Module 9: Linear Algebra III

Operational math bootcamp

Statistical Sciences
UNIVERSITY OF TORONTO

University of Toronto

July 22, 2025

Outline

- Adjoints, unitaries and orthogonal matrices
- Orthogonal decomposition
- Spectral theory
  - Eigenvalues and eigenvectors
  - Algebraic and geometric multiplicity of eigenvalues
  - Matrix diagonalization

Recall

Definition

Let V be an $\mathbb{F}$-vector space. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is called *inner product* on V if the following holds:

1. (Conjugate) symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ for all $\mathbf{x}, \mathbf{y} \in V$, where $\bar{a}$ denotes the complex conjugate for $a \in \mathbb{C}$
2. Linearity in the first argument: $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{F}$
3. Positive definiteness: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$

A vector space equipped with an inner product is called an *inner product space*.

Recall

Example

- Standard inner product on $\mathbb{R}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

  $= x^T y$

- Standard inner product on $\mathbb{C}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i \bar{y}_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

  $= x^T \bar{y}$

- On the space of polynomials $\mathbb{P}_n(\mathbb{R})$: $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x$ for $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{P}_n(\mathbb{R})$

Proposition

Let V be an inner product space. Then C-S inequality

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}$$

for all $\mathbf{x}, \mathbf{y} \in V$.

- Inner product $\mathbb{R}^n$
  Discrete version for C-S
- $\mathbb{P}_n(\mathbb{R})$
  Continuous version

Proposition

Let V be an inner product space. Then $\langle \cdot, \cdot \rangle$ induces a norm on V via $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ for all $\mathbf{x} \in V$.

Proof.

Inner product $\underset{\not\Leftarrow}{\longrightarrow}$ norm $\underset{\not\Leftarrow}{\longrightarrow}$ metric

Note: With this identification the Cauchy-Schwarz inequality can be restated as: $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$.

Example

The norm introduced by the standard inner product on $\mathbb{R}^n$ is the Euclidean distance.

Adjoint

$U, \langle , \rangle_U$ $V, \langle , \rangle_V$

Definition

Let U, V be inner product spaces and $S\colon U \to V$ be a linear map. The *adjoint* S^* of S is the linear map $S^*\colon V \to U$ defined such that

$$\langle S\mathbf{u}, \mathbf{v}\rangle_V = \langle \mathbf{u}, S^*\mathbf{v}\rangle_U \qquad \text{for all } \mathbf{u} \in U, \mathbf{v} \in V.$$

$Su, v \in V$, $u, S^*v \in U$

$S: U \to V$

we've defined adjoint S^*.

- Does such S^* exist?
- Is it unique or not?

Statistical Sciences
UNIVERSITY OF TORONTO

Proposition

Let U, V be inner product spaces and $S \colon U \to V$ be a linear map. Then S^* is unique and linear.

Proof. Assume T_1 and T_2 are both adjoint linear map.

$\forall u \in U, v \in V$

$$\langle Su, v\rangle_V = \langle u, T_1 v\rangle_U$$
$$= \langle u, T_2 v\rangle_U$$

since both T_1 and T_2 are adjoint

Then, $\langle u, T_1 v - T_2 v\rangle_U = 0$ by the linearity.

$\forall u, v$

$\Rightarrow T_1 v - T_2 v = 0 \Rightarrow T_1 v = T_2 v, \ \forall v \in V$ //.

Since the adjoint is unique, we denote it with S^*.

