SECTION VIII B: DIAGONALIZING

We have discussed at the beginning of this chapter how diagonal matrices are unusual in having simple closed forms for their powers; e.g.

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \rightarrow D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix},$$

$$k = 0, 1, 2, \ldots$$

p. 686

Almost as good is

to have

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

for some invertible P; then

$$A^2 = (PDP^{-1})(PDP^{-1}) = (PD)(P^{-1}P)(DP^{-1})$$

$$= (PD)(I)DP^{-1} = PD^2P^{-1}$$

$$A^3 = (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1}$$

$$= PD^3P^{-1}$$

$\vdots$

p. 687

TERMINOLOGY 8.18

$$\begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_m \end{bmatrix}$$ will mean

an $(m \times m)$ diagonal matrix whose diagonal entries, going down the diagonal, are $\lambda_1, \lambda_2, \ldots, \lambda_m$.

THEOREM 8.19

If $A = P \begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_m \end{bmatrix} P^{-1}$,

then

$$A^k = P \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & O \\ & & \ddots & \\ O & & & \lambda_m^k \end{bmatrix} P^{-1},$$

for $k = 1, 2, 3, \ldots$

As a corollary, we have the second of our two methods for solving Difference Equations. This method works for any initial data $\vec{x}_0$, for A as in Theorem 8.19; compare to ~~8~~.15.

p. 689

GLOBAL METHOD for solving Difference Equations 8.20

If A is as in Theorem 8.19, then, for any $\vec{x}_0$,

$$\vec{x}_n = P \begin{bmatrix} \lambda_1^n & & & \bigcirc \\ & \lambda_2^n & & \\ & & \ddots & \\ \bigcirc & & & \lambda_m^n \end{bmatrix} P^{-1} \vec{x}_0 \quad \begin{pmatrix} n = \\ 1,2,3,\ldots \end{pmatrix}$$

is the solution of

$$\vec{x}_{k+1} = A\vec{x}_k \quad (k = 0, 1, 2, \ldots)$$

p. 690

Example 8.21

If $A = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}$, then

we will soon know how to show that

$$A = P \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} P^{-1},$$

with $P = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$

(Construction of P & $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ soon to appear)

p. 691

By Theorem 8.19, for

$k = 1, 2, 3, \ldots,$

$$A^k = P \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix} P^{-1} =$$

$$\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -(2^k) & -(2^{k+1}) \\ 3^k & 3^k \end{bmatrix} =$$

$$\begin{bmatrix} (-(2^k) + 2(3^k)) & (-(2^{k+1}) + 2(3^k)) \\ (2^k - (3^k)) & (2^{k+1} - (3^k)) \end{bmatrix}$$

By 8.20, the solution of

$$\vec{x}_{k+1} = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} \vec{x}_k \quad (k = 0, 1, 2, \ldots)$$

$$\vec{x}_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$ is

$$\vec{x}_n = A^n \vec{x}_0 = P \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix} P^{-1} \begin{bmatrix} 3 \\ -1 \end{bmatrix} =$$

$$\begin{bmatrix} (-(2^n) + 2(3^n)) & (-(2^{n+1}) + 2(3^n)) \\ (2^n - (3^n)) & (2^{n+1} - (3^n)) \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} =$$

$$\begin{bmatrix} 3(-(2^n) + 2(3^n)) - (-(2^{n+1}) + 2(3^n)) \\ 3(2^n - (3^n)) - (2^{n+1} - (3^n)) \end{bmatrix} =$$

$$\begin{bmatrix} -3(2^n) + 2^{n+1} - 2(3^n) + 2(3^{n+1}) \\ 3(2^n) - 2^{n+1} + 3^n - (3^{n+1}) \end{bmatrix}$$

$(n = 1, 2, 3, \ldots)$

We need a name for the hypothesis of Theorem 8.19. First, a more general name.

DEFINITION 8.22

Matrices A and B are

similar if

$$A = PBP^{-1},$$

for some invertible matrix P.

DEFINITION 8.23

A matrix A is

diagonalizable

if it is similar to a

diagonal matrix; that is,

$$(*) \quad A = PDP^{-1},$$

for some diagonal matrix D and invertible matrix P.

