Chapter 10: Eigenvalues and diagonalization | **Section 6**

Orthogonal diagonalization

What you need to know already:	What you can learn here:
➤ What it means to diagonalize a matrix. ➤ How to diagonalize a matrix.	➤ A special kind of diagonalization that is reserved for a special kind of matrices.

By using what we have learned about eigenvalues and similarity, we can prove a really intriguing fact.

Technical fact

If $\mathbf{Q}$ is an orthogonal matrix, its ***only possible eigenvalues*** are 1 and -1.

Proof

Assume that $\mathbf{Q}$ is an orthogonal matrix, λ is one of its eigenvalues and $\mathbf{v}$ is one of its eigenvectors. Then:

$$\mathbf{Qv} = \lambda\mathbf{v} \quad \Rightarrow \quad \|\mathbf{Qv}\| = \|\lambda\mathbf{v}\|$$

But, being an orthogonal matrix, $\mathbf{Q}$ preserves lengths. Moreover, the scalar multiple λ can be separated by using its absolute value. Therefore:

$$\|\mathbf{Qv}\| = \|\lambda\mathbf{v}\| \quad \Rightarrow \quad \|\mathbf{v}\| = |\lambda|\|\mathbf{v}\| \quad \Rightarrow \quad |\lambda| = 1 \quad \Rightarrow \quad \lambda = \pm 1$$

So, orthogonal matrices are also nice in the sense that their eigenvalues are simple.

Yes, and that implies that a matrix similar to an orthogonal matrix can also have only 1 and -1 as eigenvalues, a fact that may make other computations easy.

But it turns out that another linkage between similarity and orthogonal matrices leads to interesting results.

Definition

A matrix $\mathbf{A}$ is ***orthogonally diagonalizable*** if it can be written as

$$\mathbf{A} = \mathbf{QDQ}^{-1}$$

where $\mathbf{D}$ is a diagonal matrix and $\mathbf{Q}$ is an orthogonal matrix.

So, we are looking at a matrix that is diagonalizable through an orthogonal matrix.

Yes: isn't that fun to say? But there is an easier characterization of these matrices, one that may be a big surprise for you. In fact, the big surprise I have mentioned a few sections back!

Let me start by pointing out a key property of orthogonally diagonalizable matrices.

Technical fact

Every ***orthogonally diagonalizable*** matrix is ***symmetric***.

Proof

We know that the inverse of an orthogonal matrix is its transpose and that a diagonal matrix is obviously symmetric. Therefore, if $\mathbf{A} = \mathbf{QDQ}^{-1}$:

$$\mathbf{A}^T = \left(\mathbf{QDQ}^{-1}\right)^T = \left(\mathbf{QDQ}^T\right)^T = \left(\mathbf{Q}^T\right)^T \mathbf{D}^T \mathbf{Q}^T = \mathbf{QDQ}^{-1} = \mathbf{A}$$

Therefore, the matrix is symmetric.

Too bad that this is another one of those one-way statements, or we could say that orthogonally diagonalizable matrices and symmetric matrices are the same thing.

Well, that is the surprise: they are!

Technical fact
The Spectral Theorem
for symmetric matrices

Every ***symmetric*** matrix is ***orthogonally diagonalizable***.

Wow! And how do we prove that?

That is the problem: unlike the other direction, whose proof is very easy, the proof of this statement is very technical and complicated. So, I will just give you an idea of how it works, but I will leave out the details.

But before we do that, here is an example of how to orthogonally diagonalize a symmetric matrix.

Example: $\begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}$

This matrix is symmetric, so we should be able to diagonalize it by using an orthogonal matrix. We use the usual method of finding eigenvalues and eigenvectors and then arranging them properly.

For eigenvalues, we get:

$$\begin{vmatrix} 1-\lambda & -2 & 0 \\ -2 & -\lambda & 2 \\ 0 & 2 & -1-\lambda \end{vmatrix} = -2\begin{vmatrix} 1-\lambda & 0 \\ -2 & 2 \end{vmatrix} - (1+\lambda)\begin{vmatrix} 1-\lambda & -2 \\ -2 & -\lambda \end{vmatrix}$$

$$= -4 + 4\lambda - (1+\lambda)\left(\lambda^2 - \lambda - 4\right)$$

$$= -4 + 4\lambda - \left[\lambda^2 - \lambda - 4 + \lambda^3 - \lambda^2 - 4\lambda\right] = -\lambda^3 + 9\lambda = -\lambda\left(\lambda^2 - 9\right)$$

