Bad, Good, Better, Best Matrices

Ranveer*

"...beautiful mathematics eventually tends to be useful, and useful mathematics eventually tends to be beautiful"- Anonymous

Disclaimer: none of the matrix, in fact, none of mathematical entity, idea can be bad. It is my very, very limited knowledge and experience about the matrices which characterize that some matrices are easy to work with while others are not so.

1 A one-line motivation

Given a square matrix M, how fast can you calculate its k-th power, that is, the matrix M^k for some positive integer k?

Of course, one can naively multiply M iteratively k times to get A^k. Or, we can reduce the multiplications by first calculating A^2, then multiplying two A^2 to get A^4, two A^4 to get A^8, and so on. Can we do it more efficiently? Can for some matrices we can do it more efficiently than others? We will try to find the answer, and in the process, we will know yet another beautiful use of eigenvalues and eigenvectors.

Calculating matrix powers is a need of various analyses. For example, given a network of cities, in how many ways you can commute between any two cities x, y in k-step? The answer is the (x, y)-th entry of the matrix A^k, where A is the adjacency matrix of the given network. Although computing matrix powers in not the sole reason for this write-up. There are innumerable applications where different type of matrices are involved so one should know about the nature of matrices. However, mathematics doesn't really care about the applications, it has its eternal sublime beauty, and the applications inevitably follow it.

2 Let us start

We will start with a seemingly dull and unrelated question which will turn out to be an exciting one later on. Given a square matrix M of order n, how many eigenvalues are there? How many eigenvectors are there? For those who are wondering what these two entities, eigenvalues and eigenvectors are, let us have a very brief discussion. What happens when we multiply M with a nonzero column vector v of order n? It gives some column vector Mv. Sometimes v happens to be so special that Mv is some scalar multiple of v itself. That is, $Mv = \lambda v$, for some scalar λ. This special vector v is then called as an eigenvector, and the associated scalar λ is an eigenvalue. But how to find a pair of v, λ for a square matrix M? In the equation $Mv = \lambda v$ only M is

*Email: ranveer@iiti.ac.in,

known rest two are unknowns. Fortunately, our school mate determinant comes to help us. We can write $(M - \lambda I)v = 0$, this 0 is zero-vector, that is, all its entries are zero. This implies that the determinant of the matrix $M - \lambda I$ must be zero. (If you are again wondering why it is so check this neat fact here again.) The determinant of $M - \lambda I$ gives a polynomial $p_M(\lambda)$, which we call the characteristic polynomial of M. Since $p_M(\lambda)$ has degree n, there are n roots, and these are precisely the eigenvalues of M. What about the eigenvectors? Check that if $Mv = \lambda v$, then for any nonzero scalar c we have $M(cv) = \lambda(cv)$. This means that there can be many, many (infinite) eigenvectors associated with the same eigenvalue λ.

Ok, so M has n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and suppose these are arranged in a decreasing order, that is, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let us pick an eigenvector v_i for the eigenvalue λ_i, $i = 1, \ldots, n$. So we have a set V of eigenvectors $v_1, v_2, \ldots, v_n$. Its time to look at some matrices; these will serve as a good warm up for our study.

1. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Here, $p_A(\lambda) = \lambda^2 - 1$. So A has the eigenvalues $\lambda_1 = 1, \lambda_2 = -1$. To find an eigenvector v_1 for λ_1, we should have $(A - \lambda_1 I)v_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} v_1 = 0$. We see that $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ satisfies it, so it is an eigenvector for λ_1. Similarly we can calculate that $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector for λ_2. Check v_1 and v_2 are not scaled versions of each other, that is, $v_1 \neq cv_2$ for any nonzero scalar c. We call that v_1, v_2 are independent of each other.