$(*)$ is then a

diagonalization

of A.

Example 8.24

$A = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}$, from

Example 8.21, is diagonalizable, with

diagonalization

$$A = PDP^{-1},$$

$$P = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

PLEASE NOTE that

$$A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \& \quad A\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 3\begin{bmatrix} 2 \\ -1 \end{bmatrix};$$

P has eigenvectors for columns and D has corresponding eigenvalues for diagonal entries.

p. 696

HOW TO DIAGONALIZE 8.25

If A is $(m \times m)$ and $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ is a basis for $\mathbb{R}^m$, with

$$A\vec{v}_k = \lambda_k \vec{v}_k \quad (k = 1, 2, \dots, m),$$

let

$$P \equiv [\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_m],$$

$$D \equiv [\lambda_1 \vec{e}_1 \ \ \lambda_2 \vec{e}_2 \ \dots \ \lambda_m \vec{e}_m]$$

$$= \begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_m \end{bmatrix}$$

Then P is invertible and

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

BEST WAY to get

$$\{ \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m \}:$$

Take the union of the bases for the eigenspaces;

the resulting set will automatically be linearly independent.

Partial Proof:

For $1 \le j \le m$,

$A\vec{v}_j = \lambda_j \vec{v}_j$, $P\vec{e}_j = \vec{v}_j$, thus
$P^{-1}\vec{v}_j = \vec{e}_j$, so that

$$(P^{-1}AP)\vec{e}_j = P^{-1}A\vec{v}_j =$$
$$P^{-1}(\lambda_j \vec{v}_j) = \lambda_j(P^{-1}\vec{v}_j) = \lambda_j \vec{e}_j$$
$$= D\vec{e}_j.$$

This implies that

$$P^{-1}AP = D,$$ or

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

p. 699

Examples 8.26

In each of the following, use the information given to diagonalize A.

(1) $A\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}-5\\0\\-5\end{bmatrix}$, $A\begin{bmatrix}1\\1\\2\end{bmatrix} = \begin{bmatrix}-5\\-5\\-10\end{bmatrix}$

$A\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\7\end{bmatrix}$.

SOLUTION: $\vec{v}_1 \equiv (1,0,1)$, $\vec{v}_2 \equiv (1,1,2)$, $\vec{v}_3 \equiv (0,0,1) \to A\vec{v}_1 = (-5)\vec{v}_1$, $A\vec{v}_2 = (-5)\vec{v}_2$, $A\vec{v}_3 = 7\vec{v}_3$, so define

$$P \equiv \begin{bmatrix}1 & 1 & 0\\0 & 1 & 0\\1 & 2 & 1\end{bmatrix}, \quad D \equiv \begin{bmatrix}(-5) & 0 & 0\\0 & (-5) & 0\\0 & 0 & 7\end{bmatrix}$$

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

OR we could've chosen

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix}, \text{ same } D$$

OR $$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} (-5) & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & (-5) \end{bmatrix}$$

More than one answer is possible, but in each column, the eigenvalue in D must correspond to the eigenvector in P.

We should've first checked $\text{rank}\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = 3$, to ensure $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for $\mathbb{R}^3$.

(2) A has eigenspaces

$$E_0 = \left\{ \begin{bmatrix} s+t \\ 2s \\ 0 \\ t \end{bmatrix} \middle| \ s,t \text{ real} \right\}$$

$$E_1 = \left\{ \begin{bmatrix} t \\ 3t \\ 0 \\ 0 \end{bmatrix} \middle| \ t \text{ real} \right\} \text{ and}$$

$$E_{(-1)} = \left\{ \begin{bmatrix} 0 \\ 0 \\ t \\ 0 \end{bmatrix} \middle| \ t \text{ real} \right\}$$

SOLUTION: Get bases for the eigenspaces.

p. 702

$$E_0: \begin{bmatrix} s+t \\ 2s \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix};$$

$$\boxed{\text{basis } \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}}$$

$$E_1: \begin{bmatrix} t \\ 3t \\ 0 \\ 0 \end{bmatrix} = t\begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix}; \quad \boxed{\text{basis } \left\{ \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} \right\}}$$

$$E_{(-1)}: \begin{bmatrix} 0 \\ 0 \\ t \\ 0 \end{bmatrix} = t\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; \quad \boxed{\text{basis } \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}}$$

For P, paste those eigenspace bases together:

$$P = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

corresponding eigenvalues

for D: $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & (-1) \end{bmatrix}$,

then $A = PDP^{-1}$

(3) $A \equiv \begin{bmatrix} 0 & 4 \\ -1 & 5 \end{bmatrix}$.

