7 Symmetric Matrices and Quadratic Forms

7.1 Diagonalization of Symmetric Matrices

Definition An $n \times n$ matrix A is called **symmetric** if it is equal to its transpose, i.e. if $A = A^T$.

Example The matrix

$$\begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix}$$

is symmetric. The matrix

$$\begin{bmatrix} 2 & 4 & 0 \\ 0 & 3 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

is not symmetric.

Notice that the main diagonal entries of a symmetric matrix can be arbitrary, but *non-diagonal entries come in pairs*. For example, in the first matrix above, the $(1, 2)$ and $(2, 1)$ entries are both equal to 4. In general, if $i \neq j$, then the (i, j) and (j, i) entries of a symmetric matrix are equal.

Example If possible, diagonalize the matrix

$$A = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}.$$

This has characteristic equation

$$0 = \det(A - \lambda I_2) = \begin{vmatrix} 1-\lambda & 5 \\ 5 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 25 = \lambda^2 - 2\lambda - 24 = (\lambda - 6)(\lambda + 4).$$

Hence the eigenvalues are 6 and -4. An eigenvector corresponding to 6 is given by computing as follows:

$$A - 6I_2 = \begin{bmatrix} -5 & 5 \\ 5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is a basis eigenvector. Likewise, an eigenvector for -4 is given by the following:

$$A + 4I_2 = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

is a basis eigenvector. Hence we have

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 6 & 0 \\ 0 & -4 \end{bmatrix}$$

as the diagonalization. Notice that the eigenvectors corresponding to the different eigenvalues are orthogonal:

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1 + 1 = 0.$$

This is actually a more general phenomenon.

Theorem If A is a symmetric matrix, then any two eigenvectors of A corresponding to different eigenvalues are orthogonal.

Proof Let $\bar{v}_1$ and $\bar{v}_2$ be eigenvectors that correspond to distinct eigenvalues λ_1 and λ_2, respectively. We want to show that $\bar{v}_1 \cdot \bar{v}_2$. Our goal is to show that

$$\lambda_1 \bar{v}_1 \cdot \bar{v}_2 = \lambda_2 \bar{v}_1 \cdot \bar{v}_2.$$

It then follows that

$$(\lambda_1 - \lambda_2)\bar{v}_1 \cdot \bar{v}_2 = 0.$$

Since $\lambda_1 \neq \lambda_2$, we see that $\lambda_1 - \lambda_2 \neq 0$, so the only possibility is that $\bar{v}_1 \cdot \bar{v}_2 = 0$. So how do we show that $\lambda_1 \bar{v}_1 \cdot \bar{v}_2 = \lambda_2 \bar{v}_1 \cdot \bar{v}_2$? For this, we must use A, and in particular we must use the symmetry of A:

$$\begin{aligned}
\lambda_1 \bar{v}_1 \cdot \bar{v}_2 &= (\lambda_1 \bar{v}_1) \cdot \bar{v}_2 \\
&= (A\bar{v}_1) \cdot \bar{v}_2 \\
&= (A\bar{v}_1)^T \bar{v}_2 \\
&= \bar{v}_1^T A^T \bar{v}_2 \\
&= \bar{v}_1^T A \bar{v}_2 \\
&= \bar{v}_1^T \lambda_2 \bar{v}_2 \\
&= \lambda_2 \bar{v}_1 \cdot \bar{v}_2.
\end{aligned}$$

We win.

Definition An $n \times n$ matrix A is said to be **orthogonally diagonalizable** if there exists an orthogonal matrix P (so $P^{-1} = P^T$) and a diagonal matrix D such that

$$A = PDP^T = PDP^{-1}.$$

When is a matrix orthogonally diagonalizable?

Theorem An $n \times n$ matrix A is orthogonally diagonalizable if and only if it is symmetric (i.e. $A = A^T$).

Proof We show just one direction, namely we show that if A is orthogonally diagonalizable, then $A = A^T$.

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A.$$

Now the question is: how do you go about orthogonally diagonalizing a matrix? The process is relatively straightforward, but the computations are often long and tedious. We describe the process here, and refer you to the textbook (Section 7.1 after Theorem 2) for an example. In comparison with standard diagonalization, the essential difference in this case is that we need the columns of P to be an *orthonormal* basis of eigenvectors. The process for finding an orthogonal diagonalization for a symmetric matrix A is:

1. Find the eigenvalues of A.
2. Find basis vectors for the eigenspaces corresponding to each eigenvalue found in (1).
3. Perform the Gram-Schmidt process on each of the bases for the eigenspaces found in (2) to get an *orthogonal* basis for each eigenspace.
4. Normalize all of the vectors found in (3).
5. The set of all normalized vectors is an orthonormal basis of A (consisting of eigenvectors). Use these to write P.
6. Write D.

Oftentimes you will hear the eigenvalues of a matrix referred to as its *spectrum*. This is the reason for the name of the following theorem.

Theorem (Spectral Theorem for Symmetric Matrices)
An $n \times n$ symmetric matrix A has the following properties:
(a) A has n real eigenvalues (counting multiplicities).
(b) The dimension of the eigenspaces for each eigenvalue λ is equal to the multiplicity of λ.
(c) The eigenspaces are mutually orthogonal.
(d) A is orthogonally diagonalizable.

This is a summary of several things we have already seen. It leads to what is called the Spectral Decomposition of A. Suppose A is symmetric, with $A = PDP^T$. Suppose also that the (orthonormal) columns of P are $\{\bar{u}_1, \ldots, \bar{u}_n\}$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$\begin{aligned}
A &= PDP^T \\
&= \begin{bmatrix} \bar{u}_1 & \ldots & \bar{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \bar{u}_1^T \\ \vdots \\ \bar{u}_n^T \end{bmatrix} \\
&= \lambda_1 \bar{u}_1 \bar{u}_1^T + \cdots + \lambda_n \bar{u}_n \bar{u}_n^T.
\end{aligned}$$

This last equality is called the **spectral decomposition of** A. Each $\bar{u}_i \bar{u}_i^T$ is an $n \times n$ matrix that has rank 1. Moreover, each $\bar{u}_i \bar{u}_i^T$ is a projection matrix in the sense that $\bar{u}_i \bar{u}_i^T \bar{x}$ is the projection of $\bar{x}$ onto the space spanned by $\bar{u}_i$. The moral is this: we can think of the transformation $T(\bar{x}) = A\bar{x}$ as a sum of simpler transformations $T_i(\bar{x}) = (\lambda_i \bar{u}_i \bar{u}_i^T)\bar{x}$, each of which is just projection onto $\bar{u}_i$ and followed by a scaling by λ_i.