Math 221: LINEAR ALGEBRA

Chapter 3. Determinants and Diagonalization

§3-3. Determinants and Diagonalization

Le Chen[1]

Emory University, 2020 Fall

(last updated on 10/26/2020)

[1] Slides are adapted from those by Karen Seyffarth from University of Calgary.

Why Diagonalization?

Example

Let $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$. Find A^{100}.

How can we do this efficiently?

Consider the matrix $P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$. Observe that P is invertible (why?), and that

$$P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}.$$

Furthermore,

$$P^{-1}AP = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} = D,$$

where D is a diagonal matrix.

Example (continued)

This is significant, because

$$\begin{aligned} P^{-1}AP &= D \\ P(P^{-1}AP)P^{-1} &= PDP^{-1} \\ (PP^{-1})A(PP^{-1}) &= PDP^{-1} \\ IAI &= PDP^{-1} \\ A &= PDP^{-1}, \end{aligned}$$

and so

$$\begin{aligned} A^{100} &= (PDP^{-1})^{100} \\ &= (PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)D(P^{-1}\cdots P)DP^{-1} \\ &= PDIDIDI\cdots IDP^{-1} \\ &= PD^{100}P^{-1}. \end{aligned}$$

Example (continued)

Now,

$$D^{100} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}^{100} = \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix}.$$

Therefore,

$$\begin{aligned} A^{100} &= PD^{100}P^{-1} \\ &= \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix} \left(\frac{1}{3}\right) \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2^{100} + 2 \cdot 5^{100} & 2^{100} - 2 \cdot 5^{100} \\ 2^{100} - 5^{100} & 2 \cdot 2^{100} + 5^{100} \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2^{100} + 2 \cdot 5^{100} & 2^{100} - 2 \cdot 5^{100} \\ 2^{100} - 5^{100} & 2^{101} + 5^{100} \end{bmatrix} \end{aligned}$$

Theorem (Diagonalization and Matrix Powers)

If $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$ for each $k = 1, 2, 3, \ldots$

The process of finding an invertible matrix P and a diagonal matrix D so that $A = PDP^{-1}$ is referred to as diagonalizing the matrix A, and P is called the diagonalizing matrix for A.

Problem

- When is it possible to diagonalize a matrix?
- How do we find a diagonalizing matrix?

Eigenvalues and Eigenvectors

Definition

Let A be an $n \times n$ matrix, λ a real number, and $\vec{x} \neq \vec{0}$ an n-vector. If $A\vec{x} = \lambda\vec{x}$, then λ is an eigenvalue of A, and $\vec{x}$ is an eigenvector of A corresponding to λ, or a λ-eigenvector.

Example

Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

$$A\vec{x} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\vec{x}.$$

This means that 3 is an eigenvalue of A, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to 3 (or a 3-eigenvector of A).

Suppose that A is an $n \times n$ matrix, $\vec{x} \neq 0$ an n-vector, $\lambda \in \mathbb{R}$, and that $A\vec{x} = \lambda\vec{x}$.

Then

$$\begin{aligned} \lambda\vec{x} - A\vec{x} &= \vec{0} \\ \lambda I\vec{x} - A\vec{x} &= \vec{0} \\ (\lambda I - A)\vec{x} &= \vec{0} \end{aligned}$$

Since $\vec{x} \neq \vec{0}$, the matrix $\lambda I - A$ has no inverse, and thus

$$\det(\lambda I - A) = 0.$$

Definition

The characteristic polynomial of an $n \times n$ matrix A is

$$c_A(x) = \det(xI - A).$$

Example

The characteristic polynomial of $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$ is

$$\begin{aligned} c_A(x) &= \det\left(\begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} - \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}\right) \\ &= \det\begin{bmatrix} x-4 & 2 \\ 1 & x-3 \end{bmatrix} \\ &= (x-4)(x-3) - 2 \\ &= x^2 - 7x + 10 \end{aligned}$$

Theorem (Eigenvalues and Eigenvectors of a Matrix)

Let A be an $n \times n$ matrix.

