

CISC 271 Class 5

Diagonalizable Matrices

Text Correspondence: §6.2

Main Concepts:

- *Diagonalizable: a matrix is similar to a diagonal matrix*
- *Similar matrices have the same eigenvalues*
- *Orthogonal: a matrix Q such that $QQ^T = I$*
- *Symmetric: a matrix B with $b_{ij} = b_{ji}$*
- *Symmetric matrices have only real eigenvalues*
- *Produces an especially useful decomposition*

Sample Problem, Data Analysis: When are unit eigenvectors a basis?

For many matrices that we will encounter, the eigenvectors \vec{v}_j form a special basis. Let us try to understand why.

First, a useful convention in mathematics is that an eigenvector is often assumed to be of unit magnitude. We can see immediately that any non-zero vector can be forced to be of unit magnitude, just by dividing each entry by the norm of the vector. From now on, we will assume that an eigenvector \vec{v} has the property

$$\|\vec{v}\| = 1$$

5.1 Similar Matrices

In linear algebra, when we say that two matrices are *similar*, there is a specific meaning.

Definition: similar matrices

For any matrix $A \in \mathbb{R}^{n \times n}$, and for matrix $C \in \mathbb{R}^{n \times n}$, A is *similar* to C , or $A \sim C$, is defined as:

There exists an invertible matrix $P \in \mathbb{R}^{n \times n}$ such that

$$A = P^{-1}CP \tag{5.1}$$

Because P in Definition 5.1 is invertible, we can also write

$$C = PAP^{-1}$$

5.2 Eigenvectors as a Basis

The idea of similarity is especially useful when a matrix is similar to a particular diagonal matrix. Consider any matrix $A \in \mathbb{R}^{n \times n}$ that has n linearly independent eigenvectors. For each eigenvector \vec{v}_j , we know that

$$A\vec{v}_j = \lambda_j\vec{v}_j \quad (5.2)$$

We can assemble these eigenvectors into a matrix V , which is

$$V = [\vec{v}_1 \quad \vec{v}_2 \quad \cdots \quad \vec{v}_n] \quad (5.3)$$

Applying Equation 5.2 to the columns of V , we get

$$AV = [\lambda_1\vec{v}_1 \quad \lambda_2\vec{v}_2 \quad \cdots \quad \lambda_n\vec{v}_n] \quad (5.4)$$

Consider ways that we can re-write the right-hand side of Equation 5.4. One way is to factor it, which means expressing it as the product of two simpler matrices. We can decompose the matrix as $V\Lambda$, where V is the matrix of eigenvectors and Λ is a diagonal matrix of the eigenvalues, so that

$$V\Lambda = [\vec{v}_1 \quad \vec{v}_2 \quad \cdots \quad \vec{v}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad (5.5)$$

This means that AV and $V\Lambda$ are the same. Because the eigenvectors are linearly independent, the matrix V is full rank and invertible, so V^{-1} exists. Together, these results imply that

$$\begin{aligned} AV &= V\Lambda \\ \Rightarrow AVV^{-1} &= V\Lambda V^{-1} \\ \Rightarrow A &= V\Lambda V^{-1} \end{aligned} \quad (5.6)$$

Using a closely related line of reasoning, pre-multiplying by V^{-1} , implies that

$$\Lambda = V^{-1}AV \quad (5.7)$$

A matrix A that can be converted into a diagonal matrix is called *diagonalizable*. A necessary and sufficient condition for a matrix to be diagonalizable is that its eigenvectors are a basis, which means that there are n eigenvectors that are linearly independent.

A sufficient, but not necessary, condition for a matrix to be diagonalizable is that all of the eigenvalues are distinct. This condition is not necessary because some matrices have repeated eigenvalues and also have an eigenvector basis. A 3×3 example is

$$A = \begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix} \quad \text{with} \quad \lambda_1 = 1 \quad \lambda_2 = 2 \quad \lambda_3 = 2 \quad \text{and} \quad V \sim \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

where we have multiplied the columns of V by real numbers that give us a “human-readable” set of basis vectors.

