Eigenvalues

정보컴퓨터공학부 김민환 교수

Contents

1. Eigenvalues and Eigenvectors
2. Diagonalization
3. Hermitian Matrices
4. Singular Value Decomposition
5. Quadratic Forms
6. Systems of Linear Differential Equations
7. Positive Definite Matrices

Introduction

In a small town, there are 8,000 married women and 2,000 single women and the total population remains constant. 30 percent of the married women get divorced each year and 20 % of the single women get married each year

- Let us investigate the long-range prospects if these percentages of marriage and divorces continue indefinitely into the future

$$\mathbf{w}_0 = (8000, 2000)$$

$$\mathbf{w}_1 = \begin{bmatrix} 0.7 \times 8000 + 0.2 \times 2000 \\ 0.3 \times 8000 + 0.8 \times 2000 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 8000 \\ 2000 \end{bmatrix} = \mathbf{A}\mathbf{w}_0 = \begin{bmatrix} 6000 \\ 4000 \end{bmatrix}$$

$$\mathbf{w}_2 = \mathbf{A}\mathbf{w}_1 = \mathbf{A}\mathbf{A}\mathbf{w}_0 = \mathbf{A}^2\,\mathbf{w}_0$$

$$\vdots$$

$$\mathbf{w}_n = \mathbf{A}^n\,\mathbf{w}_0$$

$$\mathbf{w}_{10} = \begin{bmatrix} 4004 \\ 5996 \end{bmatrix}, \mathbf{w}_{12} = \begin{bmatrix} 4000 \\ 6000 \end{bmatrix}, \mathbf{w}_{13} = \begin{bmatrix} 4000 \\ 6000 \end{bmatrix}, \mathbf{w}_{30} = \begin{bmatrix} 4000 \\ 6000 \end{bmatrix}$$

- What happen in case of $\mathbf{w}_0 = (10000, 0)$?

  It turns out that $\mathbf{w}_{14} = (4000, 6000)$
- Why does this process converge, and why do we seem to get the same *steady-state* vector even when we change the initial vector ?
- Let us choose a *basis* for $\mathbf{R}^2$ consisting of vectors for which the effect of the linear operator $\mathbf{A}$ is easily determined. A multiple of the steady-state vector, say, $\mathbf{x}_1 = (2,3)$, is a good choice as a basis vector

$$\mathbf{A}\mathbf{x}_1 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \mathbf{x}_1$$

We can choose another basis vector $\mathbf{x}_2 = (1,-1)$ since the effect of $\mathbf{A}$ on $\mathbf{x}_2$ is also very simple

$$\mathbf{A}\mathbf{x}_2 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \mathbf{x}_2$$

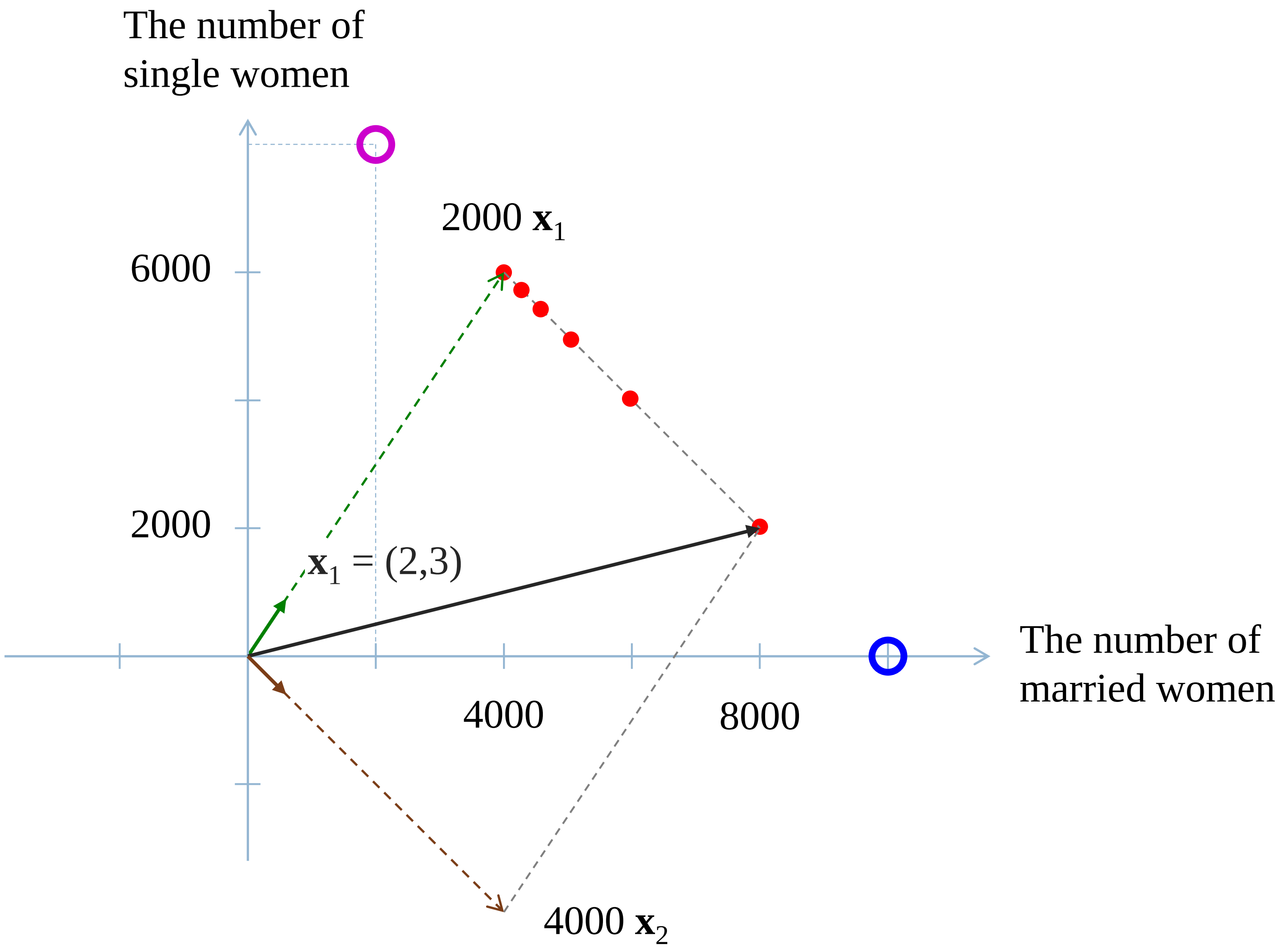
Then the initial vector $\mathbf{w}_0 = (8000, 2000)$ can be uniquely represented as $\mathbf{w}_0 = 2000\mathbf{x}_1 + 4000\mathbf{x}_2$

$$\mathbf{w}_1 = \mathbf{A}\mathbf{w}_0 = 2000\mathbf{A}\mathbf{x}_1 + 4000\mathbf{A}\mathbf{x}_2 = 2000\mathbf{x}_1 + 4000(1/2)\mathbf{x}_2$$

$$\mathbf{w}_2 = \mathbf{A}\mathbf{w}_1 = 2000\mathbf{x}_1 + 4000(1/2)^2\mathbf{x}_2$$

$$\mathbf{w}_n = \mathbf{A}^n\mathbf{w}_0 = 2000\mathbf{x}_1 + 4000(1/2)^n\mathbf{x}_2$$

- In case of $\mathbf{w}_0 = (10000, 0)$, $\mathbf{w}_0 = 2000\mathbf{x}_1 + 6000\mathbf{x}_2$


The number of single women
2000 $\mathbf{x}_1$
6000
2000
$\mathbf{x}_1$ = (2,3)
The number of married women
4000
8000
4000 $\mathbf{x}_2$

Eigenvalues & Eigenvectors

To know the definitions and how to compute eigenvalues and eigenvectors

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Def. Eigenvalue and Eigenvector

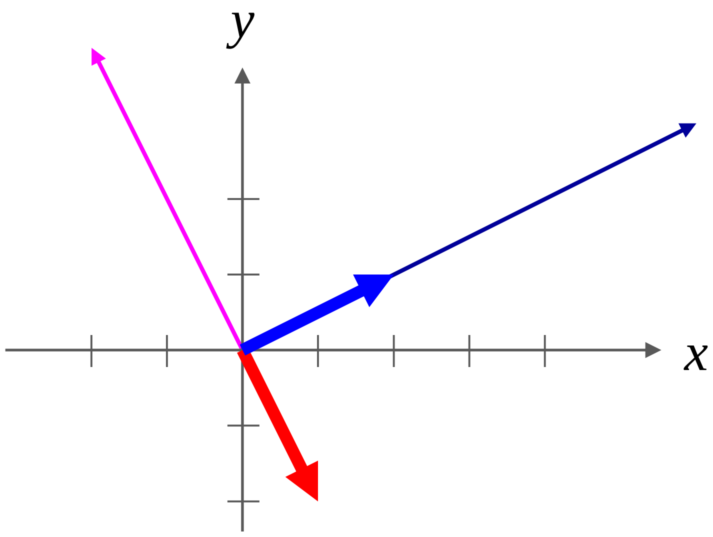
Let $\mathbf{A}$ be an $n \times n$ matrix. A scalar λ is said to be an *eigenvalue* or a *characteristic value* of $\mathbf{A}$ if there exists a *nonzero* vector $\mathbf{x}$ such that

$$\mathbf{Ax} = \lambda\mathbf{x}$$

The vector $\mathbf{x}$ is said to be an *eigenvector* or a *characteristic vector* belonging to λ. (*Eigen* in German means *proper*.)