1. The eigenvalues of A are the roots of $c_A(x)$.
2. The λ-eigenvectors $\vec{x}$ are the nontrivial solutions to $(\lambda I - A)\vec{x} = \vec{0}$.

Example (continued)

For $A = \left[\begin{array}{rr} 4 & -2 \\ -1 & 3 \end{array}\right]$, we have

$$c_A(x) = x^2 - 7x + 10 = (x-2)(x-5),$$

so A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$.

To find the 2-eigenvectors of A, solve $(2I - A)\vec{x} = \vec{0}$:

$$\left[\begin{array}{rr|r} -2 & 2 & 0 \\ 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -1 & 0 \\ -2 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Example (continued)

The general solution, in parametric form, is

$$\vec{x} = \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ where } t \in \mathbb{R}.$$

To find the 5-eigenvectors of A, solve $(5I - A)\vec{x} = \vec{0}$:

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The general solution, in parametric form, is

$$\vec{x} = \begin{bmatrix} -2s \\ s \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ where } s \in \mathbb{R}.$$

Definition

A basic eigenvector of an $n \times n$ matrix A is any nonzero multiple of a basic solution to $(\lambda I - A)\vec{x} = \vec{0}$, where λ is an eigenvalue of A.

Example (continued)

$\left[\begin{array}{r} 1 \\ 1 \end{array}\right]$ and $\left[\begin{array}{r} -2 \\ 1 \end{array}\right]$ are basic eigenvectors of the matrix

$$A = \left[\begin{array}{rr} 4 & -2 \\ -1 & 3 \end{array}\right]$$

corresponding to eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$, respectively.

Problem

For $A = \begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$, find $c_A(x)$, the eigenvalues of A, and the corresponding basic eigenvectors.

Solution

$$\det(xI - A) = \begin{vmatrix} x-3 & 4 & -2 \\ -1 & x+2 & -2 \\ -1 & 5 & x-5 \end{vmatrix} = \begin{vmatrix} x-3 & 4 & -2 \\ 0 & x-3 & -x+3 \\ -1 & 5 & x-5 \end{vmatrix}$$

$$= \begin{vmatrix} x-3 & 4 & 2 \\ 0 & x-3 & 0 \\ -1 & 5 & x \end{vmatrix} = (x-3) \begin{vmatrix} x-3 & 2 \\ -1 & x \end{vmatrix}$$

$$= (x-3)(x^2 - 3x + 2) = (x-3)(x-2)(x-1) = c_A(x).$$

Solution (continued)

Therefore, the eigenvalues of A are $\lambda_1 = 3, \lambda_2 = 2$, and $\lambda_3 = 1$.

Basic eigenvectors corresponding to $\lambda_1 = 3$: solve $(3\mathrm{I} - \mathrm{A})\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} 0 & 4 & -2 & 0 \\ -1 & 5 & -2 & 0 \\ -1 & 5 & -2 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\vec{x} = \left[\begin{array}{r} \frac{1}{2}t \\ \frac{1}{2}t \\ t \end{array}\right] = t \left[\begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{array}\right]$, $t \in \mathbb{R}$.

Choosing $t = 2$ gives us $\vec{x}_1 = \left[\begin{array}{c} 1 \\ 1 \\ 2 \end{array}\right]$ as a basic eigenvector corresponding to $\lambda_1 = 3$.

Solution (continued)

Basic eigenvectors corresponding to $\lambda_2 = 2$: solve $(2I - A)\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} -1 & 4 & -2 & 0 \\ -1 & 4 & -2 & 0 \\ -1 & 5 & -3 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\vec{x} = \begin{bmatrix} 2s \\ s \\ s \end{bmatrix} = s \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $s \in \mathbb{R}$.

