

Masters

MTH501 Final Term
Preparation
2023

Solved By Masters
MASTERS

Subscribes to Masters

Question: **41** (Marks: 2)

Determine whether the vectors $u = (0, 3, -2, 1)$, $v = (5, 2, -1, 0)$ are orthogonal with respect to Euclidean inner product.

Question 41

$$u = (0, 3, -2, 1)$$

$$v = (5, 2, -1, 0)$$

$$u^T v = \begin{bmatrix} 0 \\ 3 \\ -2 \\ 1 \end{bmatrix} \cdot (5, 2, -1, 0)$$

$$= 0 + 6 + 2 + 1$$

$$= 9$$

$9 \neq 0$ so these vectors are ~~not~~ orthogonal.

$$u^T v = 0 \text{ orthogonal}$$

Question: **42** (Marks: 2)

The matrix A is followed by a sequence $\{x_k\}$ produced by the power method. Use these data to estimate the largest eigenvalue of A .

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

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.25 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.3158 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.3298 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.3326 \end{bmatrix}$$

Because the largest entry in x_k is 1 the largest entry in Ax_k is close to λ_1 .

Example#2 Handouts page 137,138

Question: **43** (Marks: 2)

Determine whether the set of all vectors of the form $(a, 1, 1)$ is a subspace of R^3 or not. Justify your answer with appropriate reason.

Handouts page 454,455

Question: **44** (Marks: 2)

Consider a linear transformation $T : R^2 \rightarrow R^2$ defined by $T(\vec{x}, \vec{y}) = (2x - y, x)$ then find $T(\vec{u})$, where $\vec{u} = (-2, 4)$.

Handouts page 407

Question: 45 (Marks: 3)

Construct the normal equation for $\hat{x}$ of the equation $Ax = b$ where

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

Question 45

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

$$A^T A \hat{x} = A^T b$$

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

$$= \begin{bmatrix} 1+1+1 & 3+(-1)+1 \\ 3+(-1)+1 & 9+1+1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \checkmark$$

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

$$\begin{bmatrix} 5 & 1 & 0 \\ 15 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

Question: **46** (Marks: 3)

$$S = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\}$$

Check whether the set

is a basis of $\mathbb{R}^2$?

Question 46

$$\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \sim R_1 \frac{1}{3}$$

$$\begin{bmatrix} 1 & 1/3 \\ 2 & 4 \end{bmatrix} \sim 2R_1 - R_2$$

$$\begin{bmatrix} 1 & 1/3 \\ 0 & +10/3 \end{bmatrix}$$

multiply R_2 by $\frac{3}{10}$

$$\begin{bmatrix} 1 & 1/3 \\ 0 & 1 \end{bmatrix} \quad R_1 - \frac{R_2}{R_3}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (2 \text{ pivot})$$

This is basis of R_2

Question: **49** (Marks: 5)

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$$

Let
solution.

use the factorization $A=QR$ to find the

Handouts page 490

Question: **52** (Marks: 5)

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

Find the Eigenvalues and corresponding Eigenvectors of the matrix .

Handouts page 385

Paper#2

Q-41

$$u = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, w = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}$$

Let .Find

(a) $\frac{v \cdot u}{u \cdot u}$ ✓

(b) $\|w\|$ ✓

Solution:

$$(1) \frac{v \cdot u}{u \cdot u}$$

solution:

$$v \cdot u = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$= (-1)(6) + (2)(4)$$

$$v \cdot u = -6 + 8 = 2$$

$$u \cdot u = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$= (-1)(2) + (-1)(2)$$

$$= -2 - 2$$

$$= -4$$

put values

$$\frac{v \cdot u}{u \cdot u} = \frac{2}{-4} = -\frac{1}{2}$$

$$(2) \|w\|$$

solution:

$$\|w\| = \sqrt{u_1 v_1}$$

$$= \sqrt{(-1, 2) \cdot (4, 6)} = -8$$

Q42-

Let $B = \{\vec{b}_1, \vec{b}_2\}$ and $C = \{\vec{c}_1, \vec{c}_2\}$ be basis for a vector space V , suppose that

$$\begin{bmatrix} \vec{b}_1 \end{bmatrix}_C = \begin{bmatrix} 6 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \vec{b}_2 \end{bmatrix}_C = \begin{bmatrix} 9 \\ 4 \end{bmatrix} \text{ and } \vec{x} = -3\vec{b}_1 + 2\vec{b}_2. \text{ Find } \begin{bmatrix} \vec{x} \end{bmatrix}_C.$$

Solution:

$$\begin{bmatrix} \vec{x} \end{bmatrix}_C = -3 \begin{bmatrix} \vec{b}_1 \end{bmatrix}_C + 2 \begin{bmatrix} \vec{b}_2 \end{bmatrix}_C$$

now put values in equation

$$\begin{bmatrix} \vec{x} \end{bmatrix}_C = -3 \begin{bmatrix} 6 \\ -2 \end{bmatrix} + 2 \begin{bmatrix} 9 \\ -4 \end{bmatrix} \text{ answer}$$

Q44:

Let the matrix $A = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Determine whether or not $\vec{u}$ is the eigenvector of A ?

Solution:

we know that

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

$$(A - \lambda I)x = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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

$$= (2)(\lambda) + (0)(0) + (0)(0) + (-2)(\lambda)$$

$$= 2\lambda - 2\lambda$$

$$\text{then } \det(A - \lambda I)x = 0$$

$$|A - \lambda I| = |2\lambda - 2\lambda|$$

Q47:

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 12 & 0 & -5 \\ 4 & 2 & -1 \end{bmatrix}$$

The eigen values of a matrix to eigenvalue $\lambda = 2$

Solution:

Eigenvalues are given

$$-1, 1, 2$$

$$\lambda = 2$$

so for eigenvector

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

put values in this equation

$$(A - \lambda I)x = \left(\begin{bmatrix} 3 & 1 & -1 \\ 12 & 0 & -5 \\ 4 & 2 & -1 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

.

.

further simplify

Q48:

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Verify that A is diagonalizable, by computing $P^{-1}AP$

Solution:

P is given

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

we have to find P inverse = P^{-1}

$$P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

now put values in equation

$$P^{-1}AP = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

.

further simplify

Q51:

The eigenvalue of the matrix $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ is -4 . Find the corresponding eigenvector of A .

Solution:

Eigenvalues are given

$$\lambda = -4$$

so for eigenvector

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

put values in this equation

$$(A - \lambda I)x = \left(\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$(A - \lambda I)x = \left(\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$(A - \lambda I)x = (1)(4) - (6)(0) + (5)(0) - (2)(4) = 12$$

Q50-

Compute the orthogonal projection of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ onto the line through $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$ and the origin

Solution:

first we will find $y.u$ and $u.u$

$$y.u = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -4 \\ 1 \end{bmatrix} = 1 - 8 = -7$$

$$u.u = \begin{bmatrix} -4 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -4 \\ 1 \end{bmatrix} = -4 - 4 = -8$$

now put values in equation

$$\hat{y} = \frac{y.u}{u.u} u$$

$$\hat{y} = \frac{-7}{-8} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \frac{7}{8} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = .87 \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -3.5 \\ .87 \end{bmatrix}$$

now

$$y - \hat{y}$$

put values in equation

$$y - \hat{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} -3.5 \\ .87 \end{bmatrix}$$

Let $C[0, 2\pi]$ have the inner product $\int_0^{2\pi} f(t)g(t)dt$. Show that $\|\sin kt\|^2 = \pi$ for $k > 0$.

Verify that the signals 3^k and $(-3)^k$ are the solutions of the difference equation $y_{k+2} - 9y_k = 0$.

$$u = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}, v = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$$

Let $u, v, \|u\|^2, \|v\|^2$ and $\|u+v\|^2$. Do not use Pythagorean Theorem.

Show that the signals $(-1)^k, k(-1)^k$ and 5^k are linearly independent

Paper#3

Question: 42 (Marks: 2)

$$x = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix} \text{ and } y = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$$

Find the distance between

Question #42

$$x = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}, \quad y = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$$

$$\text{dis} = \|u - v\| \quad \checkmark$$

$$u - v = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix} - \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix} \quad \checkmark$$

$$= \begin{bmatrix} 0 + 4 \\ -5 + 1 \\ 2 - 8 \end{bmatrix} \rightarrow \begin{bmatrix} 4 \\ -4 \\ -6 \end{bmatrix} \quad \checkmark$$

$$\|u - v\| = \sqrt{(4)^2 + (-4)^2 + (-6)^2} \quad \checkmark$$

$$= \sqrt{16 + 16 + 36} \Rightarrow \sqrt{68} \quad \checkmark$$

Question: 43 (Marks: 2)

If order of matrix A is 9×11 with a two-dimensional null space, then $\text{rank}(A)$?

Question: 44 (Marks: 2)

Is the following set of vectors is orthogonal with respect to the Euclidean inner product on $\mathbb{R}^3$?

$$\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right)$$

Question: 45 (Marks: 3)

Construct the normal equation for x of the equation $Ax = b$

$$A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}$$

Question: 46 (Marks: 3)

Determine whether the set $S = \{u_1, u_2, u_3\}$ are an orthogonal basis for $\mathbb{R}^3$ and find the vector v as a linear combination of the given set of vectors in S , where

$$v = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Question: 47 (Marks: 3)

Verify that the signals 2^k and $(-1)^k$ are the solutions of the difference equation $y_{k+2} + 2y_{k+1} - 8y_k = 0$.

Linear Algebra (MTH501)

Question: 48 (Marks: 3)

Assume that the mapping $T: P_2 \rightarrow P_2$ defined by $T(a_0 + a_1 t + a_2 t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$ is linear transformation. Find the matrix representation of T relative to the basis $B = \{1, t, t^2\}$.

Question: 49 (Marks: 5)

Compute the least square error associated with the least square solution of the system of linear equations

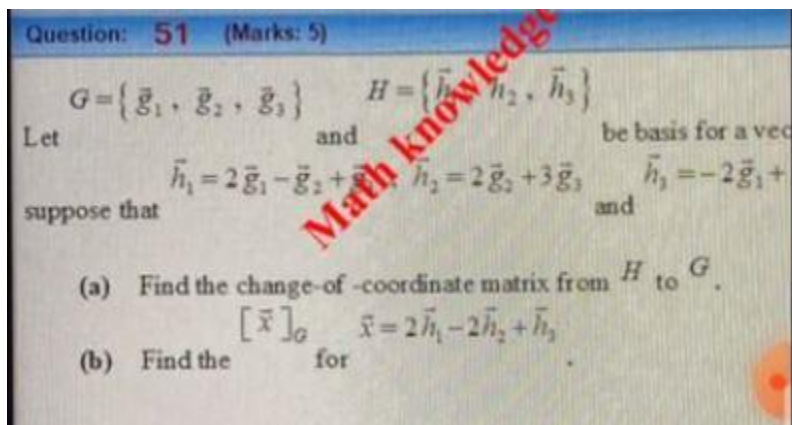
$$A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}, b = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}$$

equation $Ax = b$ where

Question: 50 (Marks: 5)

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

Determine whether the matrix A is diagonalizable. If so, find the invertible matrix P such that $P^{-1}AP = D$.



Paper#4

14 march 2022

2:30pm

40/30 Mcqs same past wala tha 10 mcs basic that past wala hi Thora SA change kr k d inya tha

Subjective paper Zara mushkil tha

1. Find characteristics equation 3 marks
2. Determine the basis of $\text{Col}(A)$ where A and B are Row equalant.. 3 marks
3. Find invertable Matrix 5 marks
4. Find basis set of sub spaced. 5 marks

Paper#4

15 feb 2022

More than half McQ's wr from past....

Lecture#25 example#3 type question

Lecture#26 example#3 type question

A matrix was given ...os matrix k transpose m number of pivot column find krny th...

Matrix given the... eigenvalues k liye characteristic equation find krni the....

u and v vector given that $\|u\|^2, \|v\|^2, \|u-v\|^2$ find k and then...

Find whether given vectors are orthogonal or not.

Aik question Fourier series s th...

Paper#5

Mth501 Final term current papers spring 2022

MCQs

Most mcqs are from junaid file

1. Which statement about the General Least Square Method is true? ✓

This method find an X that make Ax as close possible to the b.

2. Let I3 have the Education inner product. Then $u = (2, 1, 3)$, $V = (1, 7, k)$ are orthogonal for **(K= -3)**
3. The characteristics polynomial of 3 X 3 identity matrix is ----- if x is the eigen values of the given 3 x 3 identity matrix. **$(x - 1)^3$**
4. if a square matrix has orthogonal columns it also has **orthogonal rows**
5. if matrix U has orthogonal changes $UU^t =$ _____?
6. the given system $2x + 3y = 3$
 $6x + 9 = 7$ has **No solution**
7. if two row are orthogonal , they are ----- . **Linearly independent**
8. Let A be the matrix of order 2X3 and B be the matrix order 3X5, then which of the following is the order of the matrix AB? **2x5**
9. If $X - 2$ is a factor of the characteristic polynomial of matrix C then an eigenvalue of C is. **2**
10. Let A be an $n \times n$ Matrix whose entries are real. If λ is an eigenvalue of A with x a corresponding eigenvector in $\sum^n$, then **$Ax = \lambda x$**
11. Let $u = (3, -2)$, $v = (4, 5)$. For the weighted Euclidean inner product $(u, v) = 4u_1v_1 + 5u_2v_2$ **-2**
12. If r is set of real numbers which is $(\mathbb{R}^n)$?
13. Suppose that $A = \begin{bmatrix} 1.25 & -0.75 \\ -0.75 & 1.25 \end{bmatrix}$

$\begin{bmatrix} -0.75 & 1.25 \end{bmatrix}$ has eigenvalues 2 and 0.5. then Origin is a

Saddle point

14. differential operator is linear or non linear **linear**
15. if the magnitude of the eigenvalue is greater than 1 what is about origin

Subjective

1. Find eigen value
2. Tell the columns of matrix are orthonormal or not
3. If $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ then find an invertible matrix P such that $P^{-1}AP = D$ (diagonal matrix)

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

Let $W = \text{Span} \{x_1, x_2\}$, where $x_1 = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}$. Construct an orthogonal basis $\{v_1, v_2\}$ for W .

4. Let $A = \begin{bmatrix} 1 & 4 \\ 5 & 6 \end{bmatrix}$ and $\vec{c} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\vec{x}) = A\vec{x}$. Determine if $\vec{c}$ is in the range of the transformation T .
- 5.

And most paper are from 36 to 43 lecture for subjective.

Paper#6 MCQs

All mcqs are from past paper waqar files and junaaid

Subjective

1. Orthogonal or not?
2. Find image of vector
3. LU decomposition
4. Block matrix find out it is invertible or not?

Most part of the paper are from 35 to 40 lectures

Paper#7

MCQs

All Mcqs are from junaid and waqar file repeated 85%

Subjective

1. Find out whether the vectors are linearly independent or not?
2. Find out matrix is invertible or not?
3. Find Eigen values of given matrix
4. Find Orthogonal Projection of y onto $\text{span}(u, u_2)$
5. Find the value of k if $|A| = 6$ dia hua tha

Most paper from 36 to 42

80%

Lecture # 23 (Imp for Qs) ✓

$$[x]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \Rightarrow B \text{ coordinate of } x$$

Ex #1 ✓

$$b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, [x]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, x = ?$$

$$x = c_1 b_1 + c_2 b_2 \quad c_1 = -2, c_2 = 3$$

$$x = (-2) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$x = \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 \\ 6 \end{bmatrix} \quad \text{Ans}$$

Ex #3

$$P = a_0 + a_1 x + a_2 x^2, \text{ basis } S = [1, x, x^2]$$

Find coordinates vector = ?

$$a_0 + a_1 x + a_2 x^2 = c_1(1) + c_2(x) + c_3(x^2)$$

$$c_1 = a_0, c_2 = a_1, c_3 = a_2$$

$$[P]_S = (a_0, a_1, a_2)$$

To find coordinates of vector we have to write linear combination of the basis of set S .

$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n$ is equivalent

$$x = P_B [x]_B$$

→ P_B form a basis for $\mathbb{R}^n$

→ P_B is invertible

→ coordinate mapping is linear transformation

→ It preserves scalar multiplication

→ one-to-one linear transformation from a vector space V onto vector space W is **Isomorphism**.

* Iso → same, morph → structure

Lecture # 24 (NCC) ✓

→ All the vectors in Basis of vector are dimension of vector space.

→ Dimension of zero vector space = 0

→ Every finite dimensional vector space

contains a basis.

→ $m \times n \Rightarrow$ finite dimensional vector space.

→ $(-\infty, \infty)$ infinite " "

→ $\mathbb{R}^2 = \{i, j\} \Rightarrow 2$ ✓

→ $\mathbb{R}^3 = \{i, j, k\} \Rightarrow 3$ ✓

→ standard basis of $\mathbb{R}^n = (x_1 e_1 + x_2 e_2 + \dots + x_n e_n)$

→ $\dim(\mathbb{R}^n) = n$

$\dim(P_n) = n+1$

$\dim(M_{m \times n}) = mn$ ✓

→ $H \leq \dim V$

→ $P \geq 1$

Lecture # 25 (MCG) ✓

Rank: ✓

→ no of non-zero row left after row reduce echelon form are rank.

→ All possible row can become linear combination than called row space.

→ The dimension of null space called nullity of A . ✓

→ $\text{rank } A + \dim \text{Nul } A = n$ ✓

→ $\text{rank } (A) = \text{rank } (A^T)$ ✓

→ $\text{col } A = \mathbb{R}^n$ ✓

→ $\dim \text{col } A = n$ ✓

→ $\text{rank } A = n$

→ $\text{Nul } A = \{0\}$ ✓

→ $\dim \text{Nul } A = 0$ ✓

→ dimension of row space = dimension of column space.

→ $\text{nullity } (A) = n - \text{rank}$
Parameter.

→ Fundamental matrix spaces | Dimensions ★

- 1- Row space of A ✓
- 2- column space of A ✓
- 3- Null space of A ✓
- 4- Null space of A^T ✓

r ✓

r

$n - r$ ✓

$m - r$ ✓

$$\text{Nullity}(A) = n - r$$

$$\text{nullity}(A^T) = m - r$$

$$\rightarrow \dim \text{nul } A = \text{column} - \text{rank}$$

$$\rightarrow \text{Rank} = \text{Pivot column values}$$

$$\rightarrow \text{col } A^T = \text{Row}$$

Q. if a 3×8 matrix has rank = 3 then

$$\dim \text{Nul } A = ?$$

$$\dim \text{row } A = ?$$

$$\text{Rank } A^T = ?$$

$$\dim \text{nul } A = \text{column} - \text{rank} \Rightarrow 8 - 3 = 5$$

Lecture # 26 (Imp for short Q) ✓

Example #4 ✓ most important

Lecture # 27 (Imp for short) ✓

→ Discrete or digital data has been prefer more than continuous data.

→ Example # 5 important ✓

Lecture 28, 29 are imp for short questions

Lecture # 28 : (most imp for questions)

~~Characteristic~~ equation used to find fixed point

$$(I - A)x = 0$$

→ Every square matrix has at least 1 fixed point.

Theorem : mca

- A has non-trivial fixed points
- $I - A$ is singular
- $\det(I - A) = 0$

$\det(I - A) \neq 0$ trivial fixed point

$\det(I - A) = 0$ non-trivial fixed point

→ $Ax = \lambda x$ λ scalar (eigenvalue)

Example 1, 2, 3 important

Lecture # 29

→ A is invertible if $\det A \neq 0$

Properties of determinants all are imp



Lecture #30 (Imp for short Q) ✓

$$\rightarrow A^2 = P D P^{-1}, \boxed{A^2 = P D^2 P^{-1}} \checkmark$$

$$A^k = P D^k P^{-1} \quad \because \text{only effect on } D.$$

invertible matrix diagonal matrix

→ Matrix is diagonalizable if A has n linearly independent eigenvectors.

→ To avoid computing P^{-1} we can

$$AP = PD \quad \text{if matrix order is high.}$$

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

→ matrix order = eigenvalues.

→ For diagonal we write eigenvalues on diagonal and remaining values will be 0.

For example:
eigenvalues are 5, 5, -3, -3

$$D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \quad \text{diagonal matrix}$$

Lecture # 31 (imp for shoot) ✓

$$M = [T(b_1)]_c \quad [T(b_2)]_c \quad \dots \quad [T(b_n)]_c$$

To find linear transformation

$$P(t) = (a_0 + a_1 t + a_2 t^2)$$

Example # 4 important ✓

Find $T(a_0 + a_1 t + a_2 t^2)$ if $B = [1, t, t^2]$, ✓

$$[T]_B = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

Solution :-

$$[T]_B [P]_B = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

Diagram showing the matrix multiplication with red arrows indicating the dot product of rows of the transformation matrix with the column vector of coefficients. Red checkmarks are present.

$$= \begin{bmatrix} 3a_0 + 4a_1 + 0a_2 \\ 0a_0 + 5a_1 - 1a_2 \\ 1a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

Red checkmark.

$$= \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

Red checkmark.

now put these values to $T(a_0 + a_1 t + a_2 t^2)$

$$= (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$$

$$[T(x)]_B = [T]_B [x]_B$$

$$A = (I)^{-1} A I \quad A \text{ is similar to } A$$

Lecture # 32 (MCQs)

$$\rightarrow Ax = \lambda x$$

$\rightarrow$ conjugate of complex vector x will be $\bar{x}$.

$\rightarrow$ entries in $\bar{x}$ are same that was in x but it has two parts one is real and one is imaginary.

without i = real
with i = imaginary

$\rightarrow \phi$ is called argument of $A = a + bi$

ϕ is angle b/w +ve x-axis and ray from (0,0).

$$\rightarrow A = P C P^{-1}$$

$$\rightarrow P = \begin{bmatrix} \text{Re}v & \text{Im}v \end{bmatrix}$$

Lecture # 33 (MCGs) ✓

→ The graph of x_0, x_1 is called trajectory of dynamical system. ✓

$$\rightarrow x_0 = C_1 V_1 + C_2 V_2$$

→ * MCGs trajectories towards 0 = origin is attractor. ✓

→ eigenvalues smaller = greater attraction. ✓

→ trajectories away from 0 = repeller. ✓

→ eigenvalues largest = greatest repulsion. ✓

→ eigenvalue greater than 1 in magnitude than 0 will saddle point ✓

→ complex eigenvalues = vector complex ✓

Lecture # 34 (MCGs) ✓

- $x = Ax$, $x(0) = x_0$ initial value prob
- A has +ve, -ve values than origin will behave as saddle point. ✓ ✓
- eigenvalue positive = repulsion ✓ } +
eigenvalue negative = attraction ✓ } -
- eigenvalue rotation caused by sin and cosine function then origin spiral point.
- eigenvalue negative = trajectories spiral inward ✓
- eigenvalues positive = trajectories spiral outward ✓
- eigenvalue 0 = trajectories form ellipses around the origin. ✓

Lecture # 35 (Imp for short Q) ✓

- eigenvalues s- $|A - \lambda I| = 0$

Example #1

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

Example #4, 3, 2 important ✓

$$x_{k+1} = \frac{1}{\mu_k} A x_k$$

$$A x_1 = \frac{1}{\mu_0} A x_0$$

$$A x_1 = \frac{1}{\mu_1}$$

$$x_{k+1} = \frac{1}{\mu_k} A x_k$$

$$\rightarrow A x_0 \checkmark$$

$$\rightarrow x_1 = \frac{1}{\mu_0} A x_0 \checkmark$$

$$\rightarrow A x_1 \checkmark$$

$$\rightarrow x_2 = \frac{1}{\mu_1} A x_1 \checkmark$$

$$A x_2 \checkmark$$

→ continued.

→ Lecture 36 and 37 are revision of
lectures 28 to 35. ✓

Lecture 38 :- (Imp for short questions)

inner product = dot product

$$u \cdot v = u^T v$$

$$\left\{ \begin{array}{l} u \cdot v = v \cdot u \quad \text{u, v imp} \\ (u+v) \cdot w = u \cdot w + v \cdot w \\ (cu) \cdot v = c(u \cdot v) = u \cdot (cv) \\ u \cdot u \geq 0, \quad u \cdot u = 0 \text{ if } u=0 \\ u \cdot v = 0 \Rightarrow \text{orthogonal} \end{array} \right. \quad \checkmark$$

then

Pythagorean theorem

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2$$

orthogonal complement $W = W^\perp$

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos \theta$$

Lecture # 39 : MCG ✓

$u \cdot v = 0$ then orthogonal

$$c_1 u_1 + c_2 u_2 = 0$$

$$c_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3c_1 \\ c_1 \end{bmatrix} + \begin{bmatrix} -c_2 \\ 3c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$3c_1 - c_2 = 0$$

$$c_1 + 3c_2 = 0$$

$c_1 = c_2 = 0$ orthogonal set are linearly independent.

$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

$$y^n = \frac{y \cdot u}{u \cdot u} u$$

$$\rightarrow U^T U = I$$

QR factorization ✓

Find orthonormal basis by applying gram schmidt process.

$$V_1 = x_1 \quad \checkmark$$

$$V_2 = x_2 - \frac{x_2 \cdot V_1}{V_1 \cdot V_1} V_1 \quad \checkmark$$

$$V_3 = x_3 - \frac{x_3 \cdot V_1}{V_1 \cdot V_1} V_1 - \frac{x_3 \cdot V_2}{V_2 \cdot V_2} V_2$$

$\Rightarrow V_1, V_2, V_3$ are orthonormal basis ✓

$\rightarrow$ now find orthonormal ✓

$$Q = \left\{ \frac{V_1}{\|V_1\|}, \frac{V_2}{\|V_2\|}, \frac{V_3}{\|V_3\|} \right\} \quad \checkmark$$

Take transpose of Q then ✓

multiply with A . ✓

$$R = Q^T A \quad \checkmark$$

Lecture 42:

$$\|b - A\hat{x}\| \leq \|b - Ax\| \quad \checkmark$$

$$A^T A \hat{x} = A^T b \quad \checkmark$$

$$\text{least square error} = \|b - A\hat{x}\|$$

$$\hat{x} = (A^T A)^{-1} A^T b \quad \checkmark$$

$$\boxed{A^T A \hat{x} = A^T b} \quad \checkmark$$

Lecture # 43

Examples imp

→ Cauchy - Schwarz inequality

$$|(u, v)| \leq \|u\| \|v\| \quad \checkmark$$

→ Triangle inequality

$$\|u + v\| \leq \|u\| + \|v\|$$

Lecture # 44

Example #1 important
weighted least square. ✓

$$\|W y - W y\|^2 = \|W y - W A \hat{y}\|^2$$

Question 41

$$u = (0, 3, -2, 1)$$

$$v = (5, 2, -1, 0)$$

$$u^T v = \begin{bmatrix} 0 \\ 3 \\ -2 \\ 1 \end{bmatrix} \cdot (5, 2, -1, 0)$$

$$= 0 + 6 + 2 + 1$$

$$= 9$$

$9 \neq 0$ so these vectors are not orthogonal.

