Lecture I213E – Class 3

Discrete Signal Processing

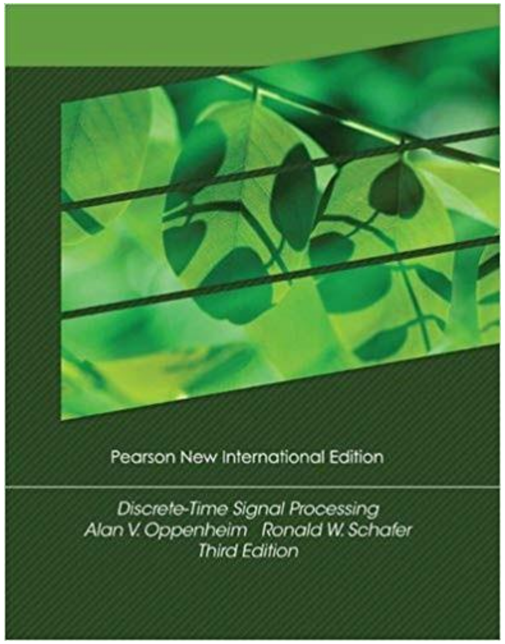
Sakriani Sakti

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY
1990

Course Materials

- ## Materials

→ Lecture notes will be uploaded before each lecture
https://jstorage-2018.jaist.ac.jp/s/PGXRrC7iFmN2FWo
Pass: dsp-i213e-2022
(Slide Courtesy of Prof. Nak Young Chong)



- ## References

→ Chi-Tsong Chen:
Linear System Theory and Design, 4th Ed., Oxford University Press, 2013.

→ Alan V. Oppenheim and Ronald W. Schafer:
Discrete-Time Signal Processing, 3rd Ed., Pearson New International Ed., 2013.

Related Courses & Prerequisite

- **Related Courses**
  - → I212 Analysis for Information Science
  - → I114 Fundamental Mathematics for Information Science

- **Prerequisite**
  - → None

Evaluation

- ## Viewpoint of evaluation

→ Students are able to understand:

- Basic principles in modeling and analysis of linear time-invariant systems
- Applications of mathematical methods and tools to different signal processing problems.

- ## Evaluation method

→ Homework, term project, midterm exam, and final exam

- ## Evaluation criteria

→ Homework/labs (30%), term project (30%)
midterm exam (15%), and final exam (25%)

Contact

- **Lecturer**

  → Sakriani Sakti

- **TA**

  Tutorial hours & Term project

  → WANG Lijun (s2010026)

  → TANG Bowen (s2110411)

  Homework

  → PUTRI Fanda Yuliana (s2110425)

- **Contact Email**

  → dsp-i213e-2022@ml.jaist.ac.jp

Schedule

- **December 8th, 2022 – February 9th, 2023**

- **Lecture Course Term 2-2**
  - → Tuesday 9:00 — 10:40
  - → Thursday 10:50 — 12:30

- **Tutorial Hours**
  - → Tuesday 13:30-15:10

Schedule

Dec

Sun	Mon	Tue	Wed	Thu	Fri	Sat
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Jan

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	2	3	4	~~5~~	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	~~24~~	25	26	27	28
29	30	31				

Feb

Sun	Mon	Tue	Wed	Thu	Fri	Sat
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28				

Lecture:
Tuesday 9:00 — 10:40
Thursday 10:50 — 12:30

Tutorial:
Tuesday 13:30 — 15:10

Course review &
term project evaluation
(on tutorial hours)

Midterm & final exam
Thursday 10:50 — 12:30

Syllabus

Class	Date	Lecture Course Tue 9:00 — 10:40 / Thr 10:50 — 12:30	Tutorial Hours Tue 13:30 — 15:10
1	12/08	Introduction to Linear Systems with Applications to Signal Processing	
2	12/13	State Space Description	○
3	12/15	Linear Algebra	
4	12/20	Quantitative Analysis (State Space Solutions) and Qualitative Analysis (Stability)	○
5	12/22	Discrete-time Signals and Systems	
X	01/05		
6	01/10	Discrete-time Fourier Analysis	
7	01/10*	Review of Discrete-time Linear Time-Invariant Signals and Systems (on Tutorial Hours)	
	01/12	Midterm Exam	
8	01/17	Sampling and Reconstruction of Analog Signals	○
9	01/19	z-Transform	
X	01/24		○
10	01/26	Discrete Fourier Transform	
11	01/31	FFT Algorithms	○
12	01/02	Implementation of Digital Filters	
13	02/07	Digital Signal Processors and Design of Digital Filters	
14	02/07*	Review of the Course and Term Project Evaluation (on Tutorial Hours)	
	02/09	Final exam	

Class 3
Linear Algebra

Linear Algebra

- ## Linear Algebra
  1. Consists mostly of studying **matrix calculus**
  2. Formalizes and gives a **geometrical interpretation** of the resolution of equation systems
  3. Creates a formal link between **matrix calculus** and the use of **linear and quadratic transformations**

Continuous (discrete)-time linear time-invariant state-space models

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad x[k+1] = Ax[k] + Bu[k]$$

$$y(t) = Cx(t) + Du(t) \qquad y[k] = Cx[k] + Du[k]$$

Vector

Vector

Definition of Vector

$\rightarrow$ A vector x is a set of numbers

$\rightarrow$ Consider an n-dimensional real linear space, denoted by R^n

$\rightarrow$ Every vector in R^n is an n-tuple of real numbers such as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$\rightarrow$ The corresponding complex space consisting of n-tuple of complex numbers is denoted by C^n

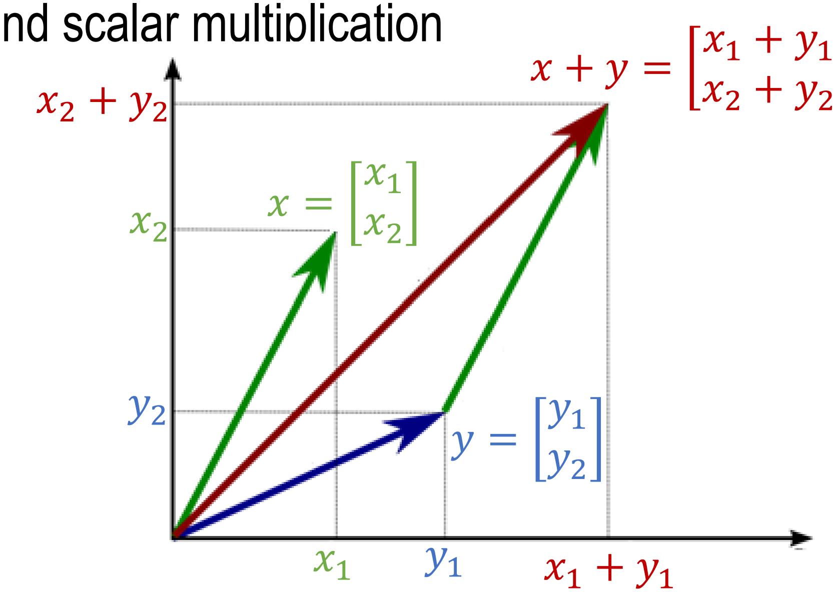
Vector Space

▪ Definition of Vector Space

→ A linear (or vector) space V is a set of elements called vectors that is closed under finite vector addition and scalar multiplication

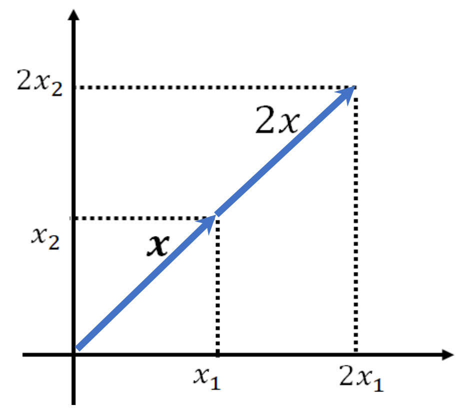
(1) **Vector addition**

Addition of any two vectors $x, y \in V$ is a vector $x + y \in V$



(2) **Scalar multiplication**

Multiplication of any vector $x \in V$ with any scalar α is a vector $\alpha x \in V$



Vector Space

▪ Definition of Vector Space

→ The operations of addition and scalar multiplication are assumed to satisfy the following axioms:

1. $x + y = y + x$ (*commutative law*)
2. $(x + y) + z = x + (y + z)$ (*associative law*)
3. There is a null vector $\mathbf{0}$ in V such that $x + \mathbf{0} = x$ for all x in V
4. $\alpha(x + y) = \alpha x + \alpha y$ (*distributive law*)
5. $(\alpha + \beta)x = \alpha x + \beta x$ (*distributive law*)
6. $(\alpha\beta)x = \alpha(\beta x)$ (*associative law*)
7. $0x = \mathbf{0}$
8. $1x = x$

Linear Dependent & Independent

Linear Dependent

→ A set of m vectors $\{x_1, x_2, \dots, x_m\}$ in R^n of vector space V is linear dependent if there exist scalars $\{a_1, a_2, \dots, a_m\}$, not all zero, such that

$$\boxed{a_1 x_1 + a_2 x_2 + \cdots + a_m x_m = 0}$$

→ Example:

$$a_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + a_2 \begin{bmatrix} -2 \\ -8 \end{bmatrix} = 0 \ \rightarrow \begin{bmatrix} 1 \\ 4 \end{bmatrix} \text{ and } \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$ are linearly dependent since they are multiples

Linear Independent

→ A set of m vectors $\{x_1, x_2, \dots, x_m\}$ in R^n of vector space V is linear independent

$$a_1 x_1 + a_2 x_2 + \cdots + a_m x_m = 0$$

only if

$$\boxed{a_1 = 0, a_2 = 0, \dots, a_m = 0,}$$

→ Example:

$$a_1 \begin{bmatrix} 9 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 18 \\ 6 \end{bmatrix} = 0 \rightarrow \begin{bmatrix} 9 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} 18 \\ 6 \end{bmatrix}$$ are linearly independent since they are not multiples

Linear Combination

Definition of Linear Combination

→ If the set of vectors is linearly dependent, then there exists at least one α_i, say, α_1, that is different from zero

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m = 0$$

$$x_1 = -\frac{1}{\alpha_1}[\alpha_2 x_2 + \cdots + \alpha_m x_m]$$

$$x_1 = \beta_2 x_2 + \cdots + \beta_m x_m \quad \text{where } \beta_k = -\frac{\alpha_k}{\alpha_1}$$

Such an expression is called a linear combination

$$a_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + a_2 \begin{bmatrix} -2 \\ -8 \end{bmatrix} = 0 \quad \rightarrow \quad 2 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} -2 \\ -8 \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$

→ The dimension of a linear space can be defined as the maximum number of linearly independent vectors in the space

Basis

Definition of Basis

→ A set of linearly independent vectors in R^n is called a basis B if every vector in R^n can be expressed as a unique combination of the set B

→ In R^n any set of n linearly independent vectors can be used as a basis

Let $\{q_1, q_2, \dots, q_n\}$ be such a set of basis
Then every vector x can be expressed uniquely as

$$x = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n$$

Define the $n \times n$ squared matrix $Q := [q_1, q_2, \dots, q_n]$

$$x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = Q\bar{x}$$

We call $\bar{x} = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ the representation of the vector x with respect to the basis $\{q_1, q_2, \dots, q_n\}$

Basis

- ## Orthonormal Basis

$\rightarrow$ We will associate with every R^n the following orthonormal basis:

$$i_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad i_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \cdots \quad , \; i_{n-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad i_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

