Contents

Chapter 6

Eigen Values, Eigen Vectors, and Matrix Diagonalization

This is Chapter 6 of "Sophomore Level Self-Teaching Webbook for Computational and Algorithmic Linear Algebra and n-Dimensional Geometry" by Katta G. Murty.

6.1 Definitions, Some Applications

Eigen values and **eigen vectors** are only defined for square matrices; hence in this chapter we will consider square matrices only.

Let $A = (a_{ij})$ be a square matrix of order n. Let I denote the unit matrix of order n. A scalar λ is said to be an **eigen value** (or **characteristic value**, or **proper value**, or **latent value**) of square matrix A if the matrix $A - \lambda I$ is singular, i.e., if the determinant $\det(A - \lambda I) = 0$.

Let $P(\lambda) = \det(A - \lambda I)$. Expansion of $\det(A - \lambda I)$ produces the polynomial in λ, $P(\lambda)$ which is known as the **characteristic polynomial** for the square matrix A. The degree of $P(\lambda)$ is n = order of the matrix A, and the leading term in it is $(-1)^n \lambda^n$.

The eigen values of A are the solutions of the **characteristic equation** for A

$$P(\lambda) = 0$$

i.e., they are the **roots of the characteristic polynomial** for A.

The set of all eigen values of A, usually denoted by the symbol $\sigma(A)$, is called **the spectrum** of A.

Since the characteristic polynomial is of degree n, it has exactly n roots by the Fundamental Theorem of Algebra. But some of these roots may be complex numbers (i.e., a number of the form $a + ib$, where a, b are real numbers and $i = \sqrt{-1}$), and some of the roots may be repeated.

Since all the entries a_{ij} in the matrix A are real numbers, all the coefficients in the characteristic polynomial are real numbers. By the fundamental theorem of algebra this implies that if $\lambda = a + ib$ is a root of the characteristic polynomial, then so is its complex conjugate $\bar{\lambda} = a - ib$. Thus the complex eigen values of A occur in complex conjugate pairs.

So, if λ_1 is an eigen value of A, then $\lambda = \lambda_1$ is a solution of $P(\lambda) = 0$. From theory of equations, this implies that $(\lambda_1 - \lambda)$ is a factor of the characteristic polynomial $P(\lambda)$. In the same way we conclude that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the various eigen values of A, then the characteristic polynomial

$$P(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda).$$

However, as mentioned earlier, some of the eigen values of A may be complex numbers, and some of the eigen values may be equal to each other. If λ_1 is an eigen value of A and $r \geq 1$ is the positive integer satisfying the property that $(\lambda_1 - \lambda)^r$ is a factor of the characteristic polynomial $P(\lambda)$, but $(\lambda_1 - \lambda)^{r+1}$ is not, then we say that r is the **algebraic multiplicity** of the eigen value λ_1 of A. If $r = 1$, λ_1 is called a **simple eigen value** of A; if $r > 1$, λ_1 is called a multiple eigen value of A with **algebraic multiplicity** r.

Thus if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are all the distinct eigen values of A including any complex eigen values, with algebraic multiplicities $r_1, \ldots, r_k$ respectively, then the characteristic polynomial

$$P(\lambda) = (\lambda_1 - \lambda)^{r_1}(\lambda_2 - \lambda)^{r_2} \ldots (\lambda_k - \lambda)^{r_k}$$

and $r_1 + \ldots + r_k = n$. In this case the spectrum of A, $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$.

If λ is an eigen value of A, since $A-\lambda I$ is singular, the homogeneous system of equations

$$(A-\lambda I)x = 0 \quad \text{or equivalently} \quad Ax = \lambda x$$

must have a nonzero solution $x \in R^n$. A column vector $x \neq 0$ satisfying this equation is called an **eigen vector** of A corresponding to its eigen value λ; and the pair (λ, x) is called an **eigen pair** for A. Notice that the above homogeneous system may have many solutions x, the set of all solutions of it is the nullspace of the matrix $(A - \lambda I)$ defined in Section 3.10 and denoted by $N(A - \lambda I)$. This nullspace $N(A - \lambda I)$ is called the **eigen space** of A corresponding to its eigen value λ. Therefore, the set of all eigen vectors of A associated with its eigen value λ is $\{x \in R^n : x \neq 0$, and $x \in N(A - \lambda I)\}$, the set of all nonzero vectors in the corresponding eigen space.

The dimension of $N(A - \lambda I)$ (as defined in Section 1.23, this is equal to the number of nonbasic variables in the final tableau when the above homogeneous system $(A - \lambda I)x = 0$ is solved by the GJ method or the G elimination method) is called the **geometric multiplicity** of the eigen value λ of A. We state the following result on multiplicities without proof (for proofs see C. Meyer [6.1]). called the **geometric multiplicity** of the eigen value λ of A.

Result 6.1.1: Multiplicity Inequality: *For simple eigen values λ of A (i.e., those with algebraic multiplicity of 1) the dimension of the eigen space $N(A - \lambda I)$ is 1; i.e., this eigen space is a straight line through the origin.*

If λ is an eigen value of A with algebraic multiplicity > 1, we will always have

$$\textit{geometric multiplicity of } \lambda \leq \textit{algebriac multiplicity of } \lambda$$

i.e., the dimension of $N(A - \lambda I) \leq$ algebriac multiplicity of λ.

