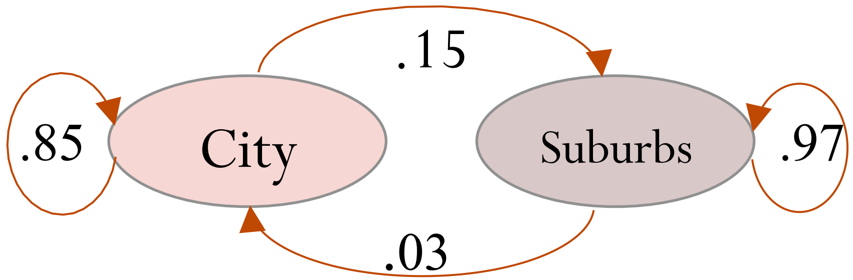
Section 5.3 Diagonalization of Matrices

Definition

An $n \times n$ matrix A is called **diagonalizable** if $A = PDP^{-1}$ for some diagonal $n \times n$ matrix D and some invertible $n \times n$ matrix P.

Example: stochastic matrix



$$\begin{array}{ll} & \text{From} \\ & \begin{array}{cc}\text{City} & \text{Suburbs}\end{array} \\ \text{To} \quad \begin{array}{r}\text{City} \\ \text{Suburbs}\end{array} & \begin{bmatrix} .85 & .03 \\ .15 & .97 \end{bmatrix} = A \end{array}$$

probability matrix of a sample person's residence movement

$$\mathbf{p} = \begin{bmatrix} 500 \\ 700 \end{bmatrix}$$: current population of the city and suburbs

$$A\mathbf{p} = \begin{bmatrix} .85 & .03 \\ .15 & .97 \end{bmatrix} \begin{bmatrix} 500 \\ 700 \end{bmatrix}$$: population distribution in the next year

$A^m\mathbf{p}$: population distribution after m years

$$A = PDP^{-1} \text{ where } P = \begin{bmatrix} -1 & 1 \\ 1 & 5 \end{bmatrix}, D = \begin{bmatrix} .82 & 0 \\ 0 & 1 \end{bmatrix}$$

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

$$\begin{aligned} A^m &= PD^mP^{-1} \\ &= \begin{bmatrix} -1 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} (.82)^m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 5 \end{bmatrix}^{-1} \\ &= \frac{1}{6} \begin{bmatrix} 1+5(.82)^m & 1-(.82)^m \\ 5-5(.82)^m & 5+(.82)^m \end{bmatrix} \end{aligned}$$

Question

Is any $n \times n$ matrix A diagonalizable?

Not all matrices are diagonalizable.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow A^2 = 0.$$

If $A = PDP^{-1}$ for some invertible P and diagonal D, then
$A^2 = PD^2P^{-1} = 0$.
$\Rightarrow\ D^2 = 0\ \Rightarrow\ D = 0$ since D is diagonal $\Rightarrow\ A = 0$. ⌦

Question

What kind of $n \times n$ matrix A is diagonalizable?

Let's observe some examples before answering this question.

	Characteristic polynomial	Eigenvalues	Eigenspaces
$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$			
$B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$			
$C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$			
$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$			

	Characteristic polynomial	Eigenvalues	Eigenspaces
$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$			
$B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$			
$C = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$			

Theorem 5.2

(a) An $n \times n$ matrix A is **diagonalizable** if and only if there is a basis for $\mathcal{R}^n$ consisting of eigenvectors of A.

(b) If P is an invertible $n \times n$ matrix and D is a diagonal $n \times n$ matrix, then $A = PDP^{-1}$ if and only if the columns of P are a basis for $\mathcal{R}^n$ consisting of eigenvectors of A and the diagonal entries of D are the eigenvalues corresponding to the respective columns of P.

Proof $A \in \mathscr{R}^{n\times n}$ is diagonalizable.

$\Leftrightarrow$ $\exists$ an invertible $P = [\ \mathbf{p}_1\ \ \mathbf{p}_2\ \ \cdots\ \ \mathbf{p}_n\ ] \in \mathscr{R}^{n\times n}$ and a diagonal $D = \text{diag}[d_1\ d_2\ \cdots\ d_n]. \in \mathscr{R}^{n\times n}$ such that $A = PDP^{-1}$.

