
1 Orthogonal Matrices and Symmetric Matrices

Recall that an $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. (see [Diagonalizable Matrices and Multiplicity](#)) Moreover, the matrix P with these eigenvectors as columns is a diagonalizing matrix for A , that is

$$P^{-1}AP \text{ is diagonal.}$$

As we have seen, the nice bases of $\mathbb{R}^n$ are the orthogonal ones, so a natural question is: which $n \times n$ matrices have n orthogonal eigenvectors, so that columns of P form an orthogonal basis for $\mathbb{R}^n$? These turn out to be precisely the *symmetric matrices* (matrices for which $A = A^T$), and this is the main result of this section.

Orthogonal Matrices

A collection of non-zero, pairwise orthogonal vectors in $\mathbb{R}^n$ is called an *orthogonal* set of vectors. An orthogonal set of vectors is called *orthonormal* if $\|\mathbf{q}\| = 1$ for each vector $\mathbf{q}$ in the set. A set of orthogonal vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ can be “normalized”, i.e. converted into an orthonormal set $\left\{ \frac{1}{\|\mathbf{v}_1\|} \mathbf{v}_1, \frac{1}{\|\mathbf{v}_2\|} \mathbf{v}_2, \dots, \frac{1}{\|\mathbf{v}_k\|} \mathbf{v}_k \right\}$. In particular, if a matrix A has n orthogonal eigenvectors, they can (by normalizing) be taken to be orthonormal. The corresponding diagonalizing matrix (we will use Q instead of P) has orthonormal columns, and such matrices are very easy to invert.

Theorem 1. The following conditions are equivalent for an $n \times n$ matrix Q .

1. Q is invertible and $Q^{-1} = Q^T$.
2. The rows of Q are orthonormal.
3. The columns of Q are orthonormal.

Proof First note that condition ?? is equivalent to $Q^T Q = I$. Let $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ denote the columns of Q . Then $\mathbf{q}_i^T$ is the i th row of Q^T , so the (i, j) -entry of $Q^T Q$ is $\mathbf{q}_i \cdot \mathbf{q}_j$. Thus $Q^T Q = I$ means that $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ if $i \neq j$ and $\mathbf{q}_i \cdot \mathbf{q}_i = 1$ if $i = j$. Hence condition ?? is equivalent to ??. The proof of the equivalence of ?? and ?? is similar. ■

Definition 1 (Orthogonal Matrices). An $n \times n$ matrix Q is called an *orthogonal matrix* if it satisfies one (and hence all) of the conditions in Theorem ??.

Example 1. The rotation matrix $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ is orthogonal for any angle θ .

Explanation. See Practice Problem ??.

Remark 1. In view of ?? and ?? of Theorem ??, *orthonormal matrix* might be a better name. But *orthogonal matrix* is standard.

It is not enough that the rows of a matrix A are merely orthogonal for A to be an orthogonal matrix. Here is an example.

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

1. Check that matrix A has rows that are orthogonal.
2. Check that matrix A has columns that are NOT orthogonal.
3. Check that matrix A has rows that are NOT orthonormal.
4. Create a matrix Q by normalizing each of the rows of A .
5. Check that Q is an orthogonal matrix.

Click the arrow to see the answer. You should get $Q = \begin{bmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, and one can check that this is orthogonal

in a number of ways.

This exploration can certainly be done by hand (although it takes some time), but it also makes for a very nice Octave exercise.

```
%Exploration from Section 9.4 Orthogonal Matrices and Symmetric Matrices
A=[2 1 1; -1 1 1; 0 -1 1]
%Check that matrix A has rows that are orthogonal.
A(1,:)*transpose(A(2,:))
A(2,:)*transpose(A(3,:))
A(1,:)*transpose(A(3,:))
%Check that matrix A has columns that are NOT orthogonal.
transpose(A(:,1))*A(:,2)
%(This is 1 of 3 calculations to do.)
%Check that matrix A in the Octave window has rows that are NOT orthonormal.
%(See the results from the first question.)
%Create a matrix Q by normalizing each of the rows of A.
q1=A(1,:)/norm(A(1,:));
q2=A(2,:)/norm(A(2,:));
q3=A(3,:)/norm(A(3,:));
Q = [q1;q2;q3]
%Check that Q is an orthogonal matrix.
Q*transpose(Q)
%(You may get numbers close to zero in some places you expect
%to get zero due to rounding error)
```

[Link to code.](#)

For more information about using Octave, please see [Octave for Chapter 9](#).

We studied the idea of closure when we studied [Subspaces of \$\mathbb{R}^n\$](#) . The next theorem tells us that orthogonal matrices are closed under matrix multiplication.

Theorem 2. 1. If P and Q are orthogonal matrices, then PQ is also orthogonal. (We say that the set of orthogonal matrices is closed under matrix multiplication.)