Choosing $s = 1$ gives us $\vec{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ as a basic eigenvector corresponding to $\lambda_2 = 2$.

Solution (continued)

Basic eigenvectors corresponding to $\lambda_3 = 1$: solve $(I - A)\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} -2 & 4 & -2 & 0 \\ -1 & 3 & -2 & 0 \\ -1 & 5 & -4 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\vec{x} = \begin{bmatrix} r \\ r \\ r \end{bmatrix} = r \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $r \in \mathbb{R}$.

Choosing $r = 1$ gives us $\vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as a basic eigenvector corresponding to $\lambda_3 = 1$. ■

Geometric Interpretation of Eigenvalues and Eigenvectors

Let A be a 2×2 matrix. Then A can be interpreted as a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Problem

How does the linear transformation affect the eigenvectors of the matrix?

Definition

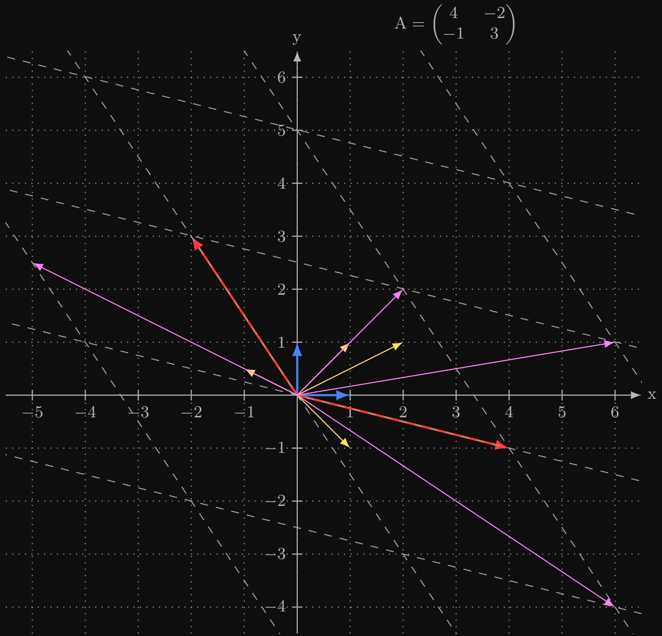
Let $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ be a nonzero vector in $\mathbb{R}^2$. Then $L_{\vec{v}}$ is the set of all scalar multiples of $\vec{v}$, i.e.,

$$L_{\vec{v}} = \mathbb{R}\vec{v} = \{t\vec{v} \mid t \in \mathbb{R}\}.$$

Example (revisited)

$A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$ has two eigenvalues: $\lambda_1 = 2$ and $\lambda_2 = 5$ with corresponding eigenvectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}$$



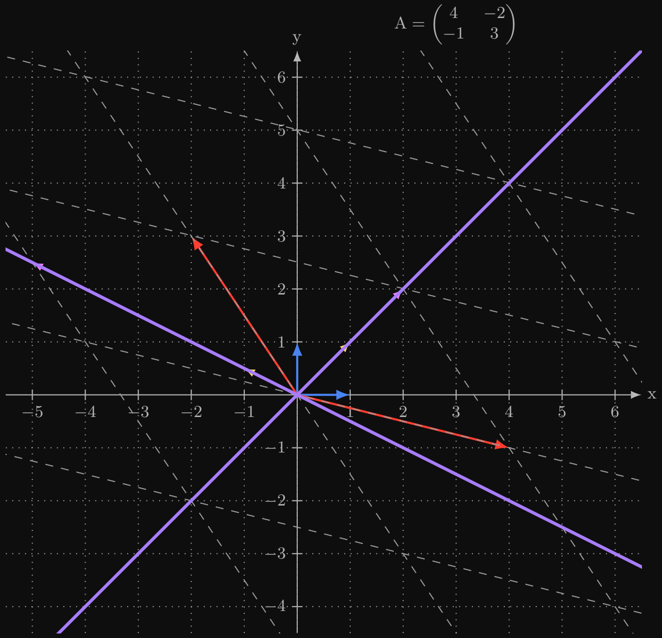
$A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$
y
x
6
5
4
3
2
1
−1
−2
−3
−4
−5
−4
−3
−2
−1
1
2
3
4
5
6