2. $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Here, $p_B(\lambda) = (1 - \lambda)^2$. So B has the eigenvalues $\lambda_1 = \lambda_2 = 1$, that is, the eigenvalues are repeated. Unlike A both the eigenvalues are the same for B. To find an eigenvector v_1 for λ_1, we should have $(B - \lambda_1 I)v_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} v_1 = 0$. We see that $v_1 = c \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for any nonzero scalar c satisfies it, so it is an eigenvector for λ_1. But now let us try to find an eigenvector v_2 for λ_2. Again we have to satisfy $(B - \lambda_1 I)v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} v_2 = 0$. Which again gives $v_1 = d \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for any nonzero scalar d. Thus v_1 and v_2 are the same vectors up to some nonzero scaling factor. That is, they are dependent vectors.

3. $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, the identity matrix. Here, again like B we have $p_C(\lambda) = (1 - \lambda)^2$. So C has two eigenvalues $\lambda_1 = \lambda_2 = 1$, the repeated eigenvalues. To find an eigenvector v_1 for λ_1, we should have $(C - \lambda_1 I)v_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} v_1 = 0$. Waaw!! we see that every nonzero vector v_1 satisfies it, that is, any nonzero vector is an eigenvector. So we can choose $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Similarly for λ_2 we can choose an eigenvector $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. This selection is special as we see v_1 and v_2 are independent.

A crucial point that is to be noted is that in B and C both have repeated eigenvalues, but for C we are able to get two independent eigenvectors, while in B, we are not able to get them.

3 The curious case of repeating eigenvalues

So till now we know that for a square matrix M of order n there are n eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and we can associate some n eigenvectors $v_1, v_2, \ldots, v_n$ to these eigenvalues such that v_i is an eigenvector for λ_i, $i = 1, \ldots, n$. We can write

$$MV = VD, \tag{1}$$

where V is the matrix with i-th column as v_i and D is a diagonal matrix with i-th diagonal entry equals λ_i (its off-diagonal entries are zero). For example, for the matrix A above, we can write

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

The next question is whether, in the equation $MV = VD$ we can extract M. Yes we can if V is an invertible matrix. In that case, we can multiply both sides of $MV = VD$ by V^{-1} to get $M = VDV^{-1}$. For example, for the matrix A, we can write

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

But what is so significant about it? Let us try to answer our motivational question. Let us try to find M^k. First see what is A^2, it is $VDV^{-1}VDV^{-1} = VDIDV^{-1} = VD^2V^{-1}$. Next, let us figure out what A^3 is. It is $A^2A = VD^2V^{-1}VDV^{-1} = VD^2IDV^{-1} = VD^3V^{-1}$. Now the following elegant form for A^k is apparent.

$$A^k = VD^kV^{-1}. \tag{2}$$

Calculating D^k is straightforward, its i-th entry will be λ_i^k. Thus once we know the eigenvalues, eigenvectors, and inverse of matrix V for A, we are done. But hold on! this is only possible when V is invertible!! Is V always invertible? No!. For matrix B above the corresponding matrix V is

$$\begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix},$$

which is not invertible.

So when is a square matrix invertible? Its when all its columns or rows are linearly independent. Thus when the eigenvectors $v_1, \ldots, v_n$ associated with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of M are linearly independent then the matrix V is invertible otherwise, not. So our query boils down to the cases when the eigenvectors $v_1, \ldots, v_n$ will be independent.

3.0.1 When eigenvalues are distinct

Suppose $\lambda_1, \ldots, \lambda_n$ all are distinct. Let us see what can be said about the invertibility of V. Without loss of generality assume that $\{v_1, \ldots, v_k\}$ is a maximal set of independent eigenvectors from the set $\{v_1, \ldots, v_n\}$.