2. If P is an orthogonal matrix, then so is $P^{-1} = P^T$.

Proof For ??, P and Q are invertible, so PQ is also invertible and

$$(PQ)^{-1} = Q^{-1}P^{-1} = Q^T P^T = (PQ)^T$$

Hence PQ is orthogonal. For ??,

$$(P^{-1})^{-1} = P = (P^T)^T = (P^{-1})^T$$

shows that P^{-1} is orthogonal. ■

Symmetric Matrices

We now shift our focus from orthogonal matrices to another important class of $n \times n$ matrices called *symmetric* matrices. A *symmetric matrix* is a matrix which is equal to its transpose. We saw a few examples of such matrices in [Transpose of a Matrix](#).

When we began our study of eigenvalues and eigenvectors, we saw numerous examples of matrices with entries that were real numbers with eigenvalues that were complex numbers. It can be shown that symmetric matrices only have real eigenvalues. We also learned that some matrices are diagonalizable while other matrices are not. It turns out that every symmetric matrix is diagonalizable. In fact, we can say more, but first we need the following definition.

Definition 2. An $n \times n$ matrix A is said to be *orthogonally diagonalizable* if an orthogonal matrix Q can be found such that $Q^{-1}AQ = Q^T A Q$ is diagonal.

We have learned earlier that when we diagonalize a matrix A , we write $P^{-1}AP = D$ for some matrix P where D is diagonal, and the diagonal entries are the eigenvalues of A . We have also learned that the columns of the matrix P are the corresponding eigenvectors of A . So when a matrix is orthogonally diagonalizable, we are able to accomplish the diagonalization using a matrix Q consisting of n eigenvectors that form an orthonormal basis for $\mathbb{R}^n$. The following remarkable theorem shows that the matrices that have this property are precisely the symmetric matrices.

Theorem 3 (Real Spectral Theorem). Let A be an $n \times n$ matrix. Then A is symmetric if and only if A is orthogonally diagonalizable.

Proof If A is orthogonally diagonalizable, then it is an easy exercise to prove that it is symmetric. You are asked to do this in Practice Problem ??.

To prove the “only if” part of this theorem, we assume A is symmetric, and we need to show it is orthogonally diagonalizable. We proceed by induction on n , the size of the symmetric matrix. If $n = 1$, A is already diagonal. If $n > 1$, assume that we know the “only if” statement holds for $(n - 1) \times (n - 1)$ symmetric matrices. Let λ_1 be an eigenvalue of A , and let $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, where $\|\mathbf{x}_1\| = 1$. Next, set $\mathbf{q}_1 = \mathbf{x}_1$, and use the Gram-Schmidt algorithm to find

an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ for $\mathbb{R}^n$. Let $Q_1 = \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix}$, so that Q_1 is an orthogonal matrix. We

have

$$\begin{aligned}
 Q_1^T A Q_1 &= \begin{bmatrix} \text{---} & \mathbf{q}_1^T & \text{---} \\ \text{---} & \mathbf{q}_2^T & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{q}_n^T & \text{---} \end{bmatrix} A \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix} \\
 &= \begin{bmatrix} \text{---} & \mathbf{q}_1^T & \text{---} \\ \text{---} & \mathbf{q}_2^T & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{q}_n^T & \text{---} \end{bmatrix} \begin{bmatrix} | & | & & | \\ A\mathbf{q}_1 & A\mathbf{q}_2 & \cdots & A\mathbf{q}_n \\ | & | & & | \end{bmatrix} \\
 &= \begin{bmatrix} \lambda_1 & B \\ \mathbf{0} & A_1 \end{bmatrix},
 \end{aligned}$$

where the block B has dimensions $1 \times (n-1)$, and the block under λ_1 is a $(n-1) \times 1$ zero matrix, because of the orthogonality of the basis vectors.

Next, using the fact that A is symmetric, we notice that

$$(Q_1^T A Q_1)^T = Q_1^T A^T (Q_1^T)^T = Q_1^T A Q_1,$$

so $Q_1^T A Q_1$ is symmetric. It follows that B is also a zero matrix and that A_1 is symmetric. Since A_1 is an $(n-1) \times (n-1)$ symmetric matrix, we may apply the inductive hypothesis, so there exists an $(n-1) \times (n-1)$ orthogonal matrix Q such that $Q^T A_1 Q = D_1$ is diagonal. We observe that $Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$ is orthogonal, and we compute:

$$\begin{aligned}
 (Q_1 Q_2)^T A (Q_1 Q_2) &= Q_2^T (Q_1^T A Q_1) Q_2 \\
 &= \begin{bmatrix} 1 & 0 \\ 0 & Q^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \\
 &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{bmatrix}
 \end{aligned}$$

is diagonal. Because $Q_1 Q_2$ is orthogonal by Theorem 11.1, this completes the proof. ■

Because the eigenvalues of a real symmetric matrix are real, Theorem 11.1 is called the *Real Spectral Theorem*, and the set of distinct eigenvalues is called the *spectrum* of the matrix. A similar result holds for matrices with complex entries (Theorem 11.2).