$u^T v = 0$ orthogonal

Question # 43

$v = (a, 1, 1)$ is subspace of $\mathbb{R}^3$

$$\|v\| = \sqrt{(a^2) + (1)^2 + (1)^2} = \sqrt{a^2 + 2}$$

$$u = \frac{1}{\|v\|} v$$

$$= \frac{1}{\sqrt{a^2 + 2}} \cdot (a, 1, 1) \Rightarrow \left(\frac{a}{\sqrt{a^2 + 2}}, \frac{1}{\sqrt{a^2 + 2}}, \frac{1}{\sqrt{a^2 + 2}} \right)$$

Lecture No.23

Coordinate System

OBJECTIVES

The objectives of the lecture are to learn about:

- Unique representation theorem.
- Coordinate of the element of a vector space relative to the basis B.
- Some examples in which B- coordinate vector is uniquely determined using basis of a vector space.
- Graphical interpretation of coordinates.
- Coordinate Mapping

Theorem

Let $B = \{b_1, b_2, \dots, b_n\}$ be a basis for a vector space V . Then for each x in V , there exist a unique set of scalars $c_1, c_2, \dots, c_n$ such that

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n \dots\dots\dots (1)$$

Proof

Since B is a basis for a vector space V , then by definition of basis every element of V can be written as a linear combination of basis vectors. That is if $x \in V$, then

$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n$. Now, we show that this representation for x is unique.

For this, suppose that we have two representations for x .

i.e.

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n \dots\dots\dots (2)$$

and

$$x = d_1 b_1 + d_2 b_2 + \dots + d_n b_n \dots\dots\dots (3)$$

We will show that the coefficients are actually equal. To do this, subtracting (3) from (2), we have

$$0 = (c_1 - d_1)b_1 + (c_2 - d_2)b_2 + \dots + (c_n - d_n)b_n.$$

Since B is a basis, it is linearly independent set. Thus the coefficients in the last linear combination must all be zero. That is

$$c_1 = d_1, \dots, c_n = d_n.$$

Thus the representation for x is unique.

Definition (B-Coordinate of x)

Suppose that the set $B = \{b_1, b_2, \dots, b_n\}$ is a basis for V and x is in V . The coordinates of x relative to basis B (**or the B-coordinate of x**) are the weights $c_1, c_2, \dots, c_n$ such that

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n.$$

Note

If $c_1, c_2, \dots, c_n$ are the **B**- coordinates of x , then the vector in R^n , $[x]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ is the

coordinate vector of x (relative to **B**) or **B**- coordinates of x .

Example 1

past paper

Consider a basis $B = \{b_1, b_2\}$ for R^2 , where $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Suppose an x in R^2 has the coordinate vector $[x]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find x .

Solution

Using above definition x is uniquely determined using coordinate vector and the basis. That is

$$\begin{aligned} x &= c_1 b_1 + c_2 b_2 \\ &= (-2)b_1 + (3)b_2 \\ &= (-2) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ x &= \begin{bmatrix} 1 \\ 6 \end{bmatrix} \end{aligned}$$

Example 2

Let $S = \{v_1, v_2, v_3\}$ be the basis for R^3 , where $v_1 = (1, 2, 1)$, $v_2 = (2, 9, 0)$, and $v_3 = (3, 3, 4)$.

- Find the coordinates vector of $v = (5, -1, 9)$ with respect to S .
- Find the vector v in R^3 whose coordinate vector with respect to the basis S is

$$[v]_S = (-1, 3, 2)$$

Solution

Since S is a basis for R^3 , Thus

$$x = c_1 v_1 + c_2 v_2 + c_3 v_3.$$

Further

$$(5, -1, 9) = c_1(1, 2, 1) + c_2(2, 9, 0) + c_3(3, 3, 4) \dots\dots\dots (A)$$

To find the coordinate vector of $\mathbf{v}$, we have to find scalars c_1, c_2, c_3 .

For this equating corresponding components in (A) gives

$$c_1 + 2c_2 + 3c_3 = 5 \quad (1)$$

$$2c_1 + 9c_2 + 3c_3 = -1 \quad (2)$$

$$c_1 + 4c_3 = 9 \quad (3)$$

Now find values of c_1, c_2 and c_3 from these equations.

From equation (3)

$$c_1 = 9 - 4c_3$$

Put this value of c_1 in equations (1) and (2)

$$9 - 4c_3 + 2c_2 + 3c_3 = 5$$

$$2c_2 - c_3 = -4 \quad (4)$$

and

$$2(9 - 4c_3) + 9c_2 + 3c_3 = -1$$

$$18 - 8c_3 + 9c_2 + 3c_3 = -1$$

$$9c_2 - 5c_3 = -19 \quad (5)$$

Multiply equation (4) by 5

$$10c_2 - 5c_3 = -20$$

Subtract equation (5) from above equation

$$10c_2 - 5c_3 = -20$$

$$\pm 9c_2 \mp 5c_3 = \mp 19$$

$$c_2 = -1$$

Put value of c_2 in equation (4) to get c_3

$$2(-1) - c_3 = -4$$

$$-2 - c_3 = -4$$

$$c_3 = 4 - 2 = 2$$

Put value of c_3 in equation (3) to get c_1

$$c_1 + 4(2) = 9$$

$$c_1 = 9 - 8 = 1$$

Thus, we obtain $c_1 = 1, c_2 = -1, c_3 = 2$

Therefore, $[\mathbf{v}]_s = (1, -1, 2)$

Using the definition of coordinate vector, we have

$$\begin{aligned}
 v &= c_1 v_1 + c_2 v_2 + c_3 v_3 \\
 &= (-1)v_1 + 3v_2 + 2v_3 \\
 &= (-1)(1, 2, 1) + 3(2, 9, 0) + 2(3, 3, 4) \\
 &= (-1 + 6 + 6, -2 + 27 + 6, -1 + 0 + 8) \\
 &= (11, 31, 7)
 \end{aligned}$$

Therefore

$$v = (11, 31, 7)$$

Example 3

Find the coordinates vector of the polynomial $p = a_0 + a_1x + a_2x^2$ relative to the basis $S = \{1, x, x^2\}$ for p_2 .

Solution

To find the coordinator vector of the polynomial p , we write it as a linear combination of the basis set S . That is

$$\begin{aligned}
 a_0 + a_1x + a_2x^2 &= c_1(1) + c_2(x) + c_3(x^2) \\
 \Rightarrow c_1 &= a_0, c_2 = a_1, c_3 = a_2
 \end{aligned}$$

Therefore

$$[p]_s = (a_0, a_1, a_2)$$

Example 4

Find the coordinates vector of the polynomial $p = 5 - 4x + 3x^2$ relative to the basis $S = \{1, x, x^2\}$ for p_2 .

Solution

To find the coordinator vector of the polynomial p , we write it as a linear combination of the basis set S . That is

$$\begin{aligned}
 5 - 4x + 3x^2 &= c_1(1) + c_2(x) + c_3(x^2) \\
 \Rightarrow c_1 &= 5, c_2 = -4, c_3 = 3
 \end{aligned}$$

Therefore

$$[p]_s = (5, -4, 3)$$

Example 5

Find the coordinate vector of A relative to the basis $S = \{A_1, A_2, A_3, A_4\}$

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}; A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}; A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}; A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Solution

To find the coordinator vector of A , we write it as a linear combination of the basis set S . That is

$$\mathbf{A} = c_1 \mathbf{A}_1 + c_2 \mathbf{A}_2 + c_3 \mathbf{A}_3 + c_4 \mathbf{A}_4$$

$$\begin{aligned} \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} &= c_1 \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -c_1 & c_1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} c_2 & c_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c_3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_4 \end{bmatrix} \\ &= \begin{bmatrix} -c_1 + c_2 + 0 + 0 & c_1 + c_2 + 0 + 0 \\ 0 + 0 + c_3 + 0 & 0 + 0 + 0 + c_4 \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -c_1 + c_2 & c_1 + c_2 \\ c_3 & c_4 \end{bmatrix}$$

$$-c_1 + c_2 = 2 \quad (1)$$

$$c_1 + c_2 = 0 \quad (2)$$

$$c_3 = -1 \quad (3)$$

$$c_4 = 3 \quad (4)$$

Adding (1) and (2), gives

$$2c_2 = 2 \Rightarrow c_2 = 1$$

Putting the value of c_2 in (2) to get c_1 , $c_1 = -1$

So $c_1 = -1, c_2 = 1, c_3 = -1, c_4 = 3$

Therefore, $[\mathbf{v}]_s = (-1, 1, -1, 3)$

Graphical Interpretation of Coordinates

A coordinate system on a set consists of a one-to-one mapping of the points in the set into $\mathbb{R}^n$. For example, ordinary graph paper provides a coordinate system for the plane when one selects perpendicular axes and a unit of measurement on each axis. Figure 1 shows the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, the vectors

$\mathbf{b}_1 (= \mathbf{e}_1)$ and $\mathbf{b}_2$ from Example 1, that is,

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$, the coordinates 1 and 6 give the location of $\mathbf{x}$ relative to the standard basis: 1 unit in the $\mathbf{e}_1$ direction and 6 units in the $\mathbf{e}_2$ direction.

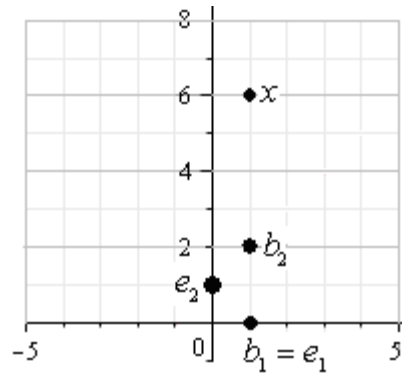


Figure 1

Figure 2 shows the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{x}$ from Figure 1. (Geometrically, the three vectors lie on a vertical line in both figures.) However, the standard coordinate grid was erased and replaced by a grid especially adapted to the basis $\mathbf{B}$ in Example 1. The coordinate vector $[\mathbf{x}]_{\mathbf{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ gives the location of $\mathbf{x}$ on this new coordinate system: -2 units in the $\mathbf{b}_1$ direction and 3 units in the $\mathbf{b}_2$ direction.

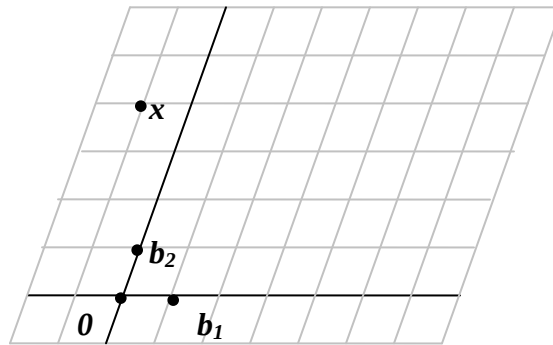


Figure 2

Example 6

In crystallography, the description of a crystal lattice is aided by choosing a basis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ for $\mathbf{R}^3$ that corresponds to three adjacent edges of one “unit cell” of the crystal. An entire lattice is constructed by stacking together many copies of one cell. There are fourteen basic types of unit cells; three are displayed in Figure 3.

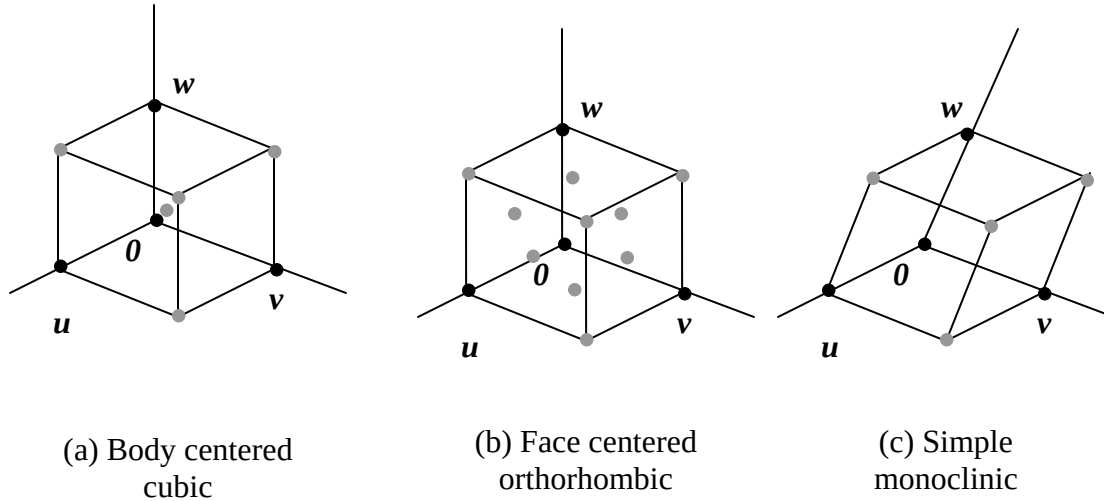


Figure 3 – Examples of unit cells

The coordinates of atoms within the crystal are given relative to the basis for the lattice.

For instance, $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$ identifies the top face-centered atom in the cell in Figure 3(b).

Coordinates in R^n When a basis B for R^n is fixed, the B -coordinate vector of a specified x is easily found, as in the next example.

Example 7 Let $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and $B = \{b_1, b_2\}$.

Find the coordinate vector $[x]_B$ of x relative to B .

Solution The B -coordinates c_1, c_2 of x satisfy

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$b_1 \qquad b_2 \qquad x$

or

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad (3)$$

$b_1 \quad b_2 \qquad x$

Now, inverse of matrix $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

From equation (3) we get

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3}(4) + \frac{1}{3}(5) \\ -\frac{1}{3}(4) + \frac{2}{3}(5) \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \end{aligned}$$

Thus, $c_1 = 3, c_2 = 2$.

(Equation (3) can also be solved by row operations on an augmented matrix. Try it yourself)

Thus $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$ and $[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

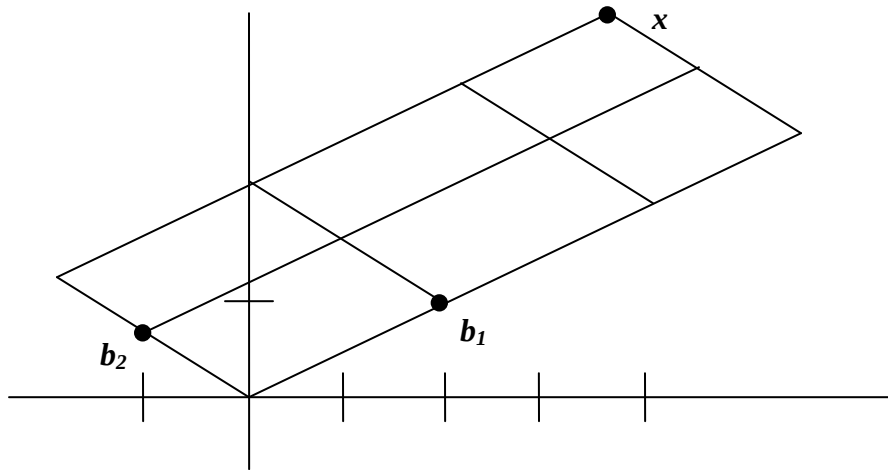


Figure 4 – The B -coordinate vector of $\mathbf{x}$ is $(3,2)$

The matrix in (3) changes the B -coordinates of a vector $\mathbf{x}$ into the standard coordinates for $\mathbf{x}$. An analogous change of coordinates can be carried out in $\mathbf{R}^n$ for a basis

$B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$.

Let $P_B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_n]$

Then the vector equation $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \dots + c_n\mathbf{b}_n$

is equivalent to $\mathbf{x} = P_B[\mathbf{x}]_B$ (4)

We call P_B the **change-of-coordinates matrix** from B to the standard basis in $\mathbf{R}^n$.

Left-multiplication by P_B transforms the coordinate vector $[\mathbf{x}]_B$ into $\mathbf{x}$. The change-of-coordinates equation (4) is important and will be needed at several points in next lectures.

Since the columns of P_B form a basis for R^n , P_B is invertible (by the Invertible Matrix Theorem). Left-multiplication by P_B^{-1} converts x into its B -coordinate vector:

$$P_B^{-1}x = [x]_B$$

The correspondence $x \rightarrow [x]_B$ produced here by P_B^{-1} , is the coordinate mapping mentioned earlier. Since P_B^{-1} is an invertible matrix, the coordinate mapping is a one-to-one linear transformation from R^n onto R^n , by the Invertible Matrix Theorem. (See also Theorem 3 in lecture 10) This property of the coordinate mapping is also true in a general vector space that has a basis, as we shall see.

The Coordinate Mapping Choosing a basis $B = \{b_1, b_2, \dots, b_n\}$ for a vector space V introduces a coordinate system in V . The coordinate mapping $x \rightarrow [x]_B$ connects the possibly unfamiliar space V to the familiar space R^n . See Figure 5. Points in V can now be identified by their new “names”.

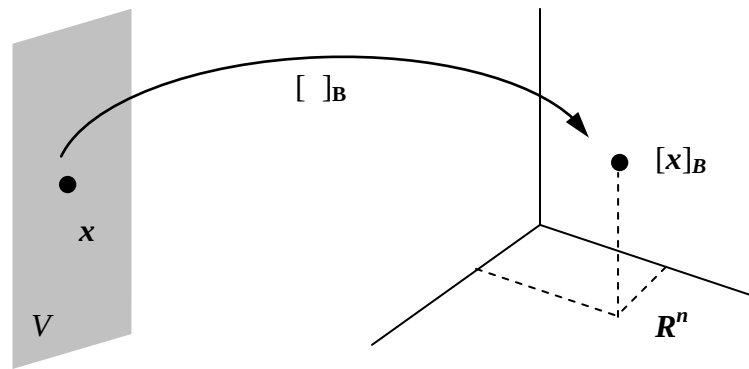


Figure 5 – The coordinate mapping from V onto R^n

Theorem 2 Let $B = \{b_1, b_2, \dots, b_n\}$ be a basis for a vector space V . Then the coordinate mapping $x \rightarrow [x]_B$ is a one-to-one linear transformation from V onto R^n .

Proof Take two typical vectors in V , say

$$u = c_1b_1 + c_2b_2 + \dots + c_nb_n$$

$$w = d_1b_1 + d_2b_2 + \dots + d_nb_n$$

Then, using vector operations, $u + w = (c_1 + d_1)b_1 + (c_2 + d_2)b_2 + \dots + (c_n + d_n)b_n$

$$\text{It follows that } [u + w]_B = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [u]_B + [w]_B$$

Thus the coordinate mapping preserves addition. If r is any scalar, then

$$ru = r(c_1b_1 + c_2b_2 + \dots + c_nb_n) = (rc_1)b_1 + (rc_2)b_2 + \dots + (rc_n)b_n$$

So
$$[r\mathbf{u}]_B = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\mathbf{u}]_B$$

Thus the coordinate mapping also preserves scalar multiplication and hence is a linear transformation. It can be verified that the coordinate mapping is one-to-one and maps V onto $\mathbf{R}^n$.

The linearity of the coordinate mapping extends to linear combinations, just as in lecture

9. If $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ are in V and if $c_1, c_2, \dots, c_p$ are scalars, then

$$[c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_p\mathbf{u}_p]_B = c_1[\mathbf{u}_1]_B + c_2[\mathbf{u}_2]_B + \dots + c_p[\mathbf{u}_p]_B \quad (5)$$

In words, (5) says that the B -coordinate vector of a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ is the same linear combination of their coordinate vectors.

The coordinate mapping in Theorem 2 is an important example of an isomorphism from V onto $\mathbf{R}^n$. In general, a one-to-one linear transformation from a vector space V onto a vector space W is called an **isomorphism** from V onto W (iso from the Greek for “the same”, and morph from the Greek for “form” or “structure”). The notation and terminology for V and W may differ, but the two spaces are indistinguishable as vector spaces. Every vector space calculation in V is accurately reproduced in W , and vice versa.

Example 8 Let B be the standard basis of the space P_3 of polynomials; that is, let $B = \{1, t, t^2, t^3\}$. A typical element p of P_3 has the form $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$. Since p is already displayed as a linear combination of the standard basis vectors, we

conclude that $[p]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Thus the coordinate mapping $p \rightarrow [p]_B$ is an isomorphism

from P_3 onto $\mathbf{R}^4$. All vector space operations in P_3 correspond to operations in $\mathbf{R}^4$.

If we think of P_3 and $\mathbf{R}^4$ as displays on two computer screens that are connected via the coordinate mapping, then every vector space operation in P_3 on one screen is exactly duplicated by a corresponding vector operation in $\mathbf{R}^4$ on the other screen. The vectors on the P_3 screen look different from those on the $\mathbf{R}^4$ screen, but they “act” as vectors in exactly the same way. See Figure 6.

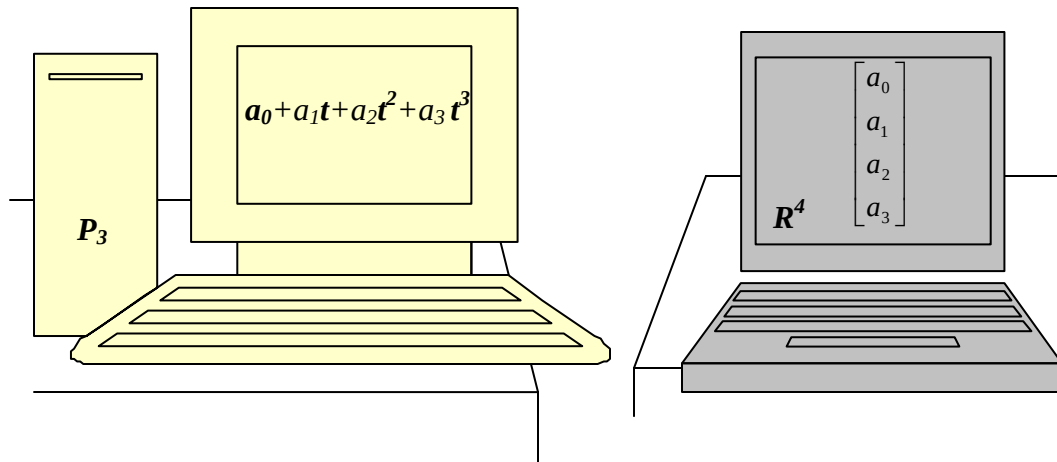


Figure 6 – The space P_3 is isomorphic to R^4

Example 9 Use coordinate vector to verify that the polynomials $1 + 2t^2$, $4 + t + 5t^2$ and $3 + 2t$ are linearly dependent in P_2 .

Solution The coordinate mapping from Example 8 produces the coordinate vectors $(1, 0, 2)$, $(4, 1, 5)$ and $(3, 2, 0)$, respectively. Writing these vectors as the columns of a matrix A , we can determine their independence by row reducing the augmented matrix

$$\text{for } Ax = 0: \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The columns of A are linearly dependent, so the corresponding polynomials are linearly dependent. In fact, it is easy to check that column 3 of A is 2 times column 2 minus 5 times column 1. The corresponding relation for the polynomials is

$$3 + 2t = 2(4 + t + 5t^2) - 5(1 + 2t^2)$$

Example 10 Let $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, and $B = \{\mathbf{v}_1, \mathbf{v}_2\}$. Then B is a

basis for $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Determine if $\mathbf{x}$ is in H and if it is, find the coordinate vector of $\mathbf{x}$ relative to B .

Solution If $\mathbf{x}$ is in H , then the following vector equation is consistent.

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

The scalars, c_1 and c_2 , if they exist, are the B – coordinates of $\mathbf{x}$.

Using row operations, we obtain $\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.

Thus $c_1 = 2$, $c_2 = 3$ and $[\mathbf{x}]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The coordinate system on H determined by B is shown in Figure 7.

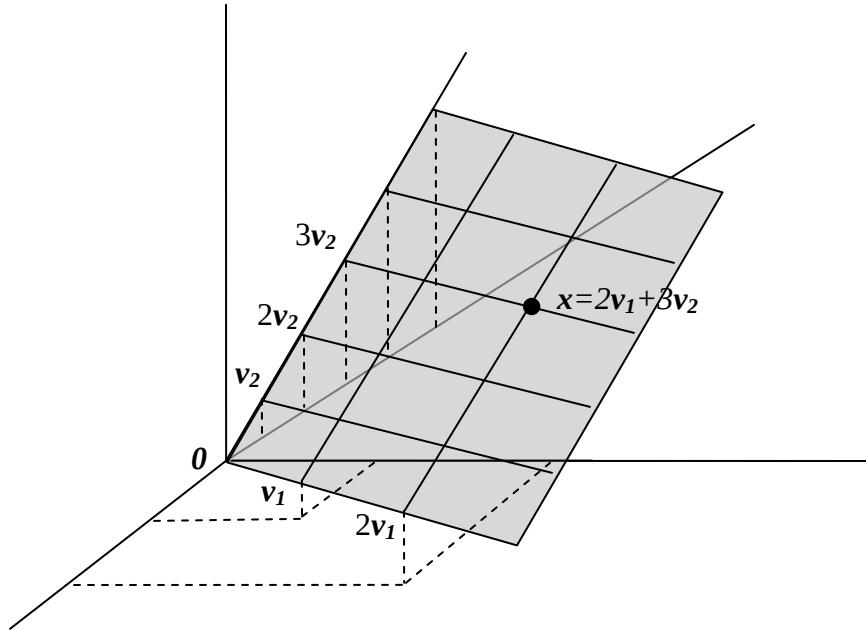


Figure 7 – A coordinate system on a plane H in $\mathbf{R}^3$

If a different basis for H were chosen, would the associated coordinate system also make H isomorphic to $\mathbf{R}^2$? Surely, this must be true. We shall prove it in the next lecture.

Example 11 Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

- Show that the set $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis of $\mathbf{R}^3$.
- Find the change-of-coordinates matrix from B to the standard basis.
- Write the equation that relates $\mathbf{x}$ in $\mathbf{R}^3$ to $[\mathbf{x}]_B$.
- Find $[\mathbf{x}]_B$, for the $\mathbf{x}$ given above.

Solution

- It is evident that the matrix $P_B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3]$ is row equivalent to the identity matrix. By the Invertible Matrix Theorem, P_B is invertible and its columns form a basis for $\mathbf{R}^3$.

- b. From part (a), the change-of-coordinates matrix is $\mathbf{P}_B = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.
- c. $\mathbf{x} = \mathbf{P}_B[\mathbf{x}]_B$.
- d. To solve part (c), it is probably easier to row reduce an augmented matrix instead of computing $\mathbf{P}_B^{-1}$. We have

$$\begin{array}{ccc} \begin{bmatrix} 1 & -3 & 3 & -8 \\ 0 & 4 & -6 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ \mathbf{P}_B & \mathbf{x} & \mathbf{I} \quad [\mathbf{x}]_B \end{array}$$

$$\text{Hence } [\mathbf{x}]_B = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

Example 12 The set $\mathbf{B} = \{1 + t, 1 + t^2, t + t^2\}$ is a basis for $\mathbf{P}_2$. Find the coordinate vector of $\mathbf{p}(t) = 6 + 3t - t^2$ relative to $\mathbf{B}$.

