§6.3
Diagonalization

Diagonalization

Theorem 6.3.1
If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct e-values of an $n \times n$ matrix, then the corresponding e-vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linearly independent.

Definition: An $n \times n$ matrix A is diagonalizable if there exists a nonsingular matrix X and a diagonal matrix D s.t. $X^{-1}AX = D$ (or, equivalently, $A = XDX^{-1}$).
We say that X diagonalizes A.

THEOREM 6.3.2
A is diagonalizable iff it has n linearly independent e-vectors.

Proof. Let $D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & \lambda_n \end{bmatrix}$ and $X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{bmatrix}$ (matrix whose columns are the linearly independent e-vectors).
Since $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$ we have
$AX = \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \ldots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \ldots \lambda_n\mathbf{x}_n \end{bmatrix} =$

$$\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & \lambda_n \end{bmatrix} = XD.$$

Thus $AX = XD$ and $D = X^{-1}AX$ (or $A = XDX^{-1}$).

Remarks:

1. If A is diagonalizable, then the column vectors of the diagonalizing matrix X are e-vectors of A and the diagonal elements of D are the corresponding e-values.
2. The diagonalization is not unique. We can change the order of the entries in D (and the order of the corresponding e-vectors) or multiply the columns of X by any nonzero number.
3. If the $n \times n$matrix A has n distinct e-values, then A is diagonalizable. If the e-values are not distinct, it may or may not be diagonalizable, depending on whether A has n linearly independent e-vectors or not.
   If A has fewer than n linearly independent e-vectors, we say that A is **DEFECTIVE**. A **DEFECTIVE** matrix **is NOT diagonalizable**
4. We define:
   (i) **Geometric Multiplicity of λ = GM** = number of independent eigenvectors for λ = dimension of nullspace of $A - \lambda I$.
   (ii) **Algebraic Multiplicity of λ = AM** = number of repetitions of λ among the eigenvalues.

   If GM <AM the matrix is defective.

EXAMPLE 1:

Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -1 & -1 \\ 3 & 1 & 1 \end{bmatrix}$. Determine whether A is diagonalizable.

Solution: We have $\det(A - \lambda I) = \lambda^2(2 - \lambda)$, thus the e-vectors are $\lambda = 0$ with algebraic multiplicity two $(\text{AM} = 2)$ and $\lambda = 2$ with algebraic multiplicity one.
To find the e-vectors belonging to the e-value $\lambda = 0$ we must find the nullspace of A.
The RREF of A is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.
We have one free variable and the solution set is $\mathbf{x} = \alpha(0, -1, 1)^T$. A basis of the eigenspace associated to $\lambda = 0$ is $(0, -1, 1)^T$. Thus the GM of $\lambda = 0$ is one. Since GM $<$ AM , the matrix is defective.

EXAMPLE 2:

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 3 \\ 1 & 1 & -1 \end{bmatrix}$. Determine whether the matrix A is diagonalizable and, if it is, find the diagonalizing matrix X and the diagonal matrix D.

Solution:
$\det(A - \lambda I) = (1 - \lambda)(\lambda^2 - 4)$ and the e-values are $\lambda_1 = 1, \quad \lambda_2 = 2$ and $\lambda_3 = -2$.
Because the e-values are distinct we know from Theorem 6.3.1 that the corresponding e-vectors are linearly independent and therefore the matrix is diagonalizable. For the e-vector corresponding to $\lambda_1 = 1$, the RREf of $A - I$ is $\begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$ and $(3, 1, 2)^T$ is a basis of the eigenspace.
For $\lambda_2 = 2$, the RREf of $A - 2I$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$ and $(0, 3, 1)^T$ is a basis of the e-space.
For $\lambda_3 = -2$, the RREf of $A + 2I$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $(0, -1, 1)^T$ is a basis of the e-space.
A possible choice for the matrices X and D is then
$X = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.
We also have $X^{-1} = \begin{bmatrix} 1/3 & 0 & 0 \\ -1/4 & 1/4 & 1/4 \\ -5/12 & -1/4 & 3/4 \end{bmatrix}$ and we can easily check that $A = XDX^{-1}$.