SOLUTION: Since we don't know eigenvectors, we must begin by finding eigenvalues.

STEP ONE: $0 = C_A(t) \equiv \det(A - tI)$

$= \det\begin{bmatrix} -t & 4 \\ -1 & (5-t) \end{bmatrix} = t(t-5) - (-4)$

$$= t^2 - 5t + 4 = (t-1)(t-4)$$

$\rightarrow$ eigenvalues 1, 4.

STEP TWO: eigenspaces

E_1: $\left[(A-I) \mid \vec{0}\right] = \left[\begin{array}{cc|c} -1 & 4 & 0 \\ -1 & 4 & 0 \end{array}\right] \xrightarrow{R_2 - R_1}$

$\left[\begin{array}{cc|c} -1 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow -x_1 + 4x_2 = 0 \rightarrow$

$$E_1 = \left\{ \begin{bmatrix} 4x_2 \\ x_2 \end{bmatrix} \middle| \ x_2 \text{ real} \right\}$$

E_4: $\left[(A-4I) \mid \vec{0}\right] = \left[\begin{array}{cc|c} -4 & 4 & 0 \\ -1 & 1 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2}$

$\left[\begin{array}{cc|c} -1 & 1 & 0 \\ -4 & 4 & 0 \end{array}\right] \xrightarrow{R_2 - 4R_1} \left[\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow$

$-x_1 + x_2 = 0 \rightarrow E_4 = \left\{ \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} \middle| \ x_2 \text{ real} \right\}$

STEP THREE: bases

E_1 : $\begin{bmatrix} 4x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 4 \\ 1 \end{bmatrix}$; basis $\left\{ \begin{bmatrix} 4 \\ 1 \end{bmatrix} \right\}$

E_4 : $\begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; basis $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

STEP FOUR: paste

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

OR

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

$\rightarrow \quad A = PDP^{-1}$

Example 8.27

Let $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 3 \end{bmatrix}$,

from Examples 8.16 (c), page 673.

Solve each of the following.

(1) $\vec{x}_{k+1} = A\vec{x}_k$, $\vec{x}_0 = (1, 0, 0)$.

(2) $\vec{x}_{k+1} = A\vec{x}_k$, $\vec{x}_0 = (0, 1, 0)$

(3) $\vec{x}_{k+1} = A\vec{x}_k$, $\vec{x}_0 = (0, 0, 1)$

(4) $\vec{x}_{k+1} = A\vec{x}_k$, $\vec{x}_0 = (0, 1, -1)$.

SOLUTIONS:

For the global method 8.20 or the local method 8.15 (we used this in Examples 8.16), we need eigenvectors and eigenvalues. In Examples 8.16 (c), we were given eigenvectors

$$(*) \quad \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

and used those to get eigenvalues:

$A\vec{v}_1 = \vec{v}_1, \quad A\vec{v}_2 = 2\vec{v}_2,$

(**) $A\vec{v}_3 = 2\vec{v}_3.$

For this problem (compare to Examples 8.16(c)), we prefer 8.20. 8.15 requires writing $\vec{x}_0$ as a linear combination of $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$; we would have to do this separately for each of the $\vec{x}_0$s in (1), (2), (3), and (4).

For the global method 8.20 we calculate A^n, then apply it to each $\vec{x}_{0}$'s

leaving only matrix multiplication.