Solution The coordinates of $\mathbf{p}(t) = 6 + 3t - t^2$ with respect to $\mathbf{B}$ satisfy

$$c_1(1 + t) + c_2(1 + t^2) + c_3(t + t^2) = 6 + 3t - t^2$$

$$c_1 + c_1t + c_2 + c_2t^2 + c_3t + c_3t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + c_1t + c_3t + c_2t^2 + c_3t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + (c_1 + c_3)t + (c_2 + c_3)t^2 = 6 + 3t - t^2$$

Equating coefficients of like powers of t , we have

$$c_1 + c_2 = 6 \text{ -----(1)}$$

$$c_1 + c_3 = 3 \text{ -----(2)}$$

$$c_2 + c_3 = -1 \text{ -----(3)}$$

Subtract equation (2) from (1) we get

$$c_2 - c_3 = 6 - 3 = 3$$

Add this equation with equation (3)

$$2c_2 = -1 + 3 = 2$$

$$\Rightarrow c_2 = 1$$

Put value of c_2 in equation (3)

$$1 + c_3 = -1$$

$$\Rightarrow c_3 = -2$$

From equation (1) we have

$$c_1 + c_2 = 6$$

$$c_1 = 6 - 1 = 5$$

Solving, we find that $c_1 = 5$, $c_2 = 1$, $c_3 = -2$, and $[\mathbf{p}]_B = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

Exercises

In exercises 1 and 2, find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_B$ and the given basis $\mathbf{B}$.

$$1. \mathbf{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}, [\mathbf{x}]_B = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \qquad 2. \mathbf{B} = \left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -7 \\ 0 \end{bmatrix} \right\}, [\mathbf{x}]_B = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$$

In exercises 3-6, find the coordinate vector $[\mathbf{x}]_B$ of $\mathbf{x}$ relative to the given basis $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$.

$$3. \mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \qquad 4. \mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 5 \\ -6 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$5. \mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$$

$$6. \mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}$$

In exercises 7 and 8, find the change of coordinates matrix from $\mathbf{B}$ to standard basis in $\mathbf{R}^n$.

$$7. \mathbf{B} = \left\{ \begin{bmatrix} 2 \\ -9 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \end{bmatrix} \right\} \qquad 8. \mathbf{B} = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 8 \\ -2 \\ 7 \end{bmatrix} \right\}$$

In exercises 9 and 10, use an inverse matrix to find $[\mathbf{x}]_{\mathbf{B}}$ for the given $\mathbf{x}$ and $\mathbf{B}$.

$$9. \mathbf{B} = \left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$10. \mathbf{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}$$

11. The set $\mathbf{B} = \{\mathbf{1} + \mathbf{t}^2, \mathbf{t} + \mathbf{t}^2, \mathbf{1} + 2\mathbf{t} + \mathbf{t}^2\}$ is a basis for $\mathbf{P}_2$. Find the coordinate vector of $\mathbf{p}(\mathbf{t}) = \mathbf{1} + 4\mathbf{t} + 7\mathbf{t}^2$ relative to $\mathbf{B}$.

12. The vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ span $\mathbf{R}^2$ but do not form a basis. Find

two different ways to express $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

13. Let $\mathbf{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Since the coordinate mapping determined by $\mathbf{B}$ is a linear transformation from $\mathbf{R}^2$ into $\mathbf{R}^2$, this mapping must be implemented by some 2×2 matrix $\mathbf{A}$. Find it.

In exercises 14-16, use coordinate vectors to test the linear independence of the sets of polynomials.

$$14. \mathbf{1} + \mathbf{t}^3, \mathbf{3} + \mathbf{t} - 2\mathbf{t}^2, -\mathbf{t} + 3\mathbf{t}^2 - \mathbf{t}^3$$

$$15. (\mathbf{t}-\mathbf{1})^2, \mathbf{t}^3 - 2, (\mathbf{t}-2)^3$$

$$16. \mathbf{3} + 7\mathbf{t}, \mathbf{5} + \mathbf{t} - 2\mathbf{t}^3, \mathbf{t} - 2\mathbf{t}^2, \mathbf{1} + 16\mathbf{t} - 6\mathbf{t}^2 + 2\mathbf{t}^3$$

17. Let $\mathbf{H} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathbf{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Show that $\mathbf{x}$ is in $\mathbf{H}$ and find the $\mathbf{B}$ -

$$\text{coordinate vector of } \mathbf{x}, \text{ for } \mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}.$$

18. Let $\mathbf{H} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathbf{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Show that $\mathbf{B}$ is a basis for $\mathbf{H}$ and $\mathbf{x}$ is

$$\text{in } \mathbf{H}, \text{ and find the } \mathbf{B}\text{-coordinate vector of } \mathbf{x}, \text{ for } \mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}.$$

Lecture 24

Dimension of a Vector Space

In this lecture, we will focus over the dimension of the vector spaces. The dimension of a vector space V is the cardinality or the number of vectors in the basis B of the given vector space. If the basis B has n (say) elements then this number n (called the dimension) is an intrinsic property of the space V . That is it does not depend on the particular choice of basis rather, all the bases of V will have the same cardinality. Thus, we can say that the dimension of a vector space is always unique. The discussion of dimension will give additional insight into properties of bases.

The first theorem generalizes a well-known result about the vector space $\mathbf{R}^n$.

Note

A vector space V with a basis B containing n vectors is isomorphic to $\mathbf{R}^n$ i.e., there exist a one-to-one linear transformation from V to $\mathbf{R}^n$.

Theorem 1 If a vector space V has a basis $B = \{b_1, \dots, b_n\}$, then any set in V containing more than n vectors must be linearly dependent.

Theorem 2 If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

Finite and infinite dimensional vector spaces

If the vector space V is spanned or generated by a finite set, then V is said to be **finite-dimensional**, and the **dimension** of V , written as $\dim V$, is the number of vectors in a basis for V . If V is not spanned by a finite set, then V is said to be **infinite-dimensional**. That is, if we are unable to find a finite set that can generate the whole vector space, then such a vector space is called infinite dimensional.

Note

- (1) The dimension of the zero vector space $\{0\}$ is defined to be zero.
- (2) Every finite dimensional vector space contains a basis.

Example 1 The n dimensional set of real numbers $\mathbf{R}^n$, set of polynomials of order n $\mathbf{P}_n$, and set of matrices of order $m \times n$ $\mathbf{M}_{mn}$ are all finite- dimensional vector spaces. However, the vector spaces $\mathbf{F} (-\infty, \infty)$, $\mathbf{C} (-\infty, \infty)$, and $\mathbf{C}^m (-\infty, \infty)$ are infinite-dimensional.

Example 2

(a) Any pair of non-parallel vectors $\mathbf{a}, \mathbf{b}$ in the xy -plane, which are necessarily linearly independent, can be regarded as a basis of the subspace $\mathbf{R}^2$. In particular the set of unit vectors $\{\mathbf{i}, \mathbf{j}\}$ forms a basis for $\mathbf{R}^2$. Therefore, $\dim(\mathbf{R}^2) = 2$.

Any set of three non coplanar vectors $\{a, b, c\}$ in ordinary (physical) space, which will be necessarily linearly independent, spans the space $\mathbf{R}^3$. Therefore any set of such vectors forms a basis for $\mathbf{R}^3$. In particular the set of unit vectors $\{i, j, k\}$ forms a basis of $\mathbf{R}^3$. This basis is called standard basis for $\mathbf{R}^3$. Therefore $\dim(\mathbf{R}^3) = 3$.

The set of vectors $\{e_1, e_2, \dots, e_n\}$ where

$$e_1 = (1, 0, 0, 0, \dots, 0),$$

$$e_2 = (0, 1, 0, 0, \dots, 0),$$

$$e_3 = (0, 0, 1, 0, \dots, 0),$$

...

...

...

$$e_n = (0, 0, 0, 0, \dots, 1)$$

is linearly independent.

Moreover, any vector $x = (x_1, x_2, \dots, x_n)$ in $\mathbf{R}^n$ can be expressed as a linear combination of these vectors as

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3 + \dots + x_n e_n.$$

Hence, the set $\{e_1, e_2, \dots, e_n\}$ forms a basis for $\mathbf{R}^n$. It is called the standard basis of $\mathbf{R}^n$, therefore $\dim(\mathbf{R}^n) = n$. Any other set of n linearly independent vectors in $\mathbf{R}^n$ will form a non-standard basis.

MCQs

(b) The set $B = \{1, x, x^2, \dots, x^n\}$ forms a basis for the vector space P_n of polynomials of degree $\leq n$. It is called the standard basis with $\dim(P_n) = n + 1$.

(c) The set of 2×2 matrices with real entries (elements) $\{u_1, u_2, u_3, u_4\}$ where

$$u_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is a linearly independent and every 2×2 matrix with real entries can be expressed as their linear combination. Therefore, they form a basis for the vector space $M_{2 \times 2}$. This basis is called the standard basis for $M_{2 \times 2}$ with $\dim(M_{2 \times 2}) = 4$.

MCQs

Note

(1) $\dim(\mathbf{R}^n) = n$ { The standard basis has n vectors }.

(2) $\dim(P_n) = n + 1$ { The standard basis has $n+1$ vectors }.

(3) $\dim(M_{m \times n}) = mn$ { The standard basis has mn vectors }.

Example 3 Let W be the subspace of the set of all (2×2) matrices defined by

$$W = \left\{ A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : 2a - b + 3c + d = 0 \right\}.$$

Determine the dimension of W .

Solution The algebraic specification for W can be rewritten as $d = -2a + b - 3c$.

Now $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Substituting the value of d, it becomes

$$A = \begin{bmatrix} a & b \\ c & -2a + b - 3c \end{bmatrix}$$

This can be written as

$$\begin{aligned} A &= \begin{bmatrix} a & 0 \\ 0 & -2a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & -3c \end{bmatrix} \\ &= a \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix} \\ &= aA_1 + bA_2 + cA_3 \end{aligned}$$

where $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, and $A_3 = \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix}$

The matrix A is in W if and only if $A = aA_1 + bA_2 + cA_3$, so $\{A_1, A_2, A_3\}$ is a spanning set for W . Now, check if this set is a basis for W or not. We will see whether $\{A_1, A_2, A_3\}$ is linearly independent or not. $\{A_1, A_2, A_3\}$ is said to be linearly independent if

$$\begin{aligned} aA_1 + bA_2 + cA_3 &= 0 \Rightarrow a=b=c=0 \text{ i.e.,} \\ a \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} a & 0 \\ 0 & -2a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & -3c \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} a & b \\ c & -2a + b - 3c \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

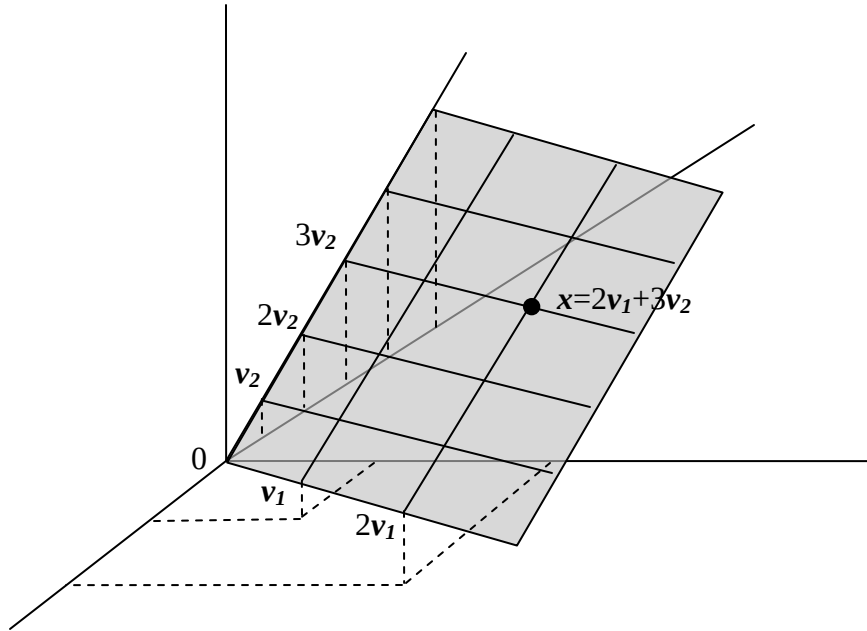
Equating the elements, we get

$$a = 0, b = 0, c = 0$$

This implies $\{A_1, A_2, A_3\}$ is a linearly independent set that spans W . Hence, it's the basis of W with $\dim(W) = 3$.

Example 4 Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Then H is the plane

studied in Example 10 of lecture 23. A basis for H is $\{\mathbf{v}_1, \mathbf{v}_2\}$, since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not multiples and hence are linearly independent. Thus, $\dim H = 2$.



A coordinate system on a plane H in R^3

Example 5 Find the dimension of the subspace

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \in R \right\}$$

Solution The representative vector of H can be written as

$$\begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} = a \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Now, it is easy to see that H is the set of all linear combinations of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Clearly, $\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$, but $\mathbf{v}_3$ is a multiple of $\mathbf{v}_2$. By the Spanning Set Theorem, we may discard $\mathbf{v}_3$ and still have a set that spans H . Finally; $\mathbf{v}_4$ is not a linear

combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is linearly independent and hence is a basis for H . Thus $\dim H = 3$.

Example 6 The subspaces of $\mathbf{R}^3$ can be classified by various dimensions as shown in Fig. 1.

0-dimensional subspaces

The only 0-dimensional subspace of $\mathbf{R}^3$ is zero space.

1-dimensional subspaces

1-dimensional subspaces include any subspace spanned by a single non-zero vector. Such subspaces are lines through the origin.

2-dimensional subspaces

Any subspace spanned by two linearly independent vectors. Such subspaces are planes through the origin.

3-dimensional subspaces

The only 3-dimensional subspace is $\mathbf{R}^3$ itself. Any three linearly independent vectors in $\mathbf{R}^3$ span all of $\mathbf{R}^3$, by the Invertible Matrix Theorem.

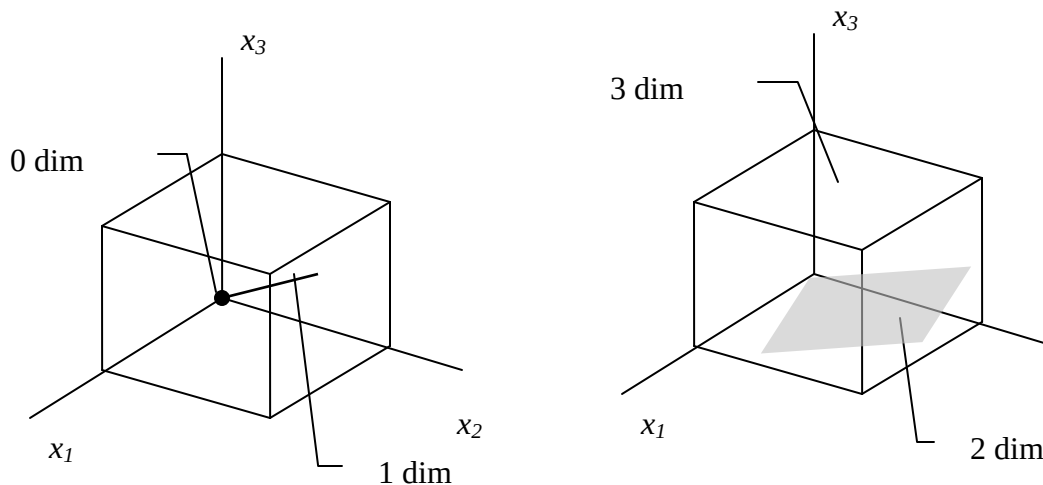


Figure 1 – Sample subspaces of $\mathbf{R}^3$

Bases for Nul A and Col A

We already know how to find vectors that span the null space of a matrix A . The discussion in Lecture 21 pointed out that our method always produces a linearly independent set. Thus the method produces a basis for Nul A .

X SKIP

Example 7 Find a basis for the null space of $A = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

Solution The null space of A is the solution space of homogeneous system

$$2x_1 + 2x_2 - x_3 + x_5 = 0$$

$$-x_1 - x_2 + 2x_3 - 3x_4 + x_5 = 0$$

$$x_1 + x_2 - 2x_3 - x_5 = 0$$

$$x_3 + x_4 + x_5 = 0$$

The most appropriate way to solve this system is to reduce its augmented matrix into reduced echelon form.

$$\left[\begin{array}{ccccc|c} 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array} \right] \quad R_4 \sim R_2, R_3 \sim R_1$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{array} \right] \quad R_3 - 2R_1, R_3 - 3R_2$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{array} \right] \quad R_3 - 3R_2$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{array} \right] \quad -\frac{1}{3}R_3$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{array} \right] \quad R_4 + R_1$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \end{bmatrix} \quad R_4 + 3R_3$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad R_1 + 2R_2$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad R_2 - R_3, R_1 - 2R_3$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the reduced row echelon form of the augmented matrix is

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

which corresponds to the system

$$\begin{aligned} 1x_1 + 1x_2 + 1x_5 &= 0 \\ 1x_3 + 1x_5 &= 0 \\ 1x_4 &= 0 \\ 0 &= 0 \end{aligned}$$

No equation of this system has a form zero = nonzero. Therefore, the system is consistent. Since the number of unknowns is more than the number of equations, we will assign some arbitrary value to some variables. This will lead to infinite many solutions of the system.

$$x_1 = -1x_2 - 1x_5$$

$$x_2 = s$$

$$x_3 = -1x_5$$

$$x_4 = 0$$

$$x_5 = t$$

The general solution of the given system is

$$x_1 = -s - t, \quad x_2 = s, \quad x_3 = -t, \quad x_4 = 0, \quad x_5 = t$$

Therefore, the solution vector can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -s-t \\ s \\ -t \\ 0 \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ -t \\ 0 \\ t \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

which shows that the vectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$ span the solution space. Since they

are also linearly independent, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\text{Nul } \mathbf{A}$.

The next two examples describe a simple algorithm for finding a basis for the column space.

Example 8 Find a basis for $\text{Col } \mathbf{B}$, where $\mathbf{B} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Solution Each non-pivot column of $\mathbf{B}$ is a linear combination of the pivot columns. In fact, $\mathbf{b}_2 = 4\mathbf{b}_1$ and $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$. By the Spanning Set Theorem, we may discard $\mathbf{b}_2$ and $\mathbf{b}_4$ and $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ will still span $\text{Col } \mathbf{B}$. Let

$$\mathbf{S} = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Since $\mathbf{b}_1 \neq \mathbf{0}$ and no vector in $\mathbf{S}$ is a linear combination of the vectors that precede it, $\mathbf{S}$ is linearly independent. Thus $\mathbf{S}$ is a basis for $\text{Col } \mathbf{B}$.

What about a matrix $\mathbf{A}$ that is not in reduced echelon form? Recall that any linear dependence relationship among the columns of $\mathbf{A}$ can be expressed in the form $\mathbf{A}\mathbf{x}$

$= \mathbf{0}$, where $\mathbf{x}$ is a column of weights. (If some columns are not involved in a particular dependence relation, then their weights are zero.) When $\mathbf{A}$ is row reduced to a matrix $\mathbf{B}$, the columns of $\mathbf{B}$ are often totally different from the columns of $\mathbf{A}$. However, the equations $\mathbf{Ax} = \mathbf{0}$ and $\mathbf{Bx} = \mathbf{0}$ have exactly the same set of solutions. That is, the columns of $\mathbf{A}$ have exactly the same linear dependence relationships as the columns of $\mathbf{B}$.

Elementary row operations on a matrix do not affect the linear dependence relations among the columns of the matrix.

Example 9 It can be shown that the matrix

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

is row equivalent to the matrix $\mathbf{B}$ in Example 8. Find a basis for Col $\mathbf{A}$.

Solution In Example 8, we have seen that $\mathbf{b}_2 = 4\mathbf{b}_1$ and $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$ so we can expect that $\mathbf{a}_2 = 4\mathbf{a}_1$ and $\mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3$. This is indeed the case. Thus, we may discard $\mathbf{a}_2$ and $\mathbf{a}_4$ while selecting a minimal spanning set for Col $\mathbf{A}$. In fact, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ must be linearly independent because any linear dependence relationship among $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ would imply a linear dependence relationship among $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$. But we know that $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ is a linearly independent set. Thus $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ is a basis for Col $\mathbf{A}$. The columns we have used for this basis are the pivot columns of $\mathbf{A}$.

Examples 8 and 9 illustrate the following useful fact.

Theorem 3 The pivot columns of a matrix $\mathbf{A}$ form a basis for Col $\mathbf{A}$.

Proof The general proof uses the arguments discussed above. Let $\mathbf{B}$ be the reduced echelon form of $\mathbf{A}$. The set of pivot columns of $\mathbf{B}$ is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since $\mathbf{A}$ is row equivalent to $\mathbf{B}$, the pivot columns of $\mathbf{A}$ are linearly independent too, because any linear dependence relation among the columns of $\mathbf{A}$ corresponds to a linear dependence relation among the columns of $\mathbf{B}$. For this same reason, every non-pivot column of $\mathbf{A}$ is a linear combination of the pivot columns of $\mathbf{A}$. Thus the non-pivot columns of $\mathbf{A}$ may be discarded from the spanning set for Col $\mathbf{A}$, by the Spanning Set Theorem. This leaves the pivot columns of $\mathbf{A}$ as a basis for Col $\mathbf{A}$.

Note Be careful to use pivot columns of $\mathbf{A}$ itself for the basis of Col $\mathbf{A}$. The columns of an echelon form $\mathbf{B}$ are often not in the column space of $\mathbf{A}$. For instance, the columns of the $\mathbf{B}$ in Example 8 all have zeros in their last entries, so they cannot span the column space of the $\mathbf{A}$ in Example 9.

Example 10 Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 7 \\ -9 \end{bmatrix}$. Determine if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbf{R}^3$. Is $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for $\mathbf{R}^2$?

Solution Let $A = [\mathbf{v}_1 \ \mathbf{v}_2]$. Row operations show that $A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \\ 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Not every

row of A contains a pivot position. So the columns of A do not span $\mathbf{R}^3$, by Theorem 4 in Lecture 6. Hence $\{\mathbf{v}_1, \mathbf{v}_2\}$ is not a basis for $\mathbf{R}^3$. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not in $\mathbf{R}^2$, they cannot possibly be a basis for $\mathbf{R}^2$. However, since $\mathbf{v}_1$ and $\mathbf{v}_2$ are obviously linearly independent, they are a basis for a subspace of $\mathbf{R}^3$, namely, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Example 11 Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} -4 \\ -8 \\ 9 \end{bmatrix}$. Find a basis for the subspace

W spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

Solution Let A be the matrix whose column space is the space spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$,

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix}$$

Reduce the matrix A into its echelon form in order to find its pivot columns.

$$\begin{aligned} A &= \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} \text{ by } R_2 + 3R_1, R_3 - 4R_1 \\ &\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } \frac{1}{4}R_2, -\frac{1}{5}R_3, R_3 - R_2 \end{aligned}$$

The first two columns of A are the pivot columns and hence form a basis of $\text{Col } A = W$. Hence $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W .

Note that the reduced echelon form of A is not needed in order to locate the pivot columns.

Procedure**Basis and Linear Combinations**

Given a set of vectors $S = \{v_1, v_2, \dots, v_k\}$ in $\mathbf{R}^n$, the following procedure produces a subset of these vectors that form a basis for $\text{span}(S)$ and expresses those vectors of S that are not in the basis as linear combinations of the basis vector.

Step1: Form the matrix A having $v_1, v_2, \dots, v_k$ as its column vectors.

Step2: Reduce the matrix A to its reduced row echelon form R , and let

$w_1, w_2, \dots, w_k$ be the column vectors of R .

Step3: Identify the columns that contain the leading entries i.e., 1's in R . The corresponding column vectors of A are the basis vectors for $\text{span}(S)$.

Step4: Express each column vector of R that does not contain a leading entry as a linear combination of preceding column vector that do contain leading entries (we will be able to do this by inspection). This yields a set of dependency equations involving the column vectors of R . The corresponding equations for the column vectors of A express the vectors which are not in the basis as linear combinations of basis vectors.

Example 12 Basis and Linear Combinations

(a) Find a subset of the vectors $v_1 = (1, -2, 0, 3)$, $v_2 = (2, -4, 0, 6)$, $v_3 = (-1, 1, 2, 0)$ and $v_4 = (0, -1, 2, 3)$ that form a basis for the space spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution (a) We begin by constructing a matrix that has v_1, v_2, v_3, v_4 as its column vectors

$$\begin{array}{cccc} \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix} & & & \\ \uparrow & \uparrow & \uparrow & \uparrow & \\ v_1 & v_2 & v_3 & v_4 & \end{array} \quad (A)$$

Finding a basis for column space of this matrix can solve the first part of our problem.

Transforming Matrix to Reduced Row Echelon Form:

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix} \begin{array}{l} \\ 2R_1 + R_2 \\ -3R_1 + R_4 \\ \end{array}$$

$$\begin{aligned}
 & \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} \begin{matrix} \\ -1R_2 \\ \\ \end{matrix} \\
 & \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ -2R_2 + R_3 \\ -3R_2 + R_4 \\ \end{matrix} \\
 & \sim \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ R_2 + R_1 \\ \\ \end{matrix}
 \end{aligned}$$

Labeling the column vectors of the resulting matrix as $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$ and $\mathbf{w}_4$ yields

$$\begin{array}{cccc}
 \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & & & \\
 \uparrow & \uparrow & \uparrow & \uparrow & \\
 \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 &
 \end{array} \quad (B)$$

The leading entries occur in column 1 and 3 so $\{\mathbf{w}_1, \mathbf{w}_3\}$ is a basis for the column space of (B) and consequently $\{\mathbf{v}_1, \mathbf{v}_3\}$ is the basis for column space of (A).

(b) We shall start by expressing $\mathbf{w}_2$ and $\mathbf{w}_4$ as linear combinations of the basis vector $\mathbf{w}_1$ and $\mathbf{w}_3$. The simplest way of doing this is to express $\mathbf{w}_2$ and $\mathbf{w}_4$ in term of basis vectors with smaller subscripts. Thus we shall express $\mathbf{w}_2$ as a linear combination of $\mathbf{w}_1$, and we shall express $\mathbf{w}_4$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_3$. By inspection of (B), these linear combinations are $\mathbf{w}_2 = 2\mathbf{w}_1$ and $\mathbf{w}_4 = \mathbf{w}_1 + \mathbf{w}_3$. We call them the dependency equations.

The corresponding relationship of (A) are $\mathbf{v}_3 = 2\mathbf{v}_1$ and $\mathbf{v}_5 = \mathbf{v}_1 + \mathbf{v}_3$.

Example 13 Basis and Linear Combinations

(a) Find a subset of the vectors $\mathbf{v}_1 = (1, -1, 5, 2)$, $\mathbf{v}_2 = (-2, 3, 1, 0)$, $\mathbf{v}_3 = (4, -5, 9, 4)$, $\mathbf{v}_4 = (0, 4, 2, -3)$ and $\mathbf{v}_5 = (-7, 18, 2, -8)$ that form a basis for the space spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors

Solution (a) We begin by constructing a matrix that has $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_5$ as its column vectors

$$\begin{array}{ccccc}
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ -1 & 3 & -5 & 4 & 18 \\ 5 & 1 & 9 & 2 & 2 \\ 2 & 0 & 4 & -3 & -8 \end{bmatrix} & & (A) \\
 \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\
 \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5
 \end{array}$$

Finding a basis for column space of this matrix can solve the first part of our problem.

Transforming Matrix to Reduced Row Echelon Form:

$$\begin{array}{ccccc}
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ -1 & 3 & -5 & 4 & 18 \\ 5 & 1 & 9 & 2 & 2 \\ 2 & 0 & 4 & -3 & -8 \end{bmatrix} & & \\
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 11 & -11 & 2 & 37 \\ 0 & 4 & -4 & -3 & 6 \end{bmatrix} & \begin{array}{l} R_1 + R_2 \\ -5R_1 + R_3 \\ -2R_1 + R_4 \end{array} & \\
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & -42 & -84 \\ 0 & 0 & 0 & -19 & -38 \end{bmatrix} & \begin{array}{l} -11R_2 + R_3 \\ -4R_2 + R_4 \end{array} & \\
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -19 & -38 \end{bmatrix} & (-1/42)R_3 & \\
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} & 19R_3 + R_4 & \\
 \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} & (-4)R_3 + R_2 & \\
 \begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} & 2R_2 + R_1 &
 \end{array}$$

Denoting the column vectors of the resulting matrix by $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, $\mathbf{w}_4$, and $\mathbf{w}_5$ yields

$$\begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{B})$$

$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow$
 $\mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{w}_3 \quad \mathbf{w}_4 \quad \mathbf{w}_5$

The leading entries occur in columns 1, 2 and 4 so that $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4\}$ is a basis for the column space of (B) and consequently $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is the basis for column space of (A).