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

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



- Note
  - Any nonzero *scalar multiple* of an eigenvector $\mathbf{v}$ is also an eigenvector

$$\mathbf{Aw} = \mathbf{A}(r\mathbf{v}) = r\mathbf{Av} = r(\lambda\mathbf{v}) = \lambda(r\mathbf{v}) = \lambda\mathbf{w}$$

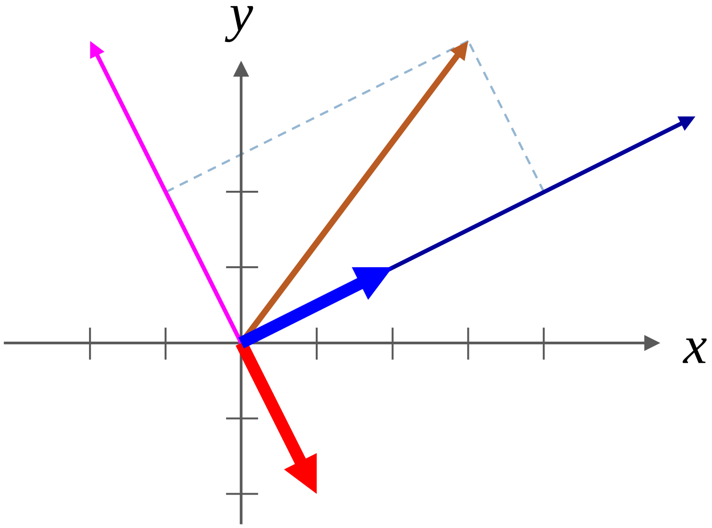
$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}\left(2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ -2 \end{bmatrix} \right) = 2(3)\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-1)(-2)\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}^k \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2(3)^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-1)(-2)^k \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4(3)^k + (-1)(-2)^k \\ 2(3)^k + 2(-2)^k \end{bmatrix}$$

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

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



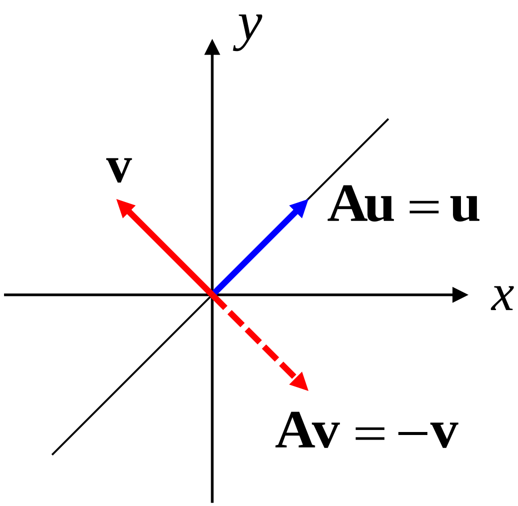
(*Ex*.1)

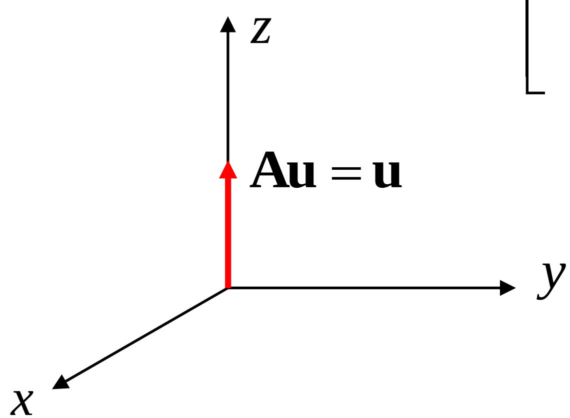
Find eigenvalues and eigenvectors of $\mathbf{A}$ geometrically

(a) $\mathbf{A}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ (reflection about line $y=x$) $\mathbf{A}(x,y)=(y,x)$

(b) $\mathbf{A}$ is a standard matrix of rotation by θ about z-axis in a counterclockwise direction

$$\mathbf{A}=\begin{bmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}$$





Th. 6.1.1

Let $\mathbf{A}$ be a square matrix

A scalar λ is an eigenvalue of $\mathbf{A}$ if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

A vector $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue λ if and only if $\mathbf{v}$ is a nontrivial solution of the system

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

$$\begin{aligned}\mathbf{Av} = \lambda\mathbf{v} &\Leftrightarrow \mathbf{Av} = \lambda\mathbf{Iv}\\ &\Leftrightarrow \mathbf{Av} - \lambda\mathbf{Iv} = \mathbf{0}\\ &\Leftrightarrow (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}\end{aligned}$$

The *nonzero* vector $\mathbf{v}$ is an eigenvector if an only if it is a *nontrivial solution* of the homogeneous linear system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$

Thus $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is, $\mathbf{A} - \lambda\mathbf{I}$ should be singular

Notice that $\mathbf{v}$ is a nonzero vector of the null space of $\mathbf{A} - \lambda\mathbf{I}$

- Terminology

  Characteristic equation : $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$

  Characteristic polynomial : $\det(\mathbf{A} - \lambda\mathbf{I})$ with degree n

  Characteristic Matrix : $\mathbf{A} - \lambda\mathbf{I}$

  Eigenspace of $\mathbf{A}$ corresponding to the eigenvalue λ :

  - Null space of $(\mathbf{A} - \lambda\mathbf{I})$, E_λ
  - E_λ is a subspace of $\mathbf{R}^n$
  - $E_\lambda = \mathrm{Span}($eigenvectors corresponding to $\lambda)$

  Algebraic Multiplicity : exponent of $(\lambda - a)$ in $\det(\mathbf{A} - \lambda\mathbf{I})$

  Geometric Multiplicity : dimension of E_a

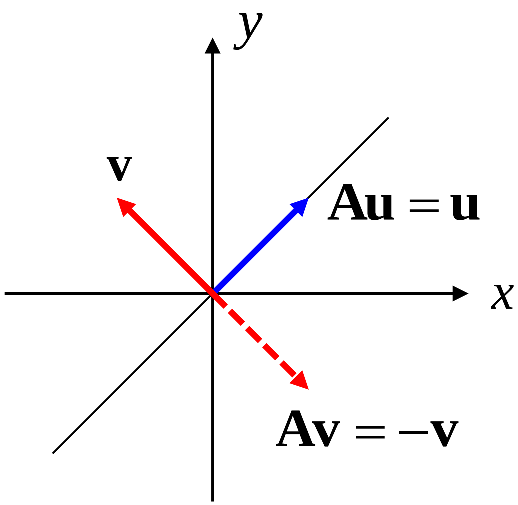
Algorithm: Computation of Eigenvalues, Eigenvectors, and Bases of Eigenspaces

INPUT : An $n \times n$ matrix $\mathbf{A}$

1. Find the eigenvalues of $\mathbf{A}$ by solving $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$
2. For each eigenvalue λ_i, solve the homogeneous linear system $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v} = \mathbf{0}$. The nontrivial solutions are the eigenvectors of $\mathbf{A}$
3. Write the general solution of $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v} = \mathbf{0}$ in Step 2 as a linear combination of vectors with coefficients the free variables. These vectors form a basis for the eigenspace $E_{\lambda i}$

(*Ex.2*)

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



$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{v}=\mathbf{0}$$

$$\det(\mathbf{A}-\lambda\mathbf{I})=\det\left(\begin{bmatrix}-\lambda & 1\\ 1 & -\lambda\end{bmatrix}\right)=\lambda^2-1=0 \quad \therefore \lambda=\pm 1$$

$$\lambda=1;\quad \begin{bmatrix}-1 & 1 & : & 0\\ 1 & -1 & : & 0\end{bmatrix}\sim\begin{bmatrix}1 & -1 & : & 0\\ 0 & 0 & : & 0\end{bmatrix} \quad \therefore E_1=\left\{r\begin{bmatrix}1\\ 1\end{bmatrix}, r\in\mathbf{R}\right\} \quad \therefore \mathbf{u}=\begin{bmatrix}1\\ 1\end{bmatrix}$$

$$\lambda=-1;\quad \begin{bmatrix}1 & 1 & : & 0\\ 1 & 1 & : & 0\end{bmatrix}\sim\begin{bmatrix}1 & 1 & : & 0\\ 0 & 0 & : & 0\end{bmatrix} \quad \therefore E_{-1}=\left\{r\begin{bmatrix}-1\\ 1\end{bmatrix}, r\in\mathbf{R}\right\} \quad \therefore \mathbf{v}=\begin{bmatrix}-1\\ 1\end{bmatrix}$$