Getting A^n begins with

diagonalizing A:

From (*), $P \equiv \begin{bmatrix} -1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$;

from (**),

$$D \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (\text{see } 8.25),$$

then $A = PDP^{-1}$ (diagonalized)

and, by (8.19),

$$(***) \quad A^n = P \begin{bmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 2^n \end{bmatrix} P^{-1} \quad (n = 0, 1, 2, \ldots)$$

For P^{-1}: $[P \mid I] =$

$$\left[\begin{array}{ccc|ccc} -1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to (R_1 \leftrightarrow R_3)$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 \end{array}\right] \to \begin{pmatrix} R_2 + R_1 \\ R_3 + R_1 \end{pmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 & 0 & 1 \end{array}\right] \to (R_3 + R_2)$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 2 \end{array}\right] \to \begin{pmatrix} R_1 - R_3 \\ R_2 - R_3 \end{pmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 & 2 \end{array}\right] \to P^{-1} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Now we must multoply a great deal in (***)

$$A^n = P\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 2^n \end{bmatrix}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} =$$

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & -1 & -1 \\ -(2^n) & 0 & -(2^n) \\ 2^n & 2^n & 2^{n+1} \end{bmatrix} =$$

$$\begin{bmatrix} (1+\cancel{2^n}-\cancel{(2^n)}) & (1-(2^n)) & (1+2^n-(2^{n+1})) \\ (1-(2^n)) & 1 & (1-(2^n)) \\ (-1+2^n) & (-1+2^n) & (-1+2^{n+1}) \end{bmatrix}$$

(i) $\vec{x}_n = A^n\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ (1-(2^n)) \\ (-1+2^n) \end{bmatrix}$

(2) $\vec{x}_n = A^n \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} (1-(2^n)) \\ 1 \\ (-1+2^n) \end{bmatrix}$

(3) $\vec{x}_n = A^n \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} (1+2^n-(2^{n+1})) \\ (1-(2^n)) \\ (-1+2^{n+1}) \end{bmatrix}$

(4) $\vec{x}_n = A^n \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} =$ (answer to (2)) $-$ (answer to (3))

$$= \begin{bmatrix} (1-2^n) - (1+2^n-(2^{n+1})) \\ 1 - (1-(2^n)) \\ (-1+2^n) - (-1+2^{n+1}) \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2^n \\ 2^n-(2^{n+1}) \end{bmatrix}$$

Now we have some bad news: not every square matrix is diagonalizable.

This will be easier to demonstrate after a general statement about similar matrices.

PROPOSITION 8.28

If A and B are similar, then their characteristic polynomials C_A and C_B are equal; in particular, A and B have the same eigenvalues.

Proof: If $A = PBP^{-1}$,

then

$C_A(t) \equiv \det(A - tI) =$

$\det(PBP^{-1} - PtIP^{-1}) =$

$\det(P(B - tI)P^{-1}) =$

$(\det P)(\det(B - tI))(\det(P^{-1}))$

$= \det(B - tI) \equiv C_B(t)$,

since $(\det P)(\det(P^{-1})) = \det(PP^{-1})$

$= \det(I) = 1$.

Equality of {eigenvalues} now follows from Theorem 8.8.

Example 8.29

We claim that $A \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not diagonalizable.

We will prove this claim "by contradiction," meaning we pretend the claim is false, then arrive at a contradiction as a consequence of our pretense.

Suppose, "for the sake of contradiction," that A is diagonalizable. Then $A = PDP^{-1}$, for some invertible P & diagonal D;

$D = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, for

some numbers a, b.

$$C_A(t) = \det \begin{bmatrix} -t & 1 \\ 0 & -t \end{bmatrix} = t^2;$$

by Proposition 8.28,

$$t^2 = C_D(t) = \det \begin{bmatrix} (a-t) & 0 \\ 0 & (b-t) \end{bmatrix}$$

$= (a-t)(b-t)$, thus $a = 0 = b$,

so that

$$A = P \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = P \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} =$$

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, a contradiction of the fact that

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

Thus our pretense is false:
A is not diagonalizable.

In diagonalizing examples we've seen eigenvalues appear more than once in the diagonal matrix similar to the matrix of interest.

There are two natural ways to define this repetition, or <u>multiplicity</u>, informally, "how many times" an eigenvalue appears.