Result 6.1.2: Linear Independence of Eigen vectors Associated With Distinct Eigen values: *Let $\lambda_1, \ldots, \lambda_h$ be some distinct eigen values of A; and $\{(\lambda_1, x^1), (\lambda_2, x^2), \ldots, (\lambda_h, x^h)\}$ a set of eigen* *pairs for A. Then the set of eigen vectors* $\{x^1, \ldots, x^h\}$ *is linearly independent.*

Suppose all the eigen values of A are real, and $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$ is this set of distinct eigen values. If, for all $i = 1$ to k, geometric multiplicity of λ_i is equal to its algebraic multiplicity; then it is possible to get a **complete set of n eigen vectors for** A which is linearly independent.

If geometric multiplicity of λ_i is $<$ its algebriac multiplicity for at least one i, then A does not have a complete set of eigen vectors which is linearly independent. Such matrices are called **deficient** or **defective matrices**.

Example 1: Let

$$A = \begin{pmatrix} 6 & -5 \\ 4 & -3 \end{pmatrix}. \text{ Then } A - \lambda I = \begin{pmatrix} 6-\lambda & -5 \\ 4 & -3-\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda)$ = determinant $(A - \lambda I) = (6-\lambda)(-3-\lambda) + 20 = \lambda^2 - 3\lambda + 2 = (2-\lambda)(1-\lambda)$. Hence the characteristic equation for A is

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

for which the roots are 2 and 1. So, the eigen values of A are 2 and 1 respectively, both are simple with algebraic multiplicity of one.

To get the eigen space of an eigen value λ_i we need to solve the homogeneous system $(A - \lambda_i I)x = 0$. For $\lambda = 1$ this system in detached coefficient form is the one in the top tableau given below. We solve it using the GJ method discussed in Section 1.23. Pivot elements are enclosed in a box; PR, PC denote the pivot row, pivot column respectively, and BV denotes the basic variable selected in the row. RC denotes redundant constraint to be eliminated.

BV	x_1	x_2		
	$\boxed{5}$	-5	0	PR
	4	-4	0	
	PC↑			
x_1	1	-1	0	
	0	0	0	RC
	Final Tableau			
x_1	1	-1	0	

Since there is one nonbasic variable in the final tableau, the dimension of the eigen space is one. The general solution of the system is $x = (\alpha, \alpha)^T$ for any α real. So, for any $\alpha \neq 0$ $(\alpha, \alpha)^T$ is an eigen vector corresponding to the eigen value 1. Taking $\alpha = 1$ leads to one eigen vector $x^1 = (1, 1)^T$.

For the eigen value $\lambda = 2$, the eigen space is determined by the following system, which we also solve by the GJ method.

BV	x_1	x_2		
	$\boxed{4}$	-5	0	PR
	4	-5	0	
	PC↑			
x_1	1	$-5/4$	0	
	0	0	0	RC
	Final Tableau			
x_1	1	$-5/4$	0	

So, the general solution of this system is $x = (5\alpha/4, \alpha)^T$ which is an eigen vector corresponding to the eigen value 2 for any $\alpha \neq 0$. In particular, taking $\alpha = 1$, yields $x^2 = (5/4, 1)^T$ as an eigen vector corresponding to 2.

So, $\{x^1, x^2\} = \{(1, 1)^T, (5/4, 1)^T\}$ is a complete set of eigen vectors for A.

Example 2: Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad \text{Then} \quad A - \lambda I = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda) =$ determinant $(A - \lambda I) = \lambda^2 + 1$. Hence the characteristic equation for A is

$$\lambda^2 + 1 = 0$$

for which the solutions are the imaginary numbers $\lambda_1 = i = \sqrt{-1}$ and $\lambda_2 = -i$. Thus the eigen values of this matrix A are the complex conjugate pair $+i$ and $-i$.

It can also be verified that the eigen vectors corresponding to

$\lambda_1 = +i$ are $(\alpha, -\alpha i)^T$

$\lambda_2 = -i$ are $(\alpha, \alpha i)^T$

for any nonzero complex number α. Thus the eigen vectors of A are both complex vectors in this example, even though the matrix A is real.

Example 3: Let

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 10 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \text{ Then } A - \lambda I = \begin{pmatrix} \lambda & 0 & -1 \\ 10 & 1-\lambda & 2 \\ 0 & 0 & 1-\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda) =$ determinant $(A - \lambda I) = -\lambda(1-\lambda)^2$. Hence the characteristic equation for A is

$$-\lambda(1-\lambda)^2 = 0$$

for which the roots are 0 and 1, with 1 being a double root. So, the eigen values of A are 0 and 1 , 0 is simple with algebraic multiplicity one, and 1 is multiple with algebraic multiplicity 2.

To get the eigen space of eigen value 0 we need to solve the homogeneous system $Ax = 0$. We solve it using the GJ method discussed in Section 1.23. Pivot elements are enclosed in a box; PR, PC denote the pivot row, pivot column respectively, and BV denotes the basic variable selected in the row. RC denotes redundant constraint to be eliminated.