(*Ex*.3) $(\mathbf{A}-\lambda\mathbf{I})\mathbf{v}=\mathbf{0}$

$$\mathbf{A}=\begin{bmatrix}0&0&1\\0&1&0\\0&0&1\end{bmatrix}\quad \det(\mathbf{A}-\lambda\mathbf{I})=\begin{vmatrix}-\lambda&0&1\\0&1-\lambda&0\\0&0&1-\lambda\end{vmatrix}=-\lambda(1-\lambda)^2=0$$

	Algebraic Multiplicity	Geometric Multiplicity
$\lambda=0$	1	1
$\lambda=1$	2	2

$$\text{For }\lambda=0;\quad \begin{bmatrix}0&0&1&:&0\\0&1&0&:&0\\0&0&1&:&0\end{bmatrix}\sim\begin{bmatrix}0&1&0&:&0\\0&0&1&:&0\\0&0&0&:&0\end{bmatrix}$$

$$E_0=\left\{\begin{bmatrix}r\\0\\0\end{bmatrix},r\in\mathbf{R}\right\}=\mathrm{Span}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right),\quad \mathbf{v}_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \{\mathbf{v}_1\}\ \text{is a basis for } E_0$$

$$\text{For }\lambda=1;\quad \begin{bmatrix}-1&0&1&:&0\\0&0&0&:&0\\0&0&0&:&0\end{bmatrix},\quad E_1=\left\{\begin{bmatrix}s\\r\\s\end{bmatrix},r,s\in\mathbf{R}\right\}=\mathrm{Span}\left(\mathbf{v}_2=\begin{bmatrix}0\\1\\0\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}1\\0\\1\end{bmatrix}\right)$$

(*Ex.4*) $\mathbf{A}=\begin{bmatrix}1&0&3\\1&-1&2\\-1&1&-2\end{bmatrix}$

	Algebraic Multiplicity	Geometric Multiplicity
λ = 0	2	1
λ = -2	1	1

$$\det(\mathbf{A}-\lambda\mathbf{I})=\begin{vmatrix}1-\lambda&0&3\\1&-1-\lambda&2\\-1&1&-2-\lambda\end{vmatrix}=(1-\lambda)\begin{vmatrix}-1-\lambda&2\\1&-2-\lambda\end{vmatrix}+3\begin{vmatrix}1&-1-\lambda\\-1&1\end{vmatrix}$$

$$=(1-\lambda)\{(1+\lambda)(2+\lambda)-2\}+3\{1-(1+\lambda)\}=-\lambda^3-2\lambda^2=-\lambda^2(\lambda+2)=0$$

$$\therefore\ \lambda=0,\quad \lambda=-2$$

For $\lambda=0$, $\begin{bmatrix}1&0&3&:&0\\1&-1&2&:&0\\-1&1&-2&:&0\end{bmatrix}\sim\begin{bmatrix}1&0&3&:&0\\0&1&1&:&0\\0&0&0&:&0\end{bmatrix}$ $\therefore E_0=\mathrm{Span}\left(\begin{bmatrix}-3\\-1\\1\end{bmatrix}\right)$

For $\lambda=-2$, $\begin{bmatrix}3&0&3&:&0\\1&1&2&:&0\\-1&1&0&:&0\end{bmatrix}\sim\begin{bmatrix}1&0&1&:&0\\0&1&1&:&0\\0&0&0&:&0\end{bmatrix}$ $\therefore E_{-2}=\mathrm{Span}\left(\begin{bmatrix}-1\\-1\\1\end{bmatrix}\right)$

Complex Eigenvalues

- Let us expand our field of scalars to the complex numbers and let us allow complex entries for our vectors and matrices
- An $n \times n$ *real* matrix $\mathbf{A}$ will have n eigenvalues, some of which may be repeated and some of which may be complex numbers.

  If $\lambda = a + bi \ (b \neq 0)$ is an eigenvalue, then $\overline{\lambda} = a - bi$ must also be an eigenvalue of $\mathbf{A}$. If $\mathbf{z}$ is an eigenvector belonging to λ, then $\mathbf{A}\overline{\mathbf{z}} = \overline{\lambda}\overline{\mathbf{z}}$

(*Ex*.5) $\mathbf{A}=\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$

$$\begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2+4=0 \quad \therefore \lambda_1 = 1+2i,\ \lambda_2 = 1-2i$$

$$\lambda = 1+2i; \quad \begin{bmatrix} -2i & 2 \\ -2 & -2i \end{bmatrix} \sim \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix} \quad E_{\lambda_1} = \left\{ r\begin{bmatrix} -i \\ 1 \end{bmatrix}, r \in \mathbf{C} \right\} \quad \mathbf{v} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$\overline{\lambda} = 1-2i; \quad \begin{bmatrix} 2i & 2 \\ -2 & 2i \end{bmatrix} \sim \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \quad E_{\lambda_2} = \left\{ r\begin{bmatrix} i \\ 1 \end{bmatrix}, r \in \mathbf{C} \right\} \quad \overline{\mathbf{v}} = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} -2i & 2 \\ -2 & -2i \end{bmatrix} \sim \begin{bmatrix} -2 & -2i \\ -2i & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

Product and Sum of Eigenvalues

For an $n \times n$ matrix $\mathbf{A}$, the characteristic polynomial of $\mathbf{A}$, $p(\lambda)$, has a form

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \Lambda & a_{1n} \\ a_{21} & a_{22} - \lambda & \Lambda & a_{2n} \\ : & : & \mathrm{O} & : \\ a_{n1} & a_{n2} & \Lambda & a_{nn} - \lambda \end{vmatrix}$$

Expanding along the first column, we get

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \underline{(a_{11} - \lambda)\det(\mathbf{M}_{11})} + \sum_{i=2}^{n} a_{i1}(-1)^{i+1}\det(\mathbf{M}_{i1})$$

where the minor $\mathbf{M}_{i1}$ do not contain the two diagonal elements $(a_{11} - \lambda)$ and $(a_{ii} - \lambda)$. Expanding $\det(\mathbf{M}_{11})$ in the same manner, we conclude that

$$(a_{11} - \lambda)(a_{22} - \lambda)\Lambda\ (a_{nn} - \lambda)$$

is the only term involving a product of more than $n - 2$ of the diagonal elements

From the above term, the lead coefficient of $p(\lambda)$ is $(-1)^n$

If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$, then

$$\begin{aligned} p(\lambda) &= \det(\mathbf{A} - \lambda\mathbf{I}) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2)\Lambda\ (\lambda - \lambda_n) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\Lambda\ (\lambda_n - \lambda) \end{aligned}$$

Thus $\lambda_1 \cdot \lambda_2 \cdots \lambda_n = p(0) = \det(\mathbf{A} - 0\mathbf{I}) = \det(\mathbf{A})$

From the term $(a_{11}-\lambda)(a_{22}-\lambda)\Lambda\ (a_{nn}-\lambda)$, the coefficient of $(-\lambda)^{n\text{-}1}$ is Σa_{ii} , which is called *trace* of $\mathbf{A}$ and is denoted by $\mathrm{tr}(\mathbf{A})$

On the one hand, we can see that the same coefficient is $\Sigma\lambda_i$ from $p(\lambda)$

$$
\begin{aligned}
p(\lambda) &= (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\Lambda\ (\lambda-\lambda_n) \\
&= (\lambda_1-\lambda)(\lambda_2-\lambda)\Lambda\ (\lambda_n-\lambda)
\end{aligned}
$$

Thus $$\boxed{\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}a_{ii} = \mathrm{tr}(\mathbf{A})}$$

(*Ex*.6)

$$\mathbf{A} = \begin{bmatrix} 5 & -18 \\ 1 & -1 \end{bmatrix} \quad \det(\mathbf{A}) = -5 + 18 = 13, \ \ \mathrm{tr}(\mathbf{A}) = 5 - 1 = 4$$

$$\begin{vmatrix} 5-\lambda & -18 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 13 = 0 \quad \therefore \lambda_1 = 2 + 3i, \lambda_2 = 2 - 3i$$

Note that

$$\lambda_1 + \lambda_2 = 4 = \mathrm{tr}(\mathbf{A}), \ \ \lambda_1 \lambda_2 = 13 = \det(\mathbf{A})$$

Th. 6.1.2

A square matrix $\mathbf{A}$ is invertible if and only if 0 is not an eigenvalue of $\mathbf{A}$

$$\mathbf{A} \text{ is invertible} \Leftrightarrow \det(\mathbf{A}) \neq 0 \Leftrightarrow \det(\mathbf{A} - 0\mathbf{I}) \neq 0$$

- Th. 6.1.3

  Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices. If $\mathbf{B}$ is similar to $\mathbf{A}$, then the two matrices have the same characteristic polynomial and, consequently, the same eigenvalues

(*Proof*)

Let $p_A(\lambda)$ and $p_B(\lambda)$ denote the characteristic polynomials of $\mathbf{A}$ and $\mathbf{B}$, respectively. If $\mathbf{B}$ is similar to $\mathbf{A}$, then there exists a nonsingular matrix $\mathbf{S}$ such that $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$.