Note that $1 \leq k \leq n$. That means we can write v_{k+1} as some linear combinations

$$v_{k+1} = \alpha_1 v_1 + \cdots + \alpha_k v_k,$$

for some scalars $\alpha_1, \ldots, \alpha_k$ not all zero. On multiplying with M on both sides it implies that

$$\lambda_{k+1} v_{k+1} = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k. \tag{3}$$

Also

$$\lambda_{k+1} v_{k+1} = \alpha_1 \lambda_{k+1} v_1 + \cdots + \alpha_k \lambda_{k+1} v_k. \tag{4}$$

Subtracting eqn (3) from eqn (4) gives

$$0 = \alpha_1(\lambda_{k+1} - \lambda_1) v_1 + \cdots + \alpha_k(\lambda_{k+1} - \lambda_k) v_k.$$

Since $v_1, \ldots, v_k$ are independent, $\alpha_i(\lambda_{k+1} - \lambda_i) = 0$, $i = 1, \ldots, k$. But as $\lambda_{k+1} \neq \lambda_i$, it implies that $\alpha_i = 0$, $i = 1, \ldots, k$. This gives a contradiction that we can write v_{k+1} as a linear combination of $v_1, \ldots, v_k$. It gives the following theorem.

Theorem 3.1. *Eigenvectors associated with distinct eigenvalues are independent.*

3.0.2 When eigenvalues are repeated

We have seen for B, C both have repeated eigenvalues, but B does not have independent eigenvectors associated with the eigenvalue while C has. In more precise terms, B has just one independent eigenvector for the eigenvalue 1, but C has two independent eigenvectors for the eigenvalue 1. In both the matrices the eigenvalue 1 is repeated two times.

The number of times an eigenvalue λ is repeated is called its algebraic multiplicity. The number of independent eigenvectors associated with an eigenvalue is called is geometric multiplicity. So, for matrix B the eigenvalue 1 has algebraic multiplicity 2 while the geometric multiplicity 1. Where as for matrix C the eigenvalue 1 has algebraic multiplicity 2 and the geometric multiplicity is also 2. Following is a fascinating fact about these two interesting numbers.

Theorem 3.2. *For any eigenvalue the geometric multiplicity can not exceed the algebraic multiplicity.*

The proof of this fascinating fact is given in Appendix 9. Theorem 3.2 has very important implications. First suppose for a matrix M there is an eigenvalue λ whose geometric multiplicity is less than the algebraic multiplicity, then it is clear that there can not be a set $v_1, \ldots v_n$ of independent eigenvectors, and V must be non-invertible. Next consider the case when, for all the eigenvalues of M the geometric multiplicities are equal to algebraic multiplicity. Let $\mu_1, \ldots, \mu_k$ are the distinct eigenvalues with multiplicities $n_1, \ldots, n_k$, respectively. Let $S_i = \{u_1^i, \ldots, u_{n_i}^i\}$ be the set of n_i independent eigenvectors for the eigenvalue μ_i, $i = 1, \ldots, k$. Owing to

Theorem 3.1 for any $j \neq i$, any eigenvector for μ_j must be independent of S_i. Since $\sum n_i = n$, the union of vectors $S_1 \cup \cdots \cup S_k$ gives n independent eigenvectors. It gives the following theorem.

Theorem 3.3. *The matrix V is invertible if and only if for all the eigenvalues the algebraic multiplicity equals the geometric multiplicity.*

Theorem 3.3 will serve as a criterion to classify the matrices into bad, good, better, best classes.

4 Bad matrices

We place the matrices which do not have a full set of independent eigenvectors into the bad class. That is, we place a matrix M in the bad class if there is no invertible matrix V such that $M = VDV^{-1}$ or equivalently, $V^{-1}MV = D$. In a decent formal language, we call M to be an nondiagonalisable or defective matrix as the diagonal form $V^{-1}MV = D$ is not possible. Matrix B is an example of such a matrix. Consider one more matrix

$$M = \begin{bmatrix} -1 & 6 & 4 \\ -5 & 10 & 6 \\ 6 & -9 & -5 \end{bmatrix}.$$

The eigenvalues of M are $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 1$. So the eigenvalue 2 has algebraic multiplicity 1 and the eigenvalue 1 has algebraic multiplicity 2. Now let us find the algebraic multiplicities. The eigenvectors of M are

$$v_1 = \begin{bmatrix} 0.53 \\ -0.27 \\ 0.80 \end{bmatrix}, \; v_2 = \begin{bmatrix} 0 \\ -0.55 \\ 0.83 \end{bmatrix}, \; v_3 = \begin{bmatrix} 0 \\ -0.55 \\ 0.83 \end{bmatrix}$$

We see that the geometric multiplicity of the eigenvalue 2 is 1, which is equal to its algebraic multiplicity. However as v_2 and v_3 are not independent, the geometric multiplicity of the eigenvalue 1 is which is less than its algebraic multiplicity. This is a bad case since, in this case, the matrix V consisting of eigenvectors v_1, v_2, and v_3 is not invertible.