A general rule is that *distinct eigenvalues imply that the eigenvectors are a basis*.

5.3 Eigenvector Basis

If a matrix $A \in \mathbb{R}^{n \times n}$ is diagonalizable, then its eigenvectors are a basis. Using the convention that eigenvectors are of unit magnitude, i.e., we require that $\|\vec{v}_j\| = 1$, we can represent any given vector \vec{u} as

$$\begin{aligned} \vec{u} &= \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n \\ &= \sum_{j=1}^n \alpha_j \vec{v}_j \end{aligned}$$

When we perform the multiplication $\vec{y} = A\vec{u}$ we get

$$\begin{aligned} A\vec{u} &= \alpha_1 A\vec{v}_1 + \alpha_2 A\vec{v}_2 + \cdots + \alpha_n A\vec{v}_n \\ &= \sum_{j=1}^n \alpha_j \lambda_j \vec{v}_j \\ &= \sum_{j=1}^n \lambda_j (\alpha_j \vec{v}_j) \end{aligned}$$

so each of the original terms $\alpha_j \vec{v}_j$ is multiplied by the eigenvalue λ_j . This is depicted in Figure 5.1 for $n = 2$ (2D, or the plane).

5.4 Orthogonal Matrices

The “best” kind of basis is an orthonormal basis, in which each vector is of unit magnitude and is orthogonal to the other vectors. It is common to refer to such vectors as columns of a matrix Q ,

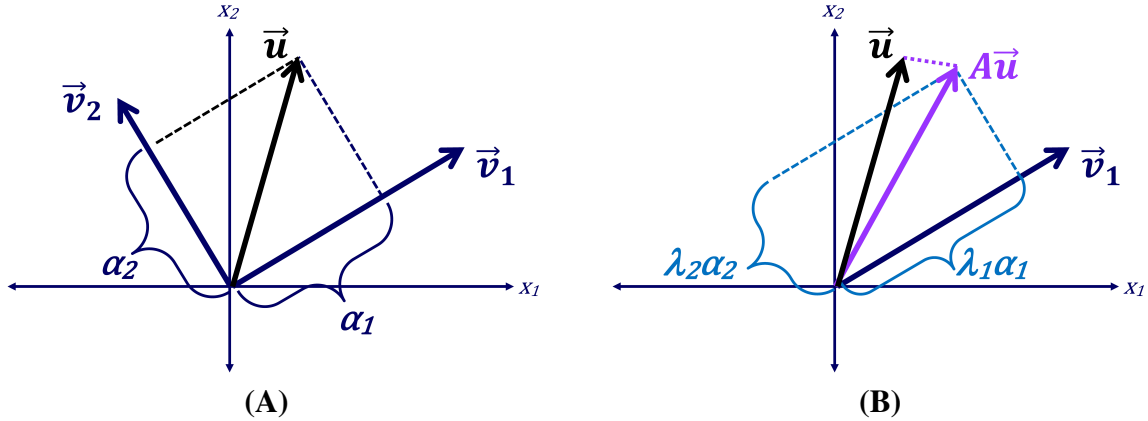


Figure 5.1: Distinct eigenvectors form a basis of a vector space, so any given vector in that space can be expressed as a weighted sum of the eigenvectors. The linear transformation of the matrix scales the eigenvectors and thus changes the given vector in a predictable manner. (A) The original vector \vec{u} is a weighted sum of two eigenvectors, so $\vec{u} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2$. (B) The eigenvectors are scaled by their corresponding eigenvalues so the original vector \vec{u} becomes $\alpha_1\lambda_1\vec{v}_1 + \alpha_2\lambda_2\vec{v}_2$. In general, both the direction and magnitude of the original vector are changed by the linear transformation of the matrix A .