BV	x_1	x_2	x_3		
	0	0	$\boxed{-1}$	0	PR
	10	1	2	0	
	0	0	1	0	
			PC↑		
x_3	0	0	1	0	
	10	$\boxed{1}$	0	0	PR
	0	0	0	0	RC
		PC↑			
	Final Tableau				
x_3	0	0	1	0	
x_2	10	1	0	0	

The general solution of this system is $x = (\alpha, -10\alpha, 0)^T$ for any α real. Taking $\alpha = 1$, an eigen vector corresponding to the eigen value 0 is $x^1 = (1, -10, 0)^T$.

For $\lambda = 1$, the eigen space is determined by the following system, which we also solve using the GJ method of Section 1.23.

BV	x_1	x_2	x_3		
	$\boxed{-1}$	0	-1	0	PR
	10	0	2	0	
	0	0	0	0	RC
	PC↑				
x_1	1	0	1	0	
	0	0	$\boxed{-8}$	0	PR
			PC↑		
	Final Tableau				
x_1	1	0	0	0	
x_3	0	0	1	0	

The general solution of this system is $x = (0, \alpha, 0)^T$ for any α real; so an eigen vector corresponding to the eigen value 1 is $(0, 1, 0)^T$. The eigen space corresponding to this eigenvalue is a straight line, a one dimensional object.

The algebraic multiplicity of the eigen value 1 is 2, but its geometric multiplicity, the dimension of the eigen space is only 1.

This example illustrates the fact that the geometric multiplicity of an eigen value, the dimension of its eigen space, can be strictly less than its algebraic multiplicity.

Example 4: Let

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \text{ Then } A - \lambda I = \begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda)$ = determinant $(A - \lambda I) = (\lambda + 1)^2(2 - \lambda)$. Hence the characteristic equation for A is

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

for which the roots are -1 and 2, with -1 being a double root. So, the eigen values of A are -1 and 2; 2 is simple with algebraic multiplicity one, and -1 is multiple with algebraic multiplicity 2.

The eigen space corresponding to eigen value -1 is determined by the following homogeneous system which we solve using the GJ method discussed in Section 1.23. Pivot elements are enclosed in a box; PR, PC denote the pivot row, pivot column respectively, and BV denotes the basic variable selected in the row. RC denotes redundant constraint to be eliminated.

BV	x_1	x_2	x_3		
	1	1	$\boxed{1}$	0	PR
	1	1	1	0	
	1	1	1	0	
			PC↑		
x_3	1	1	1	0	
	0	0	0	0	RC
	0	0	0	0	RC
	Final Tableau				
x_3	1	1	1	0	

The general solution of this system is $x = (\alpha, \beta, -(\alpha+\beta))^T$ for any α, β real. So, this eigen space has dimension 2, it is the subspace which is the linear hull of the two vectors $(1, 0, -1)^T$ and $(0, 1, -1)^T$. The set of these two vectors is a basis for this eigen space.

Thus for the matrix A in this example, both the algebraic and geometric multiplicities of its eigen value -1 are two.

The eigen vector corresponding to the eigen value 2 is obtained from the following ststem.

BV	x_1	x_2	x_3		
	-2	1	$\boxed{1}$	0	PR
	1	-2	1	0	
	1	1	-2	0	
			PC↑		
x_3	-2	1	1	0	
	$\boxed{3}$	-3	0	0	PR
	-3	3	0	0	
	PC↑				
x_3	0	-1	1	0	
x_1	1	-1	0	0	PR
	0	0	0	0	RC
	Final Tableau				
x_3	0	-1	1	0	
x_1	1	-1	0	0	

The general solution of this system is $x = (\alpha, \alpha, \alpha)^T$ for any α real; so an eigen vector corresponding to the eigen value 2 is $(1, 1, 1)^T$.

Example 5: Eigen Values of a Diagonal Matrix:

Let

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 9 \end{pmatrix}. \text{ Then } A - \lambda I = \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & -7-\lambda & 0 \\ 0 & 0 & 9-\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda) =$ determinant $(A - \lambda I) = (2-\lambda)(-7-\lambda)(9-\lambda)$. Hence the characteristic equation for A is

$$(2-\lambda)(-7-\lambda)(9-\lambda) = 0$$

for which the roots are 2, -7, and 9. So, the eigen values of A are 2, -7 and 9; the diagonal entries in the diagonal matrix A.

The eigen space corresponding to eigen value 2 is the solution set of

x_1	x_2	x_3	
0	0	0	0
0	-9	0	0
0	0	7	0

The general solution of this system is $x = (\alpha, 0, 0)^T$ for real α. Hence an eigen vector corresponding to the eigen value 2 is $(1, 0, 0)^T$, the first unit vector.

In the same way it can be verified that for any diagonal matrix, its eigen values are its diagonal entries; and an eigen vector corresponding to its ith diagonal entry is the ith unit vector.

Example 6: Eigen Values of an Upper Triangular Matrix: Let

$$A = \begin{pmatrix} -1 & -10 & 16 \\ 0 & 2 & -18 \\ 0 & 0 & -3 \end{pmatrix}. \text{ Then } A - \lambda I = \begin{pmatrix} -1-\lambda & -10 & 16 \\ 0 & 2-\lambda & -18 \\ 0 & 0 & -3-\lambda \end{pmatrix}.$$

Therefore the characteristic polynomial of A, $P(\lambda)$ = determinant $(A - \lambda I) = (-1-\lambda)(2-\lambda)(-3-\lambda)$. Hence the characteristic equation for A is

$$(-1-\lambda)(2-\lambda)(-3-\lambda) = 0$$

Its roots, the eigen values of the upper triangular matrix A, are exactly its diagonal entries.