DEFINITIONS 8.30

Suppose λ is an eigenvalue of the square matrix A.

The **geometric multiplicity** of λ is $\dim(E_\lambda)$, the dimension of the eigenspace for λ.

The **algebraic multiplicity** of λ is the number of times the factor $(\lambda - t)$ appears in the factorization of the characteristic polynomial $C_A(t)$.

Example 8.31

Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, as in Example 8.29.

We have already calculated

$$c_A(t) = t^2,$$

thus the algebraic multiplicity of the eigenvalue 0 is 2.

For the geometric multiplicity we need the eigenspace

$E_0 = \mathcal{N}(A)$: solving

$$[A|\vec{0}] = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow x_2 = 0;$$

$$E_0 = \left\{ \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \middle| \ x_1 \text{ is real} \right\},$$

so that

the geometric multiplicity of the eigenvalue 0 is $\dim(E_0) = 1$.

The difference in multiplicities of an eigenvalue for A is equivalent to A being not diagonalizable.

THEOREM 8.32

Suppose A is an $(m \times m)$ matrix. Then the following are equivalent.

p. 722

(a) A is diagonalizable.

(b) For every eigenvalue λ, the geometric multiplicity of λ equals the algebraic multiplicity of λ.

(c) The sum of the dimensions of the eigenspaces equals m.

(d) There is a basis for $\mathbb{R}^m$ consisting entirely of eigenvectors for A.

The intuition is that we need plenty of eigenvectors, enough to construct invertible P in the definition of diagonalizable (Definition 8.23)

Examples 8.33

In each of the following, diagonalize A if possible.

(i) $c_A(t) = (t+1)(t-3)^2$,

$$E_{(-1)} = \left\{ \begin{bmatrix} t \\ 0 \\ -t \end{bmatrix} \middle| \ t \text{ real} \right\},$$

$$E_3 = \left\{ \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} \middle| \ t \text{ real} \right\}$$

SOLUTION:

algebraic multiplicity of eigenvalue 3 equals $2 \neq 1$ = geometric multiplicity of eigenvalue 3, so not diagonalizable by (b) of Theorem 8.32.

OR we could've added $\dim(E_{(-1)}) + \dim(E_3) = 1+1 = 2 \neq 3$, & A is (3×3), so we could invoke (c) of Theorem 8.32.

(2) SAME as (1), except

$$E_3 = \left\{ \begin{bmatrix} s \\ t \\ 2s \end{bmatrix} \middle| \; s, t \text{ real} \right\}$$

SOLUTION: basis for

$$E_{(-1)} : \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}; \text{ for}$$

$$E_3 : \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

$$P \equiv \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 2 & 0 \end{bmatrix}, \quad D \equiv \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\rightarrow \quad A = PDP^{-1}$$

(3) A is 4×4 and has eigenspaces

$$E_2 = \{(0, s+t, 2t+s, t) \mid s, t \text{ real}\}$$

$$E_5 = \{(t, 0, 0, 0) \mid t \text{ real}\}$$

(only eigenvalues are 2 and 5).

SOLUTION: $\dim(E_2) + \dim(E_5) = 2 + 1 = 3 \neq 4$, & A is 4×4,

so $\boxed{\text{not diagonalizable}}$

Here are two special cases where it is easy to determine a matrix is diagonalizable.

THEOREM 8.34

If A is an $(m \times m)$ matrix, then A is diagonalizable if either of the following holds:

(a) A is symmetric; or

(b) A has m distinct eigenvalues.

Examples 8.35

Each of the following matrices is guaranteed by Theorem 8.34 to be diagonalizable.

(1) A is 4×4 and has eigenvalues $-2, 0, \pi, \sqrt{19}$

(2) $A = \begin{bmatrix} 7 & \sqrt{\pi} & 0 \\ \sqrt{\pi} & \sqrt{2} & 9 \\ 0 & 9 & 19 \end{bmatrix}$.

REMARK 8.36

It can be shown that an $(m \times m)$ matrix A is symmmetric if and only if $\mathbb{R}^m$ has an orthogonal basis consisting entirely of eigenvectors for A; compare to (d) of Theorem 8.32.