so each such vector is written as \vec{q}_j . The defining properties of an orthonormal basis are

$$\begin{aligned} i \neq j &\Rightarrow \vec{q}_i \cdot \vec{q}_j = 0 \\ i = j &\Rightarrow \vec{q}_i \cdot \vec{q}_j = 1 \end{aligned}$$

By using the transpose to formulate the dot product, an orthonormal basis can also be defined as

$$\begin{aligned} i \neq j &\Rightarrow \vec{q}_i^T \vec{q}_j = 0 \\ i = j &\Rightarrow \vec{q}_i^T \vec{q}_j = 1 \end{aligned}$$

From this it follows that

$$Q^T Q = I \quad \text{and} \quad Q Q^T = I \tag{5.8}$$

An orthogonal matrix, defined in Equation 5.8, commutes with itself. This is a slightly unusual property; for example, the matrix

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

does *not* commute because

$$AA^T = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad A^T A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$$

What other matrices commute with their transposes?

5.5 Symmetric Matrices

One form of matrix that is frequently encountered in applications is a *symmetric* matrix. This is a matrix B for which

$$b_{ij} = b_{ji} \tag{5.9}$$

It is easy to see that, if B is symmetric, then

$$B^T B = B B^T = B^2$$

so the matrix commutes with its transpose.

What can we say about the eigenvalues and eigenvectors of a symmetric matrix B ? As it happens, we can say a great deal. The first major result is

For a real symmetric matrix, every eigenvalue is a real number.

This can be proved by using complex conjugates, but complex analysis is beyond the scope of this course so the interested student should read the textbook and other sources for greater insight.

The second major result is

For a real symmetric matrix, eigenvectors that have distinct eigenvalues are orthogonal.

To show this, we will consider two eigenvalue/eigenvector pairs and use the transpose as a way of writing the dot product. We will be investigating whether

$$\begin{aligned} \vec{v}_1 \cdot \vec{v}_2 &\stackrel{?}{=} 0 \\ \text{or equivalently } \vec{v}_1^T \vec{v}_2 &\stackrel{?}{=} 0 \end{aligned}$$

We begin with the statement that the eigenvectors are distinct. For any pair, labeling one as λ_1 and the other as λ_2 , we have

$$\lambda_1 \neq \lambda_2$$

Consider the dot product of $B\vec{v}_1$ with \vec{v}_2 . Expanding the matrix-vector product, and factoring the scalar out of the dot product, gives

$$\begin{aligned}(B\vec{v}_1) \cdot \vec{v}_2 &= (\lambda_1\vec{v}_1) \cdot \vec{v}_2 \\ &= \lambda_1(\vec{v}_1 \cdot \vec{v}_2)\end{aligned}\tag{5.10}$$

Alternatively, we can express the dot product as a transposed multiplication; doing this, then replacing $B^T = B$ because B is symmetric, then converting back to the dot product, gives

$$\begin{aligned}(B\vec{v}_1) \cdot \vec{v}_2 &= (B\vec{v}_1)^T \vec{v}_2 \\ &= (\vec{v}_1^T B^T) \vec{v}_2 \\ &= \vec{v}_1^T (B\vec{v}_2) \\ &= \vec{v}_1^T (\lambda_2\vec{v}_2) \\ &= \vec{v}_1 \cdot (\lambda_2\vec{v}_2) \\ &= \lambda_2(\vec{v}_1 \cdot \vec{v}_2)\end{aligned}\tag{5.11}$$

Because Equation 5.10 and Equation 5.11 have the same left-hand side, the right-hand terms are equal so

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)\tag{5.12}$$

We began by assuming that $\lambda_1 \neq \lambda_2$ so, for Equation 5.12 to hold, we must have

$$\vec{v}_1 \cdot \vec{v}_2 = 0\tag{5.13}$$

This proves that, for a symmetric matrix, the eigenvectors of distinct eigenvalues are orthogonal.