The eigen space corresponding to eigen value -1 is the set of solutions of

x_1	x_2	x_3	
0	-10	16	0
0	3	-18	0
0	0	-2	0

The general solution of this system is $x = (\alpha, 0, 0)^T$ for α real. Hence an eigen vector corresponding to -1, the first diagonal entry of A, is $(1, 0, 0)^T$, the first unit vector.

In the same way it can be verified that an eigen vector corresponding to the ith diagonal entry of A is the ith unit vector, for all i.

Similarly, for any upper or lower triangular matrix, its eigen values are its diagonal entries, and its eigen vectors are the unit vectors.

Basis for the Eigen Space of a Square Matrix A Corresponding to an Eigen Value λ

Let λ be an eigen value of a square matrix A of order n. Then the eigen space corresponding to λ is $N(A - \lambda I)$, the set of all solutions of the system

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

A basis for this eigen space, as defined in Sections 1.23, 4.2, is a maximal linearly independent set of vectors in $N(A - \lambda I)$, it has the property that every point in this eigen space can be expressed as its linear combination in a unique manner.

Example 7: Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

from Example 4, and its eigen value -1. The eigen space corresponding to this eigen value is the set of all solutions of

x_1	x_2	x_3	
1	1	1	0
1	1	1	0
1	1	1	0

All the constraints in this system are the same as $x_1 + x_2 + x_3 = 0$. So, a general solution of this system is $(\alpha, \beta, -\alpha - \beta)^T$, where α, β are real valued parameters. A basis for this eigen space is the set $\{(1, 0, -1)^T, (0, 1, -1)^T\}$, the set of two vectors obtained by setting $\alpha = 1, \beta = 0$, and $\alpha = 0, \beta = 1$ respectively in the formula for the general solution.

Results on Eigen Values and Eigen Vectors

We now discuss some important results on eigen values and eigen vectors, without proofs.

Result 6.1.3: Linear independence of the set of eigen vectors corresponding to distinct eigen values: *Let* $(\lambda_1, x^1), \dots, (\lambda_k, x^k)$ *be eigen pairs for a square matrix* A*, where the eigen values* $\lambda_1, \dots, \lambda_k$ *are distinct. Then* $\{x^1, \dots, x^k\}$ *is a linearly independent set.*

Result 6.1.4: The union of bases of eigen spaces corresponding to distinct eigen values is a basis: *Let* $\{\lambda_1, \dots, \lambda_k\}$ *be a set of distinct eigen values for a square matrix* A*. For* $i = 1$ *to* k*, let* B_i *be a basis for the eigen space corresponding to* λ_i*. Then* $B = \cup_{i=1}^{k} B_i$ *is a linearly independent set.*

A Complete Set of Eigen Vectors

A **complete set of eigen vectors** for a square matrix A of order n is any linearly independent set of n eigen vectors for A. Not all square matrices have complete sets of eigen vectors. Square matrices that do not have a complete set of eigen vectors are called **deficient** or **defective matrices**.

Let A be a square matrix of order n which has distinct real eigen values $\lambda_1, \ldots, \lambda_n$. Let x^i be an eigen vector corresponding to λ_i for $i = 1$ to n. Then by Result 6.1.3, $\{x^1, \ldots, x^n\}$ is a complete set of eigen vectors for A. Thus any square matrix with all eigen values real and distinct has a complete set of eigen vectors.

Let A be a square matrix of order n with distinct real eigen values $\lambda_1, \ldots, \lambda_k$, where for $i = 1$ to k, both the algebraic and geometric multiplicities of λ_i are equal to r_i, satisfying $r_1 + \ldots + r_k = n$. Then the eigen space corresponding to λ_i has a basis B_i consisting of r_i vectors, and $B = \cup_{i=1}^{k} B_i$ is linearly independent by Result 6.1.4, and is a basis for R^n. In this case B is a complete set of eigen vectors for A. Thus if all eigen values of A are real, and for each of them its geometric multiplicity is its algebraic multiplicity, then A has a complete set of eigen values.

Thus if a square matrix A has an eigen value for which the geometric multiplicity is $<$ its algebraic multiplicity, then A is deficient in eigen vectors; i.e., it does not have a complete set of eigen vectors.

6.2 Relationship of Eigen Values and Eigen Vectors to the Diagonalizability of a Matrix

Let A be a square matrix of order n. In Section 5.4 we mentioned that the matrix diagonalization problem for matrix A is to find a square nonsingular matrix P of order n such that $P^{-1}AP$ is a diagonal matrix $D = \text{diag}(d_1, \ldots, d_n)$, say. In this case we say that the **matrix** P **diagonalizes** A, i.e.,

$$P^{-1}AP = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix} = D = \text{diag}(d_1, \ldots, d_n)$$

Then multiplying the above equation on both sides by P, we have

$$AP = PD.$$

$P_{.1}, \ldots, P_{.n}$ are the column vectors of P. Since $D = \text{diag}(d_1, \ldots, d_n)$, we have $PD = (d_1P_{.1}\vdots \ldots \vdots d_nP_{.n})$, and by the definition of matrix multiplication $AP = (AP_{.1}\vdots \ldots \vdots AP_{.n})$. So, the above equation $AP = PD$ implies

$$\begin{aligned} (AP_{.1}\vdots \ldots \vdots AP_{.n}) &= (d_1P_{.1}\vdots \ldots \vdots d_nP_{.n}) \\ \text{i.e.,} \quad AP_{.j} &= d_jP_{.j} \qquad j = 1, \ldots, n. \end{aligned}$$

Since P is invertible, $P_{.j} \neq 0$ for all j. So, the above equation $AP_{.j} = d_jP_{.j}$, i.e., $(A - d_jI)P_{.j} = 0$ implies that $(d_j, P_{.j})$ are eigen pairs for A for $j = 1$ to n.