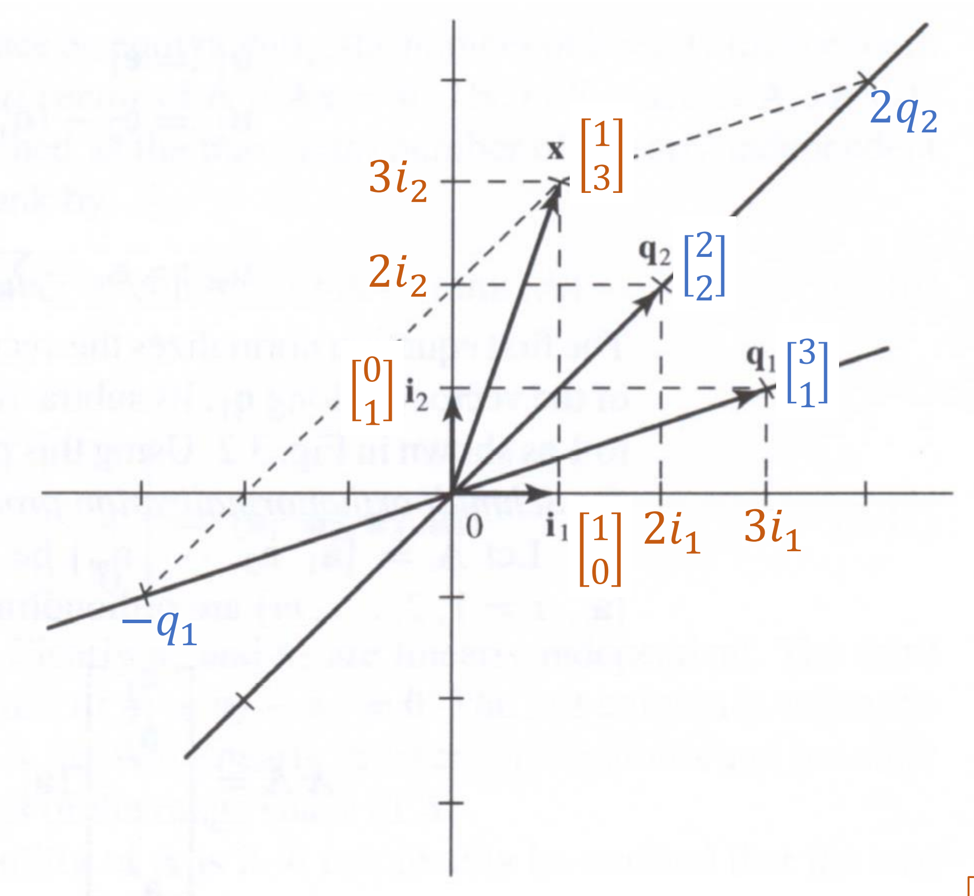
With respect to this basis, we have:

$$x := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 i_1 + x_2 i_2 + \cdots + x_n i_n = \underbrace{I_n}_{\substack{n \times n \\ \textit{unit matrix}}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Basis

- **Example:**
  → Different representation of vector x

$$x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



Representation of x
with respect to the basis $\{i_1, i_2\}$

$$x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Representation of x
with respect to the basis $\{q_1, q_2\}$

$$x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Find representation of x with basis $\{q_2, i_2\}$

Orthogonal and Orthonormal

Orthogonal

$\rightarrow$ Two vectors are orthogonal (perpendicular to each other) if their dot product is 0

$\rightarrow$ Example:

$$x = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \text{ and } y = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$x \cdot y = x_1 y_1 + x_2 y_2 = -4 + 4 = 0$$

$$x^T y = \begin{bmatrix} 4 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = x_1 y_1 + x_2 y_2 = -4 + 4 = 0$$

$$y^T x = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = y_1 x_1 + y_2 x_2 = -4 + 4 = 0$$

Orthonormal

$\rightarrow$ Two vectors are orthonormal if their dot product is 0 and their lengths are both 1

$$\text{Length} = \mathrm{norm}(x) = ||x|| = \sqrt{x \cdot x} = 1$$

$$x^T x = 1$$

Norm of Vectors

- **Norm**
  - → The concept of norm is a generalization of length or magnitude
  - → Any real-valued function of x, denoted by $\|x\|$ can be defined as a norm if it has the following properties:

1. $\|x\| \geq 0$ for every x and $\|x\| = 0$ if and only if $x=0$.

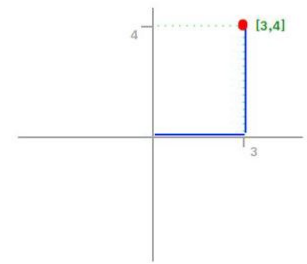
2. $\|\alpha x\| = |\alpha| \|x\|$, for any real α.

3. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ for every x_1 and x_2.
   the *triangular inequality*

Norm of Vectors

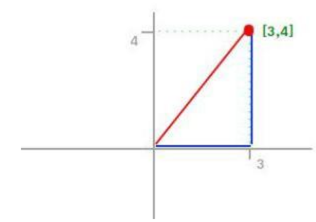
▪ Various Norms

Let $x = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$.



$$\|x\|_1 := \sum_{i=1}^{n} |x_i|$$

1-norm
the distance you have to travel between the origin (0,0) to the destination



$$\|x\|_2 := \sqrt{x^T x} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$$

2-norm
the shortest distance to go from one point to another (Euclidean norm)

$$\|x\|_\infty := \max_i |x_i|$$

infinity-norm
the largest magnitude among each element of a vector

Norm of Vectors

- **Orthonormalization**
  - → A vector x is said to be normalized if its Euclidean norm is 1 or

$$x^T x = 1$$

  - → A set of vectors x_i, $i = 1,2,\ldots,m$ is said to be orthonormal if

$$x_i^T x_j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

Constructing Orthonormal Basis

Schmidt Orthonormalization Procedure

→ Given a set of linearly independent vectors $e_1, e_2, \dots, e_m$ we can obtain an orthonormal set $q_1, q_2, \dots, q_m$ using this procedure

→ **Requirement:** $q_1, q_2, \dots, q_m$ must span the same space as $e_1, e_2, \dots, e_m$

→ **Procedure:**

Step 1: Take $u_1 := e_1$

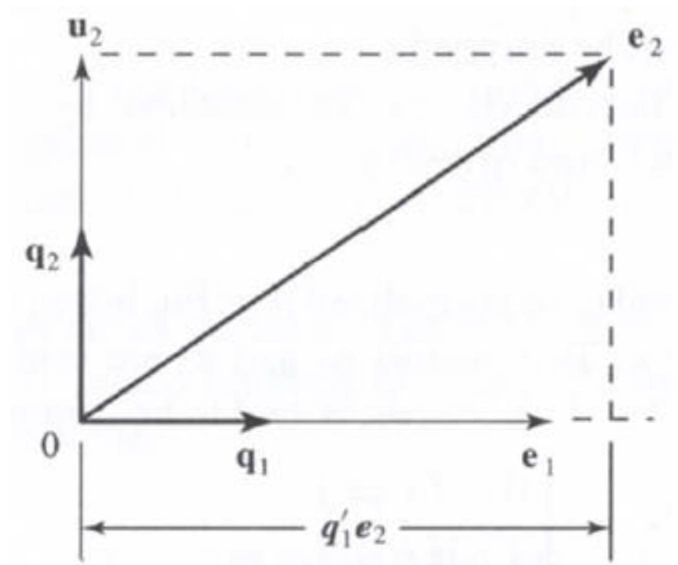
Normalized $q_1 = \frac{u_1}{\|u_1\|}$

Step 2: (a) Find the orthogonal basis:

Project e_2 along $q_1 \rightarrow ({q_1}^T e_2)q_1$

Compute residual $u_2 := e_2 - ({q_1}^T e_2)q_1$

(b) Normalized $q_2 = \frac{u_2}{\|u_2\|}$



Constructing Orthonormal Basis

Schmidt Orthonormalization Procedure

→ Given a set of linearly independent vectors $e_1, e_2, \dots, e_m$
we can obtain an orthonormal set $q_1, q_2, \dots, q_m$ using this procedure

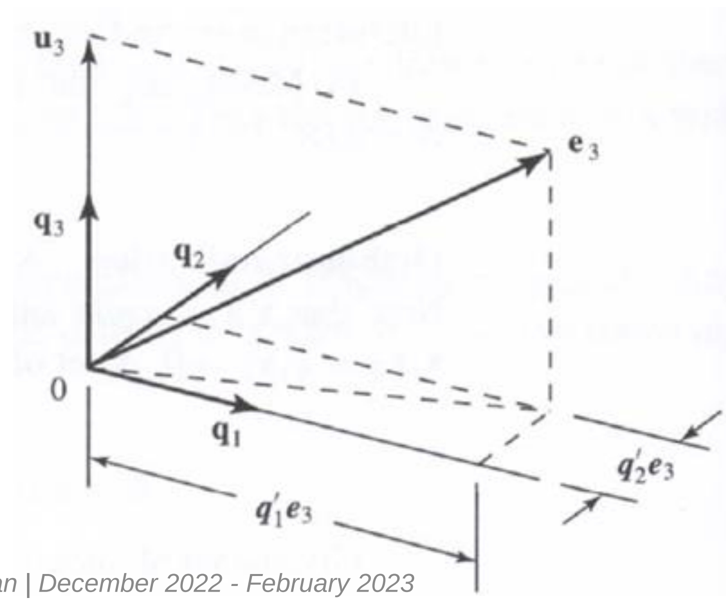
→ **Requirement:** $q_1, q_2, \dots, q_m$ must span the same space as $e_1, e_2, \dots, e_m$

→ **Procedure:**

Step 3: (a) Find the orthogonal basis:

- Project e_3 along q_1 and q_2 → $({q_1}^T e_3)q_1 + ({q_2}^T e_3)q_2$
- Compute residual $u_3 := e_3 - ({q_1}^T e_3)q_1 + \ ({q_2}^T e_3)q_2$

(b) Normalized $q_3 = \frac{u_3}{\|u_3\|}$



and so on until q_m

Matrix

Matrix

▪ Definition of Matrix

→ Rectangular display of vectors in rows and columns

$$a_1 = \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} \quad a_2 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \quad a_3 = \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}$$

$$A = [a_1 \quad a_2 \quad a_3] = \begin{bmatrix} 2 & 3 & 5 \\ 3 & 1 & -1 \\ -2 & 1 & 1 \end{bmatrix}$$

Matrix

Definition of Matrix

→ A system of m linear algebraic equations in n unknowns $x_1, x_2, \dots, x_n$ namely

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned}$$

may be represented by the matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

which in turn is represented by the simple matrix notation

$$Ax = y$$

Matrix

- **Orthonormal**

  → Let $A = a_1, a_2, \dots, a_m$ be an $\mathrm{n} \times m$ matrix with $m \leq \mathrm{n}$
  If all columns of A or $a_1, a_2, \dots, a_m$ are orthonormal,
  then

$$A^T A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_m$$

Matrix Addition and Scalar Multiplication

Addition

→ *Commutative* Law

$$A+B = B+A$$

→ *Associative* Law

$$A+(B+C) = (A+B)+C$$

Scalar Multiplication

→ *Distributive* with respect to scalar addition

$$(k_1 + k_2)A = k_1 A + k_2 A$$

→ *Distributive* over matrix addition

$$k_1(A+B) = k_1 A + k_1 B$$

→ *Associative* Law

$$k_1(k_2 A) = (k_1 k_2)A$$

Matrix Properties of Multiplication

Multiplication

→ Properties of Multiplication

1. $AB \neq BA$
2. $AB = 0$ does not imply that either A or B equals 0.
3. $AB = AC$ and $A \neq 0$ does not imply that $B = C$.