Definition

Let A be a 2×2 matrix and L a line in $\mathbb{R}^2$ through the origin. Then L is said to be A-invariant if the vector $A\vec{x}$ lies in L whenever $\vec{x}$ lies in L,
i.e., $A\vec{x}$ is a scalar multiple of $\vec{x}$,
i.e., $A\vec{x} = \lambda\vec{x}$ for some scalar $\lambda \in \mathbb{R}$,
i.e., $\vec{x}$ is an eigenvector of A.

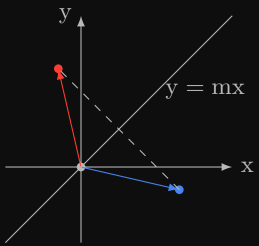
Theorem (A-Invariance)

Let A be a 2×2 matrix and let $\vec{v} \neq 0$ be a vector in $\mathbb{R}^2$. Then $L_{\vec{v}}$ is A-invariant if and only if $\vec{v}$ is an eigenvector of A.


$A = \begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix}$
y
x
−5 −4 −3 −2 −1 1 2 3 4 5 6
6 5 4 3 2 1 −1 −2 −3 −4

Problem

Let $m \in \mathbb{R}$ and consider the linear transformation $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$, i.e., reflection in the line $y = mx$.

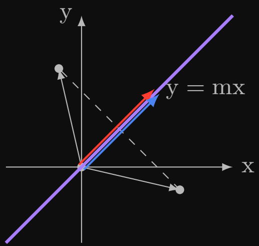


Recall that this is a matrix transformation induced by

$$A = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.$$

Find the lines that pass through origin and are A-invariant. Determine corresponding eigenvalues.

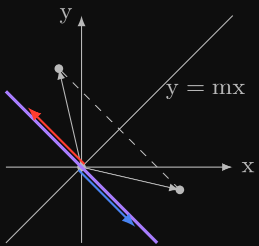
Solution



Let $\vec{x}_1 = \begin{bmatrix} 1 \\ m \end{bmatrix}$. Then $L_{\vec{x}_1}$ is A-invariant, that is, $\vec{x}_1$ is an eigenvector. Since the vector won't change, its eigenvalue should be 1. Indeed, one can verify that

$$A\vec{x}_1 = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{pmatrix} 1 \\ m \end{pmatrix} = \ldots = \begin{pmatrix} 1 \\ m \end{pmatrix} = \vec{x}_1.$$

Solution (continued)



Let $\vec{x}_2 = \begin{bmatrix} -m \\ 1 \end{bmatrix}$. Then $L_{\vec{x}_2}$ is A-invariant, that is, $\vec{x}_2$ is an eigenvector. Since the vector won't change the size, only flip the direction, its eigenvalue should be -1. Indeed, one can verify that

$$A\vec{x}_2 = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{pmatrix} -m \\ 1 \end{pmatrix} = \cdots = \begin{pmatrix} m \\ -1 \end{pmatrix} = -\vec{x}_2.$$

Example

Let θ be a real number, and $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ rotation through an angle of θ, induced by the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Claim: A has no real eigenvalues unless θ is an integer multiple of π, i.e., $\pm\pi, \pm 2\pi, \pm 3\pi$, etc.

Consequence: a line L in $\mathbb{R}^2$ is A invariant if and only if θ is an integer multiple of π.

Diagonalization

Denote an $n \times n$ diagonal matrix by

$$\mathrm{diag}(a_1, a_2, \ldots, a_n) = \begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & a_n \end{bmatrix}$$

Recall that if A is an $n \times n$ matrix and P is an invertible $n \times n$ matrix so that $P^{-1}AP$ is diagonal, then P is called a diagonalizing matrix of A, and A is diagonalizable.