Thus
$$\begin{aligned} p_B(\lambda) &= \det(\mathbf{B} - \lambda\mathbf{I}) = \det(\mathbf{S}^{-1}\mathbf{AS} - \lambda\mathbf{I}) \\ &= \det(\mathbf{S}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{S}) = \det(\mathbf{S}^{-1})\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{S}) \\ &= \det(\mathbf{S}^{-1})\det(\mathbf{S})\det(\mathbf{A} - \lambda\mathbf{I}) = p_A(\lambda) \end{aligned}$$

Eigenvalues of Linear Transformation

If $L : V \to V$ is a linear transformation and V is a vector space, then a nonzero vector $\mathbf{v}$ is an *eigenvector* of L if

$$L(\mathbf{v}) = \lambda\mathbf{v}$$

for some (possibly zero) scalar λ

We call λ an *eigenvalue* of L and we say that the eigenvector $\mathbf{v}$ belongs to (or corresponds to, or is associated with) λ

(*Ex*.7) $L:\mathbf{R}^2 \to \mathbf{R}^2,\ L(x,y)=(x+y,x+y)$

$$L\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x+y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \begin{bmatrix} x+y \\ x+y \end{bmatrix} - \lambda \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda \mathbf{I} \right) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \det\left(\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \right) = \lambda(\lambda - 2) = 0$$

$$\text{For } \lambda = 0, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \therefore \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\text{For } \lambda = 2, \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \quad \therefore \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

■

$$\begin{bmatrix} x+y \\ x+y \end{bmatrix} - \lambda \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \lambda \mathbf{I} \begin{bmatrix} x \\ y \end{bmatrix} = \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda \mathbf{I} \right) \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\mathbf{A} = [\ [L(\mathbf{v}_1)]_B\ \ [L(\mathbf{v}_2)]_B\ \ \ldots\ [L(\mathbf{v}_n)]_B\ ]$$

Th. 6.1.4

Let V be a finite-dimensional vector space.

Let $L : V \to V$ be a linear transformation with matrix $\mathbf{A}$ with respect to some basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V. Then

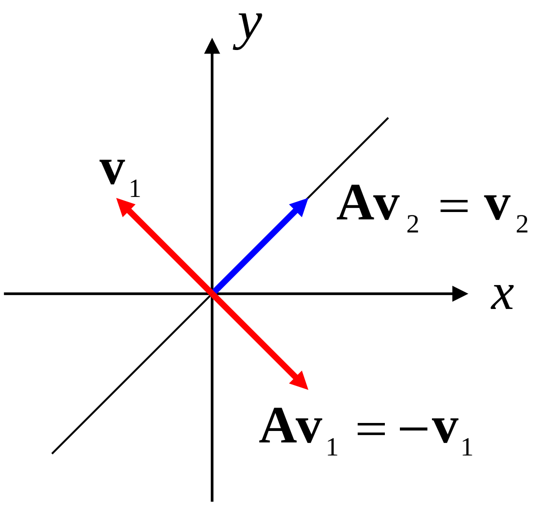
$$L(\mathbf{v}) = \lambda \mathbf{v} \quad \Leftrightarrow \quad \mathbf{A}[\mathbf{v}]_B = \lambda [\mathbf{v}]_B$$

Hence,

1. λ is an eigenvalue of L if and only if it is an eigenvalue of $\mathbf{A}$
2. $\mathbf{v}$ is an eigenvector of L if and only if $[\mathbf{v}]_B$ is an eigenvector of $\mathbf{A}$

$$[L(\mathbf{v})]_B = [\lambda \mathbf{v}]_B$$

(*Ex*.8) $L:\mathbf{R}^2 \to \mathbf{R}^2,\ T(x,y)=(y,x)$



Let $B=\{(1,0),(0,1)\}$.

$L(1,0)=(0,1) \quad \& \quad L(0,1)=(1,0) \quad \therefore \mathbf{A}=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$$\det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \lambda\mathbf{I}\right) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 - 1 = 0$$

$$\text{For } \lambda=-1,\ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},\quad [\mathbf{v}_1]_B = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \mathbf{v}_1$$

$$\text{For } \lambda=1,\ \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix},\quad [\mathbf{v}_2]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{v}_2$$

■

(*Ex*.9) $L: P_2 \to P_2, \quad L(a+bx) = (a+b) + (a+b)x$

Let $B = \{1, x\}$.

$$[L(1)]_B = [1+x]_B = (1,1)$$
$$[L(x)]_B = [1+x]_B = (1,1)$$
$$\therefore \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\det\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda \mathbf{I}\right) = \det\left(\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix}\right) = \lambda(2-\lambda) = 0$$

For $\lambda = 0$, $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad [\mathbf{v}_1]_B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_1 = -1 + x$

For $\lambda = 2$, $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \quad [\mathbf{v}_2]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = 1 + x$ ■

Summary 1

1. Def. $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ Homogeneous L.S.
   (i) Find eigenvalues from $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$
   (ii) For each λ_i, write the general solution of $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v} = \mathbf{0}$ and get an eigenvector associated with each free variable
2. $\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$, $\Sigma\lambda_i = \Sigma a_{ii}$
3. $\mathbf{A}$ is invertible $\iff \det(\mathbf{A} - 0\mathbf{I}) \neq 0$
4. Similar matrices have the same eigenvalues
5. Eigenvalues of Linear Transformation

$$L(\mathbf{v}) = \lambda\mathbf{v} \iff \mathbf{A}[\mathbf{v}]_B = \lambda[\mathbf{v}]_B$$

Diagonalization

- To know which matrices can be diagonalized and how to diagonalize them
- To know how to compute $\mathbf{A}^n$ efficiently if $\mathbf{A}$ can be diagonalized
- To know how to diagonalize a linear transformation

Usefulness of Diagonal Matrices

- A diagonal matrix *does not mix* the components of $\mathbf{x}$ in the product $\mathbf{Dx}$
- Moreover, it is very easy to compute the powers $\mathbf{D}^k$

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a \\ 3b \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} 2a & 2b & 2c \\ 3d & 3e & 3f \end{bmatrix}$$

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

Transpose

$$[\mathbf{a}_1 \ \mathbf{a}_2 \Lambda \ \mathbf{a}_n]\,\mathbf{D}_n = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2a & 3d \\ 2b & 3e \\ 2c & 3f \end{bmatrix} = [d_1\mathbf{a}_1 \ d_2\mathbf{a}_2 \Lambda \ d_n\mathbf{a}_n]$$

Th. 6.3.1

If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are *distinct* eigenvalues of an $n \times n$ matrix of $\mathbf{A}$ with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are *linearly independent*

(*Proof*)

Let r be the dimension of the subspace of spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and suppose that $r < k$. We may assume (reordering the $\mathbf{v}_i$'s and λ_i's if necessary) that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are *linearly independent*

Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}$ are linearly dependent and hence there exist scalars $c_1, c_2, \ldots, c_{r+1}$, not all zero, such that

$$c_1\mathbf{v}_1 + \Lambda + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} = \mathbf{0} \qquad (1)$$

Note that c_{r+1} must be nonzero: otherwise, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ would be linearly dependent. So $c_{r+1}\mathbf{v}_{r+1} \neq \mathbf{0}$ and hence $c_1, c_2, \ldots, c_r$ cannot all be zero

Multiplying the equation (1) by $\mathbf{A}$, we get

$$c_1\mathbf{A}\mathbf{v}_1 + \Lambda + c_r\mathbf{A}\mathbf{v}_r + c_{r+1}\mathbf{A}\mathbf{v}_{r+1} = \mathbf{0} \quad \text{or}$$

$$c_1\lambda_1\mathbf{v}_1 + \Lambda + c_r\lambda_r\mathbf{v}_r + c_{r+1}\lambda_{r+1}\mathbf{v}_{r+1} = \mathbf{0} \qquad (2)$$

From the equations (1) and (2), we get

$(2) - \lambda_{r+1}\,(1)$; $$c_1(\lambda_1 - \lambda_{r+1})\mathbf{v}_1 + \Lambda + c_r(\lambda_r - \lambda_{r+1})\mathbf{v}_r = \mathbf{0} \qquad (3)$$

Since $c_1, c_2, \ldots, c_r$ cannot all be zero and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct, all the coefficients in (3) cannot be zero. Hence the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are *linearly dependent*

This *contradicts* the independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$.