Example:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} 0 & 7 \\ 0 & 0 \end{bmatrix}, \quad AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

→ Matrix Multiplication

Associative: $(AB)C = A(BC)$

Distributive over addition: $A(B+C) = AB+AC, \ (A+B)C = AC+BC$

Matrix Properties of Transposition

- **Transposition**

1. $(A^T)^T = A.$

2. $(A+B)^T = A^T + B^T.$

3. $(kA)^T = kA^T$ for any scalar $k.$

4. $(AB)^T = B^T A^T.$

Special Matrices

- **Symmetric**

A is *symmetric* if and only if (*iff*) $A = A^T \, (i.e., a_{ij} = a_{ji})$ for all i and j.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 4 & 3 & 0 \end{bmatrix}$$

- **Skew-Symmetric**

A is *skew-symmetric iff* $A = -A^T \, (i.e., a_{ij} = -a_{ji})$ for all i and j. Note that this requires that the diagonal element be always zero.

$$B = \begin{bmatrix} 0 & 2 & -4 \\ -2 & 0 & 3 \\ 4 & -3 & 0 \end{bmatrix}$$

Special Matrices

Hermitian

A is *Hermitian* if $A = (\overline{A})^T$, where the overline indicates the complex conjugate $(i.e., a_{ij} = \overline{a}_{ji})$. *Note: The diagonal elements are always real, and a real symmetric matrix is Hermitian.*

$$\begin{bmatrix} 3 & 2+j3 \\ 2-j3 & 4 \end{bmatrix}$$

Diagonal Matrix

A square matrix $A = [a_{ij}]$ is diagonal if $a_{ij} = 0$ for $i \neq j$.

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

Identity Matrix

The identity matrix I is a diagonal matrix with ones along the main diagonal $(i.e., I = [\delta_{ij}])$. *For any* $n \times m$ *matrix* A, $AI_m = I_n A = A$.

The Kronecker delta

Partitioning

- **A Matrix**

matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

This could be written as

$$A = [B_1 \quad B_2], \quad B_1 = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, B_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

or as

$$A = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}, \quad \begin{matrix} C_1 = [1 \quad 2] \\ C_2 = [3 \quad 4] \\ C_3 = [5 \quad 6] \end{matrix}$$

or as

$$A = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad D_2 = [5 \quad 6].$$

Partitioning

- **Example:**

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} + \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} A_1 + B_1 & A_2 + B_2 \\ A_3 + B_3 & A_4 + B_4 \end{bmatrix}$$

$$AB = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = A_1 B_1 + A_2 B_2$$

$$\begin{bmatrix} 0 & 2 & \vdots & 3 \\ 1 & 4 & \vdots & 2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 1 & \vdots & 0 \end{bmatrix} \begin{bmatrix} 4 & \vdots & 2 & 1 \\ 1 & \vdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 2 & \vdots & 0 & 2 \end{bmatrix} = \begin{bmatrix} 8 & \vdots & 0 & 8 \\ 12 & \vdots & 2 & 9 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & \vdots & 0 & 1 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} + \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$$

Linear Algebraic Equations

Linear Algebraic Equations

- **A Set of Linear Equations:**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \vdots \quad & \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned}$$

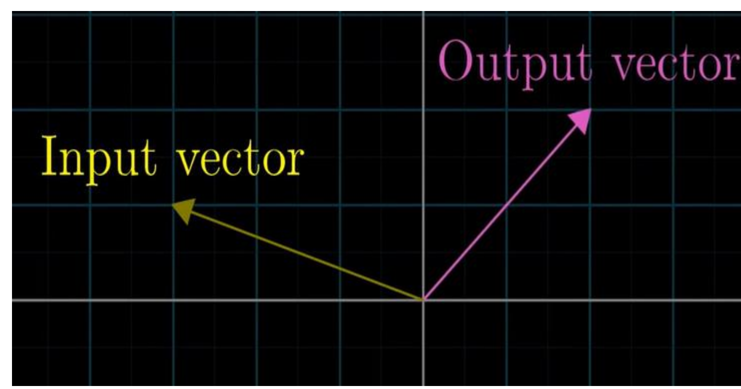
$$\underset{\text{m x n}}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\underset{\text{n x 1}}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \underset{\text{m x 1}}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

$$A\mathbf{x} = \boldsymbol{y}$$

Given

The unknown to be solved



Rank and Nullspace

- **Rank:**
  - $\rightarrow$ **Rank (A)** is defined as the dimension of the range space or, equivalently, the number of linearly independent columns in A
  - $\rightarrow$ **Rank (A)** is defined to be the nonzero rows in the row echelon form (REF) of A
  - $\rightarrow$ The rank of A can be equivalently defined as $\mathrm{rank}(A) = \dim(\mathrm{row}(A))$
  - $\rightarrow$ If A is $m \times n$, then

$$\mathrm{rank}(A) \leq \min(m, n)$$

- **Nullspace:**
  - $\rightarrow$ A vector x is called a null vector of A if $Ax = 0$.
  - $\rightarrow$ The null space of A consists of all its null vectors.
  - $\rightarrow$ The nullity is defined as the maximum number of linearly independent null vectors of A and is related to the rank by

$$\mathrm{Nullity}(A) = \text{number of columns of } A - \mathrm{rank}(A)$$

Row Echelon Form

Gaussian Elimination

$\rightarrow$ An algorithm for finding the row echelon form (REF) of a matrix

1. Interchange any two rows of the matrix
2. Multiply every entry of some row by the same nonzero scalar
3. Add a multiple of one row of the matrix to another row

$\rightarrow$ **<u>Gauss-Jordan elimination</u>**

A variant of the Gaussian elimination for finding the reduced row echelon form (RREF)

Original augmented matrix — **A row echelon form** — **The reduced row echelon form**

$$\begin{bmatrix} 1 & -2 & -1 & 3 \\ 3 & -6 & -5 & 3 \\ 2 & -1 & 1 & 0 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & -2 & -1 & 3 \\ 0 & 3 & 3 & -6 \\ 0 & 0 & 1 & 3 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Elementary row operations

Elementary row operations

Example

- **Matrix A**

$$\boxed{A = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix}}$$

$$\xrightarrow{R1\leftrightarrow R2} A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 2 & 0 \end{bmatrix} \xrightarrow{R3=R3-2R1} A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & -4 & -4 & -8 \end{bmatrix} \xrightarrow{R3=R3+4R2} A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\rightarrow$ Rank(A) = 2

$\rightarrow$ Basis of Row (A) = $\{[1 \quad 2 \quad 3 \quad 4], [0 \quad 1 \quad 1 \quad 2]\}$

Solving Linear Equations: Theorem 1

▪ Theorem 1

→ Given an $m \times n$ matrix A and an $m \times 1$ vector y,
an $n \times 1$ solution x exists in $Ax = y$
if and only if y lies in the range space of A or, equivalently,

$$\rho(A) = \rho\left([A \quad y]\right)$$

Rank(A)

where $[A \quad y]$ is an $m \times (n + 1)$ matrix with y applied to A as additional column

→ Given A, a solution exists in $Ax = y$ for every y,
if and only if A has rank m (full row rank)
(each of the rows of the matrix are linearly independent)

Solving Linear Equations: Theorem 2

- **Theorem 2: Parameterization of all solutions**
  - → Given an $m \times n$ matrix A and an $m \times 1$ vector y, let be x_p a solution of $Ax = y$ and let $\mathrm{k} := n - \rho(A)$ be the nullity of A
  - → If A has rank n (full column rank) or $\mathrm{k} = 0$, then the solution is unique
  - → If $\mathrm{k} > 0$, then for every real $\alpha_i; i = 1,2, \dots, \mathrm{k}$ the vector

$$x = x_p + \alpha_1 n_1 + \cdots + \alpha_k n_k$$

is a solution of $Ax = y$, where $\{n_1, n_1, \dots, n_k\}$ is a basis of the null space of A

Solving Linear Equations: Theorem 3

- **Theorem 3:**

Consider $Ax = y$ with A square.

1. If A is nonsingular, then the equation has a unique solution for every y and the solution equals $A^{-1}y$. In particular, the only solution of $Ax = 0$ is $x = 0$.

2. The homogeneous equation $Ax = 0$ has nonzero solutions if and only if A is singular. The number of linearly independent solutions equals the nullity of A.

Solving Linear Equations (1)

- **Matrix A**

$$Ax = y \rightarrow \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} x = \begin{bmatrix} -4 \\ -8 \\ 0 \end{bmatrix}$$

Find the solution for x

$$[A \quad y] = \begin{bmatrix} 0 & 1 & 1 & 2 & -4 \\ 1 & 2 & 3 & 4 & -8 \\ 2 & 0 & 2 & 0 & 0 \end{bmatrix} \Rightarrow [A \quad y] = \begin{array}{c} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{matrix} x_1 & x_2 & x_3 & x_4 & y \end{matrix} \end{array}$$

Free parameter

$$\left.\begin{aligned} x_3 &= \alpha_1 \\ x_4 &= \alpha_2 \end{aligned}\right\} \text{Define free parameters}$$

$$x_1 + x_3 = 0 \rightarrow x_1 = -\alpha_1$$

$$x_2 + x_3 + 2x_4 = -4 \rightarrow x_2 = -\alpha_1 - 2\alpha_2 - 4$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\alpha_1 \\ -\alpha_1 - 2\alpha_2 - 4 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \\ 0 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

Solving Linear Equations (2)

■ Solving with Inverse Matrix

$$\boxed{Ax = y} \rightarrow \boxed{x = A^{-1}y}$$

→ To get A^{-1}

(1) Using Determinant Matrix

Example:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

(2) Using Linear Row Reduction

Example:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & x & w \\ 0 & 1 & y & z \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} x & w \\ y & z \end{bmatrix}$$

Determinant and Inverse of Square Matrices

- **Determinant:**
  - → The **determinant** of a square matrix is a single number
  - → It tells immediately whether the matrix is invertible:

    $\det(A) = 0$ when A has no inverse

The determinant of a 1×1 matrix is defined as itself.
For $n = 2,3,\ldots$, the determinant of $n \times n$ square matrix $A = [a_{ij}]$ is defined recursively as, for any chosen j,

$$\det A = \sum_{i}^{n} a_{ij} c_{ij}$$

where a_{ij} denotes the entry at the ith row and jth column of A. This equation is called the *Laplace expansion*. The number c_{ij} is the *cofactor* corresponding to a_{ij} and equals $(-1)^{i+j} \det M_{ij}$, where M_{ij} is the $(n\text{-}1) \times (n\text{-}1)$ submatrix of A by deleting its ith row and jth column.

Determinant and Inverse of Square Matrices

- **Example:**

$$A = [a_{11}], \quad \det A = a_{11}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \det A = a_{11}a_{22} - a_{21}a_{12}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\det A = a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

the signs of the cofactors

Determinant and Inverse of Square Matrices

- **Example:**

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \Rightarrow A^{-1} = \frac{\text{Adj } A}{\det A} = \frac{1}{\det A}\left[c_{ij}\right]^T$$

$$m_{11} = \det\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = 0, \quad m_{12} = \det\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} = -3, \quad m_{13} = \det\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = 3$$

$$m_{21} = \det\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = -1, \quad m_{22} = \det\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = -3, \quad m_{23} = \det\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = 2$$

$$m_{31} = \det\begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix} = 1, \quad m_{32} = \det\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = -3, \quad m_{33} = \det\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} = 1$$

$$M = \begin{bmatrix} 0 & -3 & 3 \\ -1 & -3 & 2 \\ 1 & -3 & 1 \end{bmatrix} \Rightarrow C = \begin{bmatrix} 0 & 3 & 3 \\ 1 & -3 & -2 \\ 1 & 3 & 1 \end{bmatrix} \Rightarrow Adj(A) = C^T = \begin{bmatrix} 0 & 1 & 1 \\ 3 & -3 & 3 \\ 3 & -2 & 1 \end{bmatrix}$$

Determinant and Inverse of Square Matrices

- **Formula:**

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, A^{-1} = ? \qquad A^{-1} = \frac{1}{det(A)} C^T$$

$$det(A) = a(ei - fh) - b(di - fg) + c(dh - eg)$$
$$aei + bfg + cdh - ceg - bdi - afh$$

$$C = \begin{bmatrix} + \begin{vmatrix} e & f \\ h & i \end{vmatrix} & - \begin{vmatrix} d & f \\ g & i \end{vmatrix} & + \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ - \begin{vmatrix} b & c \\ h & i \end{vmatrix} & + \begin{vmatrix} a & c \\ g & i \end{vmatrix} & - \begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ + \begin{vmatrix} b & c \\ e & f \end{vmatrix} & - \begin{vmatrix} a & c \\ d & f \end{vmatrix} & + \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$$

Similarity Transformation

Similarity Transformation

Similarity Transformation

Consider an $n \times n$ matrix. If we associate with R^n the orthonormal basis $\{i_1, \ i_2, \ \ldots, i_n\}$, then the ith column of A is the representation of Ai_i with respect to the orthonormal basis.