5 Good matrices

Next, it is obvious now we are going to place the matrices, which have a full set of independent eigenvectors into the good class. That is, we place a matrix M in the bad class if there is an invertible matrix V such that $M = VDV^{-1}$. Formally we call them diagonalizable matrix. Matrix A and C are example of such a matrix. Consider one more matrix

$$M = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

The eigenvalues of M are $\lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 1$. So the eigenvalue 2 has algebraic multiplicity two and the eigenvalue 1 has algebraic multiplicity 1. Now let us find the algebraic multiplicities. The eigenvectors of M are

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \; v_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

As v_1, v_2 are two independent vectors; we see that the geometric multiplicity of the eigenvalue 2 is two, which is equal to its algebraic multiplicity. The geometric multiplicity of the eigenvalue 1 is one which is also equal to its algebraic multiplicity. This is a good case since we can write

$$M = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}^{-1}.$$

And for any integer K

$$M^k = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 1^k \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}^{-1}.$$

6 Better matrices

What do we expect next? Recall that M is good when there is an invertible matrix V, which leads to $M^k = VD^kV^{-1}$ as intermediate products gets vanish due the fact that $V^{-1}V = I$. Let us ask for more; what if $V^TV = I$? In this case we not even need to calculate the inverse of V, the inverse of V is simply its transpose V^T. But is this possible? Are there such matrices for which V happens to be like this? The answer is yes. In fact, we all have seen these matrices quite often as we frequently encounter them in our life. These are symmetric matrices!!. These matrices display very beautiful properties.

Let M be a symmetric matrix of order n. First of all, its all the eigenvalues $\lambda_1, \ldots, \lambda_n$ are real (take this as an exercise, an easy one). The next thing about M, which is of utmost importance to us is that all its eigenvectors $v_1, \ldots, v_n$ are independent. It is due to Theorem 3.3 and the following theorem.

Theorem 6.1. *For a symmetric matrix, any eigenvalue has the same geometric multiplicity and the algebraic multiplicity.*

The proof is given in Appendix 9, Theorem 9.2. In fact we can say even more; we can always choose $v_1, \ldots, v_n$ such that $v_i^Tv_j = 0$ for $i \neq j$ and $v_i^Tv_j = 1$ for $i = j$. That is, $v_1, \ldots, v_n$ can be choosen to be orthonormal vectors. Moreover this is possible if and only if M is a symmetric. The following is a beautiful theorem about the symmetric matrices, also known as spectral theorem.

Theorem 6.2. *Any square matrix M of order n is a symmetric matrix if and only if*

$$M = VDV^T, \text{ equivalently, } V^T MV = D,$$

where V is a orthogonal matrix, (that is, $V^T V = VV^T = I$) with columns $v_1, v_2, \ldots, v_n$ which are the eigenvectors of M corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, and $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

The proof of spectral theorem is given in Appendix 9, Theorem 9.3.
Both A and C are examples of such matrices. Consider one more example

$$M = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & -4 \end{bmatrix}.$$

The eigenvalues of M are $\lambda_1 = 3, \lambda_2 = 1.37, \lambda_3 = -4.37$. The eigenvectors of M are

$$v_1 = \begin{bmatrix} 0.18 \\ -0.18 \\ 0.97 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -0.68 \\ 0.68 \\ -0.25 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0.71 \\ 0.71 \\ 0 \end{bmatrix}$$

