2.2 Diagonalizability of symmetric matrices

The main theorem of this section is that every real symmetric matrix is not only diagonalizable but *orthogonally* diagonalizable. Two vectors u and v in $\mathbb{R}^n$ are *orthogonal* to each other if $u \cdot v = 0$ or equivalently if $u^T v = 0$. This is sometimes written as $u \perp v$. A matrix A in $M_n(\mathbb{R})$ is called *orthogonal* if

- $u \cdot v = 0$ if u and v are distinct columns of A (the columns of A are pairwise orthogonal to each other), and
- $u \cdot u = 1$ for each column u of A (each column of A is a vector of length 1 in $\mathbb{R}^n$).

Another way to say this is that the columns of A form an *orthonormal basis* of $\mathbb{R}^n$, which means a basis consisting of mutually orthogonal unit vectors. Note that for any matrix $B \in M_{m\times n}\mathbb{R}$, $B^T B$ is the $n \times n$ matrix whose entry in the (i, j) position is the scalar product of Columns i and j of B. Putting this together with the above definition of an orthogonal matrix, it is saying that the square matrix $A \in M_n(\mathbb{R})$ is orthogonal if and only if

$$(A^T A)_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases},$$

i.e. $A \in M_n(\mathbb{R})$ is orthogonal if and only if $A^T A = I_n$.

Definition 2.2.1. *A matrix in* $M_n(\mathbb{R})$ *is* orthogonal *if and only if its inverse is equal to its transpose.*

We note that the set of orthogonal matrices in $M_n(\mathbb{R})$ forms a group under multiplication, called the orthogonal group and written $O_n(\mathbb{R})$. The use of the term "orthogonal" for square matrices differs from its use for vectors - a vector can't just be orthogonal, it can be orthogonal to another vector, but a matrix can be orthogonal by itself. An example of an orthogonal matrix in $M_2(\mathbb{R})$ is $\begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.

The following is our main theorem of this section.

Theorem 2.2.2. *Let* A *be a symmetric matrix in* $M_n(\mathbb{R})$*. Then there exists an orthogonal matrix* P *for which* $P^T AP$ *is diagonal.*

Note that this is saying that $\mathbb{R}^n$ has a basis consisting of eigenvectors of A that are all orthogonal to each other, something that is true only for symmetric matrices. If we have a basis consisting of orthogonal eigenvectors, we can normalize its elements so that our basis consists of unit vectors as required. After we prove Theorem 2.2.2 we will deduce some consequences about positive (semi)definiteness and then look at some applications to graph spectra in the next section.

The following theorem is one of the two keys to the proof of Theorem 2.2.2, and it takes care of the case where the eigenvalues of A are distinct.

Theorem 2.2.3. *Let* A *be a real symmetric matrix. Let* λ *and* μ *be distinct eigenvalues of* A*, with respective eigenvectors* u *and* v *in* $\mathbb{R}^n$*. Then* $u^T v = 0$*.*

Note that $u^T v$ is just the ordinary scalar product of u and v (u^T is just u written as a row). So this theorem is saying that eigenvectors of a real symmetric matrix that correspond to different eigenvalues are orthogonal to each other under the usual scalar product.

Proof. The matrix product $u^T Av$ is a real number (a 1×1 matrix). We can write

$$u^T Av = u^T \mu v = \mu u^T v.$$

On the other hand, being a 1×1 matrix, $u^T Av$ is equal to its own transpose, so

$$u^T Av = (u^T Av)^T = v^T A^T (u^T)^T = v^T Au = v^T \lambda u = \lambda v^T u.$$

Now $v^T u = u^T v$ since both are equal to the scalar product $u \cdot v$ (or because they are 1×1 matrices that are transposes of each other). So what we are saying is

$$\mu u^T v = \lambda u^T v.$$

Since $\mu \neq \lambda$, it follows that $u^T v = 0$. $\square$

From Theorem 2.2.3 and Lemma 2.1.2, it follows that if the symmetric matrix $A \in M_n(\mathbb{R})$ has distinct eigenvalues, then $A = P^{-1}AP$ (or $P^T AP$) for some orthogonal matrix P. It remains to consider symmetric matrices with repeated eigenvalues. We need a few observations relating to the ordinary scalar product on $\mathbb{R}^n$.