If we select a different set of basis $\{q_1, \ q_2, \ \ldots, q_n\}$, then the matrix A has a different representation $\overline{A}$.

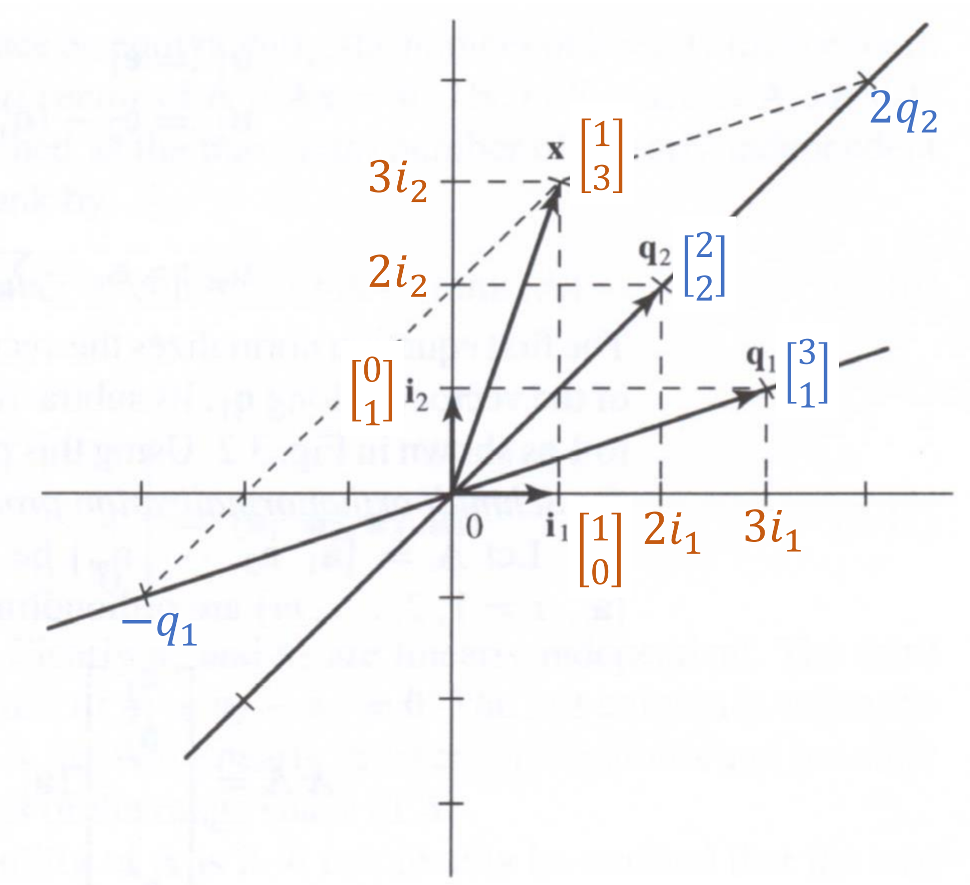
It turns out that the ith column of $\overline{A}$ is the representation of Aq_i with respect to the basis $\{q_1, \ q_2, \ \ldots, q_n\}$.

Basis

- **Example:**

  → Different representation of vector x

$$x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



Representation of x
with respect to the basis $\{i_1, i_2\}$

$$x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Representation of x
with respect to the basis $\{q_1, q_2\}$

$$x = Q \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Similarity Transformation

Let A be an $n \times n$ matrix. If there exists an $n \times 1$ vector b such that the n vectors $b, Ab, \ldots, A^{n-1}b$ are linearly independent and if

$$A^n b = \beta_1 b + \beta_2 Ab + \cdots + \beta_n A^{n-1} b$$

then the representation of A with respect to the basis $\{b,\ \ Ab,\ \ \ldots,\ \ A^{n-1}b\}$ is

$$\overline{A} = \begin{bmatrix} 0 & 0 & \cdots & 0 & \beta_1 \\ 1 & 0 & \cdots & 0 & \beta_2 \\ 0 & 1 & \cdots & 0 & \beta_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \beta_{n-1} \\ 0 & 0 & \cdots & 1 & \beta_n \end{bmatrix}$$

This matrix is said to be in a companion form.

Similarity Transformation

Example:

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$Ab = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad A^2b = A(Ab) = \begin{bmatrix} -4 \\ 2 \\ -3 \end{bmatrix}, \quad A^3b = A(A^2b) = \begin{bmatrix} -5 \\ 10 \\ -13 \end{bmatrix}$$

$$A^3b = 17b - 15Ab + 5A^2b$$

b, Ab, A^2b are linearly independent → they can be used as a basis

$$A(b) = [b \quad Ab \quad A^2b]\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad A(Ab) = [b \quad Ab \quad A^2b]\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad A(A^2b) = [b \quad Ab \quad A^2b]\begin{bmatrix} 17 \\ -15 \\ 5 \end{bmatrix}$$

$$\overline{A} = \begin{bmatrix} 0 & 0 & 17 \\ 1 & 0 & -15 \\ 0 & 1 & 5 \end{bmatrix}$$

Representation of A with respect to the basis b, Ab, A^2b

Similarity Transformation

Consider the equation

$$Ax = y$$

The square matrix A maps x in R^n into y in R^n. With respect to the basis $\{q_1, \ q_2, \ \ldots, \ q_n\}$, the equation becomes

$$\overline{A}\overline{x} = \overline{y}$$

where $\overline{x}$ and $\overline{y}$ are the representation of x and y with respect to the basis $\{q_1, \ q_2, \ \ldots, \ q_n\}$. They are related by

$$x = Q\overline{x} \quad y = Q\overline{y}$$

with $Q = [q_1 \ \ q_2 \ \ \cdots \ \ q_n]$ an $n \times n$ nonsingular matrix.

$$AQ\overline{x} = Q\overline{y} \quad \text{or} \quad Q^{-1}AQ\overline{x} = \overline{y}$$

Similarity Transformation

$$\boxed{\bar{A} = Q^{-1}AQ \quad \text{or} \quad A = Q\bar{A}Q^{-1}}$$

This is called the similarity transformation and A *and* $\bar{A}$ *are said to be similar.*

$$AQ = Q\bar{A}$$

$$A[q_1 \quad q_2 \quad \cdots \quad q_n]$$

$$= [Aq_1 \quad Aq_2 \quad \cdots \quad Aq_n] = [q_1 \quad q_2 \quad \cdots \quad q_n]\bar{A}$$

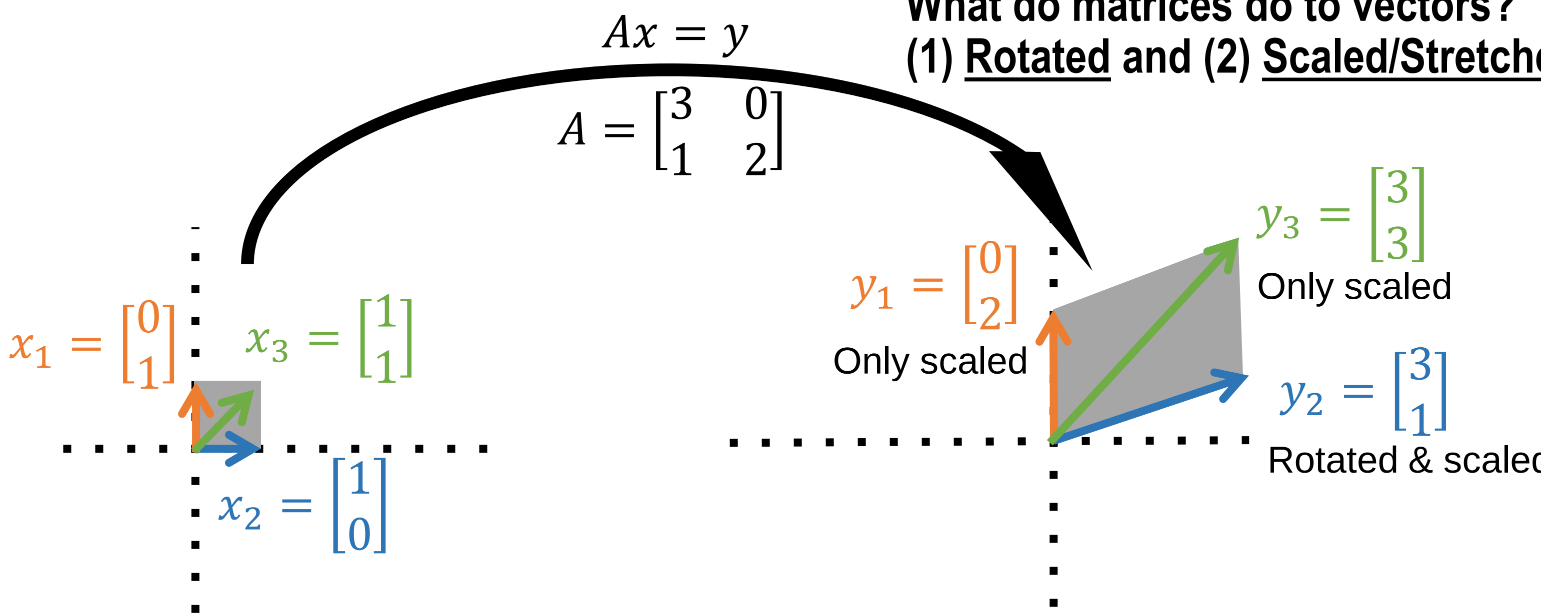
This shows that the i*th column of* $\bar{A}$ *is indeed the representation of* Aq_i *with respect to the basis* $\{q_1, \quad q_2, \quad \ldots, \quad q_n\}$.

Eigenvalues and Eigenvectors

Vector Transformation

▪ Matrix Multiplication

What do matrices do to vectors?
(1) <u>Rotated</u> and (2) <u>Scaled/Stretched</u>



$$Ax_1 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 2x_1$$
$\rightarrow$ Vector x_1 is an eigenvector with eigenvalue 2

$$Ax_2 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$
$\rightarrow$ Vector x_2 is <u>**not**</u> an eigenvector

$$Ax_3 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3x_3$$
$\rightarrow$ Vector x_3 is an eigenvector with eigenvalue 3

Eigenvector and Eigenvalue

Definition

A *real* or *complex number* λ is called an *eigenvalue* of the $n \times n$ real matrix A if there exists a *nonzero* vector x such that $Ax = \lambda x$. Any *nonzero* vector x satisfying $Ax = \lambda x$ is called a (right) *eigenvector* of A associated with eigenvalue λ.

How to find eigenvector and eigenvalue?

$$Ax = y \longrightarrow Ax = \lambda x$$

$$Ax = \lambda I x \longrightarrow (A - \lambda I)x = 0$$

$$\Delta(\lambda) = \det(A - \lambda I)$$ *The characteristic polynomial of* A

Every root of $\Delta(\lambda)$ *is eigenvalue*

The null space of $A - \lambda I$ is called the eigenspace of A associated with eigenvalue λ.