(b) We shall start by expressing $\mathbf{w}_3$ and $\mathbf{w}_5$ as linear combinations of the basis vector $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_4$. The simplest way of doing this is to express $\mathbf{w}_3$ and $\mathbf{w}_5$ in term of basis vectors with smaller subscripts. Thus we shall express $\mathbf{w}_3$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$, and we shall express $\mathbf{w}_5$ as a linear combination of $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_4$. By inspection of (B), these linear combination are $\mathbf{w}_3 = 2\mathbf{w}_1 - \mathbf{w}_2$ and $\mathbf{w}_5 = -\mathbf{w}_1 + 3\mathbf{w}_2 + 2\mathbf{w}_4$.

The corresponding relationship of (A) are $\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2$ and $\mathbf{v}_5 = -\mathbf{v}_1 + 3\mathbf{v}_2 + 2\mathbf{v}_4$.

Example 14 Basis and Linear Combinations

(a) Find a subset of the vectors $\mathbf{v}_1 = (1, -2, 0, 3)$, $\mathbf{v}_2 = (2, -5, -3, 6)$, $\mathbf{v}_3 = (0, 1, 3, 0)$, $\mathbf{v}_4 = (2, -1, 4, -7)$ and $\mathbf{v}_5 = (5, -8, 1, 2)$ that form a basis for the space spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution (a) We begin by constructing a matrix that has $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_5$ as its column vectors

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix}$$

$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow$
 $\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{v}_4 \quad \mathbf{v}_5$

(A)

Finding a basis for column space of this matrix can solve the first part of our problem. Reducing the matrix to reduced-row echelon form and denoting the column vectors of the resulting matrix by $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, $\mathbf{w}_4$, and $\mathbf{w}_5$ yields

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{B})$$

$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow$
 $\mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{w}_3 \quad \mathbf{w}_4 \quad \mathbf{w}_5$

The leading entries occur in columns 1, 2 and 4 so $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4\}$ is a basis for the column space of (B) and consequently $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is the basis for column space of (A).

(b) Dependency equations are $w_3 = 2w_1 - w_2$ and $w_5 = w_1 + w_2 + w_4$

The corresponding relationship of (A) are $v_3 = 2v_1 - v_2$ and $v_5 = v_1 + v_2 + v_4$

Subspaces of a Finite-Dimensional Space The next theorem is a natural counterpart to the Spanning Set Theorem.

Theorem 5 Let H be a subspace of a finite-dimensional vector space V . Any linearly independent set in H can be expanded, if necessary, to a basis for H . Also, H is finite-dimensional and $\dim H \leq \dim V$.

When the dimension of a vector space or subspace is known, the search for a basis is simplified by the next theorem. It says that if a set has the right number of elements, then one has only to show either that the set is linearly independent or that it spans the space. The theorem is of critical importance in numerous applied problems (involving differential equations or difference equations, for example) where linear independence is much easier to verify than spanning.

Theorem 5 (The Basis Theorem) Let V be a p -dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in V is automatically a basis for V . Any set of exactly p elements that spans V is automatically a basis for V .

The Dimensions of Nul A and Col A Since the pivot columns of a matrix A form a basis for Col A , we know the dimension of Col A as soon as we know the pivot columns. The dimension of Nul A might seem to require more work, since finding a basis for Nul A usually takes more time than a basis for Col A . Yet, there is a shortcut.

Let A be an $m \times n$ matrix, and suppose that the equation $Ax = 0$ has k free variables. From lecture 21, we know that the standard method of finding a spanning set for Nul A will produce exactly k linearly independent vectors say, $u_1, \dots, u_k$, one for each free variable. So $\{u_1, \dots, u_k\}$ is a basis for Nul A , and the number of free variables determines the size of the basis. Let us summarize these facts for future reference.

The dimension of Nul A is the number of free variables in the equation $Ax = 0$, and the dimension of Col A is the number of pivot columns in A .

Example 15 Find the dimensions of the null space and column space of

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solution Row reduce the augmented matrix $[A \ 0]$ to echelon form and obtain

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

Writing it in equations form, we get

$$x_1 - 2x_2 + 2x_3 + 3x_4 - x_5 = 0$$

$$x_3 + 2x_4 - 2x_5 = 0$$

Since the number of unknowns is more than the number of equations, we will introduce free variables here (say) x_2 , x_4 and x_5 . Hence the dimension of $\text{Nul } A$ is 3. Also $\dim \text{Col } A$ is 2 because A has two pivot columns.

Example 16 Decide whether each statement is true or false, and give a reason for each answer. Here V is a non-zero finite-dimensional vector space.

1. If $\dim V = p$ and if S is a linearly dependent subset of V , then S contains more than p vectors.
2. If S spans V and if T is a subset of V that contains more vectors than S , then T is linearly dependent.

Solution

1. False. Consider the set $\{0\}$.
2. True. By the Spanning Set Theorem, S contains a basis for V ; call that basis S' . Then T will contain more vectors than S' . By Theorem 1, T is linearly dependent.

Exercises

For each subspace in exercises 1-6, (a) find a basis and (b) state the dimension.

$$1. \left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ in } \mathbf{R} \right\}$$

$$2. \left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ in } \mathbf{R} \right\}$$

$$3. \left\{ \begin{bmatrix} a-4b-2c \\ 2a+5b-4c \\ -a+2c \\ -3a+7b+6c \end{bmatrix} : a, b, c \text{ in } \mathbf{R} \right\}$$

$$4. \left\{ \begin{bmatrix} 3a+6b-c \\ 6a-2b-2c \\ -9a+5b+3c \\ -3a+b+c \end{bmatrix} : a, b, c \text{ in } \mathbf{R} \right\}$$

$$5. \{(a, b, c) : a - 3b + c = 0, b - 2c = 0, 2b - c = 0\}$$

$$6. \{(a, b, c, d) : a - 3b + c = 0\}$$

7. Find the dimension of the subspace $\mathbf{H}$ of $\mathbf{R}^2$ spanned by

$$\begin{bmatrix} 2 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

8. Find the dimension of the subspace spanned by the given vectors.

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} -7 \\ -3 \\ 1 \end{bmatrix}$$

Determine the dimensions of $\text{Nul } \mathbf{A}$ and $\text{Col } \mathbf{A}$ for the matrices shown in exercises 9 to 12.

$$9. \mathbf{A} = \begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$10. \mathbf{A} = \begin{bmatrix} 1 & 3 & -4 & 2 & -1 & 6 \\ 0 & 0 & 1 & -3 & 7 & 0 \\ 0 & 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$11. \mathbf{A} = \begin{bmatrix} 1 & 0 & 9 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}$$

$$12. \mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$$

13. The first four Hermite polynomials are 1 , $2t$, $-2 + 4t^2$, and $-12t + 8t^3$. These polynomials arise naturally in the study of certain important differential equations in mathematical physics. Show that the first four Hermite polynomials form a basis of $\mathbf{P}_3$.

14. Let $\mathbf{B}$ be the basis of $\mathbf{P}_3$ consisting of the Hermite polynomials in exercise 13, and let $\mathbf{p}(t) = 7 - 12t - 8t^2 + 12t^3$. Find the coordinate vector of $\mathbf{p}$ relative to $\mathbf{B}$.

15. Extend the following vectors to a basis for $\mathbf{R}^5$:

$$\mathbf{v}_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

Lecture 25

Rank

With the help of vector space concepts, for a matrix several interesting and useful relationships in matrix rows and columns have been discussed.

For instance, imagine placing 2000 random numbers into a 40×50 matrix A and then determining both the maximum number of linearly independent columns in A and the maximum number of linearly independent columns in A^T (rows in A). Remarkably, the two numbers are the same. Their common value is called the rank of the matrix. To explain why, we need to examine the subspace spanned by the subspace spanned by the rows of A .

The Row Space If A is an $m \times n$ matrix, each row of A has n entries and thus can be identified with a vector in $\mathbf{R}^n$. The set of all linear combinations of the row vectors is called the row space of A and is denoted by $\text{Row } A$. Each row has n entries, so $\text{Row } A$ is a subspace of $\mathbf{R}^n$. Since the rows of A are identified with the columns of A^T , we could also write $\text{Col } A^T$ in place of $\text{Row } A$.

Example 1 Let $A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$ and $\begin{aligned} \mathbf{r}_1 &= (-2, -5, 8, 0, -17) \\ \mathbf{r}_2 &= (1, 3, -5, 1, 5) \\ \mathbf{r}_3 &= (3, 11, -19, 7, 1) \\ \mathbf{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$

The row space of A is the subspace of $\mathbf{R}^5$ spanned by $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. That is, $\text{Row } A = \text{Span } \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Naturally, we write row vectors horizontally; however, they could also be written as column vectors

Example Let

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 4 \end{bmatrix} \quad \text{and} \quad \begin{aligned} \mathbf{r}_1 &= (2, 1, 0) \\ \mathbf{r}_2 &= (3, -1, 4) \end{aligned}$$

That is $\text{Row } A = \text{Span } \{\mathbf{r}_1, \mathbf{r}_2\}$.

We could use the Spanning Set Theorem to shrink the spanning set to a basis.

Some times row operation on a matrix will not give us the required information but row reducing certainly worthwhile, as the next theorem shows

Theorem 1 If two matrices A and B are row equivalent, then their row spaces are the same. If B is in echelon form, the nonzero rows of B form a basis for the row space of A as well as B .

Theorem 2 If A and B are row equivalent matrices, then

- (a) A given set of column vectors of $\mathbf{A}$ is linearly independent if and only if the corresponding column vectors of $\mathbf{B}$ are linearly independent.
- (b) A given set of column vector of $\mathbf{A}$ forms a basis for the column space of $\mathbf{A}$ if and only if the corresponding column vector of $\mathbf{B}$ forms a basis for the column space of $\mathbf{B}$.

Example 2 (Bases for Row and Column Spaces)

Find the bases for the row and column spaces of $\mathbf{A} = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$.

Solution We can find a basis for the row space of $\mathbf{A}$ by finding a basis for the row space of any row-echelon form of $\mathbf{A}$.

Now $\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$

$$\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} -2R_1 + R_2 \\ -2R_1 + R_3 \\ R_1 + R_4 \end{array}$$

$$\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} -1R_2 + R_3 \end{array}$$

Row-echelon form of $\mathbf{A}$: $\mathbf{R} = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Here Theorem 1 implies that the non zero rows are the basis vectors of the matrix. So these bases vectors are

$$\begin{aligned} \mathbf{r}_1 &= [1 \quad -3 \quad 4 \quad -2 \quad 5 \quad 4] \\ \mathbf{r}_2 &= [0 \quad 0 \quad 1 \quad 3 \quad -2 \quad -6] \\ \mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 5] \end{aligned}$$

$\mathbf{A}$ and $\mathbf{R}$ may have different column spaces, we cannot find a basis for the column space of $\mathbf{A}$ directly from the column vectors of $\mathbf{R}$. however, it follows from the theorem (2b) if

we can find a set of column vectors of $\mathbf{R}$ that forms a basis for the column space of $\mathbf{R}$, then the corresponding column vectors of $\mathbf{A}$ will form a basis for the column space of $\mathbf{A}$.

The first, third, and fifth columns of $\mathbf{R}$ contains the leading 1's of the row vectors, so

$$\mathbf{c}'_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{c}'_3 = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{c}'_5 = \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of $\mathbf{R}$, thus the corresponding column vectors of $\mathbf{A}$

namely,
$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix} \quad \mathbf{c}_3 = \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix} \quad \mathbf{c}_5 = \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix}$$

form a basis for the column space of $\mathbf{A}$.

Example

The matrix

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

is in row-echelon form.

The vectors

$$\mathbf{r}_1 = [1 \quad -2 \quad 5 \quad 0 \quad 3]$$

$$\mathbf{r}_2 = [0 \quad 1 \quad 3 \quad 0 \quad 0]$$

$$\mathbf{r}_3 = [0 \quad 0 \quad 0 \quad 1 \quad 0]$$

form a basis for the row space of $\mathbf{R}$, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of $\mathbf{R}$.

Example 3 (Basis for a Vector Space using Row Operation)

Find bases for the space spanned by the vectors

$$\mathbf{v}_1 = (1, -2, 0, 0, 3) \quad \mathbf{v}_2 = (2, -5, -3, -2, 6)$$

$$\mathbf{v}_3 = (0, 5, 15, 10, 0) \quad \mathbf{v}_4 = (2, 6, 18, 8, 6)$$

Solution The space spanned by these vectors is the row space of the matrix

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{aligned} & \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix} \\ & \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 5 & 15 & 10 & 0 \\ 0 & 10 & 18 & 8 & 0 \end{bmatrix} \begin{array}{l} (-2)R_1 + R_2 \\ (-2)R_1 + R_4 \\ (-1)R_2 \end{array} \\ & \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & -12 & 0 \end{bmatrix} \begin{array}{l} (-5)R_2 + R_3 \\ (-10)R_2 + R_4 \end{array} \\ & \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & -12 & -12 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_{34} \\ & \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3 \end{aligned}$$

Therefore,
$$\mathbf{R} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The non-zero row vectors in this matrix are

$$\mathbf{w}_1 = (1, -2, 0, 0, 3), \mathbf{w}_2 = (0, 1, 3, 2, 0), \mathbf{w}_3 = (0, 0, 1, 1, 0)$$

These vectors form a basis for the row space and consequently form a basis for the subspace of $\mathbf{R}^5$ spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Example 4 (Basis for the Row Space of a Matrix)

Find a basis for the row space of $\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$ consisting entirely of row

vectors from $\mathbf{A}$.

Solution We find $\mathbf{A}^T$; then we will use the method of example (2) to find a basis for the column space of $\mathbf{A}^T$; and then we will transpose again to convert column vectors back to row vectors. Transposing $\mathbf{A}$ yields

$$\mathbf{A}^T = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix} \begin{array}{l} \\ 2R_1 + R_2 \\ (-3)R_1 + R_5 \\ \\ \end{array}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 5 & 10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned}
 & \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (-1)R_2 \\
 & \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} (3)R_2 + R_3 \\ (2)R_2 + R_4 \end{array} \\
 & \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (-1/12)R_3 \\
 & \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad 12R_3 + R_4
 \end{aligned}$$

Now

$$\mathbf{R} = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first, second and fourth columns contain the leading 1's, so the corresponding column vectors in $\mathbf{A}^T$ form a basis for the column space of $\mathbf{A}^T$; these are

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{bmatrix} \text{ and } \mathbf{c}_4 = \begin{bmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{bmatrix}$$

Transposing again and adjusting the notation appropriately yields the basis vectors $\mathbf{r}_1 = [1 \ -2 \ 0 \ 0 \ 3]$, $\mathbf{r}_2 = [2 \ -5 \ -3 \ -2 \ 6]$ and $\mathbf{r}_4 = [2 \ 6 \ 18 \ 8 \ 6]$ for the row space of $\mathbf{A}$.

The following example shows how one sequence of row operations on $\mathbf{A}$ leads to bases for the three spaces: Row $\mathbf{A}$, Col $\mathbf{A}$, and Nul $\mathbf{A}$.

Example 5 Find bases for the row space, the column space and the null space of the matrix

$$\mathbf{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

Solution To find bases for the row space and the column space, row reduce $\mathbf{A}$ to an

$$\text{echelon form: } \mathbf{A} \sim \mathbf{B} = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By Theorem (1), the first three rows of $\mathbf{B}$ form a basis for the row space of $\mathbf{A}$ (as well as the row space of $\mathbf{B}$). Thus Basis for Row $\mathbf{A}$:

$$\{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

For the column space, observe from $\mathbf{B}$ that the pivots are in columns 1, 2 and 4. Hence columns 1, 2 and 4 of $\mathbf{A}$ (not $\mathbf{B}$) form a basis for Col $\mathbf{A}$:

$$\text{Basis for Col } \mathbf{A}: \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Any echelon form of $\mathbf{A}$ provides (in its nonzero rows) a basis for Row $\mathbf{A}$ and also identifies the pivot columns of $\mathbf{A}$ for Col $\mathbf{A}$. However, for Nul $\mathbf{A}$, we need the reduced echelon form. Further row operations on $\mathbf{B}$ yield

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

The equation $\mathbf{Ax} = \mathbf{0}$ is equivalent to $\mathbf{Cx} = \mathbf{0}$, that is,

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 - 2x_3 + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

So $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, with x_3 and x_5 free variables. The usual calculations (discussed in lecture 21) show that

$$\text{Basis for Nul } \mathbf{A} : \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe that, unlike the bases for Col $\mathbf{A}$, the bases for Row $\mathbf{A}$ and Nul $\mathbf{A}$ have no simple connection with the entries in $\mathbf{A}$ itself.

Note

1. Although the first three rows of $\mathbf{B}$ in Example (5) are linearly independent, it is wrong to conclude that the first three rows of $\mathbf{A}$ are linearly independent. (In fact, the third row of $\mathbf{A}$ is 2 times the first row plus 7 times the second row).
2. Row operations do not preserve the linear dependence relations among the rows of a matrix.

Definition The rank of $\mathbf{A}$ is the dimension of the column space of $\mathbf{A}$.

Since Row $\mathbf{A}$ is the same as Col $\mathbf{A}^T$, the dimension of the row space of $\mathbf{A}$ is the rank of $\mathbf{A}^T$. The dimension of the null space is sometimes called the nullity of $\mathbf{A}$.

Theorem 3 (The Rank Theorem) The dimensions of the column space and the row space of an $m \times n$ matrix $\mathbf{A}$ are equal. This common dimension, the rank of $\mathbf{A}$, also equals the number of pivot positions in $\mathbf{A}$ and satisfies the equation

$$\text{rank } \mathbf{A} + \dim \text{Nul } \mathbf{A} = n$$

Example 6

- (a) If $\mathbf{A}$ is a 7×9 matrix with a two – dimensional null space, what is the rank of $\mathbf{A}$?
- (b). Could a 6×9 matrix have a two – dimensional null space?

Solution

- (a) Since $\mathbf{A}$ has 9 columns, $(\text{rank } \mathbf{A}) + 2 = 9$ and hence $\text{rank } \mathbf{A} = 7$.
- (b) No, If a 6×9 matrix, call it $\mathbf{B}$, had a two – dimensional null space, it would have to have rank 7, by the Rank Theorem. But the columns of $\mathbf{B}$ are vectors in $\mathbf{R}^6$ and so the dimension of Col $\mathbf{B}$ cannot exceed 6; that is, rank $\mathbf{B}$ cannot exceed 6.

The next example provides a nice way to visualize the subspaces we have been studying. Later on, we will learn that Row $\mathbf{A}$ and Nul $\mathbf{A}$ have only the zero vector in common and are actually “perpendicular” to each other. The same fact will apply to Row $\mathbf{A}^T (= \text{Col } \mathbf{A})$ and Nul $\mathbf{A}^T$. So the figure in Example (7) creates a good mental image for the general case.

Example 7 Let $A = \begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}$. It is readily checked that $\text{Nul } A$ is the x_2 – axis, $\text{Row } A$

A is the x_1x_3 – plane, $\text{Col } A$ is the plane whose equation is $x_1 - x_2 = 0$ and $\text{Nul } A^T$ is the set of all multiples of $(1, -1, 0)$. Figure 1 shows $\text{Nul } A$ and $\text{Row } A$ in the domain of the linear transformation $x \rightarrow Ax$; the range of this mapping, $\text{Col } A$, is shown in a separate copy of $\mathbb{R}^3$, along with $\text{Nul } A^T$.

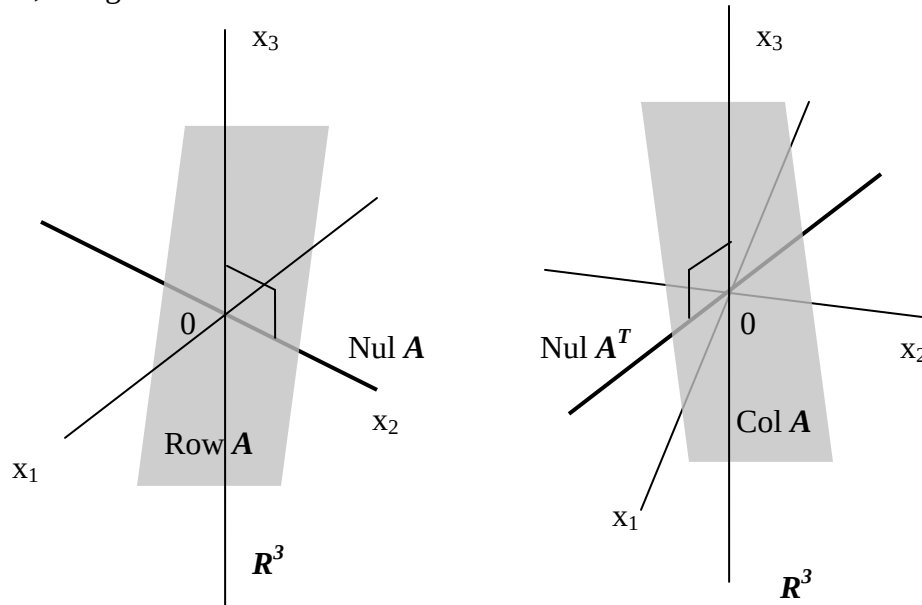


Figure 1 – Subspaces associated with a matrix A

Applications to Systems of Equations

The Rank Theorem is a powerful tool for processing information about systems of linear equations. The next example simulates the way a real-life problem using linear equations might be stated, without explicit mention of linear algebra terms such as matrix, subspace and dimension.

Example 8 A scientist has found two solutions to a homogeneous system of 40 equations in 42 variables. The two solutions are not multiples and all other solutions can be constructed by adding together appropriate multiples of these two solutions. Can the scientist be certain that an associated non-homogeneous system (with the same coefficients) has a solution?

Solution Yes. Let A be the 40×42 coefficient matrix of the system. The given information implies that the two solutions are linearly independent and span $\text{Nul } A$. So $\dim \text{Nul } A = 2$. By the Rank Theorem, $\dim \text{Col } A = 42 - 2 = 40$. Since $\mathbb{R}^{40}$ is the only subspace of $\mathbb{R}^{40}$ whose dimension is 40, $\text{Col } A$ must be all of $\mathbb{R}^{40}$. This means that every non-homogeneous equation $Ax = b$ has a solution.

Example 9

Find the rank and nullity of the matrix $A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$.

Verify that values obtained verify the dimension theorem.

Solution

$$\begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix} \quad (-1)R_1$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} \quad \begin{array}{l} (-3)R_1 + R_2 \\ (-2)R_1 + R_3 \\ (-4)R_1 + R_4 \end{array}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} \quad (-1)R_2$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} R_2 + R_3 \\ R_2 + R_4 \end{array}$$

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad 2R_2 + R_1$$

The reduced row-echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (1)$$

The corresponding system of equations will be

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$

$$x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 = 0$$

or, on solving for the leading variables,

$$x_1 = 4x_3 - 28x_4 + 37x_5 - 13x_6$$

$$x_2 = 2x_3 + 12x_4 + 16x_5 - 5x_6$$

(2)

it follows that the general solution of the system is

$$x_1 = 4r + 28s + 37t - 13u$$

$$x_2 = 2r + 12s + 16t - 5u$$

$$x_3 = r$$

$$x_4 = s$$

$$x_5 = t$$

$$x_6 = u$$

or equivalently,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r \begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (3)$$

The four vectors on the right side of (3) form a basis for the solution space, so

nullity(A) = 4. The matrix $A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$ has 6 columns,

so $\text{rank}(A) + \text{nullity}(A) = 2 + 4 = 6 = n$

Example 10 Find the rank and nullity of the matrix; then verify that the values

obtained satisfy the dimension theorem $A = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

Solution Transforming Matrix to the Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} \begin{array}{l} (-2)R_1 + R_3 \\ (-3)R_1 + R_4 \\ 2R_1 + R_5 \end{array}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} (1/3)R_2$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} (-3)R_2 + R_3 \\ (-3)R_2 + R_4 \\ (-3)R_2 + R_5 \end{array}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{aligned}
 & \left[\begin{array}{ccccc} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \\ \\ 12R_3 + R_4 \\ \\ \end{array} \\
 & \left[\begin{array}{ccccc} 1 & -3 & 0 & 2 & 11/6 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \\ (-2)R_3 + R_2 \\ (-2)R_3 + R_1 \\ \\ \end{array} \\
 & \left[\begin{array}{ccccc} 1 & 0 & 0 & 2 & 4/3 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \\ \\ (3)R_2 + R_1 \\ \\ \end{array} \quad (1)
 \end{aligned}$$

Since there are three nonzero rows (or equivalently, three leading 1's) the row space and column space are both three dimensional so rank (**A**) = 3.

To find the nullity of **A**, we find the dimension of the solution space of the linear system **Ax** = **0**. The system can be solved by reducing the augmented matrix to reduced row echelon form. The resulting matrix will be identical to (1), except with an additional last column of zeros, and the corresponding system of equations will be

$$x_1 + 0x_2 + 0x_3 + 2x_4 + \frac{4}{3}x_5 = 0$$

$$0x_1 + x_2 + 0x_3 + 0x_4 - \frac{1}{6}x_5 = 0$$

$$0x_1 + 0x_2 + x_3 + 0x_4 - \frac{5}{12}x_5 = 0$$

The system has infinitely many solutions:

$$x_1 = -2x_4 + (-4/3)x_5 \quad x_2 = (1/6)x_5$$

$$x_3 = (5/12)x_5 \quad x_4 = s$$

$$x_5 = t$$

The solution can be written in the vector form:

$$\mathbf{c}_4 = (-2, 0, 0, 1, 0) \quad \mathbf{c}_5 = (-4/3, 1/6, 5/12, 0, 1)$$

Therefore the **null space** has a basis formed by the set

$$\{(-2, 0, 0, 1, 0), (-4/3, 1/6, 5/12, 0, 1)\}$$

The nullity of the matrix is 2. Now $\text{Rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = 3 + 2 = 5 = n$

Theorem 4 If $\mathbf{A}$ is an $m \times n$, matrix, then

(a) $\text{rank}(\mathbf{A})$ = the number of leading variables in the solution of $\mathbf{Ax} = \mathbf{0}$

(b) $\text{nullity}(\mathbf{A})$ = the number of parameters in the general solution of $\mathbf{Ax} = \mathbf{0}$

Example 11 Find the number of parameters in the solution set of $\mathbf{Ax} = \mathbf{0}$ if $\mathbf{A}$ is a 5×7 matrix of rank 3.

Solution $\text{nullity}(\mathbf{A}) = n - \text{rank}(\mathbf{A}) = 7 - 3 = 4$

Thus, there are four parameters.

Example Find the number of parameters in the solution set of $\mathbf{Ax} = \mathbf{0}$ if $\mathbf{A}$ is a 4×4 matrix of rank 0.

Solution $\text{nullity}(\mathbf{A}) = n - \text{rank}(\mathbf{A}) = 4 - 0 = 4$

Thus, there are four parameters.