$\Leftrightarrow$ $AP = PD$ and $P = [\ \mathbf{p}_1\ \ \mathbf{p}_2\ \ \cdots\ \ \mathbf{p}_n\ ]$ is invertible.

$\Leftrightarrow$ $[\ A\mathbf{p}_1\ \ A\mathbf{p}_2\ \ \cdots\ \ A\mathbf{p}_n\ ] = P\text{diag}[d_1\ d_2\ \cdots\ d_n]$.

$$
\begin{aligned}
&= P\begin{bmatrix} d_1\mathbf{e}_1 & d_2\mathbf{e}_2 & \cdots & d_n\mathbf{e}_n \end{bmatrix} \\
&= \begin{bmatrix} P(d_1\mathbf{e}_1) & P(d_2\mathbf{e}_2) & \cdots & P(d_n\mathbf{e}_n) \end{bmatrix} \\
&= \begin{bmatrix} d_1(P\mathbf{e}_1) & d_2(P\mathbf{e}_2) & \cdots & d_n(P\mathbf{e}_n) \end{bmatrix} \\
&= \begin{bmatrix} d_1\mathbf{p}_1 & d_2\mathbf{p}_2 & \cdots & d_n\mathbf{p}_n \end{bmatrix}.
\end{aligned}
$$

and $P = [\ \mathbf{p}_1\ \ \mathbf{p}_2\ \ \cdots\ \ \mathbf{p}_n\ ]$ is invertible.

$\Leftrightarrow$ $A\mathbf{p}_i = d_i\mathbf{p}_i$ for $i = 1, 2, \ldots, n$ and $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ is L.I..

$\Leftrightarrow$ $\mathbf{p}_i$ is an eigenvector of A corresponding to the eigenvalue d_i and $\mathscr{B} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ is a basis for $\mathscr{R}^n$.

Steps to diagonalize a given $A \in \mathcal{R}^{n\times n}$:

$$A = \begin{bmatrix} .85 & .03 \\ .15 & .97 \end{bmatrix}$$

1. Find n eigenvalues (repeated or not) for A and form a diagonal matrix D with eigenvalues on the diagonal;

$$\det\ (A - tI_2) = \det \begin{bmatrix} .85 - t & .03 \\ .15 & .97 - t \end{bmatrix} = (t - .82)(t - 1)$$

$\Rightarrow$ eigenvalues of $A = \{\ 0.82,\ 1\ \}$, and $D = \begin{bmatrix} .82 & 0 \\ 0 & 1 \end{bmatrix}$

2. Find n L.I. eigenvectors corresponding to these eigenvalues, if possible, and form an invertible $P \in \mathcal{R}^{n\times n}$;

$$A - .82I_2 \xrightarrow[\text{echelon form}]{\text{reduced row}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{p}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$A - I_2 \xrightarrow[\text{echelon form}]{\text{reduced row}} \begin{bmatrix} 1 & -.2 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{p}_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$P = \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 5 \end{bmatrix}}_{\text{invertible}}$$

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

Every $A \in \mathcal{R}^{n \times n}$ has n eigenvalues (counting repeated ones) if complex eigenvalues are allowed. However, some matrices may not have n L.I. eigenvectors even if complex eigenvectors are allowed.

Theorem 5.3

A set of eigenvectors of a square matrix that correspond to distinct eigenvalues is linearly independent.

Theorem 5.3

A set of eigenvectors of a square matrix that correspond to distinct eigenvalues is linearly independent.

Proof Let A be an $n \times n$ matrix with eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ having corresponding distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. Suppose the set of eigenvectors are L.D.. As eigenvectors are nonzero, **Theorem 1.9** implies $\mathbf{v}_k = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{k-1}\mathbf{v}_{k-1}$ for some $k \in [2, m]$ and scalars $c_1, c_2, \ldots, c_{k-1}$.

$\Rightarrow A\mathbf{v}_k = c_1 A\mathbf{v}_1 + c_2 A\mathbf{v}_2 + \cdots + c_{k-1} A\mathbf{v}_{k-1}$

$(-\lambda_k)$

$\Rightarrow \lambda_k \mathbf{v}_k = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_{k-1}\lambda_{k-1}\mathbf{v}_{k-1}$

$\Rightarrow \mathbf{0} = c_1(\lambda_1 - \lambda_k)\,\mathbf{v}_1 + c_2(\lambda_2 - \lambda_k)\mathbf{v}_2 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1}$

$\Rightarrow c_1 = c_2 = \cdots = c_{k-1} = 0$

$\Rightarrow \mathbf{v}_k = \mathbf{0}$, a contradiction.