Example:

- **Matrix** $A = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \longrightarrow E_\lambda(A) = N(A-\lambda I) = N\left(\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$

$$= N\left(\begin{bmatrix} 3-\lambda & 0 \\ 1 & 2-\lambda \end{bmatrix}\right)$$

$$\det(A-\lambda I) = 0 \longrightarrow (3-\lambda)(2-\lambda) = 0$$

$$\lambda_1 = 2 \text{ and } \lambda_2 = 3$$

For $\lambda_1 = 2$: $E_1(A) = N(A-\lambda I) = N\left(\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right) = N\left(\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right)$

$$(A-\lambda I)x_1 = 0 \longrightarrow \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} \longrightarrow \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

<u>Eigenvector</u>: $E_1 = span\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} \longrightarrow$ Any nonzero solution $\begin{bmatrix} 0 \\ \alpha \end{bmatrix}$

For $\lambda_2 = 3$: $E_2(A) = N(A-\lambda I) = N\left(\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\right) = N\left(\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\right)$

$$(A-\lambda I)x_1 = 0 \longrightarrow \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \longrightarrow \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

<u>Eigenvector</u>: $E_2 = span\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} \longrightarrow$ Any nonzero solution $\begin{bmatrix} \alpha \\ \alpha \end{bmatrix}$

Example:

$$Ax = \lambda x$$

$$(A - \lambda I)x = 0 \quad \rightarrow \quad E_\lambda(A) = N(A - \lambda I)$$

$$A = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}, \quad \lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 3$$

$$E_1 = N\left(\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) = N\left(\begin{bmatrix} -2 & 4 & -2 \\ -3 & 3 & 0 \\ -3 & 1 & 2 \end{bmatrix} \right)$$

$$\begin{bmatrix} -2 & 4 & -2 \\ -3 & 3 & 0 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \end{bmatrix} = 0, \quad \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad t \in R \right\} = span\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Example:

$$E_2 = N\left(\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right) = N\left(\begin{bmatrix} -3 & 4 & -2 \\ -3 & 2 & 0 \\ -3 & 1 & 1 \end{bmatrix}\right)$$

$$\begin{bmatrix} -3 & 4 & -2 \\ -3 & 2 & 0 \\ -3 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = 0, \quad \begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix}, \quad E_2 = \left\{ t\begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix}, \quad t \in R \right\} = span\left\{\begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix}\right\}$$

$$E_3 = N\left(\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\right) = N\left(\begin{bmatrix} -4 & 4 & -2 \\ -3 & 1 & 0 \\ -3 & 1 & 0 \end{bmatrix}\right)$$

$$\begin{bmatrix} -4 & 4 & -2 \\ -3 & 1 & 0 \\ -3 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_{31} \\ x_{32} \\ x_{33} \end{bmatrix} = 0, \quad \begin{bmatrix} x_{31} \\ x_{32} \\ x_{33} \end{bmatrix} = \begin{bmatrix} 1/4 \\ 3/4 \\ 1 \end{bmatrix}, \quad E_3 = \left\{ t\begin{bmatrix} 1/4 \\ 3/4 \\ 1 \end{bmatrix}, \quad t \in R \right\} = span\left\{\begin{bmatrix} 1/4 \\ 3/4 \\ 1 \end{bmatrix}\right\}$$

Homework #3.1 Eigenvalue & Eigenvector (1 pt.): Due Dec. 22

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

Find the *eigenvalues* and *eigenvectors*

Use Your ID: sGFEDCBA

Nonsingular Matrix

If A is diagonal or triangular, then det A equals the product of all diagonal entries.

The determinant of any $r \times r$ submatrix of A is called a *minor* of order r. Then the rank can be defined as the largest order of all nonzero minors of A.

A square matrix is said to be *nonsingular* if its determinant is nonzero. Thus a nonsingular square matrix has full rank and all its columns (rows) are linearly independent.

The *inverse* of a nonsingular square matrix can be computed as

$$A^{-1} = \frac{\text{Adj } A}{\det A} = \frac{1}{\det A}\left[c_{ij}\right]^T$$

inv

Characteristic Polynomials

- **Characteristic Polynomials of $A \longrightarrow \Delta(\lambda) = \det(A - \lambda I)$**

$$\begin{bmatrix} 0 & 0 & 0 & -\alpha_4 \\ 1 & 0 & 0 & -\alpha_3 \\ 0 & 1 & 0 & -\alpha_2 \\ 0 & 0 & 1 & -\alpha_1 \end{bmatrix} \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} \begin{bmatrix} -\alpha_1 & 1 & 0 & 0 \\ -\alpha_2 & 0 & 1 & 0 \\ -\alpha_3 & 0 & 0 & 1 \\ -\alpha_4 & 0 & 0 & 0 \end{bmatrix}$$

These matrices all have the following characteristic polynomial:

$$\Delta(\lambda) = \lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4$$

These matrices can easily be formed from the coefficients of $\Delta(\lambda)$ *and are called companion-form matrices.*

Case1: When Eigenvalues are all Distinct

A square matrix A has different representations with respect to different sets of basis. We introduce a set of basis so that the representation will be diagonal or block diagonal.

Eigenvalue of A are all distinct: Let $\lambda_i, \ \ i = 1,2,\ldots,n$, be the eigenvalue of A and all distinct. Let q_i be an eigenvector of A associated with λ_i; that is, $Aq_i = \lambda_i q_i$. Then the set of eigen-vectors $\{q_1, \ q_2, \ \ldots, \ q_n\}$ is linearly independent and can be used as a basis. Let $\hat{A}$ be the representation of A with respect to this basis.

Eigenvector as a Basis

Then the first column of $\hat{A}$ is the representation of $Aq_1 = \lambda_1 q_1$ with respect to $\{q_1, \; q_2, \; \ldots, \; q_n\}$. From

$$Aq_1 = \lambda_1 q_1 = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

we conclude that the first column of $\hat{A}$ is $\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \end{bmatrix}^T$.
The second column of $\hat{A}$ is the representation of $Aq_2 = \lambda_2 q_2$ with respect to $\{q_1, \; q_2, \; \ldots, \; q_n\}$, that is, $\begin{bmatrix} 0 & \lambda_2 & 0 & \cdots & 0 \end{bmatrix}^T$.

Diagonal Form of Eigenvalues

Proceeding forward, we can establish

$$\hat{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

This is a diagonal matrix.

Every matrix with distinct eigenvalues has a diagonal matrix representation by using its eigenvectors as a basis. Different orderings of eigenvectors will yield different diagonal matrices for the same A.

Example: Find Eigenvectors

If we define

$$Q = [q_1 \quad q_2 \quad \cdots \quad q_n]$$

then the matrix $\hat{A}$ equals

$$\boxed{\hat{A} = Q^{-1} A Q}$$

Example

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

a=[0 0 0;1 0 2; 0 1 1];
eig(a)

$$\Delta(\lambda) = \det(\lambda I - A) = \det \begin{bmatrix} \lambda & 0 & 0 \\ -1 & \lambda & -2 \\ 0 & -1 & \lambda - 1 \end{bmatrix} = \lambda[\lambda(\lambda-1)-2] = (\lambda-2)(\lambda+1)\lambda$$

2, -1, 0

Example

Thus the representation of A w.r.t. $\{q_1, q_2, q_3\}$ is

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

```
a=[0 0 0;1 0 2; 0 1 1];
[q,d]=eig(a)
```

Normalized

$$(A-2I)q_1 = \begin{bmatrix} -2 & 0 & 0 \\ 1 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} q_1 = 0 \quad \rightarrow q_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$(A-(-1)I)q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} q_2 = 0 \quad \rightarrow q_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$$

$$(A-0I)q_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} q_3 = 0 \quad \rightarrow q_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

$$\boxed{Q = [q_1 \quad q_2 \quad q_3] = \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}}$$

$$\boxed{Q\hat{A} = AQ}$$

Case 2: When Eigenvalues are not Distinct

Eigenvalue of A are not all distinct: If A has only simple eigenvalues, it always has a diagonal-form representation. If A has repeated eigenvalues, then it may not have a diagonal form representation. It has a block-diagonal and triangular-form representation.

Consider an $n \times n$ matrix A with eigenvalue λ and multiplicity n. (A has only one distinct eigenvalue. We assume $n = 4$.) Suppose the matrix $(A - \lambda I)$ has rank $n - 1 = 3$ or, equivalently, nullity 1; then the equation

$$(A - \lambda I)q = 0$$

has only one independent solution.

An eigenvalue with multiplicity 2 or higher is called a repeated eigenvalue.
An eigenvalue with multiplicity 1 is called a simple eigenvalue.

Generalized Eigenvectors

Thus A has only one eigenvector associated with λ. We need $n - 1 = 3$ more linearly independent vectors to form a basis for R^4. The three vectors q_2, q_3, q_4 will be chosen to have the properties $(A-\lambda I)^2 q_2 = 0$, $(A-\lambda I)^3 q_3 = 0$, and $(A-\lambda I)^4 q_4 = 0$.

A vector v is called a *generalized eigenvector* of grade n if

$$(A-\lambda I)^n v = 0 \quad \text{and} \quad (A-\lambda I)^{n-1} v \neq 0$$

For $n = 4$, we define

If $n = 1$, $(A-\lambda I)v = 0$ and $v \neq 0$ and v is an ordinary eigenvector.

$$v_4 := v$$

$$v_3 := (A-\lambda I)v_4 = (A-\lambda I)v$$

$$v_2 := (A-\lambda I)v_3 = (A-\lambda I)^2 v$$

$$v_1 := (A-\lambda I)v_2 = (A-\lambda I)^3 v$$

Generalized Eigenvectors

$$(A-\lambda I)v_1 = 0, \quad v_1 \neq 0$$

$$(A-\lambda I)^2 v_2 = 0, \quad (A-\lambda I)v_2 \neq 0$$

$$(A-\lambda I)\boxed{(A-\lambda I)v_2} = 0 \quad \rightarrow \quad (A-\lambda I)v_2 = v_1$$

an eigenvector

$$(A-\lambda I)^3 v_3 = 0, \quad (A-\lambda I)^2 v_3 \neq 0$$

$$(A-\lambda I)\boxed{(A-\lambda I)^2 v_3} = 0 \quad \rightarrow \quad (A-\lambda I)^2 v_3 = (A-\lambda I)v_2$$

$$(A-\lambda I)v_3 = v_2$$

$$(A-\lambda I)^4 v_4 = 0, \quad (A-\lambda I)^3 v_4 \neq 0$$

$$(A-\lambda I)\boxed{(A-\lambda I)^3 v_4} = 0 \quad \rightarrow \quad (A-\lambda I)^3 v_4 = (A-\lambda I)^2 v_3$$

$$(A-\lambda I)v_4 = v_3$$

Generalized Eigenvectors

The *representation* of
$$
\begin{aligned}
Av_1 &= \lambda v_1 \\
Av_2 &= v_1 + \lambda v_2 \\
Av_3 &= v_2 + \lambda v_3 \\
Av_4 &= v_3 + \lambda v_4
\end{aligned}
$$

with respect to $[v_1, \quad v_2, \quad v_3, \quad v_4]$

$$
[v_1 \quad v_2 \quad v_3 \quad v_4] \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}
$$

Jordan Form of Eigenvalues

They are called a chain of generalized eigenvectors of length $n = 4$ and have the properties $(A-\lambda I)v_1 = 0,\ (A-\lambda I)^2 v_2 = 0,\ (A-\lambda I)^3 v_3 = 0,$ and $(A-\lambda I)^4 v_4 = 0.$ These vectors are automatically linearly independent and can be used as a basis.

$$
\begin{aligned}
Av_1 &= \lambda v_1 \\
Av_2 &= v_1 + \lambda v_2 \\
Av_3 &= v_2 + \lambda v_3 \\
Av_4 &= v_3 + \lambda v_4
\end{aligned}
$$

Then the representation of A with respect to the basis $\{v_1,\ v_2,\ v_3,\ v_4\}$ is

$$
J := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}
$$

Example

The matrix J has eigenvalues on the diagonal and 1 on the superdiagonal. If we reverse the order of the basis, then the 1's will appear on the subdiagonal. The matrix is called a *Jordan* block of order $n = 4$.