To show S^* is linear, we need to show

$$S^*(\alpha v_1 + \beta v_2) = \alpha S^* v_1 + \beta S^* v_2.$$

$$\langle u, S^*(\alpha v_1 + \beta v_2)\rangle_U \overset{\text{def of adjoint}}{=} \langle Su, \alpha v_1 + \beta v_2\rangle_V$$

$$= \bar{\alpha}\langle Su, v_1\rangle_V + \bar{\beta}\langle Su, v_2\rangle_V$$

$$\overset{\text{adjoint}}{=} \bar{\alpha}\langle u, S^* v_1\rangle_U + \bar{\beta}\langle u, S^* v_2\rangle_U$$

linearity and conjugate symmetry

$$= \langle u, \alpha S^* v_1 + \beta S^* v_2\rangle_U$$

$\forall u, v_1, v_2$, $\langle u, S^*(\alpha v_1 + \beta v_2)\rangle_U = \langle u, \alpha S^* v_1 + \beta S^* v_2\rangle_U$

By the uniqueness we've proved, we have

$$S^*(\alpha v_1 + \beta v_2) = \alpha S^* v_1 + \beta S^* v_2.$$

$= x^T y$

Example

Define $S\colon \mathbb{R}^3 \to \mathbb{R}^2$ by $S\mathbf{x} = (2x_1 + x_3, -x_2)$. What is the adjoint operator S^*?

$x \in \mathbb{R}^3$, $y \in \mathbb{R}^2$

$$\langle Sx, y \rangle = \left\langle \begin{pmatrix} 2x_1 + x_3 \\ -x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle$$

$$= (2x_1 + x_3) y_1 + (-x_2) y_2.$$

$$\overset{?}{=} \langle x, S^* y \rangle = x^T (S^* y)$$

$$(2x_1 + x_3) y_1 + (-x_2) y_2 = (2y_1) x_1 + (-y_2) \cdot x_2 + (y_1) \cdot x_3.$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} (2y_1, -y_2, y_1).$$

$S^* y = (2y_1, -y_2, y_1)$.

$$S^* = \begin{pmatrix} 2 & 0 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}_{3\times 2} \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{2\times 1}$$

Standard Inner Product

$T_A(x) = Ax$

adjoint operator

Proposition

Let $A \in M_{m\times n}(\mathbb{F})$ be a matrix and $T_A \colon \mathbb{F}^n \to F^m \colon \mathbf{x} \mapsto A\mathbf{x}$. Then, $T_A^*(\mathbf{x}) = A^*\mathbf{x}$, where $A^* \in M_{n\times m}(\mathbb{F})$ with $(A^*)_{ij} = \overline{A_{ji}}$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

In particular, if $\mathbb{F} = \mathbb{R}$, the adjoint of the matrix is given by its transpose,denoted A^T, and if $\mathbb{F} = \mathbb{C}$, it is given by its conjugate transpose, denoted A^*.

$\Rightarrow A^* = \bar{A}^T = \overline{A^T}$

Proof for $\mathbb{R}$: $\langle T_A(x), y \rangle = \langle x, T_A^*(y) \rangle$.

$T_A(x) = Ax \Rightarrow \langle T_A(x), y \rangle = \langle Ax, y \rangle = (Ax)^T y$

$= x^T A^T y$

$= \langle x, A^T y \rangle$

$A^T y = T_A^*(y)$

For $\mathbb{F} = \mathbb{C}$, Same logic.

Inverse = transpose, $O^TO = OO^T = I$

Definition

A matrix $O \in M_n(\mathbb{R})$ is called *orthogonal* if its inverse is given by its transpose, i.e. $O^T O = OO^T = I$.

A matrix $U \in M_n(\mathbb{C})$ is called *unitary* if the inverse is given by the conjugate transpose, i.e. $U^* U = UU^* = I$.

$U^* = \bar{U}^T = \overline{U^T}$

Let U be unitary, standard inner product in $\mathbb{C}^n$

$U^*U = I$

$\langle Ux, Uy \rangle = \langle x, U^*Uy \rangle = \langle x, y \rangle$

U does not change inner product, norm. etc.

Example

- Let $\varphi \in [0, 2\pi]$. Then

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad AA^T = I$$

is an orthogonal matrix. What does it describe geometrically? Rotation by φ.

- The following is a unitary matrix:

$$A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

$$A^* = \bar{A}^T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad A^*A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$A^* = \bar{A}^T$

Definition

Let $A \in M_n(\mathbb{F})$. We call A *self-adjoint* if $A^* = A$. In the case $\mathbb{F} = \mathbb{R}$, such an A is called *symmetric* and if $\mathbb{F} = \mathbb{C}$, such an A is called *Hermitian*.