Example 8.37

For our final example, we would like to solve the fox and rabbit Difference Equation in Examples 1.25(4) and 7.22(b).

More than this, we would like to generalize Examples 1.25(4) by having arbitrary specified rates of reproduction and (for the foxes) consumption.

As in Examples 1.25(4), for $k = 0, 1, 2, \ldots$

$r_k \equiv$ number of rabbits k years from now

$f_k \equiv$ number of foxes k years from now.

p. 731

In the following recursive model for interacting (a term probably made up by the foxes; the rabbits might have a less tolerant description) foxes and rabbits, the number F represents fox fertility, R rabbit fertility, and E (for "Eating") is the number of rabbits eaten by each fox in a year.

p. 732

$$(*)\begin{cases} r_{k+1} = R r_k - E f_k \\ f_{k+1} = F f_k \end{cases} \quad (k = 0, 1, 2, \ldots)$$

If $A \equiv \begin{bmatrix} R & -E \\ 0 & F \end{bmatrix}$, then

$$\begin{bmatrix} r_{k+1} \\ f_{k+1} \end{bmatrix} = A \begin{bmatrix} r_k \\ f_k \end{bmatrix} \quad (k = 0, 1, 2, \ldots)$$

thus

$$(**) \begin{bmatrix} r_n \\ f_n \end{bmatrix} = A^n \begin{bmatrix} r_0 \\ f_0 \end{bmatrix} \quad (n = 0, 1, 2, \ldots)$$

By counting the number of foxes and rabbits now, we can state the number of foxes and rabbits at any time in the future.

Note that Examples 1.25(4) is the special case $F = 4$, $R = 100$, $E = 360$.

CASE I: $F \neq R$.

We will leave it to the reader to calculate that the eigenvalues for A are F and R, with eigenspaces

$$E_R = \text{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\},$$

$$E_F = \text{span}\left\{\begin{bmatrix}1\\\epsilon\end{bmatrix}\right\}, \text{ where}$$

$$\epsilon \equiv \left(\frac{R-F}{E}\right).$$

Use 8.25 to diagonalize A:

$$P \equiv \begin{bmatrix}1 & 1\\0 & \epsilon\end{bmatrix}, \quad D \equiv \begin{bmatrix}R & 0\\0 & F\end{bmatrix}$$

$$\rightarrow \quad A = PDP^{-1}.$$

p. 735

We also leave it to the reader to calculate

$$P^{-1} = \frac{1}{\epsilon}\begin{bmatrix}\epsilon & -1\\ 0 & 1\end{bmatrix}.$$

Now we may get powers of A. For $n = 0, 1, 2, \ldots$

$$A^n = P\begin{bmatrix}R^n & 0\\ 0 & F^n\end{bmatrix}P^{-1} =$$

$$\frac{1}{\epsilon}\begin{bmatrix}1 & 1\\ 0 & \epsilon\end{bmatrix}\begin{bmatrix}\epsilon(R^n) & -(R^n)\\ 0 & F^n\end{bmatrix} =$$

$$\frac{1}{\epsilon}\begin{bmatrix}\epsilon(R^n) & (-(R^n)+F^n)\\ 0 & \epsilon(F^n)\end{bmatrix} =$$

$$\begin{bmatrix} R^n & \left(\frac{E}{R-F}\right)\left(F^n - (R^n)\right) \\ 0 & F^n \end{bmatrix}.$$

By (**),

$$\begin{bmatrix} r_n \\ f_n \end{bmatrix} = \begin{bmatrix} R^n r_0 + \left(\frac{E}{R-F}\right)\left(F^n - (R^n)\right) f_0 \\ F^n f_0 \end{bmatrix}$$

is the solution of (*), at least until extinction.

WHAT ABOUT Extinction?

Extinction occurs by the n^{th} year from now if and only if

$$r_n = \left[R^n r_0 + \left(\frac{E}{R-F}\right)\left(F^n - (R^n)\right) f_0\right]$$

≤ 0 if and only if $\left(\begin{array}{c}\text{after some}\\ \text{algebra}\end{array}\right)$

$$\frac{r_0}{f_0} \leq \left(\frac{E}{R-F}\right)\left(1 - \left(\frac{F}{R}\right)^n\right).$$

It is interesting that extinction depends entirely on the initial ratio of rabbits to foxes.