5.6 Skew-Symmetric Matrices

A square matrix S is *skew symmetric*, or anti-symmetric, is defined as one for which

$$s_{ij} = -s_{ji}\tag{5.14}$$

Another way to write this matrix is

$$S^T = -S$$

We can confirm that a skew-symmetric matrix S is a normal matrix by expansion:

$$SS^T = S[-S] = [-S]S = S^T S\tag{5.15}$$

From either definition, most clearly from Equation 5.14, the diagonal entries are zero so the matrix S has the form

$$S = \begin{bmatrix} 0 & s_{12} & s_{13} & \cdots & & s_{1n} \\ -s_{12} & 0 & s_{23} & \cdots & & s_{2n} \\ -s_{13} & -s_{23} & 0 & \cdots & & s_{3n} \\ \vdots & \vdots & \vdots & \ddots & & \\ & & & & 0 & s_{(n-1),n} \\ -s_{1n} & -s_{2n} & -s_{3n} & \cdots & -s_{(n-1),n} & 0 \end{bmatrix}$$

5.7 Nondiagonalizable Matrices

In linear algebra, a square matrix that is not diagonalizable is historically called a *defective matrix*. If we allow eigenvalues and eigenvectors to be complex – even when the entries of the original matrix are real – then, in a special technical sense that involves the Lebesgue measure, “almost no” matrix is defective.

A necessary condition for a matrix to be nondiagonalizable is that it must have repeated eigenvalues. For example, suppose that a matrix A is $n \times n$ and has k distinct eigenvalues; if $k < n$, then at least one eigenvalue is repeated and we say that such an eigenvalue has an *algebraic multiplicity* greater than one. This is another way of saying that there is a “multiple root” of the characteristic equation $F(\lambda) = 0$ for the matrix A .

We must be careful here, because nontrivial algebraic multiplicity is not a *sufficient* condition for a matrix to be nondiagonalizable. We must also examine the eigenvectors of the repeated eigenvalue to determine whether or not the matrix has *geometric multiplicity*, which is the dimension of the nullspace of $[A - \lambda I]$.

A specific example of a nondiagonalizable matrix is the 2×2 matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (5.16)$$

The matrix A in Equation 5.16 has a single eigenvalue $\lambda = 1$ and a single eigenvector $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

A general example of a nondiagonalizable matrix, of which Equation 5.16 is an instance, is any

matrix with a nontrivial Jordan block of size 2×2 . This is a *bidiagonal* matrix of the form

$$A = \begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{bmatrix} \quad (5.17)$$

The matrix A in Equation 5.17 has a single eigenvalue $\lambda = a$ and a single eigenvector $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

5.8 Extra Notes: Matrix Powers

Diagonalizable matrices, especially ones with distinct eigenvalues, are common in practice and have easily described matrix powers. We can raise a square matrix A to an integer power A^k using a simple recursive rule based on the identity matrix I :

$$\begin{aligned} A^0 &= I \\ A^{k+1} &= A^k I \end{aligned}$$

Because A is diagonalizable, we can see that

$$\begin{aligned} A^2 &= AA = E\Lambda E^{-1}E\Lambda E^{-1} \\ &= E\Lambda^2 E^{-1} \\ A^3 &= AA^2 = E\Lambda E^{-1}E\Lambda^2 E^{-1} \\ &= E\Lambda^3 E^{-1} \end{aligned}$$

and so on for A^k .

Consider: a diagonalizable matrix A for which each eigenvalue is non-negative. For such a matrix, we can write the eigenvalue matrix Λ in terms of a diagonal matrix D so that

$$\Lambda = DD$$

Each entry of the diagonal matrix D is, by definition, $d_{jj} = \sqrt{\lambda_j}$ and is a real number. so it is easy to compute a new matrix C that is

$$C = EDE^{-1} \tag{5.18}$$

The matrix C in Equation 5.18 is constructed so that

$$A = CC = C^2$$

This implies that, for a diagonalizable matrix A that has non-negative eigenvalues, there is a *square root matrix* C :

$$C = A^{1/2} \tag{5.19}$$

The method of solving Equation 5.19 by using the decomposition of Equation 5.18 is, numerically, not the way that MATLAB currently performs the computation of

$$C = \text{sqrtm}(A)$$

In practice there are multiple solutions for a matrix C such that $A = CC$ and MATLAB finds a solution that is numerically reasonable.