- Suppose we have n eigenvalue-eigenvector pairs:

$$A\vec{x}_j = \lambda_j \vec{x}_j\,, \quad j = 1, 2, \ldots, n$$

- Pack the above n columns vectors into a matrix:

$$\left[\, A\vec{x}_1 \mid A\vec{x}_2 \mid \cdots \mid A\vec{x}_n \,\right] = \left[\, \lambda_1\vec{x}_1 \mid \lambda_2\vec{x}_2 \mid \cdots \mid \lambda_n\vec{x}_n \,\right]$$

$$||$$

$$A\left[\, \vec{x}_1 \mid \vec{x}_2 \mid \cdots \mid \vec{x}_n \,\right] \qquad\qquad ||$$

$$\left[\, \vec{x}_1 \mid \vec{x}_2 \mid \cdots \mid \vec{x}_n \,\right] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

- By denoting:

$$\mathrm{P} = \left[\ \vec{\mathrm{x}}_1 \mid \vec{\mathrm{x}}_2 \mid \cdots \mid \vec{\mathrm{x}}_\mathrm{n}\ \right] \quad \text{and} \quad \mathrm{D} = \mathrm{diag}\left(\lambda_1, \cdots, \lambda_\mathrm{n}\right)$$

  we see that

$$\mathrm{AP} = \mathrm{PD}$$

- Hence, provided P is invertible, we have

$$\mathrm{A} = \mathrm{PDP}^{-1} \quad \text{or equivalently} \quad \mathrm{D} = \mathrm{P}^{-1}\mathrm{AP}$$

  that is, A is diagonalizable.

Theorem (Matrix Diagonalization)

Let A be an $n \times n$ matrix.

1. A is diagonalizable if and only if it has eigenvectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ so that

   $$P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix}$$

   is invertible.
2. If P is invertible, then

   $$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

   where λ_i is the eigenvalue of A corresponding to the eigenvector $\vec{x}_i$, i.e., $A\vec{x}_i = \lambda_i \vec{x}_i$.

Example

$A = \begin{bmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{bmatrix}$ has eigenvalues and corresponding basic eigenvectors

$$\lambda_1 = 3 \quad \text{and} \quad \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix};$$

$$\lambda_2 = 2 \quad \text{and} \quad \vec{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix};$$

$$\lambda_3 = 1 \quad \text{and} \quad \vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Example (continued)

Let $P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$. Then P is invertible (check this!), so by the above Theorem,

$$P^{-1}AP = \mathrm{diag}(3, 2, 1) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Remark

It is not always possible to find n eigenvectors so that P is invertible.

Example

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

Then

$$c_A(x) = \begin{vmatrix} x-1 & 2 & -3 \\ -2 & x-6 & 6 \\ -1 & -2 & x+1 \end{vmatrix} = \cdots = (x-2)^3.$$

A has only one eigenvalue, $\lambda_1 = 2$, with multiplicity three. Sometimes, one writes

$$\lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Example (continued)

To find the 2-eigenvectors of A, solve the system $(2I - A)\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ -2 & -4 & 6 & 0 \\ -1 & -2 & 3 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The general solution in parametric form is

$$\vec{x} = \left[\begin{array}{c} -2s + 3t \\ s \\ t \end{array}\right] = s\left[\begin{array}{r} -2 \\ 1 \\ 0 \end{array}\right] + t\left[\begin{array}{r} 3 \\ 0 \\ 1 \end{array}\right], s, t \in \mathbb{R}.$$

Since the system has only two basic solutions, there are only two basic eigenvectors, implying that the matrix A is not diagonalizable.

Example

Diagonalize, if possible, the matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$.