Corollary 1: Let $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_p$ be subsets of p eigenspaces of a square matrix corresponding to p distinct eigenvalues. If $\boldsymbol{S}_i$ is L.I. for all $i = 1, 2, \ldots, p$, then the set $\boldsymbol{S}_1 \cup \boldsymbol{S}_2 \cup \cdots \cup \boldsymbol{S}_p$ is L.I..

Corollary 2: If $A \in \mathcal{R}^{n\times n}$ has n distinct eigenvalues, then $\mathcal{R}^n$ has a basis consisting of eigenvectors of A, i.e., A is diagonalizable.

Definition

For $A \in \mathcal{R}^{n \times n}$, the characteristic polynomial of A may be factored into a **product of linear factors** if

$$\det(A - tI_n) = (-1)^n (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n).$$

Here, $\lambda_i, i = 1, 2, \ldots, n$ do not have to be distinct, but $\lambda_i \in \mathcal{R}$.

Test for a Diagonalizable Matrix

An n x n matrix A is diagonalizable if and only if both the following conditions are met.

1) The characteristic polynomial of A factors into a product of linear factors.
2) For each eigenvalue λ of A, the multiplicity of λ equals the dimension of the corresponding eigenspace (n – rank(A-λI_n)).

Proof "if" Follow the previous diagonalization stepts.

Condition (1) ⇒ there are n eigenvalues for Step 1.

Condition (2) and **Theorem 5.3** ⇒ there are n L.I. eigenvectors in Step 2.

"only if" If Condition (1) fails, then A has less than n eigenvalues (counting repeated ones), and (sum of all geometric multiplicities) ≤ (sum of all algebraic multiplicities) < n, which means that there are no enough L.I. eigenvectors to form a basis.

If Condition (2) fails, then there are no enough L.I. eigenvectors to form a basis.

Note: Condition (1) always holds if complex eigenvalues are allowed.

Example: The characteristic polynomial of

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

is $-(t+1)^2(t-3) \Rightarrow$ eigenvalues: 3, -1, -1 $\Rightarrow$

the eigenspaces corresponding to the eigenvalue 3 and -1 have bases

$$\mathbb{B}_1 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ and } \mathbb{B}_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}, \text{ respectively}$$

$\Rightarrow A = PDP^{-1}$, where

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Example: The characteristic polynomials of

$$A = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} -7 & -3 & -6 \\ 0 & -4 & 0 \\ 3 & 3 & 2 \end{bmatrix}$$

$$C = \begin{bmatrix} -6 & -3 & 1 \\ 5 & 2 & -1 \\ 2 & 3 & -5 \end{bmatrix} \qquad M = \begin{bmatrix} -3 & 2 & 1 \\ 3 & -4 & -3 \\ -8 & 8 & 6 \end{bmatrix}$$

are $-(t+1)(t^2+4)$, $-(t+1)(t+4)^2$, $-(t+1)(t+4)^2$, $-(t+1)(t^2-4)$, respectively

Example: The characteristic polynomials of

$$A = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} -7 & -3 & -6 \\ 0 & -4 & 0 \\ 3 & 3 & 2 \end{bmatrix}$$

$$C = \begin{bmatrix} -6 & -3 & 1 \\ 5 & 2 & -1 \\ 2 & 3 & -5 \end{bmatrix} \qquad M = \begin{bmatrix} -3 & 2 & 1 \\ 3 & -4 & -3 \\ -8 & 8 & 6 \end{bmatrix}$$

are $-(t+1)(t^2+4)$, $-(t+1)(t+4)^2$, $-(t+1)(t+4)^2$, $-(t+1)(t^2-4)$, respectively

$\Rightarrow$ A can not be diagonalized as it has only one (real) eigenvalue
B can be diagonalized as nullity of $B - (-4)I_3$ is 2
C can be not diagonalized as nullity of $C - (-4)I_3$ is 1
M can be diagonalized as M has three distinct eigenvalues -1, -2, 2

Homework Set for Section 5.3

- Section 5.3: Problems 1, 3, 5, 9, 13, 17, 29, 31, 33, 35, 41, 43, 47