Example 2. Find an orthogonal matrix Q such that $Q^{-1} A Q$ is diagonal, where $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$.

Explanation. The characteristic equation of A is:

$$\det \begin{bmatrix} 1-\lambda & 0 & -1 \\ 0 & 1-\lambda & 2 \\ -1 & 2 & 5-\lambda \end{bmatrix} = -\lambda(\lambda-1)(\lambda-6) = 0$$

Thus the eigenvalues are $\lambda = 0, 1$, and 6 , and corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$$

respectively. Moreover, by what at first appears to be remarkably good luck, these eigenvectors are *orthogonal*. We have $\|\mathbf{x}_1\|^2 = 6$, $\|\mathbf{x}_2\|^2 = 5$, and $\|\mathbf{x}_3\|^2 = 30$, so

$$P = \begin{bmatrix} \frac{1}{\sqrt{6}}\mathbf{x}_1 & \frac{1}{\sqrt{5}}\mathbf{x}_2 & \frac{1}{\sqrt{30}}\mathbf{x}_3 \end{bmatrix} = \frac{1}{\sqrt{30}} \begin{bmatrix} \sqrt{5} & 2\sqrt{6} & -1 \\ -2\sqrt{5} & \sqrt{6} & 2 \\ \sqrt{5} & 0 & 5 \end{bmatrix}$$

is an orthogonal matrix. Thus $P^{-1} = P^T$ and

$$P^T A P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Actually, the fact that the eigenvectors in Example ?? are orthogonal is no coincidence. These vectors certainly must be linearly independent (they correspond to distinct eigenvalues). We will see that the fact that the matrix is *symmetric* implies that the eigenvectors are orthogonal. To prove this we need the following useful fact about symmetric matrices.

Theorem 4. If A is an $n \times n$ symmetric matrix, then

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$$

for all columns $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

Remark 2. The converse also holds (see Practice Problem ??).

Proof Recall that $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$ for all columns $\mathbf{x}$ and $\mathbf{y}$. Because $A^T = A$, we get

$$(A\mathbf{x}) \cdot \mathbf{y} = (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \mathbf{x}^T A \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$$

■

Theorem 5. If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof Let $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \mu\mathbf{y}$, where $\lambda \neq \mu$. We compute

$$\lambda(\mathbf{x} \cdot \mathbf{y}) = (\lambda\mathbf{x}) \cdot \mathbf{y} = (A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y}) = \mathbf{x} \cdot (\mu\mathbf{y}) = \mu(\mathbf{x} \cdot \mathbf{y})$$

Hence $(\lambda - \mu)(\mathbf{x} \cdot \mathbf{y}) = 0$, and so $\mathbf{x} \cdot \mathbf{y} = 0$ because $\lambda \neq \mu$.

■

Now the procedure for diagonalizing a symmetric $n \times n$ matrix is clear. Find the distinct eigenvalues and find orthonormal bases for each eigenspace (the Gram-Schmidt algorithm may be needed when there is a repeated eigenvalue). Then the set of all these basis vectors is orthonormal (by Theorem ??) and contains n vectors. Here is an example.

Example 3. Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

Explanation. The characteristic equation is

$$\det \begin{bmatrix} 8 - \lambda & -2 & 2 \\ -2 & 5 - \lambda & 4 \\ 2 & 4 & 5 - \lambda \end{bmatrix} = -\lambda(\lambda - 9)^2 = 0$$

Hence the distinct eigenvalues are 0 and 9 are of algebraic multiplicity 1 and 2, respectively. The geometric multiplicities must be the same, for A is diagonalizable, being symmetric. It follows that $\dim(S_0) = 1$ and $\dim(S_9) = 2$. Gaussian elimination gives

$$S_0(A) = \text{span}\{\mathbf{x}_1\}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \quad \text{and} \quad S_9(A) = \text{span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

The eigenvectors in S_9 are both orthogonal to $\mathbf{x}_1$ as Theorem ?? guarantees, but not to each other. However, an orthogonal basis can be found using projections. (See [Gram-Schmidt Orthogonalization](#))

$$\{\mathbf{f}_2, \mathbf{f}_3\} \text{ of } S_9(A) \quad \text{where} \quad \mathbf{f}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{f}_3 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

Normalizing gives orthonormal vectors $\{\frac{1}{3}\mathbf{x}_1, \frac{1}{\sqrt{5}}\mathbf{f}_2, \frac{1}{3\sqrt{5}}\mathbf{f}_3\}$, so

$$Q = \begin{bmatrix} \left| \frac{1}{3}\mathbf{x}_1 \right| & \left| \frac{1}{\sqrt{5}}\mathbf{f}_2 \right| & \left| \frac{1}{3\sqrt{5}}\mathbf{f}_3 \right| \end{bmatrix} = \frac{1}{3\sqrt{5}} \begin{bmatrix} \sqrt{5} & -6 & 2 \\ 2\sqrt{5} & 3 & 4 \\ -2\sqrt{5} & 0 & 5 \end{bmatrix}$$

is an orthogonal matrix such that $Q^{-1}AQ$ is diagonal.