5.9 Extra Notes: Small Perturbations

The eigenvalue/eigenvector decomposition of a matrix can be used to analyze the numerical stability of the matrix. We can consider three examples, with extensive analysis being beyond the scope of this course.

Example: stable 2×2 matrix; consider

$$A = \begin{bmatrix} +101 & -90 \\ +110 & -98 \end{bmatrix}$$

The eigenvalues of this matrix are $\{+1, +2\}$ and the respective eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} 0.6727 \\ 0.7399 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 0.6690 \\ 0.7433 \end{bmatrix}$$

Any given vector \vec{u} will be transformed to $\vec{y} = A\vec{u}$ as

$$\begin{aligned} \vec{y} &= \lambda_1 \alpha_1 \vec{v}_1 + \lambda_2 \alpha_2 \vec{v}_2 \\ &= \alpha_1 \vec{v}_1 + 2\alpha_2 \vec{v}_2 \end{aligned}$$

so the transformed \vec{y} is relatively insensitive to small changes in \vec{u} .

Example: unstable 2×2 matrix; consider

$$A = \begin{bmatrix} +100.999 & +90.001 \\ +110 & -98 \end{bmatrix}$$

For this matrix, the eigenvalues are approximately $\{-139.2134, +142.2134\}$ and the respective unit eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} -0.3508 \\ +0.9364 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} +0.9092 \\ +0.4163 \end{bmatrix}$$

Any given vector \vec{u} will be transformed to $\vec{y} = A\vec{u}$ as, approximately,

$$\begin{aligned} \vec{y} &= \lambda_1 \alpha_1 \vec{v}_1 + \lambda_2 \alpha_2 \vec{v}_2 \\ &\approx -139\alpha_1 \vec{v}_1 + 142\alpha_2 \vec{v}_2 \end{aligned}$$

so the transformed \vec{w} is *highly sensitive* to small changes in \vec{u} .

Example: nearly singular 2×2 matrix; consider

$$A = \begin{bmatrix} +101.00 & +89.99 \\ +109.99 & +98.00 \end{bmatrix}$$

For this matrix, the eigenvalues are approximately $\{-5.0251 \times 10^{-7}, +199\}$ and the respective unit eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} -0.6652 \\ +0.7466 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} +0.6764 \\ +0.7366 \end{bmatrix}$$

Any given vector \vec{u} will be transformed to $\vec{y} = A\vec{u}$ as, approximately,

$$\begin{aligned} \vec{y} &= \lambda_1 \alpha_1 \vec{v}_1 + \lambda_2 \alpha_2 \vec{v}_2 \\ &\approx 199 \alpha_1 \vec{v}_1 \end{aligned}$$

so the transformed \vec{y} is *highly sensitive* to small changes in the direction of \vec{v}_1 and *insensitive* to changes in the direction of \vec{v}_2 .

Extra Notes: Other Normal Matrices

Is it possible for a matrix to be normal but none of orthogonal, symmetric, and skew-symmetric? The answer is positive and can be shown by example.

Consider the matrix

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

This matrix has none of the usual properties, but it is a normal matrix because

$$AA^T = A^T A = \begin{bmatrix} 5 & 2 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & 5 \end{bmatrix}$$

The eigenvalues are

$$\lambda_1 = 5 \quad \lambda_{2,3} = \pm\sqrt{3} i$$

and the eigenvectors (one real, two complex) are orthogonal. So this matrix has an orthonormal eigenvector basis but is not one of the matrix forms above.

This is a *circulant* matrix, which in general are normal matrices and which have an orthonormal eigenvector basis.

End of Extra Notes