$A^* = A \Rightarrow A^T = A$

Orthogonality and Gram-Schmidt

norm derived by the inner product

Definition

Two vectors $\mathbf{x}, \mathbf{y} \in V$ are called *orthogonal* if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, denoted $\mathbf{x} \perp \mathbf{y}$. We call them *orthonormal* if additionally the vectors are normalized, i.e. $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$. A basis $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of V is called *orthonormal basis (ONB)*, if the vectors are pairwise orthogonal and normalized.

E.g. in $\mathbb{R}^n$ and $\mathbb{C}^n$

$$e_i := (0, \cdots, \underset{\substack{\uparrow \\ i\text{-th}}}{1}, \cdots 0)^T, \quad \{e_i\}_{i=1}^n \text{ forms ONB}$$

Statistical Sciences
UNIVERSITY OF TORONTO

Proposition

Let $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ be orthonormal. Then the system of vectors is linearly independent.

Proof.

$$\begin{cases} \forall i, j, & x_i \perp x_j \iff \langle x_i, x_j \rangle = 0 \\ \forall i, & \|x_i\| = \sqrt{\langle x_i, x_i \rangle} = 1 \end{cases}$$

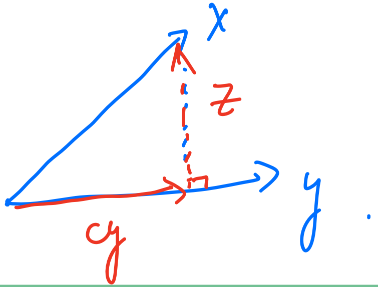
Assume $\sum_{i=1}^{n} \alpha_i x_i = 0$ (If α_i has only one solution 0 $\Rightarrow$ linearly independent).

$$0 = \left\langle \sum_{i=1}^{n} \alpha x_i, \sum_{i=1}^{n} \alpha x_i \right\rangle$$

$$= \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} \underbrace{\langle x_i, x_i \rangle}_{1} + \sum_{i \neq j} \alpha_i \overline{\alpha_j} \underbrace{\langle x_i, x_j \rangle}_{0}$$

$$= \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} = \sum_{i=1}^{n} \|\alpha_i\|^2, \quad \|\alpha_i\|^2 \geq 0$$

$$\alpha_i = 0, \ \forall i \Rightarrow \sim$$



$x = cy + z, \quad cy \perp z$

Proposition (Orthogonal Decomposition)

Let $\mathbf{x}, \mathbf{y} \in V$ with $\mathbf{y} \neq 0$. Then, there exist $c \in F$ and $\mathbf{z} \in V$ such that $\mathbf{x} = c\mathbf{y} + \mathbf{z}$ with $\mathbf{y} \perp \mathbf{z}$.

Given a basis, we can obtain an ONB from it using the Gram-Schmidt algorithm by repeating this orthogonal decomposition.

proof: Let $c = \dfrac{\overline{\langle x, y\rangle}}{\|y\|^2}$, $z = x - cy = x - \dfrac{\overline{\langle x, y\rangle}}{\|y\|^2} \cdot y$.

Next $\langle y, z\rangle = \langle y, x - cy\rangle = \langle y, x\rangle - \bar{c}\langle y, y\rangle$

$$= \langle y, x\rangle - \frac{\overline{\langle x, y\rangle}}{\|y\|^2} \cdot \langle y, y\rangle = \langle y, x\rangle - \underbrace{\overline{\langle x, y\rangle}}_{= \langle y, x\rangle} = 0$$

Proposition (Gram-Schmidt Algorithm)

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$ be a system of linearly independent vectors. Define $\mathbf{y}_1 = \mathbf{x}_1 / \|\mathbf{x}_1\|$. For $i = 2, \ldots, n$ define $\mathbf{y}_i$ inductively by

y_i

$$\mathbf{y}_i = \frac{\mathbf{x}_i - \sum_{k=1}^{i-1} \langle \mathbf{x}_i, \mathbf{y}_k \rangle \mathbf{y}_k}{\|\mathbf{x}_i - \sum_{k=1}^{i-1} \langle \mathbf{x}_i, \mathbf{y}_k \rangle \mathbf{y}_k\|}.$$

Then the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are orthonormal and

$$\mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} = \mathrm{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}.$$

The proof is omitted but can be found in the book.

