrium point is called vibration. Engineers try to avoid it due to damaging effects caused by cyclic forces created by cyclic motions, unwanted sounds, and wasting of energy. The collapse of the famous Tacoma Narrows suspension bridge was a result of unforeseen vibrations. The mass-spring-damper model, one of the basic concepts of vibrations, uses eigenvalues to calculate frequency and mode shape for vibrations. The natural frequency of the bridge is the eigenvalue of the smallest magnitude of a system that models the bridge. The engineers exploit this knowledge to ensure the stability of their constructions. [2]
- **Communication Systems:**
  Eigenvalues were used by Claude Shannon to determine the theoretical limit to how much information can be transmitted through a communication medium like your telephone line or through the air. This is done by calculating the eigenvectors and eigenvalues of the communication channel which is expressed in a matrix and then water-filling on the eigenvalues. The eigenvalues are then, in essence, the gains of the fundamental modes of the channel, which themselves are captured by the eigenvectors. [3]
- **Oil companies frequently use eigenvalue analysis to explore land for oil:**
  Oil, dirt, and other substances all give rise to linear systems which have different eigenvalues, so eigenvalue analysis can give a good indication of where oil reserves are located. Oil companies place probes around a site to pick up the waves that result from a huge truck used to vibrate the ground. The waves are changed as they pass through the different substances in the ground. The analysis of these waves directs the oil companies to possible drilling sites. [3]

VI. Conclusion

This report draws attention towards matrix exponential. We have explained the major prerequisites required to understand the matrix exponential and how it is used in solving system of linear differential equations. Finally, we have demonstrated the concept via a typical problem.

Contributors

- Chandrakishor Singh : Research & Writing on Abstract, Introduction, Prerequisites(except Eigenvalue and Eigenvector), Matrix Exponential, Application in solving System of Differential Equations
- Dablu Chauhan : Research & Writing on Other Important Applications
- Madhu Donipati : Research & Writing on Eigenvalue and Eigenvector, Editing and formatting

Acknowledgment

We would like to express our special thanks of gratitude to our teacher Dr. Mohammed Javed, who gave us the assignment which helped me to enhance our understanding in the field of digital image processing and application of linear transformation. We also want to thank Mr. Bulla Rajesh, who mentored us to complete this assignment Successfully.

References

[1] Yann Le Cun, Ido Kanter, and Sara A. Solla, "Eigenvalues of covariance matrices: Application to neural-network learning", Phys. Rev. Lett., Vol. 66, 1991.
[2] W Thomson, "Theory of vibration with applications", Upper Saddle River, NJ: Prentice-Hall, 1998.
[3] https://www.cpp.edu/ manasab/eigenvalue
[4] Introduction to Linear Algebra, 5th Edition - MIT Mathematics
[5] Leonard, Isaac. (1996). The Matrix Exponential. SIAM Review. 38. 10.1137/S0036144595286488.