Check that it is a better case since we can write

$$M = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & -4 \end{bmatrix} = \begin{bmatrix} 0.18 & -0.68 & 0.71 \\ -0.18 & 0.68 & 0.71 \\ 0.97 & -0.25 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1.37 & 0 \\ 0 & 0 & -4.37 \end{bmatrix} \begin{bmatrix} 0.18 & -0.68 & 0.71 \\ -0.18 & 0.68 & 0.71 \\ 0.97 & -0.25 & 0 \end{bmatrix}^T.$$

And for any integer K

$$M^k = \begin{bmatrix} 0.18 & -0.68 & 0.71 \\ -0.18 & 0.68 & 0.71 \\ 0.97 & -0.25 & 0 \end{bmatrix} \begin{bmatrix} 3^k & 0 & 0 \\ 0 & 1.37^k & 0 \\ 0 & 0 & -4.37^k \end{bmatrix} \begin{bmatrix} 0.18 & -0.68 & 0.71 \\ -0.18 & 0.68 & 0.71 \\ 0.97 & -0.25 & 0 \end{bmatrix}^T.$$

7 Best matrices

Consider the following good matrix (as it is a symmetric matrix)

$$M = \begin{bmatrix} 3 & 0 & -1 & -3 \\ 0 & 7 & 1 & -1 \\ -1 & 1 & 2 & 2 \\ -3 & -1 & 2 & 6 \end{bmatrix}.$$

Its eigenvalues are $\lambda_1 = 8.76, \lambda_2 = 7.13, \lambda_3 = 1.35, \lambda_4 = 0.76$. Check that all the eigenvalues of M are nonnegative. A symmetric matrix whose all the eigenvalues are nonnegative is known as a positive semidefinite matrix, and when the eigenvalues are positive it is known as a positive definite matrix. Check that M is positive definite. These matrices have very high importance for different applications and display even more beautiful properties.

Let M be a symmetric positive semidefinite matrix of order n. Using Theorem 9.3 we can write $M = VDV^T$, where V is the matrix with columns $v_1, v_2, \ldots, v_n$ which are the eigenvectors of M corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, and V is an orthogonal matrix. Since the eigenvalues are nonnegative we can write $D = D^{\frac{1}{2}} D^{\frac{1}{2}}$, where $D^{\frac{1}{2}}$ is a diagonal matrix with i-th diagonal entry equals to $\sqrt{\lambda_i}$. Let $\hat{V} = VD^{\frac{1}{2}}$, then we can write $M = VD^{\frac{1}{2}} D^{\frac{1}{2}} V^T = \hat{V}\hat{V}^T$. It implies that for any nonzero vector x, we have $x^T M x = x^T \hat{V}\hat{V}^T x = (\hat{V}^T x)^T (\hat{V}^T x) \geq 0$. This is a very useful information as it gives a beautiful quadratic form. Positive (semi) definite matrices are extremely important in optimization and machine learning problems.

Matrix C is an example of such a matrix. Another example is

$$M = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix},$$

its eigenvalues are $\lambda_1 = 3.80, \lambda_2 = 2.45, \lambda_3 = 0.75$.

8 Finding Fibonacci or Hemachandra numbers

We all know that Fibonacci or Hemachandra numbers are a sequence of numbers, $H_1 = 1, H_2 = 2$, and $H_i = H_{i-1} + H_{i-2}$, for $i \geq 3$. In a matrix form, we can write

$$\begin{bmatrix} H_{n+1} \\ H_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} H_n \\ H_{n-1} \end{bmatrix}.$$

Expanding it we can write

$$\begin{bmatrix} H_{n+1} \\ H_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} H_n \\ H_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} H_{n-1} \\ H_{n-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^3 \begin{bmatrix} H_{n-2} \\ H_{n-3} \end{bmatrix} = \cdots = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Let

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Its eigenvalues are $\lambda_1 = \phi, \lambda_2 = 1 - \phi$, where $\phi = \frac{1+\sqrt{5}}{2}$, the golden ratio. The eigenvectors are

$$v_1 = \begin{bmatrix} \phi \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 - \phi \\ 1 \end{bmatrix}.$$