Thus if the matrix P diagonalizes the square matrix A into diag$(d_1, \ldots, d_n)$ then $d_1, \ldots, d_n$ are the eigen values of A; and $P_{.j}$ is an eigen vector corresponding to d_j for each $j = 1$ to n.

Conversely, suppose a square matrix A of order n has a complete set of eigen vectors $\{P_{.1}, \ldots, P_{.n}\}$ with corresponding eigen values $\lambda_1, \ldots, \lambda_n$. Let $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Since $(\lambda_j, P_{.j})$ is an eigen pair for A for $j =$ 1 to n; we have $AP_{.j} = \lambda_jP_{.j}$ for $j = 1$ to n, i.e.

$$(AP_{.1}\vdots \ldots \vdots AP_{.n}) = (\lambda_1P_{.1}\vdots \ldots \vdots AP_{.n}) = DP$$

i.e., $AP = DP$ where P is the square matrix with the eigen vectors $P_{.1}, \ldots, P_{.n}$ as the column vectors. Since $\{P_{.1}, \ldots, P_{.n}\}$ is a complete set of eigen vectors, it is linearly independent, so P is invertible. So, $AP = PD$ implies $P^{-1}AP = D$. Thus P diagonalizes A.

Thus the square matrix A of order n is diagonalizable iff it has a complete set of eigen vectors; and a matrix which diagonalizes A is the matrix whose column vectors are a basis for R^n consisting of eigen vectors of A.

Thus the task of diagonalizing a square matrix A is equivalent to finding its eigen values and a complete set of eigen vectors for it.

We now provide some examples. In this chapter we number all the examples serially. So, continuing the numbering from the previous section, we number the next example as Example 8.

Example 8: Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

discussed in Examples 4, 7. From Examples 4, 7, we know that its eigen values are-1 and 2, and that $\{(1,0,-1)^T, (0,1,-1)^T, (1,1,1)^T\}$ is a complete set of eigen vectors for A. So

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

is a matrix that diagonalizes A. In fact $P^{-1}AP = \text{diag}(-1,-1,2)$.

Example 9: Let

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 10 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

In Example 3 we determined that its eigen values are 0 and 1, with 1 having algebraic multiplicity 2. But the geometric multiplicity of 1 is $1 <$ its algebraic multiplicity. So, this matrix does not have a complete set of eigen vectors, and hence it is not diagonalizable.

Example 10: Let

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

the matrix considered in Example 1. From Example 1 we know that its eigen values are 2, 1 and a complete set of eigen vectors for it is $\{(1,4/5)^T, (1,1)^T\}$. So, the matrix

$$P = \begin{pmatrix} 1 & 1 \\ 4/5 & 1 \end{pmatrix}$$

diagonalizes A, and $P^{-1}AP = \text{diag}(2, 1)$.

We will now present some more important results, most of them without proofs.

Result 6.2.1: Results on Eigen Values and Eigen Vectors for Symmetric Matrices:

(i): All eigen values are real: *If A is a square symmetric matrix of order n, all its eigen values are real numbers, i.e., it has no complex eigen values.*

(ii): Eigen vectors corresponding to distinct eigen values are orthogonal: *If A is a square symmetric matrix of order n, eigen vectors corresponding to distinct eigen values of A are orthogonal.*

We will provide a proof of this result because it is so simple. Since A is symmetric, $A = A^T$. Let $\lambda_1 \neq \lambda_2$ be eigen values associated with eigen vectors v^1, v^2 respectively. So, $Av^1 = \lambda_1 v^1$ and $Av^2 = \lambda_2 v^2$. So

$$\lambda_1 (v^1)^T v^2 = (\lambda_1 v^1)^T v^2 = (Av^1)^T v^2 = (v^1)^T A^T v^2 = (v^1)^T A v^2 = \lambda_2 (v^1)^T v^2$$

Hence, $\lambda_1 (v^1)^T v^2 = \lambda_2 (v^1)^T v^2$, i.e., $(\lambda_1 - \lambda_2)(v^1)^T v^2 = 0$. Since $\lambda_1 \neq \lambda_2$, $\lambda_1 - \lambda_2 \neq 0$. Hence $(v^1)^T v^2$ must be zero; i.e., v^1 and v^2 are orthogonal.

(iii): Special eigen properties of symmetric PD, PSD, ND, NSD, Indefinite matrices: *A square symmetric matrix is*

Positive definite	*iff all its eigen values are* > 0
Positive semidefinite	*iff all its eigen values are* ≥ 0
Negative definite	*iff all its eigen values are* < 0
Negative semidefinite	*iff all its eigen values are* ≤ 0
Indefinite	*iff it has both a positive and a negative eigen value.*

See Section 5.2 for definitions of these various classes of matrices.