Theorem 5 If $\mathbf{A}$ is any matrix, then $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$

Four fundamental matrix spaces

If we consider a matrix $\mathbf{A}$ and its transpose $\mathbf{A}^T$ together, then there are six vectors spaces of interest:

Row space of $\mathbf{A}$ row space of $\mathbf{A}^T$

Column space of $\mathbf{A}$ column space of $\mathbf{A}^T$

Null space of $\mathbf{A}$ null space of $\mathbf{A}^T$

However, transposing a matrix converts row vectors into column vectors and column vectors into row vectors, so that, except for a difference in notation, the row space of $\mathbf{A}^T$ is the same as the column space of $\mathbf{A}$ and the column space of $\mathbf{A}^T$ is the same as row space of $\mathbf{A}$.

This leaves four vector spaces of interest:

Row space of $\mathbf{A}$ column space of $\mathbf{A}$

Null space of $\mathbf{A}$ null space of $\mathbf{A}^T$

These are known as the **fundamental matrix spaces** associated with $\mathbf{A}$, if $\mathbf{A}$ is an $m \times n$ matrix, then the row space of $\mathbf{A}$ and null space of $\mathbf{A}$ are subspaces of $\mathbf{R}^n$ and the column space of $\mathbf{A}$ and the null space of $\mathbf{A}^T$ are subspaces of $\mathbf{R}^m$.

Suppose now that $\mathbf{A}$ is an $m \times n$ matrix of rank r , it follows from theorem (5) that $\mathbf{A}^T$ is an $n \times m$ matrix of rank r . Applying theorem (3) on $\mathbf{A}$ and $\mathbf{A}^T$ yields

$$\text{Nullity}(\mathbf{A}) = n - r, \text{nullity}(\mathbf{A}^T) = m - r$$

From which we deduce the following table relating the dimensions of the four fundamental spaces of an $m \times n$ matrix $\mathbf{A}$ of rank r .

Fundamental space	Dimension
Row space of A	r
Column space of A	r
Null space of A	$n-r$
Null space of A^T	$m-r$

Example 12 If A is a 7×4 matrix, then the rank of A is at most 4 and, consequently, the seven row vectors must be linearly dependent. If A is a 4×7 matrix, then again the rank of A is at most 4 and, consequently, the seven column vectors must be linearly dependent.

Rank and the Invertible Matrix Theorem The various vector space concepts associated with a matrix provide several more statements for the Invertible Matrix Theorem. We list only the new statements here, but we reference them so they follow the statements in the original Invertible Matrix Theorem in lecture 13.

Theorem 6 The Invertible Matrix Theorem (Continued)

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

- m. The columns of A form a basis of $\mathbf{R}^n$.
- n. $\text{Col } A = \mathbf{R}^n$.
- o. $\dim \text{Col } A = n$
- p. $\text{rank } A = n$
- q. $\text{Nul } A = \{\mathbf{0}\}$
- r. $\dim \text{Nul } A = 0$

Proof Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning. The other statements above are linked into the theorem by the following chain of almost trivial implications:

$$(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

Only the implication $(p) \Rightarrow (r)$ bears comment. It follows from the Rank Theorem because A is $n \times n$. Statements (d) and (g) are already known to be equivalent, so the chain is a circle of implications.

We have refrained from adding to the Invertible Matrix Theorem obvious statements about the row space of A , because the row space is the column space of A^T . Recall from (1) of the Invertible Matrix Theorem that A is invertible if and only if A^T is invertible. Hence every statement in the Invertible Matrix Theorem can also be stated for A^T .

Numerical Note

Many algorithms discussed in these lectures are useful for understanding concepts and making simple computations by hand. However, the algorithms are often unsuitable for large-scale problems in real life. Rank determination is a good example. It would seem easy to reduce a matrix to echelon form and count the pivots. But unless exact arithmetic is performed on a matrix whose entries are specified exactly, row operations can change the apparent rank of a matrix.

For instance, if the value of x in the matrix $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ is not stored exactly as 7 in a computer, then the rank may be 1 or 2, depending on whether the computer treats $x - 7$ as zero.

In practical applications, the effective rank of a matrix $\mathbf{A}$ is often determined from the singular value decomposition of $\mathbf{A}$.

Example 13 The matrices below are row equivalent

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Find rank $\mathbf{A}$ and $\dim \text{Nul } \mathbf{A}$.
2. Find bases for $\text{Col } \mathbf{A}$ and $\text{Row } \mathbf{A}$.
3. What is the next step to perform if one wants to find a basis for $\text{Nul } \mathbf{A}$?
4. How many pivot columns are in a row echelon form of $\mathbf{A}^T$?

Solution

1. $\mathbf{A}$ has two pivot columns, so $\text{rank } \mathbf{A} = 2$. Since $\mathbf{A}$ has 5 columns altogether, $\dim \text{Nul } \mathbf{A} = 5 - 2 = 3$.
2. The pivot columns of $\mathbf{A}$ are the first two columns. So a basis for $\text{Col } \mathbf{A}$ is

$$\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

The nonzero rows of $\mathbf{B}$ form a basis for $\text{Row } \mathbf{A}$, namely $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. In this particular example, it happens that any two rows of $\mathbf{A}$ form a basis for the row space, because the row space is two-dimensional and none of the rows of $\mathbf{A}$ is a multiple of another row. In general, the nonzero rows of an echelon form of $\mathbf{A}$ should be used as a basis for $\text{Row } \mathbf{A}$, not the rows of $\mathbf{A}$ itself.

3. For $\text{Nul } \mathbf{A}$, the next step is to perform row operations on $\mathbf{B}$ to obtain the reduced echelon form of $\mathbf{A}$.
4. $\text{Rank } \mathbf{A}^T = \text{rank } \mathbf{A}$, by the Rank Theorem, because $\text{Col } \mathbf{A}^T = \text{Row } \mathbf{A}$. So $\mathbf{A}^T$ has two pivot positions.

Exercises

In exercises 1 to 4, assume that the matrix $\mathbf{A}$ is row equivalent to $\mathbf{B}$. Without calculations, list rank $\mathbf{A}$ and $\dim \text{Nul } \mathbf{A}$. Then find bases for $\text{Col } \mathbf{A}$, $\text{Row } \mathbf{A}$, and $\text{Nul } \mathbf{A}$.

$$1. \mathbf{A} = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$2. \mathbf{A} = \begin{bmatrix} 1 & -3 & 4 & -1 & 9 \\ -2 & 6 & -6 & -1 & -10 \\ -3 & 9 & -6 & -6 & -3 \\ 3 & -9 & 4 & 9 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & -3 & 0 & 5 & -7 \\ 0 & 0 & 2 & -3 & 8 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$3. \mathbf{A} = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$4. \mathbf{A} = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

5. If a 3×8 matrix $\mathbf{A}$ has rank 3, find $\dim \text{Nul } \mathbf{A}$, $\dim \text{Row } \mathbf{A}$, and $\text{rank } \mathbf{A}^T$.
6. If a 6×3 matrix $\mathbf{A}$ has rank 3, find $\dim \text{Nul } \mathbf{A}$, $\dim \text{Row } \mathbf{A}$, and $\text{rank } \mathbf{A}^T$.
7. Suppose that a 4×7 matrix $\mathbf{A}$ has four pivot columns. Is $\text{Col } \mathbf{A} = \mathbf{R}^4$? Is $\text{Nul } \mathbf{A} = \mathbf{R}^3$? Explain your answers.
8. Suppose that a 5×6 matrix $\mathbf{A}$ has four pivot columns. What is $\dim \text{Nul } \mathbf{A}$? Is $\text{Col } \mathbf{A} = \mathbf{R}^4$? Why or why not?
9. If the null space of a 5×6 matrix $\mathbf{A}$ is 4-dimensional, what is the dimension of the column space of $\mathbf{A}$?
10. If the null space of a 7×6 matrix $\mathbf{A}$ is 5-dimensional, what is the dimension of the column space of $\mathbf{A}$?
11. If the null space of an 8×5 matrix $\mathbf{A}$ is 2-dimensional, what is the dimension of the row space of $\mathbf{A}$?
12. If the null space of a 5×6 matrix $\mathbf{A}$ is 4-dimensional, what is the dimension of the row space of $\mathbf{A}$?
13. If $\mathbf{A}$ is a 7×5 matrix, what is the largest possible rank of $\mathbf{A}$? If $\mathbf{A}$ is a 5×7 matrix, what is the largest possible rank of $\mathbf{A}$? Explain your answers.

14. If $\mathbf{A}$ is a 4×3 matrix, what is the largest possible dimension of the row space of $\mathbf{A}$? If $\mathbf{A}$ is a 3×4 matrix, what is the largest possible dimension of the row space of $\mathbf{A}$? Explain.
15. If $\mathbf{A}$ is a 6×8 matrix, what is the smallest possible dimension of $\text{Nul } \mathbf{A}$?
16. If $\mathbf{A}$ is a 6×4 matrix, what is the smallest possible dimension of $\text{Nul } \mathbf{A}$?

Lecture 26

Change of Basis

When a basis $\mathbf{B}$ is chosen for an n -dimensional vector space $\mathbf{V}$, the associated coordinate mapping onto $\mathbf{R}^n$ provides a coordinate system for $\mathbf{V}$. Each $\mathbf{x}$ in $\mathbf{V}$ is identified uniquely by its $\mathbf{B}$ -coordinate vector $[\mathbf{x}]_{\mathbf{B}}$.

In some applications, a problem is described initially using a basis $\mathbf{B}$, but the problem's solution is aided by changing $\mathbf{B}$ to a new basis $\mathbf{C}$. Each vector is assigned a new $\mathbf{C}$ -coordinate vector. In this section, we study how $[\mathbf{x}]_{\mathbf{C}}$ and $[\mathbf{x}]_{\mathbf{B}}$ are related for each $\mathbf{x}$ in $\mathbf{V}$. To visualize the problem, consider the two coordinate systems in Fig. 1.

In Fig. 1(a), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, while in Fig. 1 (b), the same $\mathbf{x}$ is shown as $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$. That is

$$[\mathbf{x}]_{\mathbf{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{x}]_{\mathbf{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Example 1 shows how to find the connection between the two coordinate vectors, provided we know how $\mathbf{b}_1$ and $\mathbf{b}_2$ are formed from $\mathbf{c}_1$ and $\mathbf{c}_2$.

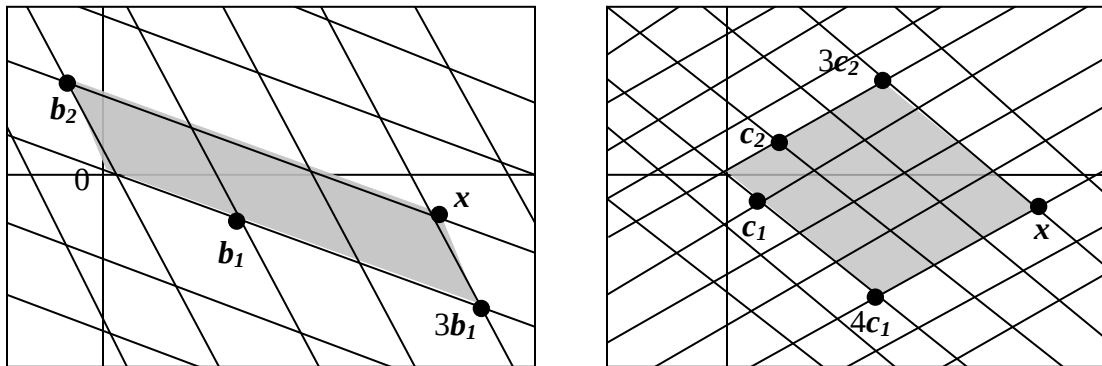


Figure 1 – Two coordinate systems for the same vector space

Example 1 Consider two bases $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ for a vector space $\mathbf{V}$, such that

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \quad \text{and} \quad \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \quad (1)$$

$$\text{Suppose that } \mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \quad (2)$$

suppose that $[\mathbf{x}]_{\mathbf{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Find $[\mathbf{x}]_{\mathbf{C}}$.

Solution Since the coordinate mapping is a linear transformation, we apply (2)

$$\begin{aligned} [\mathbf{x}]_C &= [3\mathbf{b}_1 + \mathbf{b}_2]_C \\ &= 3[\mathbf{b}_1]_C + [\mathbf{b}_2]_C \end{aligned}$$

as a matrix equation,

$$[\mathbf{x}]_C = \begin{bmatrix} [\mathbf{b}_1]_C & [\mathbf{b}_2]_C \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad (3)$$

From (1),

$$[\mathbf{b}_1]_C = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \text{ and } [\mathbf{b}_2]_C = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

$$\text{from (3)} \quad [\mathbf{x}]_C = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

The C – coordinates of $\mathbf{x}$ match those of the $\mathbf{x}$ in Fig. 1.

The argument used to derive formula (3) is easily generalized to yield the following result.

Theorem Let $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathbf{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of a vector space $\mathbf{V}$. Then there exists an $n \times n$ matrix $\mathbf{P}_{C \leftarrow B}$ such that

$$[\mathbf{x}]_C = \mathbf{P}_{C \leftarrow B} [\mathbf{x}]_B \quad (4)$$

The columns of $\mathbf{P}_{C \leftarrow B}$ are the C -coordinate vectors of the vectors in the basis $\mathbf{B}$. So

$$\mathbf{P}_{C \leftarrow B} = \begin{bmatrix} [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \end{bmatrix}$$

The matrix $\mathbf{P}_{C \leftarrow B}$ in above theorem is called the **change-of-coordinates matrix from $\mathbf{B}$ to $\mathbf{C}$** . Multiplication by $\mathbf{P}_{C \leftarrow B}$ converts $\mathbf{B}$ -coordinates into $\mathbf{C}$ -coordinates. Figure 2 illustrates the change-of-coordinates equation (4).

The columns of $\mathbf{P}_{C \leftarrow B}$ are linearly independent because they are the coordinate vectors of the linearly independent set $\mathbf{B}$. Thus $\mathbf{P}_{C \leftarrow B}$ is invertible. Left-multiplying both sides of (4)

by $\left(\mathbf{P}_{C \leftarrow B}\right)^{-1}$, we obtain

$$\left(\mathbf{P}_{C \leftarrow B}\right)^{-1} [\mathbf{x}]_C = [\mathbf{x}]_B$$

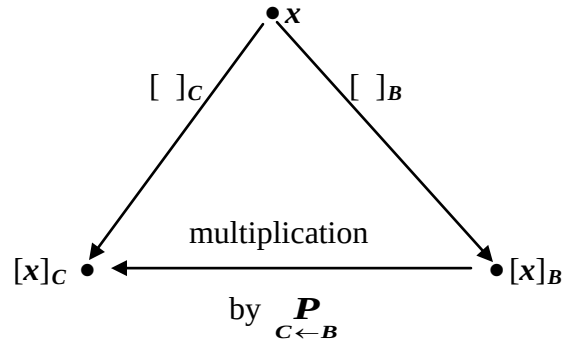


Figure 2 Two coordinate system for V .

Thus $\left(\mathbf{P}_{C \leftarrow B}\right)^{-1}$ is the matrix that converts C -coordinates into B -coordinates. That is,

$$\left(\mathbf{P}_{C \leftarrow B}\right)^{-1} = \mathbf{P}_{B \leftarrow C} \quad (5)$$

Change of Basis in $\mathbf{R}^n$ If $B = \{b_1, \dots, b_n\}$ and E is the standard basis $\{e_1, \dots, e_n\}$ in $\mathbf{R}^n$, then $[b_1]_E = b_1$, and likewise for the other vectors in B . In this case, $\mathbf{P}_{E \leftarrow B}$ is the same as the change-of-coordinates matrix $\mathbf{P}_B$ introduced in Lecture 23, namely,

$$\mathbf{P}_B = [b_1 \ b_2 \ \dots \ b_n]$$

To change coordinates between two nonstandard bases in $\mathbf{R}^n$, we need above Theorem. The theorem shows that to solve the change-of-basis problem, we need the coordinate vector of the old basis relative to the new basis.

Example 2 Let $D = \{d_1, d_2, d_3\}$ and $F = \{f_1, f_2, f_3\}$ be basis for vector space V , and suppose that $f_1 = 2d_1 - d_2 + d_3$, $f_2 = 3d_2 + d_3$, $f_3 = -3d_1 + 2d_3$.

(a) Find the change-of-coordinates matrix from F to D .

(b) Find $[x]_D$ for $x = f_1 - 2f_2 + 2f_3$.

Solution (a) $\mathbf{P}_{D \leftarrow F} = [[f_1]_D \ [f_2]_D \ [f_3]_D]$

But $[f_1]_D = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, [f_2]_D = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}, [f_3]_D = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$

$$\therefore \mathbf{P}_{D \leftarrow F} = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

(b) $[x]_D = [f_1 - 2f_2 + 2f_3]_D = [f_1]_D - 2[f_2]_D + 2[f_3]_D$

$$\begin{aligned}
&= \begin{bmatrix} [\mathbf{f}_1]_D & [\mathbf{f}_2]_D & [\mathbf{f}_3]_D \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 0 & -3 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ -7 \\ 3 \end{bmatrix}
\end{aligned}$$

Example 3 Let $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ and consider the bases for $\mathbb{R}^2$ given by $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Find the change of coordinates matrix from $\mathbf{B}$ to $\mathbf{C}$.

Solution The matrix $\mathbf{P}_{\mathbf{C} \leftarrow \mathbf{B}}$ involves the $\mathbf{C}$ – coordinate vectors of $\mathbf{b}_1$ and $\mathbf{b}_2$.

Let $[\mathbf{b}_1]_{\mathbf{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $[\mathbf{b}_2]_{\mathbf{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then, by definition,

$$[\mathbf{c}_1 \quad \mathbf{c}_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \quad \text{and} \quad [\mathbf{c}_1 \quad \mathbf{c}_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2$$

To solve both systems simultaneously augment the coefficient matrix with $\mathbf{b}_1$ and $\mathbf{b}_2$ and row reduce:

$$\begin{aligned}
[\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{b}_1 \quad \mathbf{b}_2] &= \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array} \right] \\
&\sim \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array} \right] R_2 + 4R_1 \\
&\sim \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array} \right] \frac{1}{7}R_2 \\
&\sim \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array} \right] R_1 - 3R_2 \quad (6)
\end{aligned}$$

$$[\mathbf{b}_1]_{\mathbf{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{and} \quad [\mathbf{b}_2]_{\mathbf{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

therefore

$$\mathbf{P}_{\mathbf{C} \leftarrow \mathbf{B}} = [\mathbf{b}_1]_{\mathbf{C}} \quad [\mathbf{b}_2]_{\mathbf{C}} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

past paper

Example 4

Let $b_1 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, $b_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $c_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $c_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and consider the bases for R^2 given by $B = \{b_1, b_2\}$ and $C = \{c_1, c_2\}$.

(a) Find the change – of – coordinates matrix from C to B .

(b) Find the change – of – coordinates matrix from B to C .

Solution

(a) Notice that $P_{B \leftarrow C}$ is needed rather than $P_{C \leftarrow B}$ and compute

$$[b_1 \ b_2 \mid c_1 \ c_2] = \left[\begin{array}{cc|cc} 7 & 2 & 4 & 5 \\ -2 & -1 & 1 & 2 \end{array} \right] \xrightarrow{\frac{1}{7}R_1} \left[\begin{array}{cc|cc} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ -2 & -1 & 1 & 2 \end{array} \right]$$

$$\xrightarrow{2R_1 + R_2} \left[\begin{array}{cc|cc} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ 0 & -\frac{3}{7} & \frac{15}{7} & \frac{24}{7} \end{array} \right]$$

$$\xrightarrow{-\frac{7}{3}R_2} \left[\begin{array}{cc|cc} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ 0 & 1 & -5 & -8 \end{array} \right]$$

$$\xrightarrow{-\frac{2}{7}R_2 + R_1} \left[\begin{array}{cc|cc} 1 & 0 & 2 & 3 \\ 0 & 1 & -5 & -8 \end{array} \right]$$

so $P_{B \leftarrow C} = \begin{bmatrix} 2 & 3 \\ -5 & -8 \end{bmatrix}$

(b) By part (a) and property (5) above (with B and C interchanged),

$$P_{C \leftarrow B} = (P_{B \leftarrow C})^{-1} = \frac{1}{-1} \begin{bmatrix} -8 & -3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 3 \\ -5 & -2 \end{bmatrix}.$$

Example 5 Let $b_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $b_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $c_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $c_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$ and consider the bases for

R^2 given by $B = \{b_1, b_2\}$ and $C = \{c_1, c_2\}$.

(a) Find the change – of – coordinates matrix from C to B .

(b) Find the change – of – coordinates matrix from B to C .

Solution

(a) Notice that $P_{B \leftarrow C}$ is needed rather than $P_{C \leftarrow B}$ and compute

$$[b_1 \ b_2 \mid c_1 \ c_2] = \left[\begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array} \right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array} \right]$$

so
$$\mathbf{P}_{B \leftarrow C} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

(b) By part (a) and property (5) above (with $\mathbf{B}$ and $\mathbf{C}$ interchanged),

$$\mathbf{P}_{C \leftarrow B} = (\mathbf{P}_{B \leftarrow C})^{-1} = \frac{1}{2} \begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}.$$

Example 6

- Let $\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$ and $\mathbf{G} = \{\mathbf{g}_1, \mathbf{g}_2\}$ be bases for a vector space $\mathbf{V}$, and let $\mathbf{P}$ be a matrix whose columns are $[\mathbf{f}_1]_{\mathbf{G}}$. Which of the following equations is satisfied by $\mathbf{P}$ for all $\mathbf{v}$ in $\mathbf{V}$?
 (i) $[\mathbf{v}]_{\mathbf{F}} = \mathbf{P}[\mathbf{v}]_{\mathbf{G}}$ (ii) $[\mathbf{v}]_{\mathbf{G}} = \mathbf{P}[\mathbf{v}]_{\mathbf{F}}$
- Let $\mathbf{B}$ and $\mathbf{C}$ be as in Example 1. Use the results of that example to find the change-of-coordinates matrix from $\mathbf{C}$ to $\mathbf{B}$.

Solution

- Since the columns of $\mathbf{P}$ are $\mathbf{G}$ -coordinate vectors, a vector of the form $\mathbf{P}\mathbf{x}$ must be a $\mathbf{G}$ -coordinate vector. Thus $\mathbf{P}$ satisfies equation (ii).
- The coordinate vectors found in Example 1 show that

$$\mathbf{P}_{C \leftarrow B} = \begin{bmatrix} [\mathbf{b}_1]_{\mathbf{C}} & [\mathbf{b}_2]_{\mathbf{C}} \end{bmatrix} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Hence
$$\mathbf{P}_{B \leftarrow C} = (\mathbf{P}_{C \leftarrow B})^{-1} = \frac{1}{10} \begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.6 \\ -0.1 & 0.4 \end{bmatrix}$$

Exercises

- Let $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for a vector space $\mathbf{V}$, and suppose that $\mathbf{b}_1 = 6\mathbf{c}_1 - 2\mathbf{c}_2$ and $\mathbf{b}_2 = 9\mathbf{c}_1 - 4\mathbf{c}_2$.
 (a) Find the change-of-coordinates matrix from $\mathbf{B}$ to $\mathbf{C}$.
 (b) Find $[\mathbf{x}]_{\mathbf{C}}$ for $\mathbf{x} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. Use part (a).
- Let $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for a vector space $\mathbf{V}$, and suppose that $\mathbf{b}_1 = -\mathbf{c}_1 + 4\mathbf{c}_2$ and $\mathbf{b}_2 = 5\mathbf{c}_1 - 3\mathbf{c}_2$.
 (a) Find the change-of-coordinates matrix from $\mathbf{B}$ to $\mathbf{C}$.
 (b) Find $[\mathbf{x}]_{\mathbf{C}}$ for $\mathbf{x} = 5\mathbf{b}_1 + 3\mathbf{b}_2$.
- Let $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $\mathbf{W} = \{\mathbf{w}_1, \mathbf{w}_2\}$ be bases for $\mathbf{V}$, and let $\mathbf{P}$ be a matrix whose columns are $[\mathbf{u}_1]_{\mathbf{W}}$ and $[\mathbf{u}_2]_{\mathbf{W}}$. Which of the following equations is satisfied by $\mathbf{P}$ for all $\mathbf{x}$ in $\mathbf{V}$?
 (i) $[\mathbf{x}]_{\mathbf{U}} = \mathbf{P}[\mathbf{x}]_{\mathbf{W}}$ (ii) $[\mathbf{x}]_{\mathbf{W}} = \mathbf{P}[\mathbf{x}]_{\mathbf{U}}$
- Let $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ be bases for $\mathbf{V}$, and let $\mathbf{P} = \begin{bmatrix} [\mathbf{d}_1]_{\mathbf{A}} & [\mathbf{d}_2]_{\mathbf{A}} & [\mathbf{d}_3]_{\mathbf{A}} \end{bmatrix}$. Which of the following equations is satisfied by $\mathbf{P}$ for all $\mathbf{x}$ in $\mathbf{V}$?
 (i) $[\mathbf{x}]_{\mathbf{A}} = \mathbf{P}[\mathbf{x}]_{\mathbf{D}}$ (ii) $[\mathbf{x}]_{\mathbf{D}} = \mathbf{P}[\mathbf{x}]_{\mathbf{A}}$

5. Let $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be bases for a vector space $\mathbf{V}$, and suppose that $\mathbf{a}_1 = 4\mathbf{b}_1 - \mathbf{b}_2$, $\mathbf{a}_2 = -\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, and $\mathbf{a}_3 = \mathbf{b}_2 - 2\mathbf{b}_3$.

(a) Find the change-of-coordinates matrix from $\mathbf{A}$ to $\mathbf{B}$.

(b) Find $[\mathbf{x}]_{\mathbf{B}}$ for $\mathbf{x} = 3\mathbf{a}_1 + 4\mathbf{a}_2 + \mathbf{a}_3$.

In exercises 6 to 9, let $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for $\mathbf{R}^2$. In each exercise, find the change-of-coordinates matrix from $\mathbf{B}$ to $\mathbf{C}$, and change-of-coordinates matrix from $\mathbf{C}$ to $\mathbf{B}$.

$$6. \mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

$$7. \mathbf{b}_1 = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$8. \mathbf{b}_1 = \begin{bmatrix} -6 \\ -1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

$$9. \mathbf{b}_1 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

10. In $\mathbf{P}_2$, find the change-of-coordinate matrix from the basis $\mathbf{B} = \{\mathbf{1} - 2\mathbf{t} + \mathbf{t}^2, \mathbf{3} - 5\mathbf{t} + 4\mathbf{t}^2, 2\mathbf{t} + 3\mathbf{t}^2\}$ to the standard basis $\mathbf{C} = \{\mathbf{1}, \mathbf{t}, \mathbf{t}^2\}$. Then find the $\mathbf{B}$ -coordinate vector for $-\mathbf{1} + 2\mathbf{t}$.

11. In $\mathbf{P}_2$, find the change-of-coordinates matrix from the basis $\mathbf{B} = \{\mathbf{1} - 3\mathbf{t}^2, 2 + \mathbf{t} - 5\mathbf{t}^2, \mathbf{1} + 2\mathbf{t}\}$ to the standard basis. Then write $\mathbf{t}^2$ as a linear combination of the polynomials in $\mathbf{B}$.

$$12. \text{ Let } \mathbf{P} = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -5 & 0 \\ 4 & 6 & 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$$

(a). Find a basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ for $\mathbf{R}^3$ such that $\mathbf{P}$ is the change-of-coordinates matrix from $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(b). Find a basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for $\mathbf{R}^3$ such that $\mathbf{P}$ is the change-of-coordinates matrix from $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

Lecture 27

Applications to Difference Equations

Recently Discrete or Digital data has been used preferably rather than continuous data in scientific and engineering problems. Difference equation is considered more reliable tool to analyze such type of data even if we are using a differential equation to analyze continuous process, a numerical solution is often produced from a related difference equation.