Example:

Consider a 5 X 5 matrix A with repeated eigenvalue λ_1 with multiplicity 4 and simple eigenvalue λ_2. Then there exists a nonsingular matrix Q such that

$$\hat{A} = Q^{-1}AQ$$

assumes one of the following forms

Example

$Nullity\,(A-\lambda_1 I)=1$

$$\hat{A}_1=\begin{bmatrix}\lambda_1 & 1 & 0 & 0 & 0\\ 0 & \lambda_1 & 1 & 0 & 0\\ 0 & 0 & \lambda_1 & 1 & 0\\ 0 & 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & 0 & \lambda_2\end{bmatrix}$$

$$\hat{A}_2=\begin{bmatrix}\lambda_1 & 1 & 0 & 0 & 0\\ 0 & \lambda_1 & 1 & 0 & 0\\ 0 & 0 & \lambda_1 & 0 & 0\\ 0 & 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & 0 & \lambda_2\end{bmatrix}$$

$Nullity=2$

$Nullity=2$

$$\hat{A}_3=\begin{bmatrix}\lambda_1 & 1 & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0 & 0\\ 0 & 0 & \lambda_1 & 1 & 0\\ 0 & 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & 0 & \lambda_2\end{bmatrix}$$

$$\hat{A}_4=\begin{bmatrix}\lambda_1 & 1 & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0 & 0\\ 0 & 0 & \lambda_1 & 0 & 0\\ 0 & 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & 0 & \lambda_2\end{bmatrix}$$

$Nullity=3$

$Nullity=4$

$$\hat{A}_5=\begin{bmatrix}\lambda_1 & 0 & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0 & 0\\ 0 & 0 & \lambda_1 & 0 & 0\\ 0 & 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & 0 & \lambda_2\end{bmatrix}$$

If the nullity is 4 we can find
4 linearly independent eigenvectors

If the nullity is 1 we can find
Only 1 ordinary eigenvector

Example

Consider the Jordan block of order 4. $J := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$

$$(J - \lambda I) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (J - \lambda I)^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(J - \lambda I)^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (J - \lambda I)^k = 0 \;\; \mathit{for} \;\; k \geq 4$$

This is called nilpotent.

Functions of Square Matrix

Functions of Square Matrix

Definition

→ Let A be a square matrix. If k is a positive integer, we define:

$$A^k = AA\cdots A \qquad (k \text{ terms})$$
$$A^0 = I$$

Example:

$$A^2 = AA = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 5 & 5 \end{bmatrix}$$
$$A^2 = AA = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 3^2 & 0 \\ 0 & 2^2 \end{bmatrix}$$

→ Let $f(\lambda)$ be a polynomial such that

$$\begin{aligned} f(\lambda) &= \lambda^3 + 2\lambda^2 - 6 \\ &= (\lambda + 2)(\lambda - 3) \end{aligned}$$

Then $f(A)$ is defined as

$$\begin{aligned} f(A) &= A^3 + 2A^2 - 6 \\ &= (A + 2)(A - 3) \end{aligned}$$

Solving Linear Equations (1): Diagonalization

If A is block diagonal, such as

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

where A_1 and A_2 are square matrices of any order, then it is straightforward to verify

$$A^k = \begin{bmatrix} A_1^k & 0 \\ 0 & A_2^k \end{bmatrix} \text{ and } f(A) = \begin{bmatrix} f(A_1) & 0 \\ 0 & f(A_2) \end{bmatrix}$$

Consider the similarity transformation $\hat{A} = Q^{-1}AQ$ or $A = Q\hat{A}Q^{-1}$.

Because

$$A^k = (Q\hat{A}Q^{-1})(Q\hat{A}Q^{-1})\cdots(Q\hat{A}Q^{-1}) = Q\hat{A}^kQ^{-1}$$

we have

$$f(A) = Qf(\hat{A})Q^{-1} \text{ or } f(\hat{A}) = Q^{-1}f(A)Q$$

Solving Linear Equations (2): Theorem 4

Theorem 4 (**Cayley-Hamilton theorem**)

Let

$$\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

be the characteristic polynomial of A. Then

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1} A + \alpha_n I = 0$$

A matrix satisfies its own characteristic polynomial.

Because $n_i \geq \bar{n}_i$, the characteristic polynomial contains the minimal polynomial as a factor or $\Delta(\lambda) = \boxed{\psi(\lambda)} h(\lambda)$ for some $h(\lambda)$. Because $\psi(A) = 0$, we have

Minimal polynomial

$$\Delta(A) = \psi(A)h(A) = 0 \cdot h(A) = 0.$$

This establishes the theorem.

Monic and Minimal Polynomial

A *monic* polynomial is a polynomial with 1 as its leading coefficient. The *minimal* polynomial of A is defined as the monic polynomial $\psi(\lambda)$ of least degree such that $\psi(A) = 0$.

All similar matrices have the same minimal polynomial.

We define

$$\Delta(\lambda) = \det(\lambda I - A)$$

It is a monic polynomial of degree n with real coefficients and is called the *characteristic polynomial* of A.

Example:

- **Matrix**

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \longrightarrow \Delta\lambda = \det(A - \lambda I) = 0$$

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

$$\Delta\lambda = (2-\lambda)^4 \longrightarrow \Psi(A) = (2-\lambda) \text{ or } (2-\lambda)^2 \text{ or } (2-\lambda)^3 \text{ or } (2-\lambda)^4$$

Since $\psi(\lambda)$ *must have a zero at* $\lambda = 2$, *we try* $\psi(\lambda) = (\lambda - 2)$.
We can check that $\psi(A) = A - 2I = 0$,
so $\psi(\lambda) = (\lambda - 2)$ *is correct.*

$$\Psi(A) = (A - 2) = A - 2I = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

If not, we check $\Psi(A) = (2-\lambda)^2$ or $(2-\lambda)^3$ or $(2-\lambda)^4$

Jordan Form

If the Jordan-form representation of A is available, the minimal polynomial can be read out by inspection.

Let λ_i be an eigenvalue of A with multiplicity n_i. That is, the characteristic polynomial of A is

$$\boxed{\Delta(\lambda) = \det(\lambda I - A) = \prod_i (\lambda - \lambda_i)^{n_i}}$$

The index of λ_i, denoted by $\bar{n}_i$, is defined as the largest order of all Jordan blocks associated with λ_i. Using the indices of all eigenvalues, the minimal polynomial can be expressed as

$$\boxed{\psi(\lambda) = \prod_i (\lambda - \lambda_i)^{\bar{n}_i}}$$

Minimal Polynomial of Jordan Form

$\psi_1 = (\lambda-\lambda_1)^4(\lambda-\lambda_2)$

$$\hat{A}_1 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$\psi_2 = (\lambda-\lambda_1)^3(\lambda-\lambda_2)$

$$\hat{A}_2 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$\psi_3 = (\lambda-\lambda_1)^2(\lambda-\lambda_2)$

$$\hat{A}_3 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$\psi_4 = (\lambda-\lambda_1)^2(\lambda-\lambda_2)$

$$\hat{A}_4 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$\psi_5 = (\lambda-\lambda_1)(\lambda-\lambda_2)$

$$\hat{A}_5 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$$\boxed{\Delta(\lambda) = (\lambda-\lambda_1)^4(\lambda-\lambda_2)}$$

If all eigenvalues of A are distinct, then the minimal polynomial equals the characteristic polynomial.

Solving Linear Equations (2): Theorem 4

Theorem 4 (**Cayley-Hamilton theorem**)

Let

$$\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

be the characteristic polynomial of A. Then

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1} A + \alpha_n I = 0$$

A matrix satisfies its own characteristic polynomial.

Because $n_i \geq \bar{n}_i$, the characteristic polynomial contains the minimal polynomial as a factor or $\Delta(\lambda) = \psi(\lambda)h(\lambda)$ for some $h(\lambda)$. Because $\psi(A) = 0$, we have

$$\Delta(A) = \psi(A)h(A) = 0 \cdot h(A) = 0.$$

This establishes the theorem.

Solving Linear Equations (2): Theorem 4

- **Theorem 4: Cayley Hamilton Theorem**

$$A \in R^{n\times n}$$

$$\Delta(\lambda) = \det(A-\lambda I) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) \quad \textit{All distinct}$$

$$\Delta(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_k)^{n_k} \quad \textit{Repeated}$$

$$n_1 + n_2 + \cdots + n_k = n$$

$$\Delta(A) = 0 \quad \text{Cayley-Hamilton}$$

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$$

for some positive integers m_i *that satisfy* $1 \le m_i \le n_i$.

$$\psi(A) = 0$$

$\Delta(\lambda)$ *divides* $\psi(\lambda)$ *perfectly (without any remainder).*

Solving Linear Equations (2): Theorem 4

Theorem 4: Cayley Hamilton Theorem

example 1: $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ the characteristic equation of A is

$$\Delta(\lambda) = \underbrace{(\lambda - 1)(\lambda - 3)}_{\det(A - \lambda I)} = \lambda^2 - 4\lambda + 3 = 0$$

the Cayley-Hamilton states that A *satisfies its characteristic equation*

$$\Delta(A) = 0 \quad \text{Cayley-Hamilton}$$

$$\Delta(A) = A^2 - 4A + 3I = 0$$

use this equation to write matrix powers of A

$$\begin{aligned}
A^2 &= 4A - 3I \\
A^3 &= 4A^2 - 3A &&= 4(4A - 3I) - 3A &&= 13A - 12I \\
A^4 &= 13A^2 - 12A &&= 13(4A - 3I) - 12A &&= 40A - 39I \\
&\vdots && \vdots
\end{aligned}$$

powers of A can be written as a linear combination of I and A

Solving Linear Equations (3): Theorem 5

- **Theorem 5: A has repeated Eigenvalue**

We are given $f(\lambda)$ and an $n \times n$ matrix A with characteristic polynomial

$$\Delta(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i}$$

where $n = \sum_{i=1}^{m} n_i$. Define

$$h(\lambda) := \beta_0 + \beta_1 \lambda + \cdots + \beta_{n-1} \lambda^{n-1}$$

It is a polynomial of degree $n-1$ with n unknown coefficients. These n unknowns are to be solved from the following set of n equations.

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i) \text{ for } l = 0, 1, \ldots, n_i - 1 \text{ and } i = 1, 2, \ldots, m$$

Then we have $f(A) = h(A)$ and $h(\lambda)$ is said to equal $f(\lambda)$ on the spectrum of A.

Solving Linear Equations (3): Theorem 5

Theorem 5: Example:

Let $A_1 = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ Compute $e^{A_1 t}$.

Or, equivalently, if $f(\lambda) = e^{\lambda t}$, what is $f(A_1)$?

The characteristic polynomial of A_1 is $(\lambda - 1)^2(\lambda - 2)$.

Let $h(\lambda) = \beta_0 + \beta_1\lambda + \beta_2\lambda^2$. Then

$$f(1) = h(1): \quad e^t = \beta_0 + \beta_1 + \beta_2$$

$$f'(1) = h'(1): \quad te^t = \beta_1 + 2\beta_2$$

$$f(2) = h(2): \quad e^{2t} = \beta_0 + 2\beta_1 + 4\beta_2$$

$$\beta_0 = -2te^t + e^{2t}$$

$$\beta_1 = 3te^t + 2e^t - 2e^{2t}$$

$$\beta_2 = e^{2t} - e^t - te^t$$

$$\boxed{\begin{aligned} e^{A_1 t} &= h(A_1) \\ &= (-2te^t + e^{2t})I \\ &\quad + (3te^t + 2e^t - 2e^{2t})A_1 \\ &\quad + (e^{2t} - e^t - te^t)A_1^2 \\ &= \begin{bmatrix} 2e^t - e^{2t} & 0 & 2e^t - 2e^{2t} \\ 0 & e^t & 0 \\ e^{2t} - e^t & 0 & 2e^{2t} - e^t \end{bmatrix} \end{aligned}}$$

Homework #3.2 Functions of Square Matrix (1 pt.): Due Dec. 22

Compute A^{1BC} with

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

In other words, given $f(\lambda) = \lambda^{1BC}$, compute $f(A)$.

Using both Theorem 4 and 5.