Therefore, the eigenvalues are $\lambda = -3, 0, 3$, which allows us to get the diagonal matrix:

$$\mathbf{D} = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Now we construct unit eigenvectors for each eigenvalue:

$$\lambda = -3 \;\Rightarrow\; \begin{bmatrix} 4 & -2 & 0 \\ -2 & 3 & 2 \\ 0 & 2 & 2 \end{bmatrix} \underset{\Rightarrow}{REF} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \;\Rightarrow\; \mathbf{v}_1 = s\begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}$$

$$\lambda = 0 \;\Rightarrow\; \begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix} \underset{\Rightarrow}{REF} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \;\Rightarrow\; \mathbf{v}_2 = t\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

$$\lambda = 3 \;\Rightarrow\; \begin{bmatrix} -2 & -2 & 0 \\ -2 & -3 & 2 \\ 0 & 2 & -4 \end{bmatrix} \underset{\Rightarrow}{REF} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \;\Rightarrow\; \mathbf{v}_3 = k\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

Now notice that the three basic vectors we have found are perpendicular to each other. By normalizing each of them, we can obtain an orthonormal set of vectors:

$$\mathbf{v}_1 = s\begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} \;\Rightarrow\; \mathbf{u}_1 = \frac{1}{3}\begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}$$

$$\mathbf{v}_2 = t\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \;\Rightarrow\; \mathbf{u}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{v}_3 = k\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \;\Rightarrow\; \mathbf{u}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

If we arrange these three vectors as columns of a matrix, we obtain an orthogonal one, as claimed. Therefore:

$$\mathbf{Q} = \frac{1}{3}\begin{bmatrix} -1 & 2 & -2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix} \;\Rightarrow\; \mathbf{A} = \mathbf{QDQ}^{-1}$$

Cool! But why is it called the Spectral Theorem?

Because its proof is based on finding certain characteristics of the *spectrum of eigenvalues*, that is, the entire set of eigenvalues of the matrix. It is also related to a larger set of theorems that are also called spectral for similar reasons (they relate to the entire spectrum of something)

Outline of the proof

The really difficult part of the proof is to show that a symmetric $n \times n$ matrix **A** only has real eigenvalues, that is, that all solutions of the characteristic equation are real numbers. This part of the proof requires the use of complex numbers and their theory.

Once this is done, we know that the sum of the algebraic multiplicities of the eigenvalues of **A** equals n, but what about their geometric multiplicities? It turns out that for a symmetric matrix, geometric and algebraic multiplicities are equal for each eigenvalue. This is less difficult, but still tricky.

Once we know this, we know that **A** has enough eigenvectors to make it diagonalizable, so one only needs to prove that the eigenspaces of different eigenvalues are not just independent, but orthogonal. This is another challenging step, but it turns out to be true. This, together with some steps of the modified Gram-Schmidt process, tell us that we can find a set of orthogonal eigenvectors to make up the matrix **Q**.

Although I have not given you the details of the proof, we can extract from the outline an interesting set of facts about symmetric matrices.

Technical facts

Given a symmetric, $n \times n$ matrix $\mathbf{A}$:

- It is always possible to find an orthogonal set consisting of n eigenvectors for $\mathbf{A}$.
- Any two eigenvectors of $\mathbf{A}$ corresponding to different eigenvalues are orthogonal.
- If $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ is a set of orthonormal column eigenvectors, corresponding, respectively, to the eigenvales $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ (with each eigenvalue repeated according to its multiplicity), then:

$$\mathbf{A} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

The last statement is also called the ***Spectral Decomposition Theorem***.

The first two facts follow immediately from the *Spectral Theorem*, while the last one, while not immediate, is fairly simple to prove, as follows:

Proof

Under our assumptions:

$$\mathbf{A} = \mathbf{QDQ}^{-1} \;\; ; \;\; \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \;\; ; \;\; \mathbf{Q} = [\mathbf{u}_1 \;\; \mathbf{u}_2 \cdots \mathbf{u}_n]$$

But since $\mathbf{Q}$ is orthogonal, its inverse is its transpose, so that we can write:

$$\mathbf{A} = [\mathbf{u}_1 \;\; \mathbf{u}_2 \cdots \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1{}^T \\ \mathbf{u}_2{}^T \\ \cdots \\ \mathbf{u}_n{}^T \end{bmatrix}$$

And now we only have to compute this product:

$$\mathbf{A} = [\mathbf{u}_1 \;\; \mathbf{u}_2 \cdots \mathbf{u}_n] \begin{bmatrix} \lambda_1 \mathbf{u}_1{}^T \\ \lambda_2 \mathbf{u}_2{}^T \\ \cdots \\ \lambda_n \mathbf{u}_n{}^T \end{bmatrix} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1{}^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2{}^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n{}^T$$

as claimed.