(iv): Geometric and algebraic multiplicities are equal: *If A is a square symmetric matrix of order n, for every one of its eigen values, its geometric multiplicity is equal to its algebraic multiplicity. Hence symmetric matrices can always be diagonalized.*

(v): Orthonormal set of eigen vectors: *If A is a square matrix of order n, it has an orthonormal set of n eigen vectors iff it is symmetric.*

Orthogonal Diagonalization of a Square Matrix

As discussed in Section 5.4, **orthogonal diagonalization** of a square matrix A deals with the problem of finding an orthogonal matrix P such that

$$P^{-1}AP = \quad \text{a diagonal matrix } \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \quad \text{say.}$$

When P is orthogonal $P^{-1} = P^T$. The orthogonal matrix P is said to **orthogonally diagonalize** the square matrix A if $P^{-1}AP = P^TAP$ is a diagonal matrix.

Result 6.2.2: Orthogonally diagonalizable iff symmetric: *A square matrix can be orthogonally diagonolized iff it is symmetric.*

To see this, suppose the orthogonal matrix P orthogonally diagonalizes the square matrix A. So, $P^{-1} = P^T$ and $P^{-1}AP =$ a diagonal matrix D. Since D is diagonal, $D^T = D$. Now, $P^{-1}AP = D$ implies $A = PDP^{-1} = PDP^T$. So,

$$A^T = (PDP^T)^T = PD^TP^T = PDP^T = A$$

i.e., A is symmetric.

To orthogonally diagonalize a square symmetric matrix A of order n, one needs to find its eigen values. If they are all distinct, let $(\lambda_i, P_{.i})$ be the various eigen pairs where each $P_{.i}$ is normalized to satisfy $||P_{.i}|| = 1$. Then by Result 6.2.1 (ii), $\{P_{.1}, \ldots, P_{.n}\}$ is an orthonormal set of eigen vectors for A. Let P be the matrix with $P_{.1}, \ldots, P_{.n}$ as its column vectors, it orthogonally diagonalizes A.

If the eigen values of the symmetric matrix A are not all distinct, find a basis for the eigenspace of each eigen value. Apply the Gram-Schmidt process to each of these bases to obtain an orthonormal basis for each eigen space. Then form the square matrix P whose columns are the vectors in these various orthonormal bases. Then P orthogonally diagonalizes A.

Result 6.2.3: Similar Matrices Have the Same Eigen Spectrum: *Two square matrices A, B of order n are said to be* **similar** *if there exists a nonsingular matrix P of order n satisfying $B = P^{-1}AP$. In this case we say that B is obtained by a* **similarity transformation** *on A. Since $B = P^{-1}AP$ implies $A = PBP^{-1}$, if B can be obtained by a similarity transformation on A, then A can be obtained by a similarity transformation on B. When $B = P^{-1}AP$, $det(B - \lambda I) = det(P^{-1}AP - \lambda I) = det(P^{-1}(AP - \lambda P)) = det(P^{-1})det(A - \lambda I)det(P) = det(A - \lambda I)$. Thus similar matrices all have the same charateristic polynomial, so they all have the same eigen values with the same multiplicities. However, their eigen vectors may be different.*

Result 6.2.4: Eigen values of powers of a matrix: *Suppose the square matrix A is diagonalizable, and let $P^{-1}AP = diag(\lambda_1, \ldots, \lambda_n)$. Then for any positive integer s,*

$$A^s = P^{-1}(\text{diag}(\lambda_1^s, \ldots, \lambda_n^s))P$$

Since $\lambda_1, \ldots, \lambda_n$ are real numbers, computing $\lambda_1^s, \ldots, \lambda_n^s$ is much easier than computing A^s because of the complexity of matrix multiplication.

So, calculating higher powers A^s of matrix A for large values of s can be carried out efficiently if it can be diagonalized.

Result 6.2.5: Spectral Theorem for Diagonalizable Matrices: *Let A be a square matrix of order n with spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$. Suppose A is diagonalizable, and for $i = 1$ to k, let $r_i =$ algebraic multiplicity of $\lambda_i =$ geometric multiplicity of λ_i. Then there exist square matrices of order n, $G_1, \ldots, G_k$ satisfying*

i) $G_i G_j = 0$ *for all* $i \neq j$
ii) $G_1 + \ldots + G_k = I$
iii) $\lambda_1 G_1 + \ldots + \lambda_k G_k = A$
iv) $G_i^2 = G_i$ *for all* i
v) *Rank*$(G_i) = r_i$ *for all* i.

Let X_i be the matrix of order $n \times r_i$ whose column vectors form a basis for the eigen space $N(A - \lambda_i I)$. Let $P = (X_1 \vdots X_2 \vdots \ldots \vdots X_k)$. Then P is nonsingular by Result 6.1.2. Partition P^{-1} as follows

$$P^{-1} = \begin{pmatrix} Y_1 \\ \ldots \\ Y_2 \\ \ldots \\ \vdots \\ \ldots \\ Y_k \end{pmatrix}$$

where Y_i is an $r_i \times n$ matrix. Then $G_i = X_i Y_i$ for $i = 1$ to k satisfies all the above properties. All these properties follow from the following facts

$$PP^{-1} = P^{-1}P = I$$

$$A = PDP^{-1} = (X_1 \vdots X_2 \vdots \ldots \vdots X_k) \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 I_{r_2} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_k I_{r_k} \end{pmatrix} \begin{pmatrix} Y_1 \\ \ldots \\ Y_2 \\ \ldots \\ \vdots \\ \ldots \\ Y_k \end{pmatrix}$$

where I_{r_i} is the unit matrix of order r_i for $i = 1$ to k.