Intuition: generalization of previous page.

Recall: connection between matrices and linear maps

$$\text{Matrix} \xRightarrow{\text{define.}} \text{linear map}$$

Multiplication by a matrix defines a linear map

Let $A \in M_{m\times n}$ be a fixed matrix. Then, we can define a linear map $T_A : \mathbb{F}^n \to \mathbb{F}^m$ via $T_A(\mathbf{v}) = A\mathbf{v}$, where we recall matrix vector multiplication $(A\mathbf{v})_i = \sum_{k=1}^n A_{ik}v_k$ for $i = 1, \dots, m$.

$$\text{linear map} \Longrightarrow \text{Matrix}$$

Given a bases for U and V, $T : U \to V$ can be written as a matrix

Let $T \in \mathcal{L}(U, V)$ where U and V are vector spaces. Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ be bases for U and V respectively. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ with entries A_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$ defined by

$$T\mathbf{u}_k = A_{1k}\mathbf{v}_1 + \cdots + A_{mk}\mathbf{v}_m.$$

Eigenvalues

linear map (matrix).

Definition

Given an operator $A\colon V \to V$ and $\lambda \in \mathbb{F}$, λ is called an *eigenvalue* of A if there exists a non-zero vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ such that

operator

$$A\mathbf{v} = \lambda \mathbf{v}.$$

scalar.

We call such $\mathbf{v}$ an *eigenvector* of A with eigenvalue λ. We call the set of all eigenvalues of A *spectrum* of A and denote it by $\sigma(A)$.

Motivation in terms of linear maps: Let $T\colon V \to V$ be a linear map, where V is a vector space. We would like to describe the action of this linear map in a particularly "nice" way: such that T acts only by scaling, i.e. $T\mathbf{v}_i = \lambda_i \mathbf{v}_i$ where $\lambda_i \in \mathbb{F}$ for $i = 1, \ldots, n$.

Finding eigenvalues

"Assume A is linear map"

Note: here we will assume $\mathbb{F} = \mathbb{C}$, so that we are working on an algebraically closed field.

Identity operator.

- Rewrite $A\mathbf{v} = \lambda\mathbf{v}$ as $(A-\lambda I)v=0 \iff v \in \ker(A-\lambda I)$
- Thus, if λ is an eigenvalue, we can find the corresponding eigenvectors by finding the null space of $A - \lambda I$.
- The subspace null$(A - \lambda I)$ is called the *eigenspace*
- To find the eigenvalues of A, one must find the scalars λ such that null$(A - \lambda I)$ contains non-trivial vectors (i.e. not $\mathbf{0}$)

  λ is eigenvalues $\iff$ $A-\lambda I$ is not injective
- Recall: We saw that $T \in \mathcal{L}(U, V)$ is injective if and only if null $T = \{\mathbf{0}\}$.
- Thus λ is an eigenvalue if and only if $A - \lambda I$ is not invertible.
- Recall: $|A| \neq 0$ if and only if A is invertible.
- Thus λ is an eigenvalue if and only if

  $|A-\lambda I|=0$

Recall $A-\lambda I$ is a special linear map s.t. $U=V$

$A-\lambda I$ is surjective

Idea: eigenvalue is a abstract concept, we want connect it with a computable criteria.

T is invertible $\Leftrightarrow$ T is injective

Theorem

The following are equivalent

1. $\lambda \in \mathbb{F}$ *is an eigenvalue of* A,
2. $(A - \lambda I)\mathbf{v} = 0$ *has a non-trivial solution,* directly from def
3. $|A - \lambda I| = 0$.

if λ ..., $Av = \lambda v \Rightarrow (A - \lambda I)v = 0$.