Definition 2.2.4. *Let* U *be a subspace of* $\mathbb{R}^n$. *Then the* orthogonal complement of U, *denoted* $U^\perp$, *is defined by*

$$U^\perp = \{v \in \mathbb{R}^n : v \cdot u = 0 \; \forall \; u \in U\}.$$

Notes

1. For example, if $U = \langle e_1, e_2 \rangle$ in $\mathbb{R}^n$, then $U^\perp = \langle e_3, \dots, e_n \rangle$.
2. It is easily checked that $U^\perp$ is a subspace of $\mathbb{R}^n$, not just a subset.
3. For any subspace U of $\mathbb{R}^n$, $U \cap U^\perp = \{0\}$, since element of $U \cap U^\perp$ must be orthogonal to itself under the usual scalar product. However the scalar product of any non-zero vector in $\mathbb{R}^n$ with itself is the sum of the squares of its entries, which is a positive real number.
4. Suppose that U has dimension k and let $\{u_1, \dots, u_k\}$ be a basis of u. Let A_U be the $k \times n$ matrix that has $u_1^T, \dots, u_k^T$ as its k rows. Then A_U has rank k since its rows are linearly independent, and by definition $U^\perp$ is just the right nullspace of A_U. It follows from the rank-nullity theorem that the dimension of $U^\perp$ is $n - k$.
5. Suppose that $\{u_1, \dots, u_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$ whose elements are mutually orthogonal, so that $u_i \cdot u_j = 0$ whenever $i \neq j$. Let $U = \langle u_1, \dots, u_k \rangle$. If $k < n$, let $v_{k+1} \in U^\perp$ and note that $\{u_1, \dots, u_k, v_{k+1}\}$ is a linearly independent set, since $v_{k+1} \notin U$. If the span of these $k+1$ elements is still not all of $\mathbb{R}^n$, we can add an element of $\langle u_1, \dots, u_k, v_{k+1} \rangle^\perp$ to obtain a larger linearly independent set of mutually orthogonal vectors in $\mathbb{R}^n$. Continuing in this way we can extend $\{u_1, \dots, u_k\}$ to a basis of $\mathbb{R}^n$ consisting of mutually orthogonal elements (we can normalize these if we wish to obtain an orthonormal basis). We have the following useful fact: *any linearly independent set of mutually orthogonal unit vectors in* $\mathbb{R}^n$ *can be extended to an orthonormal basis of* $\mathbb{R}^n$.

The following lemma is the last ingredient needed for the proof of Theorem 2.2.2. This lemma would not be true without the hypothesis that A is symmetric. When you are studying the proof, make sure that you are attentive to how the symmetry of A is used. Note the statement that U is A-invariant means that $Au \in U$ whenever $u \in U$.

Lemma 2.2.5. *Let* $A \in M_n(\mathbb{R})$ *be symmetric and suppose that* U *is an* A*-invariant subspace of* $\mathbb{R}^n$. *Then* $U^\perp$ *is also* A*-invariant.*

Proof. Suppose that $v \in U^\perp$. We need to show that $Av \in U^\perp$ also, i.e. that $u^T Av = 0$ for all $u \in U$. So let $u \in U$ and observe that

$$(u^T Av)^T = v^T A^T u = v^T Au.$$

Since $Au \in U$ and $v \in U^\perp$, we know that $v^T Au = 0$. Thus $u^T Av = 0$ also, for all $u \in U$. This means exactly that $Av \in U^\perp$, as required. $\square$

We are now ready to complete the proof of Theorem 2.2.2.

Proof. The proof proceeds by induction on n, but Lemma 2.2.5 is the key ingredient. The case $n = 1$ is trivial, since all 1×1 matrices are diagonal.