Therefore, r must equal k ■

Def. Diagonalizable

An $n \times n$ matrix $\mathbf{A}$ is said to be *diagonalizable* if there exists a nonsingular matrix $\mathbf{X}$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D}$$

We say that $\mathbf{X}$ diagonalizes $\mathbf{A}$

- *Not all* square matrices can be diagonalized
- The matrices $\mathbf{X}$ and $\mathbf{D}$ are *not unique*

□ Th. 6.3.2

An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n *linearly independent* eigenvectors

(*Proof*)

Suppose that the matrix $\mathbf{A}$ has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Let λ_i be the eigenvalue of $\mathbf{A}$ corresponding to $\mathbf{v}_i$ for each i. (Some of the λ_i's may be equal.) Let $\mathbf{X}$ be the matrix whose jth column vector is $\mathbf{v}_j$ for $j = 1, \ldots, n$. It follows that $\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$ is the jth column vector of $\mathbf{AX}$

Thus

$$\mathbf{AX} = \mathbf{A}[\,\mathbf{v}_1\ \mathbf{v}_2\ \Lambda\ \ \mathbf{v}_n\,] = [\,\mathbf{Av}_1\ \mathbf{Av}_2\ \Lambda\ \ \mathbf{Av}_n\,]$$

$$= [\,\lambda_1\mathbf{v}_1\ \lambda_2\mathbf{v}_2\ \Lambda\ \ \lambda_n\mathbf{v}_n\,]$$

$$= [\,\mathbf{v}_1\ \mathbf{v}_2\ \Lambda\ \ \mathbf{v}_n\,]\begin{bmatrix} \lambda_1 & 0 & \Lambda & 0 \\ 0 & \lambda_2 & \Lambda & 0 \\ : & : & \mathrm{O} & 0 \\ 0 & 0 & \Lambda & \lambda_n \end{bmatrix} = \mathbf{XD}$$

Since $\mathbf{X}$ has n linearly independent column vectors, it follows that $\mathbf{X}$ is nonsingular and hence $\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX}$

Conversely, suppose that $\mathbf{A}$ is diagonalizable. Then there exists a *nonsingular* matrix $\mathbf{X}$ such that $\mathbf{AX} = \mathbf{XD}$.

If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the column vectors of $\mathbf{X}$, then

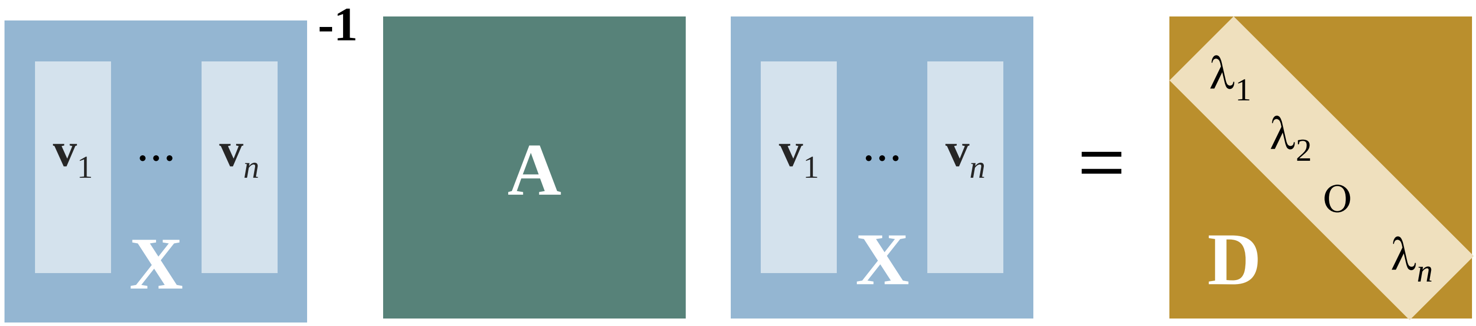
$$\mathbf{AX} = [\,\mathbf{Ax}_1\ \mathbf{Ax}_2\ \Lambda\ \ \mathbf{Ax}_n\,] = [\,d_1\mathbf{x}_1\ d_2\mathbf{x}_2\ \Lambda\ \ d_n\mathbf{x}_n\,] = \mathbf{XD}$$

That is, $\mathbf{Ax}_j = d_j\,\mathbf{x}_j$ for each j.

Thus, for each j, d_j is an *eigenvalue* of $\mathbf{A}$ and $\mathbf{x}_j$ is an *eigenvector* belonging to d_j. Since the column vectors of the nonsingular matrix $\mathbf{X}$ are linearly independent, $\mathbf{A}$ has n linearly independent eigenvectors $\blacksquare$

Remarks

1. If $\mathbf{A}$ is diagonalizable, then the column vectors of the diagonalizing matrix $\mathbf{X}$ are eigenvectors of $\mathbf{A}$ and the diagonal elements of $\mathbf{D}$ are the corresponding eigenvalues of $\mathbf{A}$



2. The diagonalizing matrix $\mathbf{X}$ are *not unique*. Reordering the columns of $\mathbf{X}$ or multiplying them by nonzero scalars will produce a new diagonalizing matrix
3. If an $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues, then $\mathbf{A}$ is diagonalizable. If the eigenvalues are *not* distinct, then $\mathbf{A}$ *may or may not* be diagonalizable, depending on whether $\mathbf{A}$ has n linearly independent eigenvectors
4. If $\mathbf{A}$ is diagonalizable, then $\mathbf{A}$ can be factored into a product $\mathbf{XDX}^{-1}$

It follows from remark 4 that

$$\mathbf{A}^k = (\mathbf{XDX}^{-1})^k = (\mathbf{XDX}^{-1})(\mathbf{XDX}^{-1})\Lambda\ (\mathbf{XDX}^{-1})$$

$$= \mathbf{XD}^k\mathbf{X}^{-1} = \mathbf{X}\begin{bmatrix} (\lambda_1)^k & 0 & \Lambda & 0 \\ 0 & (\lambda_2)^k & \Lambda & 0 \\ : & : & \mathrm{O} & : \\ 0 & 0 & \Lambda & (\lambda_n)^k \end{bmatrix}\mathbf{X}^{-1}$$

Once we have a factorization $\mathbf{A} = \mathbf{XDX}^{-1}$, it is easy to compute powers of $\mathbf{A}$

(*Ex*.12) Compute $\mathbf{A}^{100}$ and $\mathbf{B}^{100}$

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

(*Solution*)

The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = -4$. Their corresponding eigenvectors are $\mathbf{v}_1 = (3,1)$ and $\mathbf{v}_2 = (1,2)$. $\mathbf{A}$ is diagonalizable and it can be factored into $\mathbf{XDX}^{-1}$. Thus

$$\mathbf{A}^{100} = \mathbf{XD}^{100}\mathbf{X}^{-1} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1^{100} & 0 \\ 0 & (-4)^{100} \end{bmatrix} \begin{bmatrix} 2/5 & -1/5 \\ -1/5 & 3/5 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 6 - 4^{100} & -3 + 3 \cdot 4^{100} \\ 2 - 2 \cdot 4^{100} & -1 + 6 \cdot 4^{100} \end{bmatrix}$$

Meanwhile, the eigenvalues of $\mathbf{B}$ are both equal to 1 and any eigenvector corresponding to $\lambda = 1$ must be a multiple of $\mathbf{v}_1 = (1,0)$. Thus $\mathbf{B}$ cannot be diagonalizable

But, fortunately, we can compute $\mathbf{B}^{100}$ easily

$$\mathbf{B}^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{B}^3 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

$$\vdots$$

$$\mathbf{B}^{100} = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}$$

$$\det(\mathbf{B} - \lambda\mathbf{I}) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix}$$

$$= (1-\lambda)^2 = 0$$

$$\lambda = 1;\ \mathbf{B} - \mathbf{I} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\therefore E_1 = \left\{ r\begin{bmatrix} 1 \\ 0 \end{bmatrix},\ r \in \mathbf{R} \right\}$$

■

Diagonalization of Linear Transformation

- Let $T: V \rightarrow V$ be a linear transformation of a finite-dimensional vector space V
- If there exists a basis B of V such that the matrix of T with respect to B is *diagonal*, we say that T is diagonalizable and that B diagonalizes T
- It turns out that the vectors of B are eigenvectors of T
- The diagonalizable T would be easy to evaluate

Th. 6.3.3

Let $T: V \to V$ be a linear transformation of a finite-dimensional vector space V. Then

1. T is diagonalizable if and only if V has a basis of eigenvectors of T
2. T is diagonalizable if and only if the matrix of T with respect to any basis of V is diagonalizable
3. If T is *diagonalized* by B, then the vectors of B are eigenvectors of T

(*Ex*.13)

Show that the linear transformation $T : P_2 \to P_2$, $T(a + bx) = b + ax$ is diagonalizable. Find a basis B of P_2 that diagonalizes T. Evaluate T using B