Statistical Sciences
UNIVERSITY OF TORONTO

Characteristic polynomial

$|A - \lambda I|$, it can be viewed as a function λ.

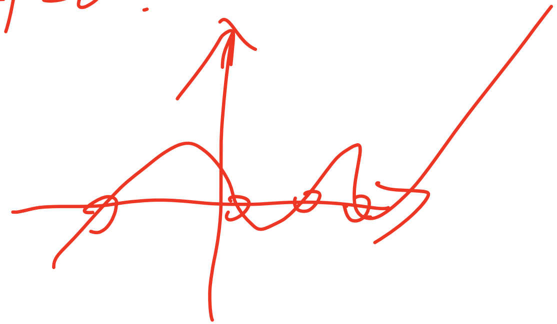
Definition

If A is an $n \times n$ matrix, $p_A(\lambda) = |A - \lambda I|$ is a polynomial of degree n called the *characteristic polynomial* of A.

To find the eigenvectors of A, one needs to find the roots of the characteristic polynomial.

To find eigenvalues, we need to find λ s.t. $|A - \lambda I| = 0$.

$\Leftrightarrow$ find λ s.t. $p_A(\lambda) = 0$.



Example

Find the eigenvalues of

$$A = \begin{bmatrix} 4 & -2 \\ 5 & -3 \end{bmatrix}.$$

$$|A - \lambda I| = \begin{vmatrix} 4-\lambda & -2 \\ 5 & -3-\lambda \end{vmatrix} = (4-\lambda)(-3-\lambda) - (-2)\cdot 5$$

$$= \lambda^2 - \lambda - 12 + 10 = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$$

$$\Rightarrow \lambda = -1, \ 2$$

Multiplicity

$p(\lambda) = \lambda^{2} (\lambda - 1)^{3} (\lambda - 5)^{4}.$

Definition

The multiplicity of the root λ in the characteristic polynomial is called the *algebraic multiplicity* of the eigenvalue λ. The dimension of the eigenspace $\mathrm{null}(A - \lambda I)$ is called the *geometric multiplicity* of the eigenvalue λ.

# of vectors contains in

the basis of such subspaces.

Definition (Similar matrices)

Square matrices A and B are called *similar* if there exists an invertible matrix S such that

$n \times n$ $n \times n$.

$$A = SBS^{-1}. \iff S^{-1}AS = B.$$

Similar matrices have the same characteristic polynomials and hence the same eigenvalues (see exercise).

Theorem

Suppose A is a square matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors corresponding to these eigenvalues. Then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

Proof. Prove by induction.

① $n=1$, we only have one v_1, trivially.

② Suppose $v_1, \cdots v_k$, are linearly independent

③ prove $\{v_1, \cdots v_k, v_{k+1}\}$ is also linearly independent.

Let $\sum_{i=1}^{k+1} \alpha_i v_i = 0$ (1)

Multiply A to both sides.

$\sum_{i=1}^{k+1} \alpha_i A v_i = 0 \iff \sum_{i=1}^{k+1} \alpha_i a_i v_i = 0$ (2)

Next, multiply λ_{k+1} to (1)

$$\sum_{i=1}^{k+1} \alpha_i \lambda_{k+1} v_i = 0 \qquad (3)$$

By (2) and (3)

$$\sum_{i=1}^{k+1} \alpha_i (\lambda_{k+1} - \lambda_i) v_i = 0 \iff \sum_{i=1}^{k} \alpha_i (\lambda_{k+1} - \lambda_i) v_i = 0. \qquad (4)$$

By induction hypothesis, $v_1, \cdots v_k$ are linearly independent.

Thus, $\alpha_i \underbrace{(\lambda_{k+1} - \lambda_i)}_{\neq 0 \text{ by assumption}} = 0$, $\forall i = 1, \cdots k$.