EXAMPLE 3: a diagonalizable matrix where the e-values are not all distinct

Let $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{bmatrix}$. We have $\det(A - \lambda I) = \lambda^2(2 - \lambda)$, thus $\lambda = 0$ is an e-value with AM $= 2$ and $\lambda = 2$ is an e-value with AM $= 1$.
The RREF of A is $\begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. There are two free variables and a basis for the eigenspace is given by $\{(-2, 1, 0)^T, (1, 0, 1)^T\}$ (thus GM $= 2$).
The RREF of $A - 2I$ is $\begin{bmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{bmatrix}$ and a basis for the eigenspace is given by $\{(1, 2, 3)^T\}$.
Thus A is diagonalizable and a possible choice for the diagonalizing matrix is

$X = \begin{bmatrix} -2 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$ and $D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $X^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ -3/2 & -3 & 5/2 \\ 1/2 & 1 & -1/2 \end{bmatrix}$ and we can easily check that $A = XDX^{-1}$.

Powers of A

If A is diagonalizable, it is easier to evaluate powers of A.
$A = XDX^{-1} \Rightarrow A^2 = (XDX^{-1})(XDX^{-1}) = XD(X^{-1}X)DX^{-1} = XD^2X^{-1}$
and, **in general,**

$$A^k = XD^kX^{-1}.$$

EXAMPLE: Let $M = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}$. Find formulas for the entries of M^n.

Solution: The matrix M has e-values $\lambda = 3$ with corresponding e-vector $\mathbf{x}_1 = (1, -1)^T$ and $\lambda = 2$ with e-vector $\mathbf{x}_2 = (2, -1)^T$.
$X = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, $X^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$
$M^n = XD^nX^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2^{n+1} - 3^n & 2^{n+1} - 2(3^n) \\ -2^n + 3^n & -2^n + 2(3^n) \end{bmatrix}$.

We can use the diagonalization to define A^p with p a non integer number: $A^p = XD^pX^{-1}$.

EXAMPLE:
Let $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$. Find $A^{1/2} = \sqrt{A}$.

Solution: First diagonalize A. The eigenvalues are 1 and 9 with corresponding e-vectors $(-1, 1)^T$ and $(1, 1)^T$ respectively.
Then $A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ and
$\sqrt{A} = X\sqrt{D}X^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.
We can check the answer by evaluating $(\sqrt{A})^2 = A$.

Similar Matrices

Definition: B is **SIMILAR** to A if there exists a non singular matrix S such that $B = S^{-1}AS$.

For example the matrix A and the matrix D of the eigenvalues are similar, with $S = X$, however we can find infinitely many examples of similar matrices by choosing any nonsingular matrix as our S.

THEOREM:
Two similar matrices have the same e-values.

Proof: Let A and B be two similar matrices with $B = S^{-1}AS$. Then
$\det(B - \lambda I) = \det(S^{-1}AS - \lambda I) = \det(S^{-1}AS - \lambda S^{-1}S) = \det(S^{-1}(A - \lambda I)S) =$

$$= \det(S^{-1})\det(A - \lambda I)\det(S) = \frac{1}{\det(S)}\det(A - \lambda I)\det(S) = \det(A - \lambda I).$$

The two matrices have the same characteristic polynomial and therefore the same e-values. $\blacksquare$.

THEOREM:
If A is diagonalizable and B is similar to A, then B is diagonalizable and the diagonalizing matrix is $S^{-1}X$.

Proof. Let $A = XDX^{-1}$ and $B = S^{-1}AS$, then
$B = S^{-1}(XDX^{-1})S = (S^{-1}X)D(X^{-1}S) = (S^{-1}X)D(S^{-1}X)^{-1}$. ■

SYMMETRIC MATRICES
Key facts:

1. A symmetric matrix has *real eigenvalues* and *perpendicular eigenvectors*.
2. Diagonalization becomes $A = QDQ^T$ with an orthogonal matrix Q (the columns of Q are the unit eigenvectors).
3. All symmetric matrices are diagonalizable, even with repeated eigenvalues.