(*Solution*) $T(\mathbf{v}) = \lambda\mathbf{v} \quad \Leftrightarrow \quad \mathbf{A}[\mathbf{v}]_S = \lambda[\mathbf{v}]_S$

Because $T(1) = x$ and $T(x) = 1$, the matrix of T with respect to the standard basis $\{1, x\}$ is $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$\mathbf{A}$ has two distinct eigenvaluses, 1 and -1, and two corresponding eigenvectors $\underline{(1,1)\text{ and }(-1,1)}$, respectively. Thus, from the 2nd statement in Th.6.3.3, $\mathbf{A}$ is diagonalizable and so is T

The eigenvectors of T are $1+x$ and $-1+x$. Therefore, $B = \{1+x, -1+x\}$ diagonalizes T

$$[T(1+x)]_B = [1+x]_B = (1,0)$$
$$[T(-1+x)]_B = [1-x]_B = (0,-1)$$
$$\therefore \mathbf{A}_B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in D_2$$

T may be evaluated by using B as follows;

$$\begin{aligned} T(a+bx) &= T\left(\frac{a+b}{2}(1+x) + \frac{b-a}{2}(-1+x)\right) \\ &= \frac{a+b}{2}T(1+x) + \frac{b-a}{2}T(-1+x) \\ &= \frac{a+b}{2}\cdot 1 \cdot (1+x) + \frac{b-a}{2}\cdot(-1)\cdot(-1+x) \\ &= b + ax \end{aligned}$$

(*Ex*.14)

Show that differentiation $d/dx : P_2 \rightarrow P_2$ is not diagonalizable

(*Solution*)

Because $d/dx(1) = 0$ and $d/dx(x) = 1$, the matrix of d/dx with respect to the standard basis $\{1, x\}$ is $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ which is not diagonalizable

So d/dx is not diagonalizable. Hence, P_2 has no basis of eigenvectors of d/dx

$$\lambda^2 = 0, \quad E_0 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

Application: Markov Chains

A *stochastic process* is any sequence of experiments for which the outcome at any stage depends on chance. A *Markov process* is a stochastic process with the following properties:

1. The set of possible outcomes or states is finite
2. The probability of the next outcome depends only on the previous outcome
3. The probabilities are constant over time

(*Ex*.15)   Automobile Leasing

Table 1. *Transition Probabilities for Vehicle Leasing*

Current Lease				Next Lease
Sedan	Sports Car	Minivan	SUV	
0.80	0.10	0.05	0.05	Sedan
0.10	0.80	0.05	0.05	Sports Car
0.05	0.05	0.80	0.10	Minivan
0.05	0.05	0.10	0.80	SUV

There are *four* types of outcomes and the probability of each outcome can be estimated by reviewing records of previous leases

Suppose that initially there are 200 sedans leased and 100 of each of the other three types of vehicles

$$\mathbf{A} = \begin{bmatrix} 0.80 & 0.10 & 0.05 & 0.05 \\ 0.10 & 0.80 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.80 & 0.10 \\ 0.05 & 0.05 & 0.10 & 0.80 \end{bmatrix} \quad \mathbf{x}_0 = \begin{bmatrix} 200 \\ 100 \\ 100 \\ 100 \end{bmatrix} \quad \mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \begin{bmatrix} 180 \\ 110 \\ 105 \\ 105 \end{bmatrix}$$

$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n$ for $n = 1, 2, \ldots$

The vectors $\mathbf{x}_i$ produced in this manner are referred to as *state vectors*, and the *sequence of state vectors* is called a *Markov chain*

In general, a matrix is said to be *stochastic* if its entries are nonnegative and the entries in each column add up to 1. Thus each column of a stochastic matrix can be viewed as a *probability vector*

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix}^{-1}$$

$$\begin{aligned} \mathbf{x}_n &= \mathbf{X}\mathbf{D}^n\mathbf{X}^{-1}\mathbf{x}_0 \\ &= \mathbf{X}\mathbf{D}^n\mathbf{X}^{-1}(0.40, 0.20, 0.20, 0.20) \;\; ; \;\; \text{probability vector } \mathbf{x}_0 \\ &= 0.25(1,1,1,1) - 0.05\underline{(0.8)}^n(-1,-1,1,1) + 0.10\underline{(0.7)}^n(1,-1,0,0) \end{aligned}$$

As n increases, $\mathbf{x}_n$ approaches the *steady-state vector*

$$\mathbf{x} = (0.25, 0.25, 0.25, 0.25)$$

■

- Th. 6.3.3

  If $\lambda_1 = 1$ is a *dominant* (absolutely largest) eigenvalue of a stochastic matrix $\mathbf{A}$, then the Markov chain with $\mathbf{A}$ converges to a steady-state vector $\mathbf{x}$ that satisfies

  1. $\mathbf{x}$ is a probability vector
  2. $\mathbf{x}$ is an eigenvector belonging to λ_1

$$\mathbf{x}_n = 0.25(1,1,1,1) - 0.05(0.8)^n(-1,-1,1,1) + 0.10(0.7)^n(1,-1,0,0)$$

Application: Discrete Dynamical Systems

A dynamic system is an equation involving a time-dependent vector quantity $\mathbf{x}(t)$. In a discrete dynamical system we write $\mathbf{x}_k$ for $\mathbf{x}(t)$

A *first-order discrete homogeneous dynamical system* is a vector equation of the form $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$

Then $\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^2\mathbf{x}_{k-2} = \ldots$ Hence, $\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0$

If $\mathbf{A}$ is *diagonalizable* by a basis of eigenvectors $\mathbf{v}_1,\ldots,\mathbf{v}_n$ of $\mathbf{A}$, then the system can be written for an initial vector $\mathbf{x}_0 = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$

$$\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0 = c_1\lambda_1^k\mathbf{v}_1 + \ldots + c_n\lambda_n^k\mathbf{v}_n$$

(*Ex*.16) A stretch of desert in northwestern Mexico is populated mainly by two species of animals: *coyotes* and *roadrunners*. The populations of coyotes and roadrunners in $t+1$ year, $c(t+1)$ and $r(t+1)$, can be modeled as follows

$$c(t+1) = 0.86c(t) + 0.08r(t)$$
$$r(t+1) = -0.12c(t) + 1.14r(t)$$

Describe the long-term behavior of populations for

(a) $c_0 = 200$ and $r_0 = 600$

(b) $c_0 = 600$ and $r_0 = 300$

(c) $c_0 = r_0 = 1000$

/ kaióuti / , / róudrʌnə(r) /

$$\begin{bmatrix} c(t+1) \\ r(t+1) \end{bmatrix} = \begin{bmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{bmatrix} \begin{bmatrix} c(t) \\ r(t) \end{bmatrix} \quad \Rightarrow \quad \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t \Rightarrow \mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0$$

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \Rightarrow \lambda_1 = 1.1,\ \lambda_2 = 0.9 \text{ and } \mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(a) For $\mathbf{x}_0 = (200, 600)$,

$$\begin{bmatrix} c(t) \\ r(t) \end{bmatrix} = \mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 = \mathbf{A}^t (200\mathbf{v}_1) = 200(1.1)^t \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

So, both populations will grow exponentially, by 10% each year.

(b) For $\mathbf{x}_0 = (600, 300)$,

$$\begin{bmatrix} c(t) \\ r(t) \end{bmatrix} = \mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 = \mathbf{A}^t (300\mathbf{v}_2) = 300(0.9)^t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