In this lecture we will study some fundamental properties of linear difference equation that are considered best tool in Linear Algebra.

Discrete-Time Signals Let $\mathcal{S}$ is the space of discrete-time signals. A signal in $\mathcal{S}$ is a function defined only on the integers and is visualized as a sequence of numbers, say, $\{y_k\}$.

Digital signals obviously arise in electrical and control systems engineering, but discrete-data sequences are also generated in biology, physics, economics, demography and many other areas, wherever a process is measured, or sampled, at discrete time intervals.

When a process beings at a specific time, it is sometimes convenient to write a signal as a sequence of the form $(y_0, y_1, y_2, \dots)$. The term y_k for $k < 0$ either are assumed to be zero or are simply omitted.

Discrete time signals are functions defined on integers that are sequences or it takes only a discrete set of values.

For Example a Radio Station broadcast's weather report once in a day so the sampling period is discrete, but sampling may be uniform or constant accordingly.

Example 1 The crystal clear sounds from a compact disc player are produced from music that has been sampled at the rate of 44,100 times per second. At each measurement, the amplitude of the music signal is recorded as a number, say, y_k . The original music is composed of many different sounds of varying frequencies, yet the sequence $\{y_k\}$ contains enough information to reproduce all the frequencies in the sound up to about 20,000 cycles per second, higher than the human ear can sense.

Linear Independence in the Space $\mathcal{S}$ of Signals To simplify notation, we consider a set of only three signals in $\mathcal{S}$, say, $\{u_k\}, \{v_k\}$ and $\{w_k\}$. They are linearly independent precisely when the equation

$$c_1 u_k + c_2 v_k + c_3 w_k = 0 \text{ for all } k \quad (1)$$

implies that $c_1 = c_2 = c_3 = 0$. The phrase “for all k ” means for all integers positive, negative and zero. One could also consider signals that start with $k = 0$, for example, in a case “for all k ” would mean for all integers $k \geq 0$.

Suppose c_1, c_2, c_3 satisfy (1). Then the equation (1) holds for any three consecutive values of k , say, $k, k + 1$ and $k + 2$. Thus (1) implies that

$$c_1 u_{k+1} + c_2 v_{k+1} + c_3 w_{k+1} = 0 \text{ for all } k$$

and
$$c_1 u_{k+2} + c_2 v_{k+2} + c_3 w_{k+2} = 0 \text{ for all } k$$

Hence c_1, c_2, c_3 satisfy

$$\begin{bmatrix} \mathbf{u}_k & \mathbf{v}_k & \mathbf{w}_k \\ \mathbf{u}_{k+1} & \mathbf{v}_{k+1} & \mathbf{w}_{k+1} \\ \mathbf{u}_{k+2} & \mathbf{v}_{k+2} & \mathbf{w}_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ for all } k \quad (2)$$

The coefficient matrix in this system is called the **Casorati matrix** of the signals, and the determinant of the matrix is called the **Casoratian** of $\{\mathbf{u}_k\}, \{\mathbf{v}_k\}$, and $\{\mathbf{w}_k\}$. If for at least one value of k the Casorati matrix is invertible, then (2) will imply that $c_1 = c_2 = c_3 = 0$, which will prove that the three signals are linearly independent.

Example 2 Verify that $1^k, (-2)^k$ and 3^k are linearly independent signals.

Solution The Casorati matrix is $\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$

Row operations can show fairly easily that this matrix is always invertible. However; it is faster to substitute a value for k – say, $k = 0$ – and row reduce the numerical matrix:

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 1 & 4 & 9 \end{bmatrix} R_2 - R_1 \\ &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} R_3 - R_1 \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix} R_3 + R_2 \end{aligned}$$

The Casorati matrix is invertible for $k = 0$. So $1^k, (-2)^k$ and 3^k are linearly independent.

If a Casorati matrix is not invertible, the associated signals being tested may or may not be linearly dependent. However, it can be shown that if the signals are all solutions of the same homogeneous difference equation (described below), then either the Casorati matrix is invertible for all k and the signals are linearly independent, or else for all k the Casorati matrix is not invertible and the signals are linearly dependent.

Activity

Verify that $2^k, (-1)^k$ and 0^k are linearly independent signals.

Linear Difference Equations Given scalars $a_0, \dots, a_n$, with a_0 and a_n nonzero, and given a signal $\{z_k\}$, the equation

$$a_0 y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = z_k \text{ for all } k \quad (3)$$

is called a **linear difference equation** (or **linear recurrence relation**) of order n . For simplicity, a_0 is often taken equal to 1. If $\{z_k\}$ is the zero sequence, the equation is **homogeneous**; otherwise, the equation is **non-homogeneous**.

In simple words, **an equation which expresses a value of a sequence as a function of the other terms in the sequence is called a difference equation.**

In particular an equation which expresses the value a_n of a sequence $\{a_n\}$ as a function of the term a_{n-1} is called a **first order difference equation.**

Example 3 In digital signal processing, a difference equation such as (3) above describes a linear filter and $a_0, \dots, a_n$ are called the **filter coefficients**. If $\{y_k\}$ is treated as the input and $\{z_k\}$ the output, then the solutions of the associated homogeneous equation are the signals that are filtered out and transformed into the zero signal. Let us feed two different signals into the filter

$$0.35y_{k+2} + 0.5y_{k+1} + 0.35y_k = z_k$$

Here 0.35 is an abbreviation for $\sqrt{2}/4$. The first signal is created by sampling the continuous signal $y = \cos(\pi t/4)$ at integer values of t , as in Fig. 3 (a). The discrete signals is $\{y_k\} = \{\dots \cos(0), \cos(\pi/4), \cos(2\pi/4), \cos(3\pi/4) \dots\}$

For simplicity, write ± 0.7 in place of $\pm\sqrt{2}/2$, so that

$$\{y_k\} = \{\dots, 1, 0.7, 0, -0.7, -1, -0.7, 0, 0.7, 1, 0.7, 0, \dots\}$$

$\uparrow$
 $k = 0$

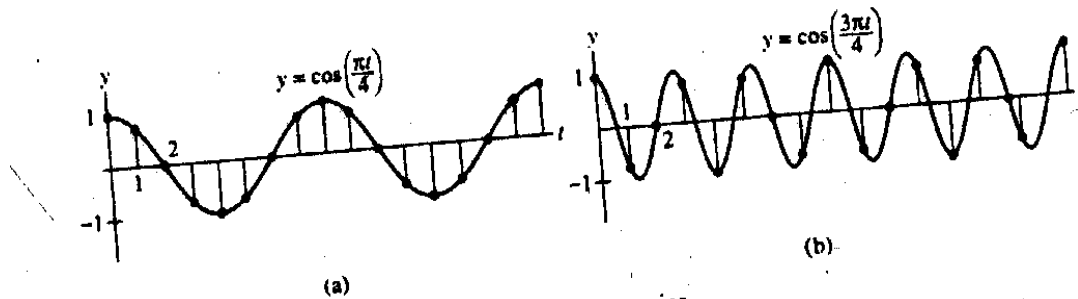


Figure Discrete signals with different frequencies

The following table shows a calculation of the output sequence $\{z_k\}$, where $0.35(0.7)$ is an abbreviation for $(\sqrt{2}/4) \cdot (\sqrt{2}/2) = 0.25$. The output is $\{y_k\}$, shifted by one term.

Table Computing the output of a filter

K	y_k	y_{k+1}	y_{k+2}	$0.35y_k + 0.5y_{k+1} + 0.35y_{k+2} = z_k$
0	1	0.7	0	$0.35(1) + 0.5(0.7) + 0.35(0) = 0.7$
1	0.7	0	-0.7	0
2	0	-0.7	-1	-0.7
3	-0.7	-1	-0.7	-1
4	-1	-0.7	0	-0.7
5	-0.7	0	0.7	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

A different input signal is produced from the higher frequency signal $y = \cos(3\pi t/4)$. Sampling at the same rate as before produces a new input sequence:

$$\{w_k\} = \{\dots, 1, -0.7, 0, 0.7, -1, 0.7, 0, -0.7, 1, -0.7, 0, \dots\}$$

$\uparrow$
 $k = 0$

When $\{w_k\}$ is fed into the filter, the output is the zero sequence. The filter, called a **low pass filter**, lets $\{y_k\}$ pass through, but stops the higher frequency $\{w_k\}$.

In many applications, a sequence $\{z_k\}$ is specified for the right side of a difference equation (3) and a $\{y_k\}$ that satisfies (3) is called a solution of the equation. The next example shows how to find solution for a homogeneous equation.

Example 4 Solutions of a homogeneous difference equation often have the form $y_k = r^k$ for some r . Find some solutions of the equation

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \text{ for all } k \quad (4)$$

Solution Substitute r^k for y_k in the equation and factor the left side:

$$r^{k+3} - 2r^{k+2} - 5r^{k+1} + 6r^k = 0 \quad (5)$$

By synthetic division

$$r^k (r^3 - 2r^2 - 5r + 6) = 0$$

	1	-2	-5	6
1	0	1	-1	-6
	1	-1	-6	0

So we get

$$r^k [(x-1)(x^2 - x - 6)] = 0$$

$$r^k [(x-1)(x^2 - 3x + 2x - 6)] = 0$$

$$r^k [(x-1)\{x(x-3) + 2(x-3)\}] = 0$$

$$r^k (r-1)(r+2)(r-3) = 0 \quad (6)$$

since (5) is equivalent to (6), r^k satisfies the difference equation (4) if and only if r satisfies (6). Thus 1^k , $(-2)^k$ and 3^k are all solutions of (4). For instance, to verify that 3^k is a solution of (4), compute

$$3^{k+3} - 2 \cdot 3^{k+2} - 5 \cdot 3^{k+1} + 6 \cdot 3^k = 3^k (27 - 18 - 15 + 6) = 0 \text{ for all } k$$

In general, a nonzero signal r^k satisfies the homogeneous difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = 0 \text{ for all } k$$

if and only if r is a root of the **auxiliary equation**

$$r^n + a_1 r^{n-1} + \dots + a_{n-1} r + a_n \cdot 1 = 0$$

We will not consider the case when r is a repeated root of the auxiliary equation. When the auxiliary equation has a complex root, the difference equation has solutions of the form $s^k \cos kw$ and $s^k \sin kw$, for constants s and w . This happened in Example 3.

Solution Sets of Linear Difference Equations

Given $a_1, \dots, a_n$, consider the mapping $T: S \rightarrow S$ that transforms a signal $\{y_k\}$ into a signal $\{w_k\}$ given by

$$w_k = y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k$$

It is readily checked that T is a linear transformation. This implies that the solution set of the homogeneous equation

$$y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = 0 \text{ for all } k$$

is the kernel of T (kernel is the set of signals that T maps into the zero signal) and hence the solution set is a subspace of S . Any linear combination of solutions is again a solution.

Theorem If $a_n \neq 0$ and if $\{z_k\}$ is given, the equation

$$y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = z_k \text{ for all } k \quad (7)$$

has a unique solution whenever $y_0, \dots, y_{n-1}$ are specified.

Proof If $y_0, \dots, y_{n-1}$ are specified, use (7) to define

$$y_n = z_0 - [a_1 y_{n-1} + \dots + a_n y_0]$$

And now that $y_1, \dots, y_n$ are specified, use (7) to define y_{n+1} . In general, use the recursion relation

$$y_{n+k} = z_k - [a_1 y_{k+n-1} + \dots + a_n y_k] \quad (8)$$

to define y_{n+k} for $k \geq 0$. To define y_k for $k < 0$, use the recursion relation

$$y_k = \frac{1}{a_n} z_k - \frac{1}{a_n} [y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1}] \quad (9)$$

This produces a signal that satisfies (7). Conversely, any signal that satisfies (7) for all k certainly satisfies (8) and (9) so the solution of (7) is unique.

Theorem The set H of all solutions of the n th-order homogeneous linear difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = 0 \text{ for all } k \quad (10)$$

is an n -dimensional vector space.

Proof We explained earlier why H is a subspace of S . For $\{y_k\}$ in H , let $F\{y_k\}$ be the vector in $\mathbf{R}^n$ given by $(y_0, y_1, \dots, y_{n-1})$. It is readily verified that $F: H \rightarrow \mathbf{R}^n$ is a linear transformation. Given any vector $(y_0, y_1, \dots, y_{n-1})$ in $\mathbf{R}^n$, the previous theorem says that there is a unique signal $\{y_k\}$ in H such that $F\{y_k\} = (y_0, y_1, \dots, y_{n-1})$. It means that F is a one-to-one linear transformation of H onto $\mathbf{R}^n$; that is, F is an isomorphism. Thus $\dim H = \dim \mathbf{R}^n = n$.

Past Paper

Example 5 Find a basis for the set of all solutions to the difference equation

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \text{ for all } k$$

Solution

Generally it is difficult to identify directly that a set of signals spans the solution space, but this problem is being resolved by two above key theorems, from the last theorem the solution space is exactly three – dimensional and the Basis Theorem describes that a

linearly independent set of n vectors in an n – dimensional space is automatically a basis. So 1^k , $(-2)^k$ and 3^k form a basis for the solution space.

The standard way to describe the “general solution” of (10) is to exhibit a basis for the subspace of all solutions. Such a basis is usually called a **fundamental set of solutions** of (10). In practice, if we can find n linearly independent signals that satisfy (10), they will automatically span the n -dimensional solution space, as we saw in above example.

Non-homogeneous Equations The general solution of the non-homogeneous difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{for all } k \quad (11)$$

can be written as one particular solution of (11) plus an arbitrary linear combination of a fundamental set of solutions of the corresponding homogeneous equation (10). This fact is analogous to the result in lecture 7 about how the solution sets of $Ax = b$ and $Ax = 0$ are parallel. Both results have the same explanation: The mapping $x \rightarrow Ax$ is linear, and the mapping that transforms the signal $\{y_k\}$ into the signal $\{z_k\}$ in (11) is linear.

Example 6 Verify that the signal $y_k = k^2$ satisfies the difference equation

$$y_{k+2} - 4y_{k+1} + 3y_k = -4k \quad \text{for all } k \quad (12)$$

Then find a description of all solutions of this equation.

Solution Substitute k^2 for y_k in the left side of (12):

$$(k+2)^2 - 4(k+1)^2 + 3k^2 = (k^2 + 4k + 4) - 4(k^2 + 2k + 1) + 3k^2 = -4k$$

So k^2 is indeed a solution of (12). The next step is to solve the homogeneous equation

$$y_{k+2} - 4y_{k+1} + 3y_k = 0 \quad (13)$$

The auxiliary equation is

$$r^2 - 4r + 3 = r^2 - 3r - r + 3 = r(r-3) - 1(r-3) = (r-1)(r-3) = 0$$

The roots are $r = 1, 3$. So two solutions of the homogeneous difference equation are 1^k and 3^k . They are obviously not multiples of each other, so they are linearly independent signals. (The Casorati test could have been used, too.) By the last Theorem, the solution space is two – dimensional, so 3^k and 1^k form a basis for the set of solutions of (13). Translating that set by a particular solution of the non-homogeneous equation (12), we obtain the general solution of (12):

$$k^2 + c_1 1^k + c_2 3^k, \quad \text{or} \quad k^2 + c_1 + c_2 3^k$$

Reduction to Systems of First-Order Equations A modern way to study a homogeneous n th-order linear difference equation is to replace it by an equivalent system of first order difference equations, written in the form $x_{k+1} = Ax_k$ for $k = 0, 1, 2, \dots$ Where the vectors x_k are in $\mathbf{R}^n$ and A is an $n \times n$ matrix.

Example 7 Write the following difference equation as a first order system:

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{for all } k$$

Solution For each k , set $x_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$

The difference equation says that $y_{k+3} = 2y_{k+2} + 5y_{k+1} - 6y_k$, so

$$\mathbf{x}_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 + y_{k+1} + 0 \\ 0 + 0 + y_{k+2} \\ -6y_k + 5y_{k+1} + 2y_{k+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

That is, $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$ for all k , where $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$

In general, the equation $y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = 0$ for all k can be written as $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$ for all k , where

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{bmatrix}$$

Example 8 It can be shown that the signals 2^k , $3^k \sin \frac{k\pi}{2}$, and $3^k \cos \frac{k\pi}{2}$ are solutions of $y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$. Show that these signals form a basis for the set of all solutions of the difference equation.

Solution Examine the Casorati matrix:

$$C(k) = \begin{bmatrix} 2^k & 3^k \sin \frac{k\pi}{2} & 3^k \cos \frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} \sin \frac{(k+1)\pi}{2} & 3^{k+1} \cos \frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} \sin \frac{(k+2)\pi}{2} & 3^{k+2} \cos \frac{(k+2)\pi}{2} \end{bmatrix}$$

Set $k = 0$ and row reduce the matrix to verify that it has three pivot positions and hence is invertible:

$$\begin{aligned} C(0) &= \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 4 & 0 & -9 \end{bmatrix} \mathbf{R}_2 - 2\mathbf{R}_1 \\ &\sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix} \mathbf{R}_3 - 4\mathbf{R}_1 \end{aligned}$$

The Casorati matrix is invertible at $k = 0$, so signals are linearly independent. Since there are three signals, and the solution space $\mathbf{H}$ of the difference equation has three-dimensions (Theorem 2), the signals form a basis for $\mathbf{H}$, by the Basis Theorem.

Exercises

Verify that the signals in exercises 1 and 2 are solution of the accompanying difference equation.

1. $2^k, (-4)^k; y_{k+2} + 2y_{k+1} - 8y_k = 0$

2. $3^k, (-3)^k; y_{k+2} - 9y_k = 0$

Show that the signals in exercises 3 to 6 form a basis for the solution set of accompanying difference equation.

3. $2^k, (-4)^k; y_{k+2} + 2y_{k+1} - 8y_k = 0$

4. $3^k, (-3)^k; y_{k+2} - 9y_k = 0$

5. $(-3)^k, k(-3)^k; y_{k+2} + 6y_{k+1} + 9y_k = 0$

6. $5^k \cos \frac{k\pi}{2}, 5^k \sin \frac{k\pi}{2}; y_{k+2} + 25y_k = 0$

In exercises 7 to 10, assume that the signals listed are solutions of the given difference equation. Determine if the signals form a basis for the solution space of the equation.

7. $1^k, 3^k \cos \frac{k\pi}{2}, 3^k \sin \frac{k\pi}{2}; y_{k+3} - y_{k+2} + 9y_{k+1} - 9y_k = 0$

8. $(-1)^k, k(-1)^k, 5^k; y_{k+3} - 3y_{k+2} - 9y_{k+1} - 5y_k = 0$

9. $(-1)^k, 3^k; y_{k+3} + y_{k+2} - 9y_{k+1} - 9y_k = 0$

10. $1^k, (-1)^k; y_{k+4} - 2y_{k+2} + y_k = 0$

In exercises 11 and 12, find a basis for the solution space of the difference equation.

11. $y_{k+2} - 7y_{k+1} + 12y_k = 0$

12. $16y_{k+2} + 8y_{k+1} - 3y_k = 0$

In exercises 13 and 14, show that the given signal is a solution of the difference equation. Then find the general solution of that difference equation.

13. $y_k = k^2; y_{k+2} + 3y_{k+1} - 4y_k = 10k + 7$

14. $y_k = 2 - 2k; y_{k+2} - (9/2)y_{k+1} + 2y_k = 3k + 2$

Write the difference equations in exercises 15 and 16 as first order systems, $\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k$, for all k .

15. $y_{k+4} - 6y_{k+3} + 8y_{k+2} + 6y_{k+1} - 9y_k = 0$

16. $y_{k+3} - (3/4)y_{k+2} + (1/16)y_k = 0$

Lecture 28

Eigenvalues and Eigenvectors

In this lecture we will discuss linear equations of the form $Ax = x$ and, more generally, equations of the form $Ax = \lambda x$, where λ is a scalar. Such equations arise in a wide variety of important applications and will be a recurring theme in the rest of this course.

Fixed Points

A fixed point of an $n \times n$ matrix A is a vector x in $\mathbf{R}^n$ such that $Ax = x$. Every square matrix A has at least one fixed point, namely $x = 0$. We call this the trivial fixed point of A .

The general procedure for finding the fixed points of a matrix A is to rewrite the equation $Ax = x$ as $Ax = Ix$ or, alternatively, as

$$(I - A)x = 0 \quad (1)$$

Since this can be viewed as a homogeneous linear system of n equations in n unknowns with coefficient matrix $I - A$, we see that the set of fixed points of an $n \times n$ matrix is a subspace of $\mathbf{R}^n$ that can be obtained by solving (1).

The following theorem will be useful for ascertaining the nontrivial fixed points of a matrix.

Theorem 1

If A is an $n \times n$ matrix, then the following statements are equivalent.

- (a) A has nontrivial fixed points.
- (b) $I - A$ is singular.
- (c) $\det(I - A) = 0$.

Example 1

In each part, determine whether the matrix has nontrivial fixed points; and, if so, graph the subspace of fixed points in an xy -coordinate system.

$$(a) \quad A = \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix} \quad (b) \quad A = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$$

Solution

(a) The matrix has only the trivial fixed point since.

$$(I - A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -6 \\ -1 & -1 \end{pmatrix}$$

$$\det(I - A) = \det \begin{pmatrix} -2 & -6 \\ -1 & -1 \end{pmatrix} = (-1)(-2) - (-1)(-6) = -4 \neq 0$$

(b) The matrix has nontrivial fixed points since

$$(I - A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}$$

$$\det(I - A) = \det \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} = 0$$

The fixed points $\mathbf{x} = (x, y)$ are the solutions of the linear system $(I - A)\mathbf{x} = \mathbf{0}$, which we can express in component form as

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

A general solution of this system is

$$x = 2t, y = t \quad (2)$$

which are parametric equations of the line $y = \frac{1}{2}x$. It follows from the corresponding vector form of this line that the fixed points are

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2t \\ t \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad (3)$$

As a check, $\mathbf{Ax} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2t \\ t \end{bmatrix} = \begin{bmatrix} 2t \\ t \end{bmatrix} = \mathbf{x}$

so every vector of form (3) is a fixed point of $\mathbf{A}$.

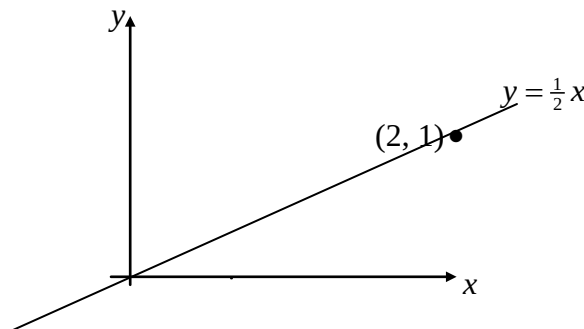


Figure 1

Eigenvalues and Eigenvectors

In a fixed point problem one looks for nonzero vectors that satisfy the equation $\mathbf{Ax} = \mathbf{x}$. One might also consider whether there are nonzero vectors that satisfy such equations as

$$\mathbf{Ax} = 2\mathbf{x}, \mathbf{Ax} = -3\mathbf{x}, \mathbf{Ax} = \sqrt{2}\mathbf{x}$$

or, more generally, equations of the form $\mathbf{Ax} = \lambda\mathbf{x}$ in which λ is a scalar.

Definition If $\mathbf{A}$ is an $n \times n$ matrix, then a scalar λ is called an **eigenvalue** of $\mathbf{A}$ if there is a nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$. If λ is an eigenvalue of $\mathbf{A}$, then every nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$ is called an **eigenvector** of $\mathbf{A}$ corresponding to λ .

Example 2

Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $u = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Are u and v eigenvectors of A ?

Solution

$$Au = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4u$$

$$Av = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Thus u is an eigenvector corresponding to an eigenvalue -4 , but v is not an eigenvector of A , because Av is not a multiple of v .

Example 3

Show that 7 is an eigenvalue of $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, find the corresponding eigenvectors.

Solution

The scalar 7 is an eigenvalue of A if and only if the equation

$$Ax = 7x \tag{A}$$

has a nontrivial solution. But (A) is equivalent to $Ax - 7x = 0$, or

$$(A - 7I)x = 0 \tag{B}$$

To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

The columns of $A - 7I$ are obviously linearly dependent, so (B) has nontrivial solutions. Thus 7 is an eigenvalue of A . To find the corresponding eigenvectors, use row operations:

$$\begin{aligned} & \begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \\ & \sim \begin{bmatrix} 1 & -1 & 0 \\ 5 & -5 & 0 \end{bmatrix} (-1R_1 - R_2) \\ & \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} (R_2 - 5R_1) \end{aligned}$$

The general solution has the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

The equivalence of equations (A) and (B) obviously holds for any λ in place of $\lambda = 7$. Thus λ is an eigenvalue of A if and only if the equation

$$(A - \lambda I)x = 0 \tag{C}$$

has a nontrivial solution.

Eigen space

The set of all solutions of $(A - \lambda I)x = 0$ is just the null space of the matrix $A - \lambda I$. So this set is a subspace of $\mathbf{R}^n$ and is called the eigenspace of A corresponding to λ . The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ .

Example 3 shows that for matrix A in Example 2, the eigenspace corresponding to $\lambda = 7$ consists of all multiples of $(1, 1)$, which is the line through $(1, 1)$ and the origin. From Example 2, one can check that the eigenspace corresponding to $\lambda = -4$ is the line through $(6, -5)$. These eigenspaces are shown in Fig. 1, along with eigenvectors $(1, 1)$ and $(3/2, -5/4)$ and the geometric action of the transformation $x \rightarrow Ax$ on each eigenspace.

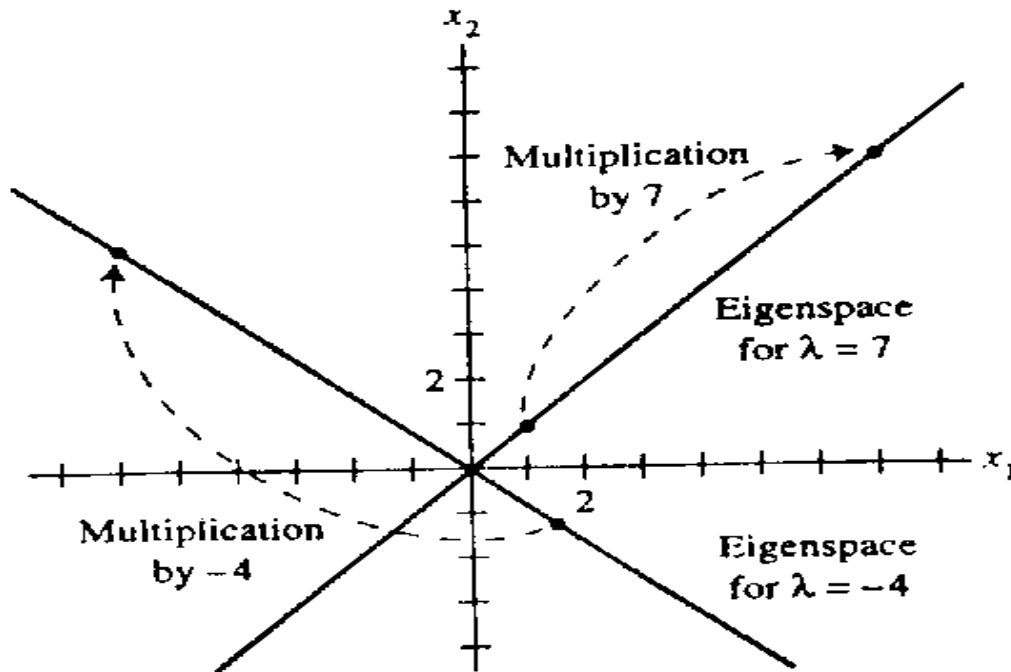


FIGURE 2 Eigenspaces for $\lambda = -4$ and $\lambda = 7$.