$\Rightarrow \alpha_i = 0, \quad \forall i = 1, \cdots, k.$

Then, $\sum_{i=1}^{k+1} \alpha_i v_i = 0 \Rightarrow \alpha_{k+1} \cdot v_{k+1} = 0 \Rightarrow \alpha_{k+1} = 0$

$\Rightarrow \alpha_i = 0, \quad \forall i = 1, \cdots, k+1$

$\Rightarrow v_1, v_2, \cdots, v_{k+1}$ are ~~~

$A_{n\times n}$ there is not multiplicity,

Corollary

If a $A \in M_n(\mathbb{C})$ *has* n *distinct eigenvalues, then* A *is diagonalizable. That is there exists an invertible matrix* $S \in M_n(\mathbb{C})$ *such that* $A = SDS^{-1}$*, where* D *is a diagonal matrix with the eigenvalues of* A *in the diagonal.*

$\Leftrightarrow \quad S^{-1}AS = D$

Let V_i and λ_i be eigenvalues and eigen vectors repestively.

Let $S = (v_1, v_2, \cdots v_n)$, $D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots \lambda_n)$

is invertible since $v_1, \cdots, v_n$ are linearly independent.

$A \cdot S = (Av_1, Av_2 \cdots Av_n) = (\lambda_1 v_1, \cdots, \lambda_n v_n) = SD.$

S is invertible $\Longrightarrow$ $S^{-1}AS = D$ or $A = SDS^{-1}$

Theorem

Let $A : V \to V$ be an operator with n eigenvalues. A is diagonalizable if and only if for each eigenvalue λ, the geometric multiplicity of λ and the algebraic multiplicity of λ are the same.

$\ker(A - \lambda I)$.

$p_A(\lambda)$.

Example: a diagonalizable matrix

$$A = \begin{bmatrix} 1 & 2 \\ 8 & 1 \end{bmatrix} \text{ is diagonalizable.}$$

$$p_A(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 8 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 1 - 16 = \lambda^2 - 2\lambda - 15$$

$$= \underbrace{(\lambda - 5)(\lambda + 3)}$$

distinct eigenvalues

$\Rightarrow$ diagonalizable.

Example continued

Eigenvectors.

$\lambda = -3$, $\underbrace{\begin{pmatrix} 4 & 2 \\ 8 & 4 \end{pmatrix}}_{A-\lambda I} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. $\Leftrightarrow$ $2x + y = 0$.

eigenvectors: $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$

$\lambda = 5$, $\begin{pmatrix} -4 & 2 \\ 8 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\Leftrightarrow$ $2x - y = 0$

eigenvectors: $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Example continued

$$S = \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix} \Rightarrow S^{-1} = \frac{1}{4}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}.$$

$$\lambda = -3 \quad 5$$

$$D = \begin{pmatrix} -3 & 0 \\ 0 & 5 \end{pmatrix}$$

$$S^{-1}AS = \frac{1}{4}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}$$

$$= \frac{1}{4}\begin{pmatrix} 2-8=-6 & 4-1=3 \\ 2+8=10 & 4+1=5 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}$$

$$= \frac{1}{4}\begin{pmatrix} -12 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ 0 & 5 \end{pmatrix}$$

$$= D$$

Example: a matrix that is not diagonalizable

① $p_A(\lambda)$

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

② algebraic multiplicity $\neq$ geometric ~.

$\lambda = 1$ is a eigenvalues.

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (\lambda-1)^2 \Rightarrow \text{Algebraic} \sim = 2$$

$$(A - \lambda I) = (A - I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \ker(A - I). \iff$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$\iff y = 0 \qquad \Rightarrow \dim \ker(A - I) = 1$$

$$\Rightarrow \text{geometric} \sim = 1$$

References

Howard Anton and Chris Rorres. Elementary Linear Algebra. 11th ed. Wiley, 2014

Axler S. *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Springer, 2015. Available from: https://link.springer.com/book/10.1007/978-3-319-11080-6

Treil S. *Linear Algebra Done Wrong*. 2017. Available from: https://www.math.brown.edu/streil/papers/LADW/LADW.html