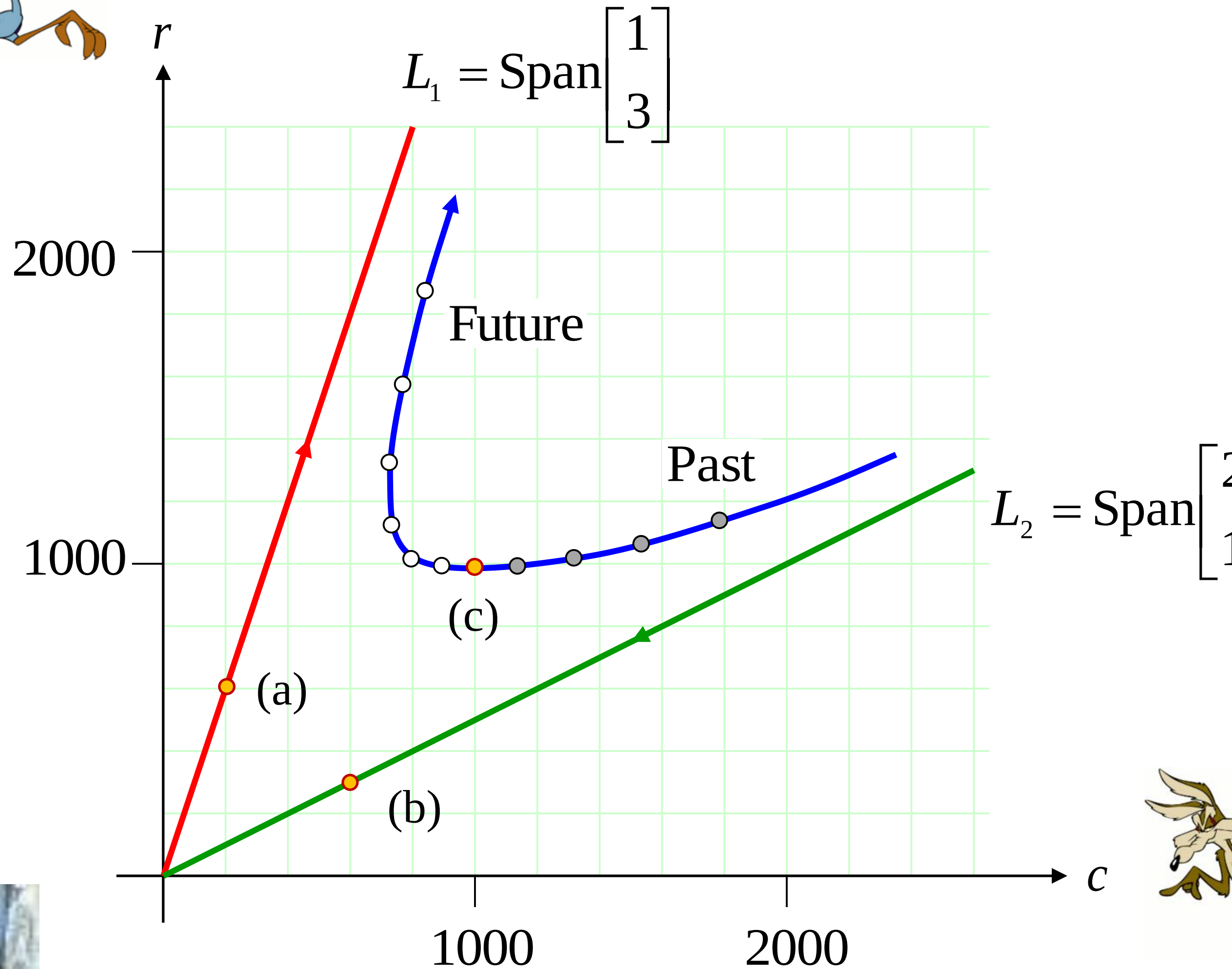
Both populations decline 10% each year.

(c) For $\mathbf{x}_0 = (1000, 1000)$,

$$\begin{bmatrix} c(t) \\ r(t) \end{bmatrix} = \mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 = \mathbf{A}^t (200\mathbf{v}_1 + 400\mathbf{v}_2)$$

$$= 200(1.1)^t \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 400(0.9)^t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 200(1.1)^t + 800(0.9)^t \\ 600(1.1)^t + 400(0.9)^t \end{bmatrix}$$



r
$L_1 = \text{Span}\begin{bmatrix}1\\3\end{bmatrix}$
2000
1000
Future
Past
$L_2 = \text{Span}\begin{bmatrix}2\\1\end{bmatrix}$
(c)
(a)
(b)
1000
2000
c

Summary 2

1. Eigenvectors for distinct eigenvalues are linearly indep.
2. Def. A square matrix $\mathbf{A}$ is *diagonalizable* if $\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D}$
3. An $n \times n$ matrix $\mathbf{A}$ is diagonalizable $\Leftrightarrow$ $\mathbf{A}$ has n linearly independent eigenvectors that are column vectors of $\mathbf{X}$
4. Def. $T : V \rightarrow V$ is diagonalizable if there exist a basis B of V such that the matrix of T, $\mathbf{A}_B$, is diagonal
5. If T is diagonalized by B, then the vectors of B are eigenvectors of T

Hermitian Matrices

To study matrices with complex entries and look at the complex analogues of symmetric and orthogonal matrices

1. Eigenvalues and Eigenvectors
2. Diagonalization
3. Hermitian Matrices

Complex Inner Products

Let $\mathbf{C}^n$ denote the vector space of all n-tuples of complex numbers. The set $\mathbf{C}$ of all complex numbers will be taken as our field of scalars

If $\alpha = a + bi$ is a *complex scalar*, the *length* of α is given by

$$|\alpha| = \sqrt{\bar{\alpha}\alpha} = \sqrt{a^2 + b^2} \quad \text{where} \quad \bar{\alpha} = a - bi$$

The *length of a vector* $\mathbf{z} = (z_1, \ldots, z_n)$ in $\mathbf{C}^n$ is given by

$$\|\mathbf{z}\| = \left(|z_1|^2 + |z_2|^2 + \Lambda + |z_n|^2\right)^{1/2} = \left(\bar{z}_1 z_1 + \bar{z}_2 z_2 + \Lambda + \bar{z}_n z_n\right)^{1/2}$$

$$= \left(\bar{\mathbf{z}}^T \mathbf{z}\right)^{1/2} \equiv \left(\mathbf{z}^H \mathbf{z}\right)^{1/2}$$

Def. Complex Inner Product

Let V be a vector space *over the complex numbers*. An *inner product* on V is an operation that assigns, to each pair of vectors $\mathbf{z}$ and $\mathbf{w}$ in V, a *complex number* $<\mathbf{z}, \mathbf{w}>$ satisfying the following conditions:

1. $<\mathbf{z}, \mathbf{z}> \geq 0$, with equality if and only if $\mathbf{z} = \mathbf{0}$
2. $<\mathbf{z}, \mathbf{w}> = \overline{<\mathbf{w}, \mathbf{z}>}$ for all $\mathbf{z}$ and $\mathbf{w}$ in V

   conjugate symmetry
3. $<\alpha \mathbf{z} + \beta \mathbf{w}, \mathbf{u}> = \alpha <\mathbf{z}, \mathbf{u}> + \beta <\mathbf{w}, \mathbf{u}>$

- An inner product on $\mathbf{C}^n$
  - It can be defined by $<\mathbf{z}, \mathbf{w}> = \mathbf{w}^H\mathbf{z}$
  - The complex inner product space $\mathbf{C}^n$ is similar to the real inner product space $\mathbf{R}^n$. The main difference is that in the complex case it is necessary to conjugate before transposing when taking an inner product

$\mathbf{R}^n$	$\mathbf{C}^n$
$<\mathbf{x},\mathbf{y}> = \mathbf{y}^T\mathbf{x}$	$<\mathbf{z},\mathbf{w}> = \mathbf{w}^H\mathbf{z}$
$\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$	$\mathbf{z}^H\mathbf{w} = \overline{\mathbf{w}^H\mathbf{z}}$
$\lVert\mathbf{x}\rVert^2 = \mathbf{x}^T\mathbf{x}$	$\lVert\mathbf{z}\rVert^2 = \mathbf{z}^H\mathbf{z}$

(*Ex*.17) For two complex vector $\mathbf{z}$ and $\mathbf{w}$ in $\mathbf{C}^n$, compute their complex inner product and the length of each

$$\mathbf{z} = \begin{bmatrix} 5+i \\ 1-3i \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 2+i \\ -2+3i \end{bmatrix}$$

(*Solution*)

$$\mathbf{w}^H \mathbf{z} = [\ 2-i\ , -2-3i\ ] \begin{bmatrix} 5+i \\ 1-3i \end{bmatrix} = (2-i)(5+i) + (-2-3i)(1-3i) = 0$$

$$\|\mathbf{z}\|^2 = \mathbf{z}^H \mathbf{z} = |5+i|^2 + |1-3i|^2 = 26+10 = 36 \quad \therefore \|\mathbf{z}\| = 6$$

$$\|\mathbf{w}\|^2 = \mathbf{w}^H \mathbf{w} = |2+i|^2 + |-2+3i|^2 = 5+13 = 18 \quad \therefore \|\mathbf{w}\| = 3\sqrt{2}$$

Notations for Complex Matrices

- Let $\mathbf{M} = (m_{ij})$ be an $m \times n$ complex matrix with $m_{ij} = a_{ij} + i\,b_{ij}$ for each i and j. We may write $\mathbf{M}$ in the form

$$\mathbf{M} = \mathbf{A} + i\,\mathbf{B}$$

  where $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ have real entries
- We define the *conjugate* of $\mathbf{M}$ by $\overline{\mathbf{M}} = \mathbf{A} - i\mathbf{B}$
- The transpose of $\overline{\mathbf{M}}$ will be denoted by $\mathbf{M}^H$
- The vector space of all $m \times n$ matrices with complex entries is denoted by $\mathbf{C}^{m\times n}$

- If $\mathbf{A}$ and $\mathbf{B}$ are elements of $\mathbf{C}^{m\times n}$ and $\mathbf{C} \in \mathbf{C}^{n\times r}$, then the following rules are easily verified