Use Your ID: sGFEDCBA

Solving Linear Equations (4): Power Series

- **Infinite Power Series:**

Suppose $f(\lambda)$ can be expressed as the power series

$$f(\lambda) = \sum_{i=0}^{\infty} \beta_i \lambda^i$$

with the radius of convergence ρ. If all the eigenvalues of A have magnitude less than ρ, then $f(A)$ can be defined as

$$f(A) = \sum_{i=0}^{\infty} \beta_i A^i$$

- **Example:**

$$f(\lambda) = f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1) + \frac{f''(\lambda_1)}{2!}(\lambda - \lambda_1)^2 + \cdots$$

$$f(\hat{A}) = f(\lambda_1)I + f'(\lambda_1)(\hat{A} - \lambda_1 I) + \cdots + \frac{f^{(n-1)}(\lambda_1)}{(n-1)!}(\hat{A} - \lambda_1 I)^{n-1} + \cdots$$

zero

Solving Linear Equations (4): Power Series

- **Example:**

Because the Taylor series

$$e^{\lambda t} = 1 + \lambda t + \frac{\lambda^2 t^2}{2!} + \cdots + \frac{\lambda^n t^n}{n!} + \cdots$$

converges for all finite λ and t, we have

$$e^{At} = I + tA + \frac{t^2 A^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k$$

Solving Linear Equations (4): Power Series

- **Find Solution where** $A = \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix}$
- Using infinite power series:

$$e^{At} = I + At + \frac{t^2}{2!}A^2 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k$$

$$\begin{aligned}
e^{At} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix} t + \frac{t^2}{2!} \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix} + \cdots \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -t \\ t & -2t \end{bmatrix} + \frac{t^2}{2} \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix} + \cdots \\
&= \begin{bmatrix} 1 - 0t - \frac{t^2}{2} - \cdots & 0 - t + \frac{2t^2}{2} + \cdots \\ 0 + t - \frac{2t^2}{2} + \cdots & 1 - 2t + \frac{3t^2}{2} + \cdots \end{bmatrix} \\
&= \begin{bmatrix} e^{-t} + te^{-t} & -te^{-t} \\ te^{-t} & e^{-t} - te^{-t} \end{bmatrix}
\end{aligned}$$

Solving Linear Equations (4): Power Series

- **Some Useful Properties**

$$e^0 = I$$

$$e^{A(t_1+t_2)} = e^{At_1}e^{At_2}$$

$$[e^{At}]^{-1} = e^{-At}$$

$$\frac{d}{dt}e^{At} = \sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}A^k$$

$$= A\left(\sum_{k=0}^{\infty}\frac{1}{k!}t^kA^k\right) = \left(\sum_{k=0}^{\infty}\frac{1}{k!}t^kA^k\right)A$$

$$\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$$

$e^{(A+B)t} \neq e^{At}e^{Bt}$ *The equality holds only if A and B commute or $AB=BA$.*

Solving Linear Equations (4): Power Series

Some Useful Properties

$$\sin A = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots \qquad \sinh A = A + \frac{A^3}{3!} + \frac{A^5}{5!} + \cdots$$

$$\cos A = I - \frac{A^2}{2!} + \frac{A^4}{4!} - \cdots \qquad \cosh A = I + \frac{A^2}{2!} + \frac{A^4}{4!} + \cdots$$

$$\sin^2 A + \cos^2 A = I$$

$$\sin A = \frac{e^{jA} - e^{-jA}}{2j}, \qquad \cos A = \frac{e^{jA} + e^{-jA}}{2}$$

$$\cosh^2 A - \sinh^2 A = I$$

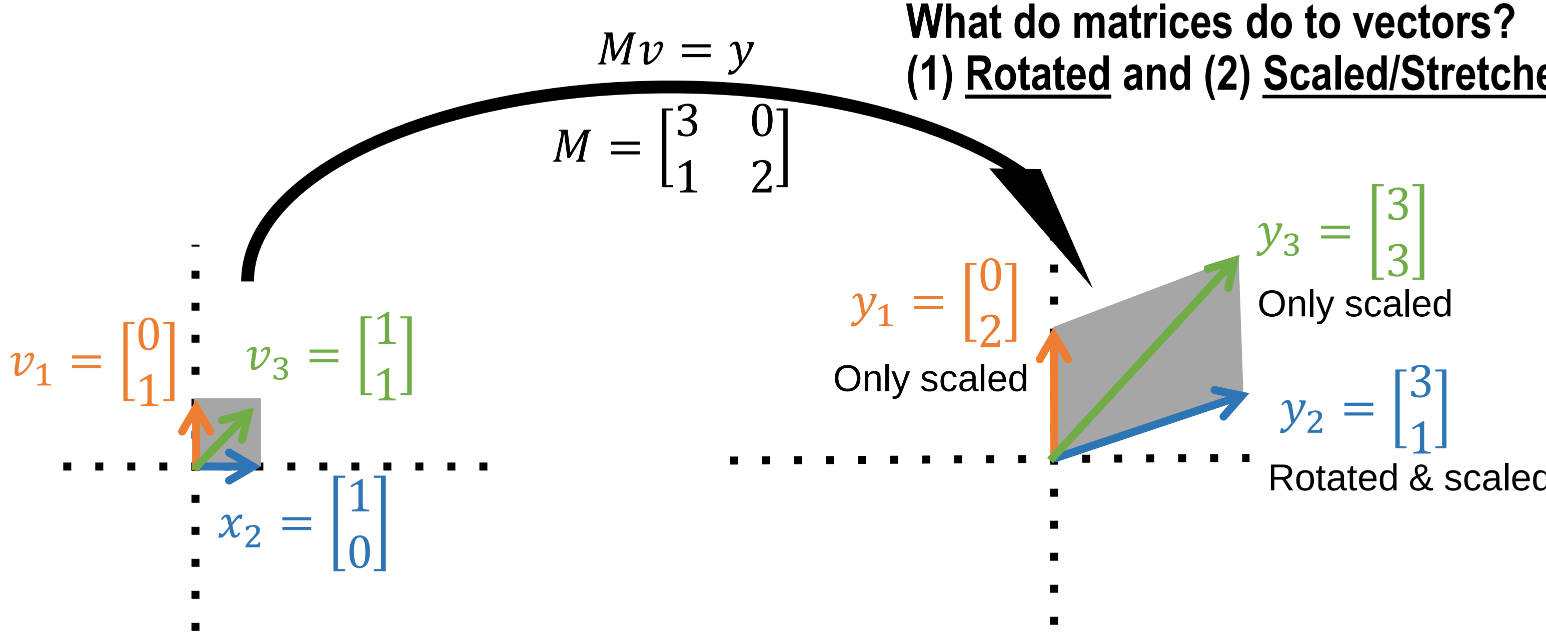
$$\sinh A = \frac{e^{A} - e^{-A}}{2}, \qquad \cosh A = \frac{e^{A} + e^{-A}}{2}$$

Eigenvalue Decomposition (EVD)

Vector Transformation

Matrix Multiplication

What do matrices do to vectors?
(1) <u>Rotated</u> and (2) <u>Scaled/Stretched</u>



$M v_1 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 2v_1$ $\rightarrow$ Vector v_1 is an eigenvector with eigenvalue 2

$M v_2 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ $\rightarrow$ Vector v_2 is **<u>not</u>** an eigenvector

$M v_3 = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3v_3$ $\rightarrow$ Vector v_3 is an eigenvector with eigenvalue 3

Eigenvalue Decomposition

- **Square Matrix** $n \times n$

$$M = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \longrightarrow (3-\lambda)(2-\lambda) = 0$$

$$\lambda_1 = 2 \text{ and } \lambda_2 = 3$$

$$v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and } v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\left.\begin{aligned} Mv &= \lambda v \\ A\begin{bmatrix} 0 \\ 1 \end{bmatrix} &= 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ A\begin{bmatrix} 1 \\ 1 \end{bmatrix} &= 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}\right\} \longrightarrow A \begin{array}{c} \begin{array}{cc} v_1 & v_2 \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{array}{cc} v_1 & v_2 \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

$$MQ = QD$$

$$MQQ^{-1} = QDQ^{-1}$$

$$M = QDQ^{-1}$$ **Eigenvalue Decomposition**

$$\underbrace{\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}}_{\text{Original matrix}} = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}}_{\text{Eigen-vectors matrix}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}}_{\text{Eigen-values matrix}} \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}}_{\text{Inverse of Eigenvectors matrix}}$$

Eigenvalue Decomposition

An $n \times n$ real matrix M is said to be *symmetric* if its transpose equals itself. The scalar function $x^T Mx$, where x is an $n \times 1$ real vector and $M^T = M$, is called a *quadratic form*.

All eigenvalues of symmetric M are real.

Every symmetric matrix can be diagonalized using a similarity transformation even it has repeated eigenvalue λ. There exists a Q that consists of all linearly independent eigenvectors of M such that

$$M = QDQ^{-1}$$

where D is a diagonal matrix with real eigenvalues of M on the diagonal. Q can be selected as an *orthogonal* matrix. $(Q^{-1} = Q^T)$

Eigenvalue Decomposition

Theorem 6

For every real symmetric matrix M, there exists an orthogonal matrix Q such that

$$M = QDQ^T \quad \text{or} \quad D = Q^T MQ$$

where D is a diagonal matrix with the eigenvalues of M, which are all real, on the diagonal.

A symmetric matrix M is said to be *positive definite*, denoted by $M > 0$, if $x^T Mx > 0$ for every nonzero x. It is *positive semidefinite*, denoted by $M \geq 0$, if $x^T Mx \geq 0$ for every nonzero x. If $M > 0$, then $x^T Mx = 0$ if and only if $x = 0$. If $M \geq 0$ but not $M > 0$, then there exists a nonzero x such that $x^T Mx = 0$.

Eigenvalue Decomposition

Theorem 7

A symmetric $n \times n$ matrix M is positive definite (positive semi-definite) if and only if any one of the following conditions holds.

1. Every eigenvalue of M is positive (zero or positive).
2. All the *leading* principal minors of M are positive (all the principal minors of M are zero or positive).
3. There exists an $n \times n$ nonsingular matrix N (an $n \times n$ singular matrix N or an $m \times n$ matrix N with $m < n$) such that $M = N^T N$.

Example

- **Square & Symmetric Matrix $n \times n$**

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \longrightarrow (\lambda-1)(\lambda-3) = 0 \longrightarrow \lambda_1 = 1 \text{ and } \lambda_2 = 3$$

$$x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } q_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

The vector is usually normalized. It a real square matrix whose columns and rows are orthonormal vectors.

$$Mx = \lambda x$$

$$M \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$M \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\longrightarrow M \begin{array}{c} \begin{array}{cc} q_1 & q_2 \end{array} \\ \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \end{array} = \begin{array}{c} \begin{array}{cc} q_1 & q_2 \end{array} \\ \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \end{array} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

$$MQ = QD$$

$$MQQ^{-1} = QDQ^{-1}$$

$$M = QDQ^{-1}$$

$$M = QDQ^T$$

Eigenvalue Decomposition

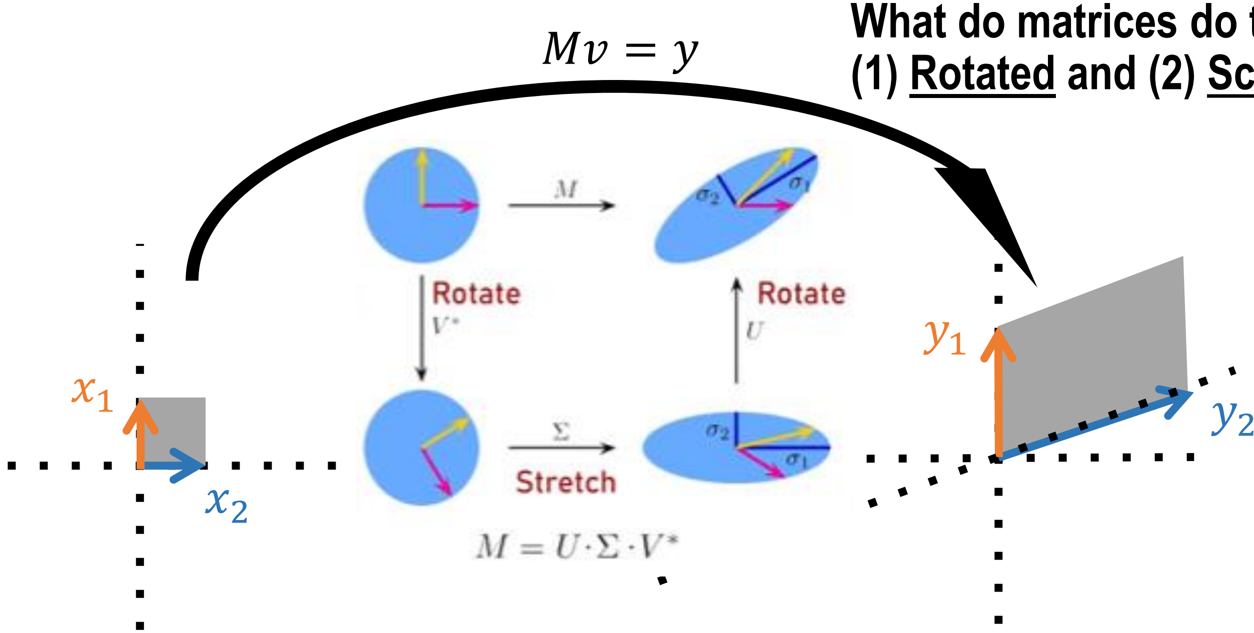
$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Singular Value Decomposition (SVD)