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-1 & 0 & -1 \\ 0 & x-1 & 0 \\ 0 & 0 & x+3 \end{vmatrix} = (x-1)^2(x+3).$$

A has eigenvalues $\lambda_1 = 1$ of multiplicity two; $\lambda_2 = -3$ of multiplicity one.

Example (continued)

Eigenvectors for $\lambda_1 = 1$: solve $(I - A)\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\vec{x} = \left[\begin{array}{c} s \\ t \\ 0 \end{array}\right]$, $s, t \in \mathbb{R}$ so basic eigenvectors corresponding to $\lambda_1 = 1$ are

$$\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right], \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right]$$

Example (continued)

Eigenvectors for $\lambda_2 = -3$: solve $(-3\mathrm{I} - \mathrm{A})\vec{\mathrm{x}} = \vec{0}$.

$$\left[\begin{array}{rrr|r} -4 & 0 & -1 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\vec{\mathrm{x}} = \left[\begin{array}{c} -\frac{1}{4}\mathrm{t} \\ 0 \\ \mathrm{t} \end{array}\right]$, $\mathrm{t} \in \mathbb{R}$ so a basic eigenvector corresponding to $\lambda_2 = -3$ is

$$\left[\begin{array}{r} -1 \\ 0 \\ 4 \end{array}\right]$$

Example (continued)

Let

$$P = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & 0 & 0 \end{bmatrix}.$$

Then P is invertible, and

$$P^{-1}AP = \mathrm{diag}(-3, 1, 1) = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Theorem (Matrix Diagonalization Test)

A square matrix A is diagonalizable if and only if every eigenvalue λ of multiplicity m yields exactly m basic eigenvectors, i.e., the solution to $(\lambda I - A)\vec{x} = \vec{0}$ has m parameters.

A special case of this is:

Theorem (Distinct Eigenvalues and Diagonalization)

An $n \times n$ matrix with distinct eigenvalues is diagonalizable.

Example

Show that $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is not diagonalizable.

First,

$$c_A(x) = \begin{vmatrix} x-1 & -1 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & x-2 \end{vmatrix} = (x-1)^2(x-2),$$

so A has eigenvalues $\lambda_1 = 1$ of multiplicity two; $\lambda_2 = 2$ (of multiplicity one).

Example (continued)

Eigenvectors for $\lambda_1 = 1$: solve $(I - A)\vec{x} = \vec{0}$.

$$\left[\begin{array}{rrr|r} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, $\vec{x} = \begin{bmatrix} s \\ 0 \\ 0 \end{bmatrix}$, $s \in \mathbb{R}$.

Since $\lambda_1 = 1$ has multiplicity two, but has only one basic eigenvector, A is not diagonalizable.

Linear Dynamical Systems

Definition

A linear dynamical system consists of

- an $n \times n$ matrix A and an n-vector $\vec{v}_0$;
- a matrix recursion defining $\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots$ by $\vec{v}_{k+1} = A\vec{v}_k$; i.e.,

$$\begin{array}{ccl} \vec{v}_1 & = & A\vec{v}_0 \\ \vec{v}_2 & = & A\vec{v}_1 = A(A\vec{v}_0) = A^2\vec{v}_0 \\ \vec{v}_3 & = & A\vec{v}_2 = A(A^2\vec{v}_0) = A^3\vec{v}_0 \\ \vdots & \vdots & \vdots \\ \vec{v}_k & = & A^k\vec{v}_0. \end{array}$$

Remark

Linear dynamical systems are used, for example, to model the evolution of populations over time.

If A is diagonalizable, then

$$P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of A.

Thus $A = PDP^{-1}$, and $A^k = PD^kP^{-1}$. Therefore,

$$\vec{v}_k = A^k\vec{v}_0 = PD^kP^{-1}\vec{v}_0.$$

Example

Consider the linear dynamical system $\vec{v}_{k+1} = A\vec{v}_k$ with

$$A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}, \quad \text{and} \quad \vec{v}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Find a formula for $\vec{v}_k$.