1. $(\mathbf{A}^H)^H = \mathbf{A}$
2. $(\alpha\mathbf{A} + \beta\mathbf{B})^H = \overline{\alpha}\,\mathbf{A}^H + \overline{\beta}\,\mathbf{B}^H$
3. $(\mathbf{AC})^H = \mathbf{C}^H\mathbf{A}^H$

click

Def. Hermitian Matrix

A matrix $\mathbf{M}$ is said to be *Hermitian* if $\mathbf{M} = \mathbf{M}^H$

$$\mathbf{M} = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix}, \quad \mathbf{M}^H = \begin{bmatrix} 3 & \overline{2-i} \\ \overline{2+i} & 4 \end{bmatrix}^T = \mathbf{M}$$

- If $\mathbf{M}$ is a matrix with real entries, $\mathbf{M}^H = \mathbf{M}^T$. Thus a real symmetric matrix $\mathbf{A}$ is Hermitian

Th. 6.4.1

The eigenvalues of a Hermitian matrix are all real. Furthermore, eigenvectors belonging to distinct eigenvalues are orthogonal

(*Proof*)

Let λ be an eigenvalue of a Hermitian matrix $\mathbf{A}$ and let $\mathbf{x}$ be an eigenvector belonging to λ. If $\alpha = \mathbf{x}^H\mathbf{A}\mathbf{x}$, then α is real since

$$\overline{\alpha} = (\overline{\alpha})^T = \alpha^H = (\mathbf{x}^H\mathbf{A}\mathbf{x})^H = (\mathbf{A}\mathbf{x})^H(\mathbf{x}^H)^H = \mathbf{x}^H\mathbf{A}^H\mathbf{x} = \mathbf{x}^H\mathbf{A}\mathbf{x} = \alpha$$

and hence $\lambda = \alpha / \|\mathbf{x}\|^2$ is real since

$$\alpha = \mathbf{x}^H\mathbf{A}\mathbf{x} = \lambda\mathbf{x}^H\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

If $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors belonging to distinct eigenvalues λ_1 and λ_2, respectively, then

$$(\mathbf{A}\mathbf{x}_1)^H\mathbf{x}_2 = \mathbf{x}_1^H\mathbf{A}^H\mathbf{x}_2 = \mathbf{x}_1^H\mathbf{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

and

$$(\mathbf{A}\mathbf{x}_1)^H\mathbf{x}_2 = (\mathbf{x}_2^H\mathbf{A}\mathbf{x}_1)^H = (\lambda_1\mathbf{x}_2^H\mathbf{x}_1)^H = \lambda_1\mathbf{x}_1^H\mathbf{x}_2$$

Consequently,

$$\lambda_1\mathbf{x}_1^H\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

and since $\lambda_1 \neq \lambda_2$, it follows that

$$< \mathbf{x}_2, \mathbf{x}_1 > = \mathbf{x}_1^H\mathbf{x}_2 = 0 \qquad \blacksquare$$

Def. Unitary Matrix

An $n \times n$ matrix $\mathbf{U}$ is said to be *unitary* if its column vectors form an orthonormal set in $\mathbf{C}^n$

- Thus, $\mathbf{U}$ is unitary if and only if $\mathbf{U}^H \mathbf{U} = \mathbf{I}$
- If $\mathbf{U}$ is unitary, then $\mathbf{U}$ must have rank n since the column vectors are orthonormal
- It follows that $\mathbf{U}^{-1} = \mathbf{I}\mathbf{U}^{-1} = \mathbf{U}^H \mathbf{U}\,\mathbf{U}^{-1} = \mathbf{U}^H$
- A *real* unitary matrix is an *orthogonal matrix*

□ Corollary 6.4.2

If the eigenvalues of a Hermitian matrix $\mathbf{A}$ are distinct, then there exists a unitary matrix $\mathbf{U}$ that diagonalizes $\mathbf{A}$

□ Th. 6.4.3 *Schur's Theorem*

For each $n \times n$ matrix $\mathbf{A}$, there exists a unitary matrix $\mathbf{U}$ such that $\mathbf{U}^H\mathbf{A}\mathbf{U}$ is upper triangular

□ Th. 6.4.4 *Spectral Theorem*

If $\mathbf{A}$ is Hermitian, then there exists a unitary matrix $\mathbf{U}$ that diagonalizes $\mathbf{A}$

(*Proof*)

By the Schur's theorem, there is a unitary matrix such that $\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{T}$, where $\mathbf{T}$ is upper triangular. Furthermore,

$$\mathbf{T}^H = (\mathbf{U}^H\mathbf{A}\mathbf{U})^H = \mathbf{U}^H\mathbf{A}^H\mathbf{U} = \mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{T}$$

Therefore, $\mathbf{T}$ is Hermitain and consequently must be diagonal

★

Corollary 6.4.5

Spectral Theorem – Real Symmetric Matrices

If $\mathbf{A}$ is a real symmetric matrix, then there is an orthogonal matrix $\mathbf{Q}$ that diagonalizes $\mathbf{A}$; that is, $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D}$, where $\mathbf{D}$ is diagonal

➔ *orthogonally diagonalizable*

Summary 3

1. (*Spectral Theorem*) If $\mathbf{A}$ is a real symmetric matrix, then there is an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{D}$.
2. Def. *Hermitian* matrix $\mathbf{M}$ if $\mathbf{M} = \mathbf{M}^H$ $(\equiv \overline{\mathbf{M}}^T)$
3. Def. An $n \times n$ matrix $\mathbf{U}$ is *unitary* if its column vectors form an orthonormal set in $\mathbf{C}^n$, that is, $\mathbf{U}^H\mathbf{U} = \mathbf{I}$

   (If real matrix $\mathbf{Q}$ is orthogonal, then $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$)
4. A Hermitian (or real symmetric) matrix has all real eigenvalues and its eigenvectors belonging to distinct eigenvalues are orthogonal

★

(*Ex*.18) Find an orthogonal matrix **U** that diagonalizes **A**

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$$

(*Solution*)

The eigenvalues of **A** are -1 and 5.

The eigenvectors belonging to $\lambda_1 = \lambda_2 = -1$

$$\mathbf{x}_1 = (1,0,1), \quad \mathbf{x}_2 = (-2,1,0)$$

$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} = \mathbf{0}$

form a basis for the eigenspace $\mathrm{Null}(\mathbf{A} + \mathbf{I})$.

Notice that a vector $\mathbf{x} \in \mathrm{Null}(\mathbf{A} + \mathbf{I})$ can also be a eigenvector corresponding to the eigenvalue -1, since

$$(\mathbf{A} + \mathbf{I})(a\mathbf{x}_1 + b\mathbf{x}_2) = \mathbf{0}$$

$$\Leftrightarrow \mathbf{A}(a\mathbf{x}_1 + b\mathbf{x}_2) = (-1)(a\mathbf{x}_1 + b\mathbf{x}_2)$$

$\mathbf{x}_1 = (1,0,1), \quad \mathbf{x}_2 = (-2,1,0)$

Thus the following vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, orthogonalized through the Gram-Schmidt process can also be eigenvectors

$$\mathbf{u}_1 = \mathbf{x}_1 / \|\mathbf{x}_1\| = (1,0,1)/\sqrt{2}$$

$$\mathbf{p} = (\mathbf{x}_2^T \mathbf{u}_1)\,\mathbf{u}_1 = -\sqrt{2}\,\mathbf{u}_1 = (-1,0,-1)$$

$$\mathbf{u}_2 = (\mathbf{x}_2 - \mathbf{p}) / \|\mathbf{x}_2 - \mathbf{p}\| = (-1,1,1)/\sqrt{3}$$

The eigenspace corresponding to $\lambda_3 = 5$ is spanned by $\mathbf{x}_3 = (-1,-2,1)$. Then the matrix $\mathbf{X} = [\ \mathbf{u}_1\ \ \mathbf{u}_2\ \ \mathbf{x}_3\ ]$ diagonalizes $\mathbf{A}$

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \mathbf{D}$$

Since $\mathbf{x}_3$ must be orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$ by the theorem 6.4.1, we need only normalize

$$\mathbf{u}_3 = \mathbf{x}_3 / \|\mathbf{x}_3\| = (-1,-2,1)/\sqrt{6}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set and the matrix $\mathbf{U}$ = [$\mathbf{u}_1$ $\mathbf{u}_2$ $\mathbf{u}_3$] diagonalizes $\mathbf{A}$

$$\mathbf{U} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix} \qquad \mathbf{U}^T\mathbf{A}\mathbf{U} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \mathbf{D}$$

■

real symmetric

det($\mathbf{A}$ - $\lambda\mathbf{I}$) = $(\lambda\text{-}a)\,(\lambda\text{-}b)^2\,(\lambda\text{-}c)^3 = 0$ $\rightarrow$ eigenvectors $\mathbf{v}_1, \{\mathbf{v}_2, \mathbf{v}_3\}, \{\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}$

$\rightarrow$ $\mathbf{u}_1, \{\mathbf{u}_2, \mathbf{u}_3\}, \{\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6\}$ $\rightarrow$ $\mathbf{U} = [\, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6 \,]$