The decomposition of A into $\lambda_1 G_1 + \ldots + \lambda_k G_k$ is known as the **spectral decomposition** of A, and the G_is are called the **spectral projectors** associated with A.

Example 11: Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

discussed in Example 8. Its eigen values are $\lambda_1 = -1$ with an algebraic and geometric multiplicity of two, and $\lambda_2 = 2$ with an algebraic and geometric multiplicity of one. A basis for the eigen space correponding to -1 is $\{(1, 0, -1)^T, (0, 1, -1)^T\}$; and an eigen vector corresponding to 2 is $(1, 1, 1)^T$. So, using the notation in Result 6.2.5, we have

$$X_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}, X_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 0 & 1 & \vdots & 1 \\ -1 & -1 & \vdots & 1 \end{pmatrix}.$$

We compute P^{-1} and obtain

$$P^{-1} = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ \dots & \dots & \dots \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \quad \begin{matrix} Y_1 = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \end{pmatrix}, \\ \\ Y_2 = \begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix}. \end{matrix}$$

So

$$G^1 = X_1 Y_1 = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}$$

$$G_2 = X_2 Y_2 = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

It can be verified that $\text{rank}(G_1) = 2 =$ algebraic multiplicity of the eigen value -1, and $\text{rank}(G_2) = 1 =$ algebraic multiplicity of the eigen value 2. Also, we verify that

$G_1 + G_2 = I =$ unit matrix of order 3

$2G_1 + G_2 = A$, the spectral decomposition of A.

G_1, G_2 are the spectral projectors associated with A.

Result 6.2.6: Eigen pairs for A and A^{-1} *If (λ, x) is an eigen pair for a nonsingular square matrix A, then $(1/\lambda, x)$ is an eigen pair for A^{-1}.*

By hypothesis we have $Ax = \lambda x$, and since A is nonsingular, we have $\lambda \neq 0$. Multiplying both sides of the equation $Ax = \lambda x$ on the left by $(1/\lambda)A^{-1}$ we get $(A^{-1} - (1/\lambda)I)x = 0$. Therefore $((1/\lambda), x)$ is an eigen pair for A^{-1}.

Result 6.2.7: Sylvester's Law of Inertia: *Let A be a square symmetric matrix of order n. Let C be any nonsingular square matrix of order n, and $B = C^T AC$. Then we say that B is obtained from A by a* **congruent transformation**, *and that A, B are* **congruent**. *Since C is nonsingular, $B = C^T AC$ implies $A = (C^T)^{-1}BC^{-1} = (C^{-1})^T B(C^{-1})$, so if B can be obtained by a congruent transformation on A, then A can be obtained by a congruent transformation on B.*

The **inertia** *of a real symmetric matrix is defined to be the triple (ρ, ν, ζ) where*

> *ρ, ν, ζ are the numbers of positive, negative, and zero eigen values of the matrix respectively.*

Sylvester's law states that two symmetric matrices are congruent iff they have the same inertia.

6.3 An Illustrative Application of Eigen Values, Eigen Vectors, and Diagonalization in Differential Equations

Eigen values, eigen vectors, and diagonalization of square matrices find many applications in all branches of science. Here we provide one illustrative application in differential equations, which are equations involving functions and their derivatives. Mathematical models involving differential equations appear very often in studies in chemical engineering and other branches of engineering, physics, chemistry, biology, pharmacy, and other sciences.

Let $u_1(t), \ldots, u_n(t)$ be n real valued differentiable functions of a single real variable t; and let $u_i'(t)$ denote the derivative of $u_i(t)$. Suppose the formulas for these functions are unknown, but we know that these functions satisfy the following equations.

$$
\begin{aligned}
u_1'(t) &= a_{11}u_1(t) + \ldots + a_{1n}u_n(t) \\
\vdots \quad &\quad \vdots \\
u_n'(t) &= a_{n1}u_1(t) + \ldots + a_{nn}u_n(t) \\
\text{and} \quad u_i(0) &= c_i \quad i = 1, \ldots, n
\end{aligned}
$$

where $A = (a_{ij})$ is a given square matrix of order n, and $c = (c_1, \ldots, c_n)^T$ is a given vector of function values corresponding to $t = 0$.

A system like this is known as a **homogeneous system of first order linear differential equations with constant coefficients**. Solving this system means to find the formulas for the functions $u_1(t)$, $\ldots$, $u_n(t)$ satisfying these equations. Let

$$
A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}, u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix}, u'(t) = \begin{pmatrix} u_1'(t) \\ \vdots \\ u_n'(t) \end{pmatrix}.
$$

Then the system can be written in matrix form as

$$
u'(t) = Au(t), \quad u(0) = c.
$$

For simplicity, we only consider the case where the matrix A is diagonalizable. Let the spectrum of A be $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$, and let the spectral projector associated with λ_i be G_i for $i = 1$ to k. Then it can be shown that the unique solution of this system is

$$
u(t) = e^{\lambda_1 t}v^1 + \ldots + e^{\lambda_k t}v^k
$$

where $v^i = G_i c$ for $i = 1$ to k.