First, $c_A(x) = (x-2)(x+1)$, so A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -1$, and thus is diagonalizable.

Solve $(2I - A)\vec{x} = \vec{0}$:

$$\left[\begin{array}{cc|c} 0 & 0 & 0 \\ -3 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

has general solution $\vec{x} = \begin{bmatrix} s \\ s \end{bmatrix}$, $s \in \mathbb{R}$, and basic solution $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Example (continued)

Solve $(-I - A)\vec{x} = \vec{0}$:

$$\left[\begin{array}{rr|r} -3 & 0 & 0 \\ -3 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

has general solution $\vec{x} = \left[\begin{array}{c} 0 \\ t \end{array}\right]$, $t \in \mathbb{R}$, and basic solution $\vec{x}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$.

Thus, $P = \left[\begin{array}{rr} 1 & 0 \\ 1 & 1 \end{array}\right]$ is a diagonalizing matrix for A,

$$P^{-1} = \left[\begin{array}{rr} 1 & 0 \\ -1 & 1 \end{array}\right], \quad \text{and} \quad P^{-1}AP = \left[\begin{array}{rr} 2 & 0 \\ 0 & -1 \end{array}\right].$$

Example (continued)

Therefore,

$$\begin{aligned}
\vec{v}_k &= A^k \vec{v}_0 \\
&= PD^kP^{-1}\vec{v}_0 \\
&= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}^k \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^k & 0 \\ 0 & (-1)^k \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 2^k & 0 \\ 2^k & (-1)^k \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 2^k \\ 2^k - 2(-1)^k \end{bmatrix}
\end{aligned}$$

Remark

Often, instead of finding an exact formula for $\vec{v}_k$, it suffices to estimate $\vec{v}_k$ as k gets large.

This can easily be done if A has a dominant eigenvalue with multiplicity one: an eigenvalue λ_1 with the property that

$$|\lambda_1| > |\lambda_j| \text{ for } j = 2, 3, \ldots, n.$$

Suppose that

$$\vec{v}_k = PD^kP^{-1}\vec{v}_0,$$

and assume that A has a dominant eigenvalue, λ_1, with corresponding basic eigenvector $\vec{x}_1$ as the first column of P.
For convenience, write $P^{-1}\vec{v}_0 = \left[\begin{array}{cccc} b_1 & b_2 & \cdots & b_n \end{array}\right]^T$.

Then

$$
\begin{aligned}
\vec{v}_k &= PD^kP^{-1}\vec{v}_0 \\
&= \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\
&= b_1\lambda_1^k\vec{x}_1 + b_2\lambda_2^k\vec{x}_2 + \cdots + b_n\lambda_n^k\vec{x}_n \\
&= \lambda_1^k\left(b_1\vec{x}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\vec{x}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\vec{x}_n\right)
\end{aligned}
$$

Now, $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$ for $j = 2, 3, \ldots n$, and thus $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$.

Therefore, for large values of k, $\vec{v}_k \approx \lambda_1^k b_1\vec{x}_1$.

Example

If

$$A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}, \quad \text{and} \quad \vec{v}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

estimate $\vec{v}_k$ for large values of k.

In our previous example, we found that A has eigenvalues 2 and -1. This means that $\lambda_1 = 2$ is a dominant eigenvalue; let $\lambda_2 = -1$.

As before $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a basic eigenvector for $\lambda_1 = 2$, and $\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a basic eigenvector for $\lambda_2 = -1$, giving us

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Example (continued)

$$P^{-1}\vec{v}_0 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

For large values of k,

$$\vec{v}_k \approx \lambda_1^k b_1 \vec{x}_1 = 2^k(1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2^k \\ 2^k \end{bmatrix}$$

Let's compare this to the formula for $\vec{v}_k$ that we obtained earlier:

$$\vec{v}_k = \begin{bmatrix} 2^k \\ 2^k - 2(-1)^k \end{bmatrix}$$