NOTE that, if
$F > R$ (foxes more fertile than rabbits), then extinction is inevitable, since then
$\left|\left(\frac{E}{R-F}\right)\left(1-\left(\frac{F}{R}\right)^n\right)\right|$ gets arbitrarily large, as n gets large.

If $F < R$, then $\left(\frac{E}{R-F}\right)\left(1-\left(\frac{F}{R}\right)^n\right)$ is increasing to
$\left(\frac{E}{R-F}\right)$ as n increases to ∞,

thus

Extinction occurs eventually

if and only if

$$\frac{r_0}{f_0} < \left(\frac{E}{R-F}\right).$$

So long as rabbits are more fertile than foxes, extinction can be avoided by starting out with enough rabbits (relative to the number of foxes).

CASE II: $F = R$.

We will leave it to the reader to show that

$$A = \begin{bmatrix} F & -E \\ 0 & F \end{bmatrix}$$

is not diagonalizable.

But a pattern for powers of A does emerge:

$$A^2 = \begin{bmatrix} F^2 & -2EF \\ 0 & F^2 \end{bmatrix}, \quad A^3 = \begin{bmatrix} F^3 & -3EF^2 \\ 0 & F^3 \end{bmatrix},$$

...

the pattern turns out to be

$$A^n = \begin{bmatrix} F^n & -nEF^{n-1} \\ 0 & F^n \end{bmatrix} =$$

$F^{n-1}\begin{bmatrix} F & -nE \\ 0 & F \end{bmatrix}$, so that, by (**),

$$\begin{bmatrix} r_n \\ f_n \end{bmatrix} = F^{n-1}\begin{bmatrix} F & -nE \\ 0 & F \end{bmatrix}\begin{bmatrix} r_0 \\ f_0 \end{bmatrix}$$

$$= F^{n-1}\begin{bmatrix} Fr_0 - nEf_0 \\ Ff_0 \end{bmatrix}$$

This implies that $r_n \le 0$ (extinction by n^{th} year) if and only if

$$\frac{r_0}{f_0} \le \frac{nE}{F},$$

which is guaranteed to happen eventually, for any r_0, f_0.

In CASE II, eventual extinction is guaranteed. It is not sufficient to have equal fertility of rabbits and foxes, if one wishes to avoid extinction; we must have rabbit fertolity (strictly) greater than fox fertolity.

Specific Examples 8.3B

(1) In Examples 1.25(4)

((*) of Example 8.37 with

$F = 4$, $R = 100$, $E = 360$),

$$\left(\frac{E}{R-F}\right) = 3.75,$$ thus

extinction occurs eventually

if and only if

$$\left(\frac{r_0}{f_0}\right) < 3.75 \; ;$$

if there are initially 100 foxes, you need 375 or more rabbits initially to avoid extinction.

(2) $r_{k+1} = 2r_k - f_k$

$f_{k+1} = 3f_k$

is guaranteed for eventual extinction, regardless of r_0 & f_0, since $F = 3 \geq 2 = R$ in ($*$) of Example 8.37.

(3) $r_{k+1} = 6r_k - 2f_k$

$f_{k+1} = 4f_k$

((*) with $R = 6$, $E = 2$ & $F = 4$)

will avoid extinction if and only if $\frac{r_0}{f_0} \geq 1$;

that is, the number of rabbits is not less than the number of foxes, since

$$\left(\frac{E}{R-F}\right) = 1$$

p. 746

(4) $r_{k+1} = 6r_k - 200f_k$

$f_{k+1} = 4f_k$

((*) with $R = 6$, $E = 200$
d $F = 4$)

will avoid extinction if
and only if $\frac{r_0}{f_0} \geq 100$,

since $\left(\frac{E}{R-F}\right) = 100$.

(4) has the same fertilities as (3), but hungrier foxes, hence we need more rabbits initially.