Example 12: Consider the following linear differential equation involving three functions $u_1(t), u_2(t), u_3(t)$ and their derivatives $u_1'(t), u_2'(t), u_3'(t)$.

$$\begin{array}{rcccc} u_1'(t) = & & u_2(t) & +u_3(t) \\ u_2'(t) = & u_1(t) & & +u_3(t) \\ u_3'(t) = & u_1(t) & +u_2(t) & \\ (u_1(0), u_2(0), u_3(0))^T = & & (12, -15, 18)^T & \end{array}$$

Letting $u(t) = (u_1(t), u_2(t), u_3(t))^T$, $u'(t) = (u_1'(t), u_2'(t), u_3'(t))^T$, $c = (12, -15, 18)^T$, the system in matrix notation is

$$\begin{aligned} u'(t) &= Au(t) \\ u(0) &= c \end{aligned}$$

where

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

the matrix discussed in Example 8, 11. Its eigen values are $\lambda_1 = -1$ and $\lambda_2 = 2$, and the associated spectral projectors are

$$G^1 = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}, G_2 = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

from Example 11. So, in the notation given above, we have

$$v^1 = G^1 c = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} \begin{pmatrix} 12 \\ -15 \\ 18 \end{pmatrix} = \begin{pmatrix} 7 \\ -20 \\ 13 \end{pmatrix}$$

$$v^2 = G_2 c = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 12 \\ -15 \\ 18 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix}.$$

So, from the above, we know that the unique solution of this system is

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} = e^{-t} \begin{pmatrix} 7 \\ -20 \\ 13 \end{pmatrix} + e^{2t} \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 7e^{-t} + 5e^{2t} \\ -20e^{-t} + 5e^{2t} \\ 13e^{-t} + 5e^{2t} \end{pmatrix}.$$

i.e., $u_1(t) = 7e^{-t} + 5e^{2t}$, $u_2(t) = -20e^{-t} + 5e^{2t}$, $u_3(t) = 13e^{-t} + 5e^{2t}$. It can be verified that these functions do satisfy the given differential equations.

6.4 How to Compute Eigen Values ?

Computing the eigen values of a matrix A of order n using the main definition boils down to finding the roots of a polynomial equation of degree n. Finding the roots of a polynomial equation of degree ≥ 5 is not a simple task, it has to be carried out using iterative methods.The most practical eigen value computation method is the QR iteration algorithm, an iterative method that computes the QR factors (see Section 4.12) of a matrix in every iteration. Providing a mathematical description of this algorithm, and the details of the work needed to convert it into a practical implementation, are beyond the scope of this book. For details see the references given below.

Exercises

6.4.1: Obtain the characteristic polynomial of the following matrices. Using it obtain their eigen values, and then the corresponding eigen vectors and eigen spaces.

$$\begin{pmatrix} 3 & 2 \\ 8 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1/3 \\ 31/3 & 5 \end{pmatrix}, \begin{pmatrix} 5 & 1 & 2 \\ 6 & 5 & 3 \\ 6 & 2 & 6 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 2 \\ 2 & 4 & 4 \\ 3 & 3 & 8 \end{pmatrix}.$$

6.4.2: Compute eigen values and eigen vectors of the following matrices.

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{pmatrix}.$$

6.4.3: Let A be a square matrix of order n in which all diagonal entries are $= \alpha$, and all off-diagonal entries are $= \beta$. Show that it has an eigen value $\alpha - \beta$ with algebraic multiplicity $n-1$, and $\alpha + (n-1)\beta$ as the only other eigen value.

6.4.4: If λ is an eigen value of A associated with the eigen vector x, show that λ^k is an eigen value of A^k associated with the same eigen vector x for all positive integers k.

6.4.5: If A, B are square matrices of order n, show that the eigen values of AB, BA are the same.

6.5 References

[6.1] C. P. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, PA, 2000.
[6.2] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.

Index

For each index entry we provide the section number where it is defined or discussed first. Clicking on the light blue section number takes you to the page where this entry is defined or discussed.

List of Results

1. Result 6.1.1: Multiplicity inequality

2. Result 6.1.2: Linear independence of eigen values associated with distinct eigen values

3. Result 6.1.3: Linear independence of the set of eigen vectors corresponding to distinct eigen values

4. Result 6.1.4: The union of bases of eigen spaces corresponding to distinct eigen values is a basis

5. Result 6.2.1: Results on Eigen Values and Eigen Vectors for Symmetric Matrices

   (i): All eigen values are real

   (ii): Eigen vectors corresponding to distinct eigen values are orthogonal

   (iii): Special eigen properties of symmetric PD, PSD, ND, NSD, Indefinite matrices

   (iv): Geometric and algebraic multiplicities are equal

   (v): Orthonormal set of eigen vectors

6. Result 6.2.2: Orthogonally diagonalizable iff symmetric

7. Result 6.2.3: Similar Matrices Have the Same Eigen Spectrum

8. Result 6.2.4: Eigen values of powers of a matrix

9. Result 6.2.5: Spectral Theorem for Diagonalizable Matrices
10. Result 6.2.6: Eigen pairs for A and A^{-1}
11. Result 6.2.7: Sylvester's Law of Inertia.