Let $\lambda_1, \ldots, \lambda_k$ be the *distinct* eigenvalues of A, and let u_i be an eigenvector (of length 1) corresponding to λ_i. Note that $k \geqslant 1$ since A has at least one eigenvalue. If $k = n$, then by Theorem 2.2.3 and Lemma 2.1.2, there is nothing to do. So we assume that $k < n$ and write $U = \langle u_1, \ldots, u_k \rangle \subseteq \mathbb{R}^n$. Then U is A-invariant, since Au_i is a scalar multiple of u_i for each i. Moreover, the u_i are mutually orthogonal by Theorem 2.2.3, and $\dim U = k$ by Lemma 2.1.2.

Now as in item 5. in the notes above, we can extend $\{u_1, \ldots, u_k\}$ to an orthonormal basis $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$, where $U^\perp = \{v_{k+1}, \ldots, v_n\}$. Let Q be the orthogonal matrix whose columns are $u_1, \ldots, u_k, v_{k+1}, \ldots, v_n$. Then $Q^{-1}AQ$ is symmetric, since $Q^{-1} = Q^T$. Moreover, because $u_1, \ldots, u_k$ are eigenvectors of A and because $U^\perp$ is A-invariant, the matrix Q^TAQ has $\lambda_1, \ldots, \lambda_k$ in the first k diagonal positions, has a symmetric $(n-k) \times (n-k)$ block A_1 in the lower right, and is otherwise full of zeros.

By the induction hypothesis, there exists an orthogonal matrix $Q_1 \in M_{n-k}(\mathbb{R})$ for which $Q_1^{-1}A_1Q_1$ is diagonal. Let $P \in M_n(\mathbb{R})$ be the orthogonal matrix that has I_k in the upper left $k \times k$ block, Q_1 in the lower right $(n-k) \times (n-k)$ block, and zeros elsewhere. Then

$$P^{-1}Q^{-1}AQP = (QP)^{-1}A(QP)$$

is diagonal. Moreover QP is orthogonal since

$$(QP)^{-1} = P^{-1}Q^{-1} = P^TQ^T = (QP)^T.$$

So A is orthogonally diagonalizable as required. □

Two consequences of Theorem 2.2.2 are the following two characterizations of symmetric positive semidefinite matrices.

Theorem 2.2.6. *Let A be a symmetric matrix in $M_n(\mathbb{R})$. Then the following conditions are equivalent.*

1. *A is positive semidefinite.*
2. *All eigenvalues of A are non-negative.*
3. *$A = BB^T$ for some $B \in M_n(\mathbb{R})$.*

We have seen some of the implications of this theorem already in Section 2.1, where we proved that 1. $\Longrightarrow$ 2 and 3. $\Longrightarrow$ 1. We complete the proof by using Theorem 2.2.2 to show that 2. $\Longrightarrow$ 3.

Proof. First assume 2., that the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A are all non-negative. Then, by Theorem 2.2.2, the matrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ satisfies

$$D = P^TAP,$$

for some orthogonal matrix $A \in M_n(\mathbb{R})$. Then $A = PDP^T$. Let D_1 be the diagonal matrix in $M_n(\mathbb{R})$ whose diagonal entries are the non-negative square roots in $\mathbb{R}$ of $\lambda_1, \ldots, \lambda_n$. Then D_1 is symmetric and $D_1^2 = D$. We use this to deduce 3. as follows:

$$A = PDP^T = P(D_1)^2P^T = (PD_1)(D_1P^T) = (PD_1)(D_1^TP^T) = (PD_1)(PD_1)^T.$$

Thus A satisfies 3., and we now have the implications 1. $\Longrightarrow$ 2., 2. $\Longrightarrow$ 3. and 3. $\Longrightarrow$ 1, which means that any of the three conditions of Theorem 2.2.6 follows from any of the others. □

We will look at some consequences for graphs in the next section.