Example: $\begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}$

We have seen in the previous example that this matrix has eigenvalues $\lambda = -3, 0, 3$ with corresponding orthonormal eigenvectors:

$$\mathbf{u}_1 = \frac{1}{3}\begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} \quad \mathbf{u}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \quad \mathbf{u}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

Therefore, the Spectral Decomposition Theorem allows us to state that:

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

$$= -3\left(\frac{1}{3}\begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}\right)\left(\frac{1}{3}[-1 \;\; -2 \;\; 2]\right) +$$

$$+0\left(\frac{1}{3}\begin{bmatrix}2\\1\\2\end{bmatrix}\right)\left(\frac{1}{3}\begin{bmatrix}2 & 1 & 2\end{bmatrix}\right)+$$

$$+3\left(\frac{1}{3}\begin{bmatrix}-2\\2\\1\end{bmatrix}\right)\left(\frac{1}{3}\begin{bmatrix}-2 & 2 & 1\end{bmatrix}\right)$$

By adding these three matrices we get:

$$=-\frac{1}{3}\begin{bmatrix}-1\\-2\\2\end{bmatrix}\begin{bmatrix}-1 & -2 & 2\end{bmatrix}+\frac{1}{3}\begin{bmatrix}-2\\2\\1\end{bmatrix}\begin{bmatrix}-2 & 2 & 1\end{bmatrix}$$

$$=-\frac{1}{3}\begin{bmatrix}1 & 2 & -2\\2 & 4 & -4\\-2 & -4 & 4\end{bmatrix}+\frac{1}{3}\begin{bmatrix}4 & -4 & -2\\-4 & 4 & 2\\-2 & 2 & 1\end{bmatrix}=\frac{1}{3}\begin{bmatrix}3 & -6 & 0\\-6 & 0 & 6\\0 & 6 & -3\end{bmatrix}$$

And isn't that true?

Now you got me confused: how is this spectral decomposition useful? The calculations it requires are fairly long…

Its value lies in two features. First, it allows us to write a matrix in terms of orthogonal vectors, a fact that turns out to be useful in further development of matrix theory and linear algebra in general. Second, while the calculations may be long, they are easy, since at every step we are just doing a single product or a single addition, without combining them through a dot product.

But I must admit that its true value only comes up in later applications. Doesn't that make you want to take a second linear algebra course?

I am not sure about that!

Well, in the worst-case scenario that this is the last fact about matrices you'll ever learn, it is now part of your knowledge and you are in the company of a very small group of people educated enough to know it!

Congratulations!

Summary

- Symmetric matrices are the one and only matrices that can be diagonalized through an orthogonal matrix.
- This implies that symmetric matrices can be decomposed into a sum of products of orthogonal vectors.

Common errors to avoid

- This is a very theoretical section, but with a very simple and practical punch line! Enjoy the punch line, but do not ignore the theoretical part, even if it looks strange and complicated. Wrestling with them will do your intellectual health much good!

Learning questions for Section LA 10-6

Review questions:

1. Describe why the *Spectral Theorem* is so surprising and difficult to prove.

Memory questions:

1. What are the possible eigenvalues for an orthogonal matrix?
2. What does the *Spectral theorem* state?
3. Which matrices are orthogonally diagonalizable?
4. Are all symmetric matrices diagonalizable?

Computation questions:

1. Construct an orthogonal diagonalization of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$.

2. Construct a symmetric matrix whose eigenvalues are –1, 0 and 1 with corresponding eigenvectors $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

3. Construct a symmetric matrix whose eigenvalues are 1, 2, 3 and whose corresponding eigenvectors are $\begin{bmatrix} 1 & 2 & 2 \end{bmatrix}, \begin{bmatrix} 2 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 4 & -5 \end{bmatrix}$ respectively.

4. Determine one pair of values (x, y), if any, for which $\mathbf{Q} = \begin{bmatrix} x & 3y & x & 3y \\ x & -5y & y & y \\ 3y & x/3 & y & -5y \\ x & y & -5y & y \end{bmatrix}$ is an orthogonal matrix, and one pair for which it is orthogonally diagonalizable. I expect you to determine such values through an organized procedure, NOT by trial and error.

Theory questions:

1. If **Q** is an orthogonal matrix, what can the product of its eigenvalues be?
2. What is special about symmetric matrices with respect to diagonalization?
3. What feature of the matrix **A** tells us that it is not orthogonally diagonalizable?
4. Can an upper triangular matrix be orthogonally diagonalizable?
5. Does every orthogonal matrix have eigenvalues?
6. Assuming that a matrix **A** is orthogonally diagonalizable, when may the Gram-Schmidt process be needed to complete such diagonalization?
7. Which theorem relates symmetric and orthogonal matrices?

Proof questions:

1. Prove that if **Q** is an orthogonal matrix and λ is one of its eigenvalues, then $\lambda^n \mathbf{Q}$ is also orthogonal for any whole number n.
2. Prove that if **A** and **B** are two orthogonal matrices of the same size, then the square of the product of all eigenvalues of **BA** is 1.

Templated questions:

1. Construct a simple symmetric matrix and obtain its orthogonal diagonalization.

What questions do you have for your instructor?