As v_1 and v_2 are independent, we can write

$$\begin{aligned} M^k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k &= \begin{bmatrix} \phi & 1-\phi \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi^k & 0 \\ 0 & (1-\phi)^k \end{bmatrix} \begin{bmatrix} \phi & 1-\phi \\ 1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \phi & 1-\phi \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi^k & 0 \\ 0 & (1-\phi)^k \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \phi-1 \\ -1 & \phi \end{bmatrix}. \end{aligned}$$

Thus,

$$\begin{aligned} \begin{bmatrix} H_{n+1} \\ H_n \end{bmatrix} &= \frac{1}{\sqrt{5}} \begin{bmatrix} \phi & 1-\phi \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi^{n-1} & 0 \\ 0 & (1-\phi)^{n-1} \end{bmatrix} \begin{bmatrix} 1 & \phi-1 \\ -1 & \phi \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \phi & 1-\phi \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi^{n-1} & 0 \\ 0 & (1-\phi)^{n-1} \end{bmatrix} \begin{bmatrix} \phi \\ \phi-1 \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \phi^{n+1} - (1-\phi)^{n+1} \\ \phi^n - (1-\phi)^n \end{bmatrix}. \end{aligned}$$

Which gives

$$H_n = \frac{\phi^n - (1-\phi)^n}{\sqrt{5}}.$$

9 Appendix

Theorem 9.1. *The geometric multiplicity of an eigenvalue cannot exceed the algebraic multiplicity.*

Proof. Let M be a square matrix of order n having an eigenvalue λ with geometric multiplicity r. That means there are r independent eigenvectors $u_1, u_2, \ldots, u_r$ corresponding to λ. To these eigenvectors by including some more suitable $n-r$ independent vectors we can form a basis $B = \{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ for $\mathbb{R}^n$. Let us consider the matrix

$$W = [u_1 \ u_2 \ \ldots \ u_r \ u_{r+1} \ \ldots \ u_n].$$

Then

$$MW = W \begin{bmatrix} \lambda I_r & X \\ O & Y \end{bmatrix}, \tag{5}$$

for some matrices $X \in \mathbb{R}^{r\times(n-r)}$, $Y \in \mathbb{R}^{(n-r)\times(n-r)}$.

By equation 5 we can write

$$(M - xI_n)W = W \begin{bmatrix} (\lambda - x)I_r & X \\ O & Y - xI_{n-r} \end{bmatrix}.$$

Taking determinants on both sides, we have

$$\det(M - xI_n)\det W = \det W \det \begin{bmatrix} (\lambda - x)I_r & X \\ O & Y - xI_{n-r} \end{bmatrix}.$$

Since $\det W$ is nonzero, we have

$$\det(M - xI_n) = \det \begin{bmatrix} (\lambda - x)I_r & X \\ O & Y - xI_{n-r} \end{bmatrix}.$$

$\det(M - xI_n)$ is a characteristic polynomial; thus the eigenvalues are the roots of the polynomial

$$(\lambda - x)^r \det(Y - xI_{n-r}),$$

thus λ has algebraic multiplicity at least r as λ can appear as roots of polynomial $\det(Y - xI_{n-r})$ too.

□

Theorem 9.2. *For a symmetric matrix, any eigenvalue has the same geometric multiplicity and the algebraic multiplicity.*

Proof. **The proof of Theorem 9.1 specialized for symmetric matrices:** We can choose the basis B to be a set of orthonormal vectors. So we have $WW^T = I$, we can write

$$M = W \begin{bmatrix} \lambda I_r & X \\ O & Y \end{bmatrix} W^T.$$

Next, as M is symmetric we have,

$$M = W \begin{bmatrix} \lambda I_r & X \\ O & Y \end{bmatrix} W^T = W \begin{bmatrix} \lambda I_r & O \\ X & Y \end{bmatrix} W^T.$$