Example 4 Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

Find a basis for the corresponding eigenspace where eigen value of matrix is 2.

Solution Form $\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$ and row reduce the

augmented matrix for $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} R_2 - R_1$$

$$\sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_3 - R_1$$

At this point we are confident that 2 is indeed an eigenvalue of $\mathbf{A}$ because the equation $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ has free variables. The general solution is

$$2x_1 - x_2 + 6x_3 = 0 \dots\dots(a)$$

Let $x_2 = t$, $x_3 = s$ then

$$2x_1 = t - 6s$$

$$x_1 = \left(\frac{1}{2}\right)t - 3s$$

then

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t/2 - 3s \\ t \\ s \end{bmatrix} = \begin{bmatrix} t/2 \\ t \\ 0 \end{bmatrix} + \begin{bmatrix} -3s \\ 0 \\ s \end{bmatrix} = t \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

By back substitution the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, x_2 \text{ and } x_3 \text{ free}$$

The eigenspace, shown in Fig. 2, is a two – dimensional subspace of $\mathbf{R}^3$. A basis is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ is a basis.}$$

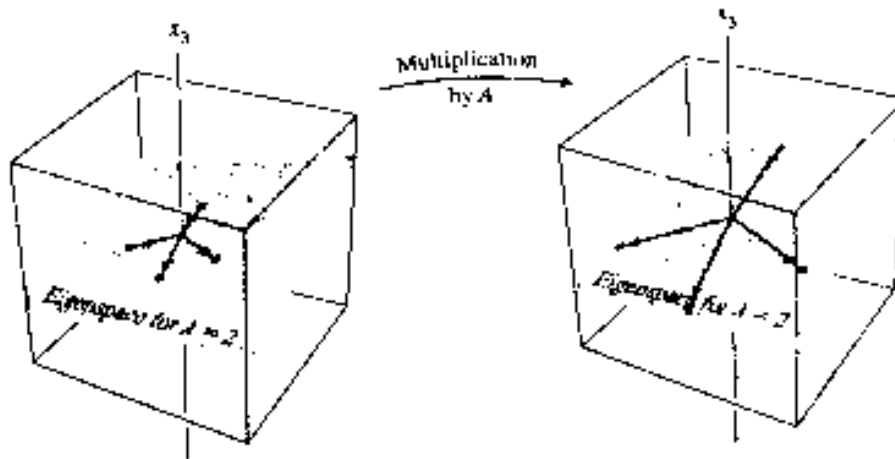


FIGURE 3 A acts as a dilation on the eigenspace.

The most direct way of finding the eigenvalues of an $n \times n$ matrix A is to rewrite the equation $Ax = \lambda x$ as $Ax = \lambda Ix$, or equivalently, as

$$(\lambda I - A)x = 0 \quad (4)$$

and then try to determine those values of λ , if any, for which this system has nontrivial solutions. Since (4) have nontrivial solutions if and only if the coefficient matrix $\lambda I - A$ is singular, we see that the eigenvalues of A are the solutions of the equation

$$\det(\lambda I - A) = 0 \quad (5)$$

Equation (5) is known as characteristic equation. Also, if λ is an eigenvalue of A , then equation (4) has a nonzero solution space, which we call the eigenspace of A corresponding to λ . It is the nonzero vectors in the eigenspace of A corresponding to λ that are the eigenvectors of A corresponding to λ .

The above discussion is summarized by the following theorem.

Theorem If A is an $n \times n$ matrix and λ is a scalar, then the following statements are equivalent.

- (i) λ is an eigenvalue of A .
- (ii) λ is a solution of the equation $\det(\lambda I - A) = 0$.
- (iii) The linear system $(\lambda I - A)x = 0$ has nontrivial solutions.

Eigenvalues of Triangular Matrices If A is an $n \times n$ triangular matrix with diagonal entries $a_{11}, a_{22}, \dots, a_{nn}$, then $\lambda I - A$ is a triangular matrix with diagonal entries $\lambda - a_{11}, \lambda - a_{22}, \dots, \lambda - a_{nn}$. Thus, the characteristic polynomial of A is

$$\det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$$

which implies that the eigenvalues of A are

$$\lambda_1 = a_{11}, \lambda_2 = a_{22}, \dots, \lambda_n = a_{nn}$$

Thus, we have the following theorem.

Theorem If A is a triangular matrix (upper triangular, lower triangular, or diagonal) then the eigenvalues of A are the entries on the main diagonal of A .

Example 5 (Eigenvalues of Triangular Matrices)

By inspection, the characteristic polynomial of the matrix $A = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ -1 & -\frac{2}{3} & 0 & 0 \\ 7 & \frac{5}{8} & 6 & 0 \\ \frac{4}{9} & -4 & 3 & 6 \end{bmatrix}$ is

$p(\lambda) = (\lambda - \frac{1}{2})(\lambda + \frac{2}{3})(\lambda - 6)^2$. So the distinct eigenvalues of A are $\lambda = \frac{1}{2}$, $\lambda = -\frac{2}{3}$, and $\lambda = 6$.

Eigenvalues of Powers of a Matrix Once the eigenvalues and eigenvectors of a matrix A are found, it is a simple matter to find the eigenvalues and eigenvectors of any positive integer power of A . For example, if λ is an eigenvalue of A and x is a corresponding eigenvector, then $A^2x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda(\lambda x) = \lambda^2x$, which shows that λ^2 is an eigenvalue of A^2 and x is a corresponding eigenvector. In general we have the following result.

Theorem If λ is an eigenvalue of a matrix A and x is a corresponding eigenvector, and if k is any positive integer, then λ^k is an eigenvalue of A^k and x is a corresponding eigenvector.
Some problems that use this theorem are given in the exercises.

A Unifying Theorem Since λ is an eigenvalue of a square matrix A if and only if there is a nonzero vector x such that $Ax = \lambda x$, it follows that $\lambda = 0$ is an eigenvalue of A if and only if there is a nonzero vector x such that $Ax = 0$. However, this is true if and only if $\det(A) = 0$, so we list the following

Theorem If A is an $n \times n$ matrix, then the following statements are equivalent.

- (a) The reduced row echelon form of A is I_n .
- (b) A is expressible as a product of elementary matrices.
- (c) A is invertible.
- (d) $Ax = 0$ has only the trivial solution.
- (e) $Ax = b$ is consistent for every vector b in R^n .
- (f) $Ax = b$ has exactly one solution for every vector b in R^n .
- (g) The column vectors of A are linearly independent.
- (h) The row vectors of A are linearly independent.
- (i) $\det(A) \neq 0$.
- (j) $\lambda = 0$ is not an eigenvalue of A .

Example 6

(1) Is 5 an eigenvalue of $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?

(2) If $\mathbf{x}$ is an eigenvector for A corresponding to λ , what is $A^3\mathbf{x}$?

Solution

(1) The number 5 is an eigenvalue of A if and only if the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution. Form

$$A - 5I = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

and row reduce the augmented matrix:

$$\begin{aligned} & \begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \\ & \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \quad R_2 - 3R_1 \\ & \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} \quad R_3 - 2R_1 \\ & \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix} \quad R_3 - 2R_2 \end{aligned}$$

At this point it is clear that the homogeneous system has no free variables. Thus $A - 5I$ is an invertible matrix, which means that 5 is not an eigenvalue of A .

(2). If $\mathbf{x}$ is an eigenvector for A corresponding to λ , then $A\mathbf{x} = \lambda\mathbf{x}$ and so

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Again $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. The general pattern, $A^k\mathbf{x} = \lambda^k\mathbf{x}$, is proved by induction.

Exercises

1. Is $\lambda = 2$ an eigenvalue of $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$?

2. Is $\begin{bmatrix} -1+\sqrt{2} \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$? If so, find the eigenvalue.

3. Is $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}$? If so, find the eigenvalue.

4. Is $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 3 & 7 \\ 5 & 6 & 5 \end{bmatrix}$? If so, find the eigenvalue.

5. Is $\lambda = 4$ an eigenvalue of $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? If so, find one corresponding eigenvector.

6. Is $\lambda = 3$ an eigenvalue of $\begin{bmatrix} 1 & 2 & 2 \\ 3 & -2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$? If so, find one corresponding eigenvector.

In exercises 7 to 12, find a basis for the eigenspace corresponding to each listed eigenvalue.

7. $\mathbf{A} = \begin{bmatrix} 4 & -2 \\ -3 & 9 \end{bmatrix}, \lambda = 10$

8. $\mathbf{A} = \begin{bmatrix} 7 & 4 \\ -3 & -1 \end{bmatrix}, \lambda = 5$

9. $\mathbf{A} = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & -1 \\ -2 & 0 & 1 \end{bmatrix}, \lambda = 2, 3$

10. $\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -3 & 0 \\ 4 & -13 & 1 \end{bmatrix}, \lambda = 2$

11. $\mathbf{A} = \begin{bmatrix} 4 & 2 & 3 \\ -1 & 1 & -3 \\ 2 & 4 & 9 \end{bmatrix}, \lambda = 3$

12. $\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \lambda = 4$

Find the eigenvalues of the matrices in Exercises 13 and 14.

13. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 5 \\ 0 & 0 & -1 \end{bmatrix}$

14. $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

15. For $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, find one eigenvalue, with no calculation. Justify your answer.

16. Without calculation, find one eigenvalue and two linearly independent vectors of

$\mathbf{A} = \begin{bmatrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & 5 & 5 \end{bmatrix}$. Justify your answer.

Lecture 29

The Characteristic Equation

The Characteristic equation contains useful information about the eigenvalues of a square matrix $\mathbf{A}$. It is defined as

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

Where λ is the eigenvalue and $\mathbf{I}$ is the identity matrix. We will solve the Characteristic equation (also called the characteristic polynomial) to work out the eigenvalues of the given square matrix $\mathbf{A}$.

Example 1 Find the eigenvalues of $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

Solution In order to find the eigenvalues of the given matrix, we must solve the matrix equation

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

for the scalar λ such that it has a nontrivial solution (since the matrix is non singular). By the Invertible Matrix Theorem, this problem is equivalent to finding all λ such that the matrix $\mathbf{A} - \lambda \mathbf{I}$ is not invertible, where

$$\mathbf{A} - \lambda \mathbf{I} = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix}.$$

By definition, this matrix $\mathbf{A} - \lambda \mathbf{I}$ fails to be invertible precisely when its determinant is zero. Thus, the eigenvalues of $\mathbf{A}$ are the solutions of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \det \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix} = 0.$$

Recall that $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$

$$\begin{aligned} \text{So } \det(\mathbf{A} - \lambda \mathbf{I}) &= (2 - \lambda)(-6 - \lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21 \\ \lambda^2 + 4\lambda - 21 &= 0, \\ (\lambda - 3)(\lambda + 7) &= 0, \end{aligned}$$

so the eigenvalues of $\mathbf{A}$ are 3 and -7 .

Example 2 Compute $\det \mathbf{A}$ for $\mathbf{A} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$

Solution

Firstly, we will reduce the given matrix in echelon form by applying elementary row operations

$$\text{by } R_2 - 2R_1$$

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix},$$

$$\text{by } R_2 \leftrightarrow R_3$$

$$\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix},$$

$$\text{by } R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

which is an upper triangular matrix. Therefore,

$$\det A = (1)(-2)(-1) \\ = 2.$$

Theorem 1 Properties of Determinants

Let A and B be two matrices of order n then

(a) A is invertible if and only if $\det A \neq 0$.

(b) $\det AB = (\det A)(\det B)$.

(c) $\det A^T = \det A$.

(d) If A is triangular, then $\det A$ is the product of the entries on the main diagonal of A .

(e) A row replacement operation on A does not change the determinant.

(f) A row interchange changes the sign of the determinant.

(g) A row scaling also scales the determinant by the same scalar factor.

MCS

Note These Properties will be helpful in using the characteristic equation to find eigenvalues of a matrix A .

Example 3 (a) Find the eigenvalues and corresponding eigenvectors of the matrix

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

(b) Graph the eigenspaces of A in an xy -coordinate system.

Solution (a) The eigenvalues will be worked out by solving the characteristic equation of A . Since

$$\lambda \mathbf{I} - \mathbf{A} = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{bmatrix}.$$

The characteristic equation $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$ becomes

$$\begin{vmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{vmatrix} = 0.$$

Expanding and simplifying the determinant, it yields

$$\lambda^2 - 3\lambda - 10 = 0,$$

or

$$(\lambda + 2)(\lambda - 5) = 0. \quad (1)$$

Thus, the eigenvalues of $\mathbf{A}$ are $\lambda = -2$ and $\lambda = 5$.

Now, to work out the eigenspaces corresponding to these eigenvalues, we will solve the system

$$\begin{bmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2)$$

for $\lambda = -2$ and $\lambda = 5$. Here are the computations for the two cases.

(i) Case $\lambda = -2$

In this case Eq. (2) becomes

$$\begin{bmatrix} -3 & -3 \\ -4 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which can be written as

$$-3x - 3y = 0,$$

$$-4x - 4y = 0 \Rightarrow x = -y.$$

In parametric form,

$$x = -t, y = t. \quad (3)$$

Thus, the eigenvectors corresponding to $\lambda = -2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -t \\ t \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \quad (4)$$

It can be verified as

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} -t \\ t \end{bmatrix} = \begin{bmatrix} 2t \\ -2t \end{bmatrix} = -2 \begin{bmatrix} -t \\ t \end{bmatrix} = -2\mathbf{x}$$

Thus,

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

(ii) Case $\lambda = 5$

In this case Eq. (2) becomes

$$\begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which can be written as

$$4x - 3y = 0$$

$$-4x + 3y = 0 \Rightarrow x = \frac{3}{4}y.$$

In parametric form,

$$x = \frac{3}{4}t, y = t. \quad (5)$$

Thus, the eigenvectors corresponding to $\lambda = 5$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{3}{4} \\ 1 \end{bmatrix}. \quad (6)$$

It can be verified as

$$\mathbf{Ax} = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = \begin{bmatrix} \frac{15}{4}t \\ 5t \end{bmatrix} = 5 \begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = 5\mathbf{x}.$$

(b) The eigenspaces corresponding to $\lambda = -2$ and $\lambda = 5$ can be sketched from the parametric equations (3) and (5) as shown in figure 1(a).

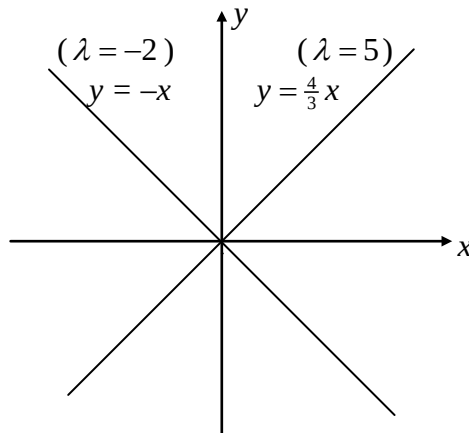


Figure 1(a)

It can also be drawn using the vector equations (4) and (6) as shown in Figure 1(b). When an eigenvector $\mathbf{x}$ in the eigenspace for $\lambda = 5$ is multiplied by $\mathbf{A}$, the resulting vector has the same direction as $\mathbf{x}$ but the length is increased by a factor of 5 and when an eigenvector $\mathbf{x}$ in the eigenspace for $\lambda = -2$ is multiplied by $\mathbf{A}$, the resulting vector is oppositely directed to $\mathbf{x}$ and the length is increased by a factor of 2. In both cases, multiplying an eigenvector by $\mathbf{A}$ produces a vector in the same eigenspace.

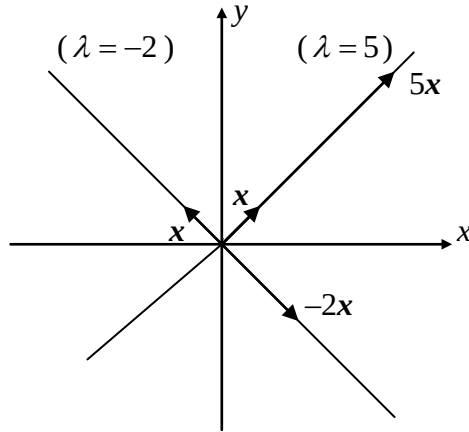


Figure 1(b)

Eigenvalues of an $n \times n$ matrix

Eigen values of an $n \times n$ matrix can be found in the similar fashion. However, for the higher values of n , it is more convenient to work them out using various available mathematical software. Here is an example for a 3×3 matrix.

Example 4 Find the eigen values of the matrix $A = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -4 & -17 & 8 \end{bmatrix}$

Solution

$$\begin{aligned} \det(\lambda I - A) &= \det \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} - \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -4 & -17 & 8 \end{bmatrix} \\ &= \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & -1 \\ 4 & 17 & \lambda - 8 \end{vmatrix} \\ &= \lambda^3 - 8\lambda^2 + 17\lambda - 4, \end{aligned} \quad (7)$$

which yields the characteristic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0 \quad (8)$$

To solve this equation, firstly, we will look for integer solutions. This can be done by using the fact that if a polynomial equation has integer coefficients, then its integer solutions, if any, must be divisors of the constant term of the given polynomial. Thus, the only possible integer solutions of Eq.(8) are the divisors of -4 , namely $\pm 1, \pm 2$, and ± 4 . Substituting these values successively into Eq. (8) yields that $\lambda = 4$ is an integer solution. This implies that $\lambda - 4$ is a factor of Eq.(7), Thus, dividing the polynomial by $\lambda - 4$ and rewriting Eq.(8), we get

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0.$$

Now, the remaining solutions of the characteristic equation satisfy the quadratic equation

$$\lambda^2 - 4\lambda + 1 = 0.$$

Solving the above equation by the quadratic formula, we get the eigenvalues of A as

$$\lambda = 4, \quad \lambda = 2 + \sqrt{3}, \quad \lambda = 2 - \sqrt{3}$$

Example 5 Find the characteristic equation of $A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Solution Clearly, the given matrix is an upper triangular matrix. Forming $A - \lambda I$, we get

$$\det(A - \lambda I) = \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

Now using the fact that determinant of a triangular matrix is equal to product of its diagonal elements, the characteristic equation becomes

$$(5 - \lambda)^2(3 - \lambda)(1 - \lambda) = 0.$$

Expanding the product, we can also write it as

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0.$$

Here, the eigenvalue 5 is said to have multiplicity 2 because $(\lambda - 5)$ occurs two times as a factor of the characteristic polynomial. In general, the **(algebraic)** multiplicity of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

Note

From the above mentioned examples, it can be easily observed that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n called the characteristic polynomial of A .

Example 6 The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.

Solution

In order to find the eigenvalues, we will factorize the polynomial as

$$\begin{aligned} & \lambda^6 - 4\lambda^5 - 12\lambda^4 \\ &= \lambda^4(\lambda^2 - 4\lambda - 12) \\ &= \lambda^4(\lambda - 6)(\lambda + 2) \end{aligned}$$

The eigenvalues are 0 (multiplicity 4), 6 (multiplicity 1) and -2 (multiplicity 1). We could also list the eigenvalues in Example 6 as 0, 0, 0, 0, 6 and -2 , so that the eigenvalues are repeated according to their multiplicities

Activity

Work out the eigenvalues and eigenvectors for the following square matrix.

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}.$$

Similarity

Let A and B be two $n \times n$ matrices, A is said to be similar to B if there exist an invertible matrix P such that

$$P^{-1}AP = B,$$

or equivalently,

$$A = PBP^{-1}.$$

Replacing Q by P^{-1} , we have

$$Q^{-1}BQ = A.$$

So B is also similar to A . Thus, we can say that A and B are similar.

Similarity transformation

The act of changing A into $P^{-1}AP$ is called a similarity transformation.

The following theorem illustrates use of the characteristic polynomial and it provides the foundation for several iterative methods that approximate eigenvalues.

Theorem 2

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof If $B = P^{-1}AP$, then

$$B - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Using the multiplicative property (b) of Theorem 1, we compute

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}).\det(A - \lambda I).\det(P) \end{aligned} \quad (A)$$

Since

$$\begin{aligned} \det(P^{-1}).\det(P) &= \det(P^{-1}P) \\ &= \det I \\ &= 1, \end{aligned}$$

Eq. (A) implies that

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

Hence, both the matrices have the same characteristic polynomials and therefore, same eigenvalues.

Note It must be clear that Similarity and row equivalence are two different concepts. (If A is row equivalent to B , then $B = EA$ for some invertible matrix E .) Row operations on a matrix usually change its eigenvalues.

Application to Dynamical Systems

Dynamical system is the one which evolves with the passage of time. Eigenvalues and eigenvectors play a vital role in the evaluation of a dynamical system. Let's consider an example of a dynamical system.

Example 7 Let $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analyze the long term behavior of the dynamical

system defined by $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \dots$), with $\mathbf{x}_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$.

Solution The first step is to find the eigenvalues of A and a basis for each eigenspace. The characteristic equation for A is

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ 0 &= \det \begin{bmatrix} 0.95 - \lambda & 0.03 \\ 0.05 & 0.97 - \lambda \end{bmatrix} = (.95 - \lambda)(.97 - \lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

By the quadratic formula

$$\lambda = \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} = \frac{1.92 \pm .08}{2} = 1 \text{ or } .92$$

Firstly, the eigenvectors will be found as given below.

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

For $\lambda = 1$

$$\begin{aligned} \left[\begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0, \\ \begin{pmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0, \end{aligned}$$

which can be written as

$$-0.05x_1 + 0.03x_2 = 0$$

$$0.05x_1 - 0.03x_2 = 0 \Rightarrow x_1 = \frac{0.03}{0.05}x_2 \text{ or } x_1 = \frac{3}{5}x_2.$$

In parametric form, it becomes

$$x_1 = \frac{3}{5}t \text{ and } x_2 = t.$$

For $\lambda = 0.92$

$$\left[\begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} - \begin{pmatrix} 0.92 & 0 \\ 0 & 0.92 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

$$\begin{pmatrix} 0.03 & 0.03 \\ 0.05 & 0.05 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

It can be written as

$$0.03x_1 + 0.03x_2 = 0$$

$$0.05x_1 + 0.05x_2 = 0 \Rightarrow x_1 = -x_2$$

In parametric form, it becomes

$$x_1 = t \text{ and } x_2 = -t$$

Thus, the eigenvectors corresponding to $\lambda = 1$ and $\lambda = .92$ are multiples of

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ respectively.}$$

The next step is to write the given $\mathbf{x}_0$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$. This can be done because $\{\mathbf{v}_1, \mathbf{v}_2\}$ is obviously a basis for $\mathbf{R}^2$. So there exists weights c_1 and c_2 such that

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (1)$$

$$\text{In fact, } \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = [\mathbf{v}_1 \quad \mathbf{v}_2]^{-1} \mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .60 \\ .40 \end{bmatrix}$$

Here,

$$\begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} = \frac{1}{\begin{vmatrix} 3 & 1 \\ 5 & -1 \end{vmatrix}} \text{Adj} \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix} = -\frac{1}{8} \begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix}$$

Therefore,

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{-8} \begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \quad (2)$$

Because $\mathbf{v}_1$ and $\mathbf{v}_2$ in Eq.(1) are eigenvectors of $\mathbf{A}$, with $\mathbf{A}\mathbf{v}_1 = \mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2 = (.92)\mathbf{v}_2$, $\mathbf{x}_k$ can be computed as

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{A}\mathbf{x}_0 = c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2 \quad (\text{Using linearity of } \mathbf{x} \rightarrow \mathbf{A}\mathbf{x}) \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 \quad (\mathbf{v}_1 \text{ and } \mathbf{v}_2 \text{ are eigenvectors.}) \end{aligned}$$

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = c_1\mathbf{A}\mathbf{v}_1 + c_2(.92)\mathbf{A}\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2.$$

Continuing in the same way, we get the general equation as

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \dots).$$

Using c_1 and c_2 from Eq.(2),

$$\mathbf{x}_k = .125 \begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \dots) \quad (3)$$

This explicit formula for $\mathbf{x}_k$ gives the solution of the difference equation $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$.

As $k \rightarrow \infty$, $(.92)^k$ tends to zero and $\mathbf{x}_k$ tends to $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$.

Example 8 Find the characteristic equation and eigenvalues of $\mathbf{A} = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

Solution The characteristic equation is

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda \mathbf{I}) = \det \begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix} \\ &= (1-\lambda)(2-\lambda) - (-4)(4), \\ &= \lambda^2 - 3\lambda + 18, \end{aligned}$$

which is a quadratic equation whose roots are given as

$$\begin{aligned} \lambda &= \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} \\ &= \frac{3 \pm \sqrt{-63}}{2} \end{aligned}$$

Thus, we see that the characteristic equation has no real roots, so $\mathbf{A}$ has no real eigenvalues. $\mathbf{A}$ is acting on the real vector space $\mathbf{R}^2$ and there is no non-zero vector $\mathbf{v}$ in $\mathbf{R}^2$ such that $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$ for some scalar λ .

Exercises

Find the characteristic polynomial and the eigenvalues of matrices in exercises 1 to 12.

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

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

3.
$$\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$$

4.
$$\begin{bmatrix} 3 & -4 \\ 4 & 8 \end{bmatrix}$$

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

6.
$$\begin{bmatrix} 7 & -2 \\ 2 & 3 \end{bmatrix}$$

7.
$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 6 & 0 \end{bmatrix}$$

8.
$$\begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$$

9.
$$\begin{bmatrix} 4 & 0 & 0 \\ 5 & 3 & 2 \\ -2 & 0 & 2 \end{bmatrix}$$

10.
$$\begin{bmatrix} -1 & 0 & 1 \\ -3 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

11.
$$\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$$

12.
$$\begin{bmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{bmatrix}$$

For the matrices in exercises 13 to 15, list the eigenvalues, repeated according to their multiplicities.

13.
$$\begin{bmatrix} 4 & -7 & 0 & 2 \\ 0 & 3 & -4 & 6 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

14.
$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 8 & -4 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 1 & -5 & 2 & 1 \end{bmatrix}$$

15.
$$\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$$

16. It can be shown that the algebraic multiplicity of an eigenvalue λ is always greater than or equal to the dimension of the eigenspace corresponding to λ . Find h in the matrix $\mathbf{A}$ below such that the eigenspace for $\lambda=5$ is two-dimensional:

$$\mathbf{A} = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & h & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Lecture 30

Diagonalization

Diagonalization is a process of transforming a vector A to the form $A = \mathbf{PDP}^{-1}$ for some invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$. In this lecture, the factorization enables us to compute A^k quickly for large values of k which is a fundamental idea in several applications of linear algebra. Later, the factorization will be used to analyze (and decouple) dynamical systems.

The “ D ” in the factorization stands for diagonal. Powers of such a D are trivial to compute.

Example 1 If $\mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, then $\mathbf{D}^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$ and

$$\mathbf{D}^3 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

In general, $\mathbf{D}^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}$ for $k \geq 1$

The next example shows that if $A = \mathbf{PDP}^{-1}$ for some invertible $\mathbf{P}$ and diagonal $\mathbf{D}$, then it is quite easy to compute A^k .

