Diagonalization

Recall that if a $n \times n$ matrix D is a *diagonal matrix*, then D has zero entries everywhere, except possibly along the main diagonal (that is, an $n \times n$ diagonal matrix D has entries $d_{ij} = 0$, where $i \neq j$). An $n \times n$ matrix A is **diagonalizable** if there exists a diagonal matrix D and invertible matrix P such that $A = PDP^{-1}$ (that is, a matrix A is diagonalizable if A is similar to some diagonal matrix D).

Let A be any $n \times n$ matrix, let X be any invertible $n \times n$ matrix with columns $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, and let D be any diagonal matrix with diagonal entries $\{\lambda_1, \dots, \lambda_n\}$. Then it follows that the matrices AX and XD are defined by the following.

$$\begin{aligned} AX &= A \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{x}_1 & \cdots & A\mathbf{x}_n \end{bmatrix} \end{aligned} \qquad \begin{aligned} XD &= X \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{x}_1 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix} \end{aligned}$$

Observe that if A is diagonalizable, where $A = XDX^{-1}$, then by right-multiplying $A = XDX^{-1}$ by X, it follows that $AX = XDX^{-1}X = XDI = XD$ and hence by the preceding expressions for the matrices AX and XD, it follows that $\begin{bmatrix} A\mathbf{x}_1 & \cdots & A\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{x}_1 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix}$, which further implies that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ for each vector $\mathbf{x}_i$ and scalar λ_i. Also observe that since X is an invertible matrix, it follows from the Invertible Matrix Theorem that the columns $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of X are linearly independent and, since every vector $\mathbf{x}_i$ in $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is nonzero, the expressions for AX and XD show that $\{\lambda_1, \dots, \lambda_n\}$ are the eigenvalues of A that correspond respectively to $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, the eigenvectors of A.

Theorem (The Diagonalization Theorem).

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. If X and D are both $n \times n$ matrices, where X is invertible and D is diagonal, then $A = XDX^{-1}$ if and only if the columns of X are n linearly independent eigenvectors of A. In such a case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in X.

Diagonalizing Matrices

If A is an $n \times n$ matrix, then diagonalizing A amounts to determining whether there exists an invertible $n \times n$ matrix X and $n \times n$ diagonal matrix D such that $A = XDX^{-1}$, where the columns of X are a linearly independent set $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of eigenvectors with corresponding eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ from the diagonal of D. Therefore, a given $n \times n$ matrix A can be diagonalized by first using the characteristic equation $\det(A - \lambda I) = 0$ to obtain $\{\lambda_1, \dots, \lambda_n\}$, the set of all eigenvalues of A and, with the eigenvalues $\{\lambda_1, \dots, \lambda_n\}$, a set of linearly independent eigenvectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ can then be obtained by producing a basis for the eigenspace $\mathrm{Nul}(A - \lambda_i I)$ corresponding to each eigenvalue λ_i.

Example. Diagonalize the matrix A defined below, if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

TBF ■

Recall that if $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ are eigenvectors corresponding to n distinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ of an $n \times n$ matrix A, then the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is linearly independent. Therefore, since the set of eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is linearly independent in such a case, it follows from the Diagonalization Theorem that an $n \times n$ matrix A with distinct n eigenvalues is diagonalizable.

Theorem (Distinct Eigenvalues and Diagonalization).

An $n \times n$ matrix A with n distinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ is diagonalizable.

Note. An $n \times n$ matrix A may have n indistinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ and be diagonalizable if A has n linearly independent eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. That is, an $n \times n$ matrix A may have an eigenvalue λ_i of $\mathrm{mul}(\lambda_i) > 1$ and still have n linearly independent eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

Matrices with Indistinct Eigenvalues