This Implies that $X = O$. So

$$MW = W \begin{bmatrix} \lambda I_r & O \\ O & Y \end{bmatrix}. \tag{6}$$

As we have seen

$$\det(M - xI_n) = (\lambda - x)^r \det(Y - xI_{n-r}).$$

Suppose the algebraic connectivity of λ is more than r, that is, λ is also a root of $\det(Y - xI_{n-r})$. In other words, λ is also an eigenvalue for Y, and let v be a corresponding eigenvector. Now consider a vector $\begin{bmatrix} 0 \\ v \end{bmatrix} \in \mathbb{R}^n$. By equation 6 we get

$$
\begin{aligned}
MW\begin{bmatrix}0\\ v\end{bmatrix} &= W\begin{bmatrix}\lambda I_r & O\\ O & Y\end{bmatrix}\begin{bmatrix}0\\ v\end{bmatrix}\\
&= \lambda W\begin{bmatrix}0\\ v\end{bmatrix}.
\end{aligned}
$$

That means that $W\begin{bmatrix}0\\ v\end{bmatrix}$ is also an eigenvector of M associated with the eigenvalue λ. Observe that the vector $W\begin{bmatrix}0\\ v\end{bmatrix}$ is a linear combination of the vectors $u_{r+1}, \dots, u_n$. But as it is an eigenvector for M, it must be a linear combination of the vectors $u_1, \dots, u_r$. But this makes the set of vectors $u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_n$ a dependent set which is a contradiction as this set is a basis for $\mathbb{R}^n$. So the algebraic connectivity of λ cannot be more than r contrary to our assumption. Hence for symmetric matrices, the algebraic multiplicity and the geometric multiplicity of any eigenvalue are the same. $\square$

Theorem 9.3. *Let M be a symmetric matrix of order n. Then*

$$M = VDV^T, \textit{ equivalently, } V^T MV = D,$$

where V is a orthogonal matrix, (that is, $V^T V = VV^T = I$) with columns $v_1, v_2, \dots, v_n$ which are the eigenvectors of M corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$, and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Proof. We prove the theorem using induction on n. The case $n = 1$ is true, since $M = 1M1^T$. Assume that the theorem is true when the order of matrix is $n - 1$.
Consider the case when the order of M is n. Using the eigenvector v_1, let us make an orthonormal basis of $\mathbb{R}^n$, suppose this be $S = \{v_1, x_2, \dots, x_n\}$. Note that since M is symmetric, for any $i = 2, \dots, i = n$

$$(Mx_i)^T v_1 = x_i^T M^T v_1 = x_i^T M v_1 = \lambda_1 x_i^T v_1 = 0. \tag{7}$$

Now, let Q be an orthogonal matrix whose columns are $v_1, x_2, \dots, x_n$. Then using 7

$$
\begin{aligned}
Q^T MQ &= \begin{bmatrix} - & v_1^T & - \\ - & x_2^T & - \\ \vdots & \vdots & \vdots \\ - & x_n^T & - \end{bmatrix} \begin{bmatrix} | & | & \dots & | \\ \lambda_1 v_1 & Mx_2 & \dots & Mx_n \\ | & | & \dots & | \end{bmatrix}\\
&= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{M} \end{bmatrix}.
\end{aligned}
$$

Note that $\hat{M}$ is a symmetric matrix of order $n-1$. By induction hypothesis, there exists an orthogonal matrix

$\hat{Q}$ such that $\hat{Q}^T \hat{M} \hat{Q} = \hat{D}$ is a diagonal matrix. Considering the matrix

$$P = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix},$$

we see that the product

$$P^T Q^T M Q P = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix}^T \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{M} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{D} \end{bmatrix} \tag{9}$$

is a diagonal matrix. As $P^T Q^T Q P = I$, matrix QP is an orthogonal matrix. By (8) QP contains the eigenvectors of M and the corresponding eigenvalues are in the diagonal matrix $P^T Q^T M Q P$. It completes the proof. □