It is worth noting that other, more convenient, diagonalizing matrices Q exist. For example, $\mathbf{y}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{y}_3 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ lie in $S_9(A)$ and they are orthogonal. Moreover, they both have norm 3 (as does $\mathbf{x}_1$), so

$$\hat{Q} = \begin{bmatrix} \left| \frac{1}{3}\mathbf{x}_1 \right| & \left| \frac{1}{3}\mathbf{y}_2 \right| & \left| \frac{1}{3}\mathbf{y}_3 \right| \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$$

is a nicer orthogonal matrix with the property that $\hat{Q}^{-1}A\hat{Q}$ is diagonal.

Theorem 6. Let A be an $n \times n$ matrix. A has an orthonormal set of n eigenvectors if and only if A is orthogonally diagonalizable.

Proof Let $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ be orthonormal eigenvectors of A with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. We must

show A is orthogonally diagonalizable. Let $Q = \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix}$ so that Q is orthogonal. We have

$$AQ = \begin{bmatrix} | & | & & | \\ A\mathbf{q}_1 & A\mathbf{q}_2 & \cdots & A\mathbf{q}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{q}_1 & \lambda_2\mathbf{q}_2 & \cdots & \lambda_n\mathbf{q}_n \\ | & | & & | \end{bmatrix} = QD,$$

where D is the diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$. But then $Q^T A Q = D$, proving this half of the theorem.

For the converse, if A is orthogonally diagonalizable, then by Theorem ?? it is symmetric. But then Theorem ?? tells us that eigenvectors corresponding to distinct eigenvalues are orthogonal. Because A is (orthogonally) diagonalizable, we know the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity. This implies that we can use Gram-Schmidt on each eigenspace of dimension > 1 to get a full set of n orthogonal eigenvectors. ■

If we are willing to replace “diagonal” by “upper triangular” in the real spectral theorem, we can weaken the requirement that A is symmetric to insisting only that A has real eigenvalues.

Theorem 7 (Schur Triangularization Theorem). If A is an $n \times n$ matrix with n real eigenvalues, an orthogonal matrix Q exists such that $Q^T A Q$ is upper triangular.

Remark 3. There is also a lower triangular version of this theorem.

Proof See Practice Problem ?? ■

The eigenvalues of an upper triangular matrix are displayed along the main diagonal. Because A and $Q^T A Q$ have the same determinant and trace whenever Q is orthogonal (for they are similar matrices), Theorem ?? gives:

Corollary 1. If A is an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (possibly not all distinct), then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$ and $\text{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Remark 4. This corollary remains true even if the eigenvalues are not real.

Practice Problems

Problem 2 Normalize the columns to make the following matrix orthogonal. $A = \begin{bmatrix} 1 & -4 \\ 2 & 2 \end{bmatrix}$

$$\frac{1}{\sqrt{\dots}} \begin{bmatrix} \dots & \dots \\ \dots & \dots \end{bmatrix}$$

Problem 3 Normalize the columns to make the following matrix orthogonal. $A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

Problem 4 For each matrix A , find an orthogonal matrix Q such that $Q^{-1}AQ$ is diagonal.

Question 4.1

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$Q = \frac{1}{\sqrt{\dots\dots}} \begin{bmatrix} 1 & -1 \\ \dots\dots & \dots\dots \end{bmatrix}$$

Question 4.2

$$A = \begin{bmatrix} 3 & 0 & 7 \\ 0 & 5 & 0 \\ 7 & 0 & 3 \end{bmatrix}$$

Problem 5 Suppose A is orthogonally diagonalizable. Prove that A is symmetric. (This is the easy direction of the "if and only if" in Theorem ??.)

Problem 6 Prove the converse of Theorem ??:

If $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$ for all n -columns $\mathbf{x}$ and $\mathbf{y}$, then A is symmetric.

Problem 7 Let Q be an orthogonal matrix.

1. Show that $\det Q = 1$ or $\det Q = -1$.
2. Give 2×2 examples of Q such that $\det Q = 1$ and $\det Q = -1$.

Text Source

This section was adapted from Section 8.2 of Keith Nicholson's *Linear Algebra with Applications*. (CC-BY-NC-SA)

W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, p. 424