Vector Transformation

▪ Matrix Multiplication

What do matrices do to vectors?
(1) <u>Rotated</u> and (2) <u>Scaled/Stretched</u>



$$Mv_1 = y_1 \longrightarrow Mv_1 = \sigma_1 u_1$$
$$Mv_2 = y_2 \longrightarrow Mv_2 = \sigma_2 u_2$$

$$\longrightarrow M[v_1 \quad v_2] = [u_1 \quad u_2]\begin{bmatrix}\sigma_1 & 0 \\ 0 & \sigma_1\end{bmatrix}$$

v_1 and v_2 are orthogonal
y_1 and y_2 are orthogonal

$$MV = U\Sigma$$
$$M = U\Sigma V^{-1}$$

σ **magnitude and** u **unit direction**

Singular Value Decomposition

Singular Value Decomposition

Definition

Let H be an $m \times n$ real matrix. Define $M := H^T H$. Clearly M is $n \times n$, symmetric, and semidefinite. Thus all eigenvalues of M are real and nonnegative (zero or positive). Let r be the number of its positive eigenvalues. Then the eigenvalues of $M = H^T H$ can be arranged as

$$\lambda_1^2 \geq \lambda_2^2 \geq \cdots \lambda_r^2 > 0 = \lambda_{r+1} = \cdots = \lambda_n$$

Let $\bar{n} := \min(m, n)$. Then the set

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_{\bar{n}}$$

is called the singular values of H. The singular values are usually arranged in descending order in magnitude.

Singular Value Decomposition

Theorem 9 (Singular-value decomposition)

Every $m \times n$ matrix H can be transformed into the form

$$H = RSQ^T$$

with $R^T R = RR^T = I_m$, $Q^T Q = QQ^T = I_n$, and S being $m \times n$ with the singular values of H on the diagonal.

The columns of Q are orthonormalized eigenvectors of $H^T H$. The columns of R are orthonormalized eigenvectors of HH^T. Once R, S, and Q are computed, the rank of H equals the number of nonzero singular values. If the rank of H is r, the first r columns of R are an orthonormal basis of the range space of H. The $(n - r)$ columns of Q are an orthonormal basis of the null space of H.

Singular Value Decomposition

- **Any Matrix** $m \times n$

$$\underset{m \times n}{H} = \underset{m \times m}{R}\ \underset{m \times n}{S}\ \underset{n \times n}{Q^T}$$

R: Orthonormal Set (Independent) → Eigenvectors of HH^T

$(m \times n)(\mathrm{n} \times \mathrm{m})$

$(m \times m)$

S: Diagonal → Singular values of H

Q^T: Orthonormal Set (Independent) → Eigenvectors of $H^T H$

$(n \times m)(\mathrm{m} \times \mathrm{n})$

$(n \times n)$

Rotate **Stretch** **Rotate**

EVD versus SVD

Consider the EVD $M = \mathrm{Q}D\mathrm{Q}^{-1}$ and SVD $H = RS\mathrm{Q}^{T}$.

- The vectors in the EVD matrix Q are **not necessarily orthogonal**, so the change of basis isn't a simple rotation. On the other hand, the vectors in the matrices R and Q in the SVD are **orthonormal**, so they do represent rotations (and possibly flips).

- In the SVD, the nondiagonal matrices R and Q are **not necessarily the inverse** of one another. They are usually not related to each other at all. In the EVD the nondiagonal matrices Q and Q^{-1} are **inverses** of each other.

- In the SVD the entries in the diagonal matrix S are all **real and nonnegative**. In the EVD, the entries of D can be **any complex number** - negative, positive, imaginary.

- The SVD **always exists for any sort of rectangular or square ma**trix, whereas the EVD can **only exists for square matrices**, and even among square matrices sometimes it doesn't exist.

Example:

$$H = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$HH^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} \qquad H^T H = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$$

eigenvalues: $\lambda_1 = 25$, $\lambda_2 = 9$

eigenvectors

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad u_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

eigenvalues: $\lambda_1 = 25$, $\lambda_2 = 9$, $\lambda_3 = 0$

eigenvectors

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix} \quad v_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$$

The singular values are the square root of positive eigenvalues, i.e. 5 and 3.

Therefore, the SVD composition is

$$H = RSQ^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}$$

Homework #3.3 SVD (1 pt.): Due Dec. 22

Find the SVD of the matrix $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

Least-Squares Solution

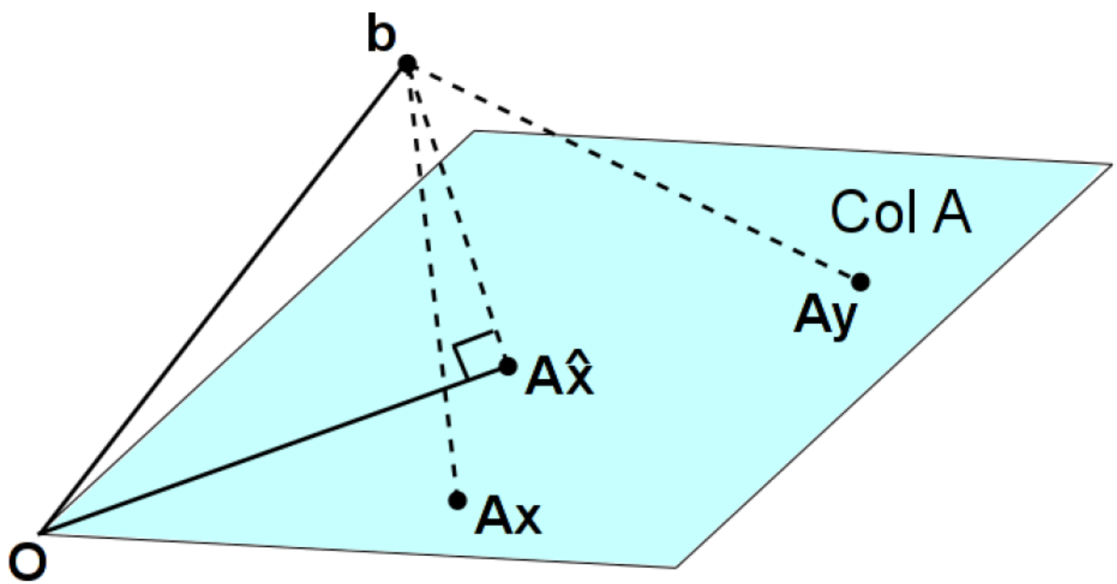
Least-Squares Solutions

Problem: What do we do when the matrix equation $A\mathbf{x} = \mathbf{b}$ has no solution $\mathbf{x}$?
Such inconsistent systems $A\mathbf{x} = \mathbf{b}$ often arise in applications, sometimes with large coefficient matrices.

Answer: Find $\hat{\mathbf{x}}$ such that $A\hat{\mathbf{x}}$ is as close as possible to $\mathbf{b}$.

In this situation $A\hat{\mathbf{x}}$ is an *approximation* to $\mathbf{b}$. The **general least squares problem** is to find an $\hat{\mathbf{x}}$ that makes $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ as small as possible.

Least-Squares Solutions



We seek $\hat{\mathbf{x}}$ such that $A\hat{\mathbf{x}}$ is the closest point to $\mathbf{b}$ in Col A.

Equivalently, we need to find $\hat{\mathbf{x}}$ with the property that $A\hat{\mathbf{x}}$ is the orthogonal projection of $\mathbf{b}$ onto $\text{Col}(A)$.

Example

$$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$$

Find x for $Ax = b$ → No solution

To solve the normal equations $A^T A\hat{\mathbf{x}} = A^T \mathbf{b}$, we first compute the relevant matrices:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

So we need to solve $\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \hat{\mathbf{x}} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$

Example

$$\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \hat{\mathbf{x}} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 11 & 14 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 3 & 11 & 14 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 8 & 8 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

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

This gives $\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow A\hat{\mathbf{x}} = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$

Note that $\begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$ is the closest point in Col A to $\mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$

Norm of Matrices

Norm of Vectors

Norm

→ The concept of norm is a generalization of length or magnitude

→ Any real-valued function of x, denoted by $\|x\|$ can be defined as a norm if it has the following properties:

1. $\|x\| \geq 0$ for every x and $\|x\| = 0$ if and only if $x=0$.

2. $\|\alpha x\| = |\alpha| \|x\|$, for any real α.

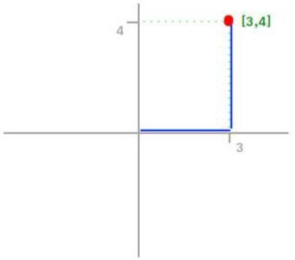
3. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ for every x_1 and x_2.
   the *triangular inequality*

Norm of Vectors

▪ Various Norms

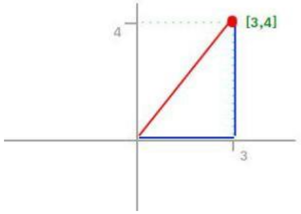
Let $x = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$.

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|$$



1-norm
the distance you have to travel between the origin (0,0) to the destination

$$\|x\|_2 := \sqrt{x^T x} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$$



2-norm
the shortest distance to go from one point to another (Euclidean norm)

$$\|x\|_\infty := \max_i |x_i|$$

infinity-norm
the largest magnitude among each element of a vector

Norm of Matrices

Let A be an $m \times n$ matrix. The norm of A can be defined as

$$\| A \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|} = \sup_{\|x\|=1} \| Ax \|$$

supremum or the least upper bound

For different $\| x \|$, we have different $\| A \|$.

If the 1-norm $\| x \|_1$ is used, then

$$\| A \|_1 = \max_j \left(\sum_{i=1}^{m} | a_{ij} | \right) = \textit{ largest column absolute sum}$$

the ijth element of A

Norm of Matrices

If the Euclidean norm $\| x \|_2$ is used, then

$$\| A \|_2 = \textit{largest singular value of } A$$

$$= (\textit{largest eigenvalue of } A^T A)^{1/2}$$

If the infinite-norm $\| x \|_\infty$ is used, then

$$\| A \|_\infty = \max_i \left(\sum_{j=1}^{m} | a_{ij} | \right) = \textit{largest row absolute sum}$$

$\boxed{\text{Example}}$

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

$$\| A \|_1 = 3 + |-1| = 4, \quad \| A \|_2 = 3.7, \quad \| A \|_\infty = 3 + 2 = 5$$

Norm of Matrices

The norm of matrices has the following properties:

$$
\begin{aligned}
\| Ax \| &\leq \| A \| \| x \| \\
\| A + B \| &\leq \| A \| + \| B \| \\
\| AB \| &\leq \| A \| \| B \|
\end{aligned}
$$

Thank you