Example 2 Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for A^k , given that $A = \mathbf{PDP}^{-1}$, where

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \text{ and } \mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Solution The standard formula for the inverse of a 2×2 matrix yields

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

By associative property of matrix multiplication,

$$A^2 = (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1}) = \mathbf{PD} \underbrace{(\mathbf{P}^{-1}\mathbf{P})}_I \mathbf{DP}^{-1} = \mathbf{PDIDP}^{-1} = \mathbf{PDDP}^{-1}$$

where I is the identity matrix.

$$= \mathbf{PD}^2 \mathbf{P}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Again, $A^3 = (\mathbf{PDP}^{-1})A^2 = (\mathbf{PD} \underbrace{\mathbf{P}^{-1}\mathbf{P}}_I) \mathbf{D}^2 \mathbf{P}^{-1} = \mathbf{PDD}^2 \mathbf{P}^{-1} = \mathbf{PD}^3 \mathbf{P}^{-1}$

Thus, in general, for $k \geq 1$, $A^k = \mathbf{PD}^k \mathbf{P}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$,

$$\begin{aligned}
&= \begin{bmatrix} 5^k & 3^k \\ -5^k & -2 \cdot 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}, \\
&= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}.
\end{aligned}$$

Activity

Work out C^4 , given that $C = PDP^{-1}$ where

$$P = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Remarks

A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D . The next theorem gives a characterization of diagonalizable matrices and tells how to construct a suitable factorization.

Theorem 1 The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A . In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P .

In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbf{R}^n$. We call such a basis an eigenvector basis.

Proof First, observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ and if D is any diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$ then

$$AP = A[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n] = [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \dots \quad A\mathbf{v}_n], \quad (1)$$

$$\text{while} \quad PD = P \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = [\lambda_1 \mathbf{v}_1 \quad \lambda_2 \mathbf{v}_2 \quad \dots \quad \lambda_n \mathbf{v}_n] \quad (2)$$

Suppose now that A is diagonalizable and $A = PDP^{-1}$. Then right-multiplying this relation by P , we have $AP = PD$. In this case, (1) and (2) imply that

$$[A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \dots \quad A\mathbf{v}_n] = [\lambda_1 \mathbf{v}_1 \quad \lambda_2 \mathbf{v}_2 \quad \dots \quad \lambda_n \mathbf{v}_n] \quad (3)$$

Equating columns, we find that

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2, \dots, A\mathbf{v}_n = \lambda_n \mathbf{v}_n \quad (4)$$

Since P is invertible, its columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ must be linearly independent. Also, since these columns are nonzero, Eq.(4) shows that $\lambda_1, \dots, \lambda_n$ are eigenvalues and $\mathbf{v}_1, \dots, \mathbf{v}_n$ are

corresponding eigenvectors. This argument proves the “only if” parts of the first and second statements along with the third statement, of the theorem.

Finally, given any n eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ use them to construct the columns of $\mathbf{P}$ and use corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ to construct $\mathbf{D}$. By Eqs. (1) – (3). $\mathbf{AP} = \mathbf{PD}$. This is true without any condition on the eigenvectors. If, in fact, the eigenvectors are linearly independent, then $\mathbf{P}$ is invertible (by the Invertible Matrix Theorem), and $\mathbf{AP} = \mathbf{PD}$ implies that $\mathbf{A} = \mathbf{PDP}^{-1}$.

Diagonalizing Matrices

Example 3 Diagonalize the following matrix, if possible $\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$

Solution To diagonalize the given matrix, we need to find an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$ which can be done in following four steps.

Step 1 Find the eigenvalues of $\mathbf{A}$.

The characteristic equation becomes

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda \mathbf{I}) = -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$ (multiplicity 2)

Step 2 Find three linearly independent eigenvectors of $\mathbf{A}$. Since $\mathbf{A}$ is a 3×3 matrix and we have obtained three eigen values, we need three eigen vectors. This is the critical step. If it fails, then above Theorem says that $\mathbf{A}$ cannot be diagonalized. Now we will produce basis for these eigen values.

Basis vector for $\lambda = 1$:

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

$$\begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

After applying few row operations on the matrix $(\mathbf{A} - \lambda \mathbf{I})$, we get

$$\begin{bmatrix} 0 & 1 & 1 \\ 3 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which can be written as

$$x_2 + x_3 = 0$$

$$3x_1 + 3x_2 = 0$$

In parametric form, it becomes

$$x_1 = t, x_2 = -t, x_3 = t$$

Thus, the basis vector for $\lambda = 1$ is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

Basis vector for $\lambda = -2$

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

$$\begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which can be written as

$$3x_1 + 3x_2 + 3x_3 = 0$$

$$-3x_1 - 3x_2 - 3x_3 = 0$$

$$3x_1 + 3x_2 + 3x_3 = 0$$

In parametric form, it becomes

$$x_1 = -s - t, \quad x_2 = s, \quad x_3 = t$$

Now,

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} -s - t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix}, \\ &= s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \\ &= x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Thus, the basis for $\lambda = -2$ is $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

We can check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

Step 3 Check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

Construct $\mathbf{P}$ from the vectors in step 2. The order of the vectors is not important. Using

the order chosen in step 2, form $\mathbf{P} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Step 4 Form $\mathbf{D}$ from the corresponding eigen values. For this purpose, the order of the eigen values must match the order chosen for the columns of $\mathbf{P}$. Use the eigen value $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

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

Now, we need to check do $\mathbf{P}$ and $\mathbf{D}$ really work. To avoid computing $\mathbf{P}^{-1}$, simply verify that $\mathbf{AP} = \mathbf{PD}$. This is equivalent to $\mathbf{A} = \mathbf{PDP}^{-1}$ when $\mathbf{P}$ is invertible. We compute

$$\begin{aligned} \mathbf{AP} &= \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \\ \mathbf{PD} &= \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \end{aligned}$$

Example 4 Diagonalize the following matrix, if possible.

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution The characteristic equation of $\mathbf{A}$ turns out to be exactly the same as that in example 3 i.e.,

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda \mathbf{I}) \\ &= \begin{vmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} \\ &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigen values are $\lambda = 1$ and $\lambda = -2$ (multiplicity 2). However, when we look for eigen vectors, we find that each eigen space is only one – dimensional.

$$\begin{aligned} \text{Basis for } \lambda = 1: \quad \mathbf{v}_1 &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \\ \text{Basis for } \lambda = -2: \quad \mathbf{v}_2 &= \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

There are no other eigen values and every eigen vector of A is a multiple of either v_1 or v_2 . Hence it is impossible to form a basis of $\mathbf{R}^3$ using eigenvectors of A . By above Theorem, A is not diagonalizable.

Theorem 2 An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

The condition in Theorem 2 is sufficient but not necessary i.e., it is not necessary for an $n \times n$ matrix to have n distinct eigen values in order to be diagonalizable. Example 3 serves as a counter example of this case where the 3×3 matrix is diagonalizable even though it has only two distinct eigen values.

Example 5 Determine if the following matrix is diagonalizable.

$$A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

Solution In the light of Theorem 2, the answer is quite obvious. Since the matrix is triangular, its eigen values are obviously 5, 0, and -2 . Since A is a 3×3 matrix with three distinct eigen values, A is diagonalizable.

Matrices Whose Eigenvalues Are Not Distinct

If an $n \times n$ matrix A has n distinct eigen values, with corresponding eigen vectors $v_1, \dots, v_n$ and if $P = [v_1 \dots v_n]$, then P is automatically invertible because its columns are linearly independent, by Theorem 2 of lecture 28. When A is diagonalizable but has fewer than n distinct eigen values, it is still possible to build P in a way that makes P automatically invertible, as shown in the next theorem.

Theorem 3 Let A be an $n \times n$ matrix whose distinct eigen values are $\lambda_1, \dots, \lambda_p$.

- For $1 \leq k \leq p$, the dimension of the eigen space for λ_k is less than or equal to the multiplicity of the eigen value λ_k
- The matrix A is diagonalizable if and only if the sum of the dimensions of the distinct eigen spaces is equal to n , and this happens if and only if the dimension of the eigen space for each of λ_k equals the multiplicity of λ_k .
- If A is diagonalizable and B_k is basis for the eigen space corresponding to λ_k for each k , then the total collection of vectors in the sets $B_1, \dots, B_p$ form an eigenvector basis for $\mathbf{R}^n$.

Example 6 Diagonalize the following matrix, if possible.

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}$$

Solution Since $\mathbf{A}$ is triangular matrix, the eigenvalues are 5 and -3 , each with multiplicity 2. Using the method of lecture 28, we find a basis for each eigen space.

$$\text{Basis for } \lambda = 5 : \mathbf{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -3 : \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The set $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is linearly independent, by Theorem 3. So the matrix $\mathbf{P} = [\mathbf{v}_1 \dots \mathbf{v}_4]$ is invertible, and $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, where

$$\mathbf{P} = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{D} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

Example 7

- (1) Compute $\mathbf{A}^8$ where $\mathbf{A} = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$
- (2) Let $\mathbf{A} = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of $\mathbf{A}$. Use this information to diagonalize $\mathbf{A}$.
- (3) Let $\mathbf{A}$ be a 4 x 4 matrix with eigenvalues 5, 3, and -2 , and suppose that you know the eigenspace for $\lambda=3$ is two-dimensional. Do you have enough information to determine if $\mathbf{A}$ is diagonalizable?

Solution

Here, $\det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$.

The eigen values are 2 and 1, and corresponding eigenvectors are

$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Next, form

$$\mathbf{P} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \mathbf{P}^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

$$\begin{aligned} \text{Since } \mathbf{A} &= \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \mathbf{A}^8 = \mathbf{P}\mathbf{D}^8\mathbf{P}^{-1} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

$$(2) \text{ Here, } \mathbf{A}\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1, \text{ and}$$

$$\mathbf{A}\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

Clearly, $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors for the eigenvalues 1 and 3, respectively. Thus

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \text{ where } \mathbf{P} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \text{ and } \mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

(3) Yes $\mathbf{A}$ is diagonalizable. There is a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the eigen space corresponding to $\lambda=3$. Moreover, there will be at least one eigenvector for $\lambda=5$ and one for $\lambda=-2$ say $\mathbf{v}_3$ and $\mathbf{v}_4$. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is linearly independent and $\mathbf{A}$ is diagonalizable, by Theorem 3. There can be no additional eigen vectors that are linearly independent from $\mathbf{v}_1$ to $\mathbf{v}_4$ because the vectors are all in $\mathbf{R}^4$. Hence the eigenspaces for $\lambda=5$ and $\lambda=-2$ are both one-dimensional.

Exercise

In exercises 1 and 2, let $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ and compute $\mathbf{A}^4$.

$$1. \mathbf{P} = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$2. \mathbf{P} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

In exercises 3 and 4, use the factorization $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ to compute $\mathbf{A}^k$, where k represents an arbitrary positive integer.

$$3. \begin{bmatrix} a & 0 \\ 3(a-b) & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$$

$$4. \begin{bmatrix} -2 & 12 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix}$$

In exercises 5 and 6, the matrix $\mathbf{A}$ is factored in the form $\mathbf{P}\mathbf{D}\mathbf{P}^{-1}$. Use the Diagonalization Theorem to find the eigenvalues of $\mathbf{A}$ and a basis for each eigenspace.

$$5. \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & -3/4 \\ 1/4 & -1/2 & 1/4 \end{bmatrix}$$

$$6. \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} -2 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 4 \\ -1 & 0 & -2 \end{bmatrix}$$

Diagonalize the matrices in exercises 7 to 18, if possible.

$$7. \begin{bmatrix} 3 & -1 \\ 1 & 5 \end{bmatrix}$$

$$8. \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

$$9. \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$$

$$10. \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

$$11. \begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$$

$$12. \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$$

$$13. \begin{bmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{bmatrix}$$

$$14. \begin{bmatrix} 0 & -4 & -6 \\ -1 & 0 & -3 \\ 1 & 2 & 5 \end{bmatrix}$$

$$15. \begin{bmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$16. \begin{bmatrix} -7 & -16 & 4 \\ 6 & 13 & -2 \\ 12 & 16 & 1 \end{bmatrix}$$

$$17. \begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$18. \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$$

Lecture 31

Eigenvectors and Linear Transformations

The goal of this lecture is to investigate the relationship between eigenvectors and linear transformations. Previously, we have learned how to find the eigenvalues and eigenvectors. We shall see that the matrix factorization $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ is also a type of linear transformations for some invertible matrix $\mathbf{P}$ and some diagonal matrix $\mathbf{D}$.

The Matrix of a Linear Transformation

Let $\mathbf{V}$ be an n -dimensional vector space, $\mathbf{W}$ an m -dimensional vector space, and $\mathbf{T}$ any linear transformation from $\mathbf{V}$ to $\mathbf{W}$. To associate a matrix with $\mathbf{T}$, we choose (ordered) bases $\mathbf{B}$ and $\mathbf{C}$ for $\mathbf{V}$ and $\mathbf{W}$, respectively.

Given any $\mathbf{x}$ in $\mathbf{V}$, the coordinate vector $[\mathbf{x}]_{\mathbf{B}}$ is in $\mathbf{R}^n$ and the coordinate vector of its image, $[\mathbf{T}(\mathbf{x})]_{\mathbf{C}}$, is in $\mathbf{R}^m$, as shown in Fig. 1.

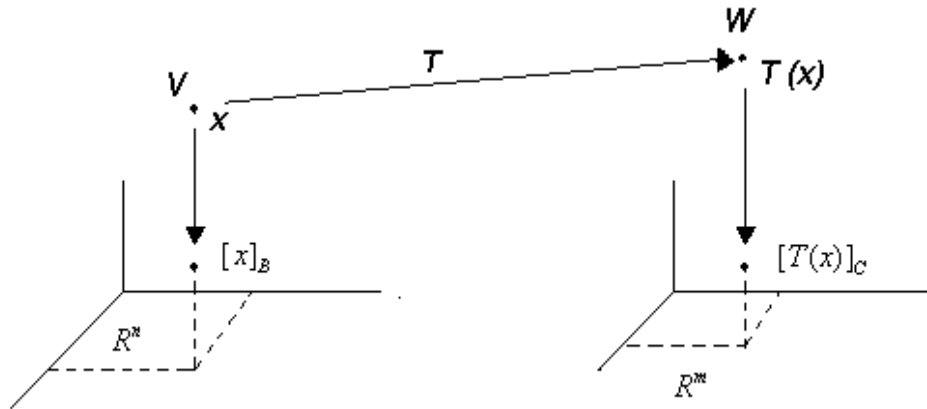


Figure 1 A linear transformation from $\mathbf{V}$ to $\mathbf{W}$

Let $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be the basis for $\mathbf{V}$. If $\mathbf{x} = r_1\mathbf{b}_1 + \dots + r_n\mathbf{b}_n$,

$$\text{then } [\mathbf{x}]_{\mathbf{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

and $\mathbf{T}$ is a linear transformation

$$\mathbf{T}(\mathbf{x}) = \mathbf{T}(r_1\mathbf{b}_1 + \dots + r_n\mathbf{b}_n) = r_1\mathbf{T}(\mathbf{b}_1) + \dots + r_n\mathbf{T}(\mathbf{b}_n) \quad (1)$$

Using the basis C in W , we can rewrite (1) in terms of C -coordinate vectors:

$$[T(x)]_C = r_1[T(b_1)]_C + \cdots + r_n[T(b_n)]_C \quad (2)$$

Since C -coordinate vectors are in $\mathbf{R}^m$, the vector equation (2) can be written as a matrix equation, namely

$$[T(x)]_C = M[x]_B \quad (3)$$

where $M = [[T(b_1)]_C \quad [T(b_2)]_C \quad \cdots \quad [T(b_n)]_C]$ (4)

The matrix M is a matrix representation of T , called the **matrix for T relative to the bases B and C** .

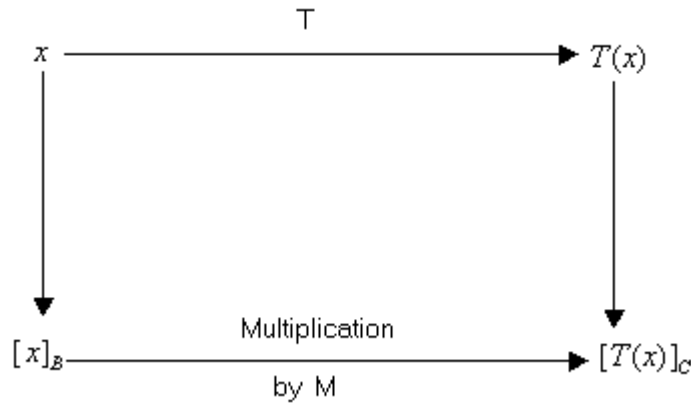


Figure 2

Example 1

Suppose that $B = \{b_1, b_2\}$ is a basis for V and $C = \{c_1, c_2, c_3\}$ is a basis for W . Let $T : V \rightarrow W$ be a linear transformation with the property that

$$T(b_1) = 3c_1 - 2c_2 + 5c_3 \text{ and } T(b_2) = 4c_1 + 7c_2 - c_3$$

Find the matrix M for T relative to B and C .

Solution

Since $M = [[T(b_1)]_C \quad [T(b_2)]_C \quad \cdots \quad [T(b_n)]_C]$ and here

$$[T(b_1)]_C = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \text{ and } [T(b_2)]_C = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Hence,

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

Linear Transformations from V into V

In general when W is the same as V and the basis C is the same as B , the matrix M in (4) is called the **matrix for T relative to B** , or simply the **B -matrix for T** , is denoted by $[T]_B$. See Fig. 3.

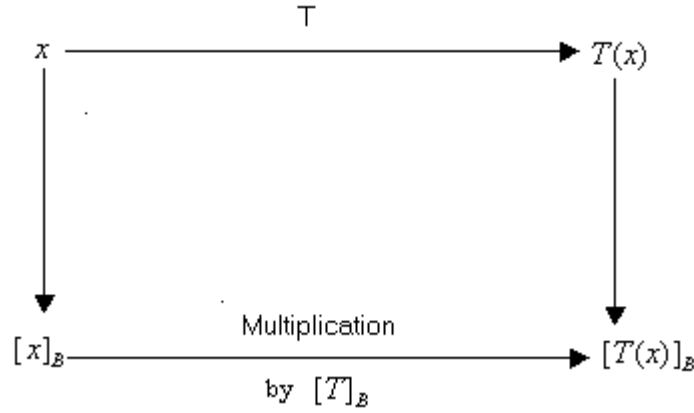


Figure 3

The B -matrix of $T : V \rightarrow V$ satisfies

$$[T(x)]_B = [T]_B [x]_B, \quad \text{for all } x \text{ in } V \quad (5)$$

Example 2

The mapping $T : P_2 \rightarrow P_2$ defined by $T(a_0 + a_1t + a_2t^2) = a_1 + 2a_2t$ is a linear transformation.

- Find the B -matrix for T , when B is the basis $\{1, t, t^2\}$.
- Verify that $[T(p)]_B = [T]_B [p]_B$ for each p in P_2 .

Solution

(a) We have to find the B -matrix of $T : V \rightarrow V$ satisfies

$$[T(x)]_B = [T]_B [x]_B, \quad \text{for all } x \text{ in } V$$

Since $T(a_0 + a_1t + a_2t^2) = a_1 + 2a_2t$ therefore

$$T(1) = 0, \quad T(t) = 1, \quad T(t^2) = 2t$$

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, [T(t)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, [T(t^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

(b) For a general $p(t) = a_0 + a_1t + a_2t^2$, we have

$$[T(p)]_B = [a_1 + 2a_2t]_B = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_B [p]_B$$

See figure 4.

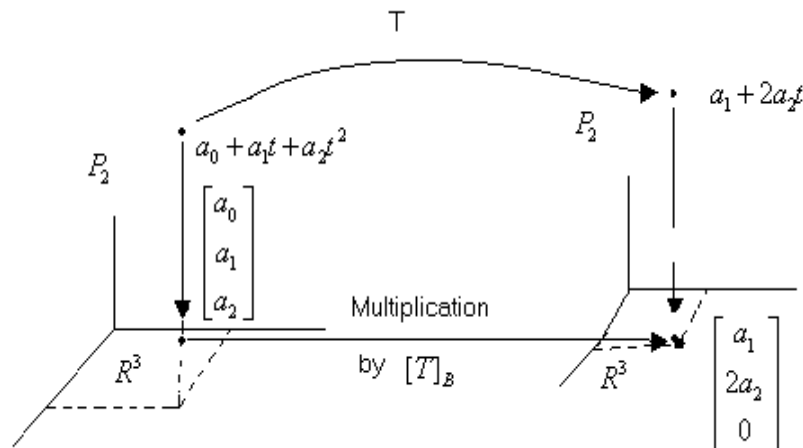


Figure 4 Matrix representation of a linear transformation

Linear Transformations on R^n

In an applied problem involving R^n , a linear transformation T usually appears first as a matrix transformation, $x \rightarrow Ax$. If A is diagonalizable then there is a basis B for R^n consisting of eigenvectors of A . Theorem below shows that in this case, the B -matrix of T is diagonal. Diagonalizing A amounts to finding a diagonal matrix representation of $x \rightarrow Ax$.

Theorem: Diagonal Matrix Representation

Suppose $A = PDP^{-1}$, where D is diagonal $n \times n$ matrix. If B is the basis for $\mathbf{R}^n$ formed from the columns of P , then D is the B -matrix of the transformation $\mathbf{x} \rightarrow A\mathbf{x}$.

Proof Denote the columns of P by $\mathbf{b}_1, \dots, \mathbf{b}_n$, so that $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $P = [\mathbf{b}_1 \dots \mathbf{b}_n]$. In this case, P is the change-of-coordinates matrix P_B discussed in lecture 23, where $P[\mathbf{x}]_B = \mathbf{x}$ and $[\mathbf{x}]_B = P^{-1}\mathbf{x}$

If $T(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x}$ in $\mathbf{R}^n$, then

$$\begin{aligned} [T]_B &= [[T(\mathbf{b}_1)]_B \dots [T(\mathbf{b}_n)]_B] && \text{Definition of } [T]_B \\ &= [[A\mathbf{b}_1]_B \dots [A\mathbf{b}_n]_B] && \text{Since } T(\mathbf{x}) = A\mathbf{x} \\ &= [P^{-1}A\mathbf{b}_1 \dots P^{-1}A\mathbf{b}_n] && \text{Change of Coordinates} \\ &= P^{-1}A[\mathbf{b}_1 \dots \mathbf{b}_n] && \text{Matrix multiplication} \\ &= P^{-1}AP && \end{aligned} \tag{6}$$

Since $A = PDP^{-1}$, we have $[T]_B = P^{-1}AP = D$.

Example 3

Define $T: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a basis B for $\mathbf{R}^2$ with the property that the B -matrix of T is a diagonal matrix.

Solution

Since the eigenvalues of matrix A are 5 and 3 and the eigenvectors corresponding to eigenvalue 5 is $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and the eigenvectors corresponding to eigenvalue 3 is $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$

Therefore $A = PDP^{-1}$, where $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$

By the above theorem D matrix is the B matrix of T when $B = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$.

Similarity of Matrix Representations

We know that A is similar to C if there is an invertible matrix P such that $A = PCP^{-1}$, therefore if A is similar to C , then C is the B -matrix of the transformation $\mathbf{x} \rightarrow A\mathbf{x}$ when the basis B is formed from the columns of P by the theorem above. (Since in the proof, the information that D is a diagonal matrix was not used).

The factorization $A = PCP^{-1}$ is shown in Fig. 5.

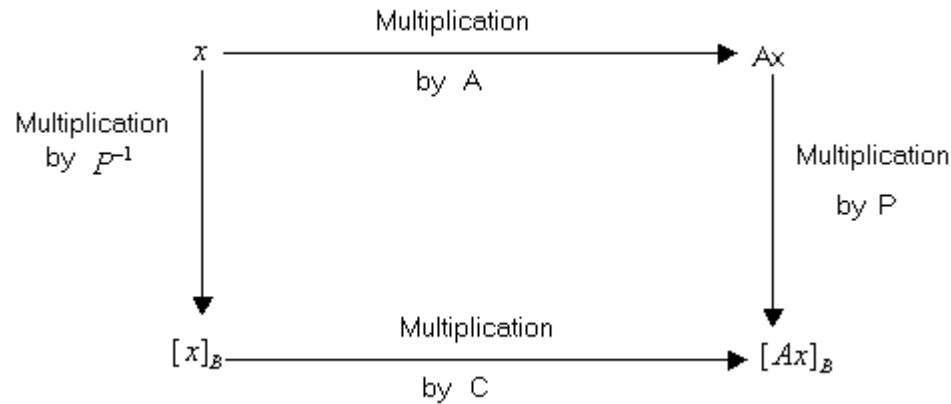


Figure 5 Similarity of two matrix representations: $A = PCP^{-1}$

Conversely, if $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is defined by $T(x) = Ax$, and if B is any basis for $\mathbf{R}^n$, then the B -matrix of T is similar to A . In fact, the calculations in (6) show that if P is the matrix whose columns come from the vectors in B , then $[T]_B = P^{-1}AP$. Thus, the set of all matrices similar to a matrix A coincides with the set of all matrix representations of the transformation $x \rightarrow Ax$.

An efficient way to compute a B -matrix $P^{-1}AP$ is to compute AP and then to row reduce the augmented matrix $[P \quad AP]$ to $[I \quad P^{-1}AP]$. A separate computation of P^{-1} is unnecessary.

Example 4

Find $T(a_0 + a_1t + a_2t^2)$, if T is the linear transformation from P_2 to P_2 whose matrix

relative to $B = \{1, t, t^2\}$ is $[T]_B = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$

Solution

Let $p(t) = a_0 + a_1t + a_2t^2$ and compute

$$[T(p)]_B = [T]_B[p]_B = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

So $T(p) = (3a_0 + 4a_1)1 + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

Example 5

Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ be $n \times n$ matrices. Verify that

(a) $\mathbf{A}$ is similar to $\mathbf{A}$.

(b) If $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to $\mathbf{C}$, then $\mathbf{A}$ is similar to $\mathbf{C}$.

Solution

(a) $\mathbf{A} = (\mathbf{I})^{-1} \mathbf{A} \mathbf{I}$, so $\mathbf{A}$ is similar to $\mathbf{A}$.

(b) By hypothesis, there exist invertible matrices $\mathbf{P}$ and $\mathbf{Q}$ with the property that $\mathbf{B} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}$ and $\mathbf{C} = \mathbf{Q}^{-1} \mathbf{B} \mathbf{Q}$. Substituting, and using facts about the inverse of a product we have

$$\mathbf{C} = \mathbf{Q}^{-1} \mathbf{B} \mathbf{Q} = \mathbf{Q}^{-1} (\mathbf{P}^{-1} \mathbf{A} \mathbf{P}) \mathbf{Q} = (\mathbf{PQ})^{-1} \mathbf{A} (\mathbf{PQ})$$

This equation has the proper form to show that $\mathbf{A}$ is similar to $\mathbf{C}$.

Lecture 32

Eigenvalues and Eigenvectors

Definition

A complex scalar λ satisfies

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

if and only if there is a nonzero vector x in $\mathbb{C}^n$ such that

$$Ax = \lambda x.$$

We call λ a (complex) eigenvalue and x a (complex) eigenvector corresponding to λ .

Example 1

If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $x \rightarrow Ax$ on $\mathbb{R}^2$ rotates the plane counterclockwise through a quarter-turn.

Solution

The action of A is periodic since after four quarter-turns, a vector is back where it started. Obviously, no nonzero vector is mapped into a multiple of itself, so A has no eigenvectors in $\mathbb{R}^2$ and hence no real eigenvalues. In fact, the characteristic equation of A is $\lambda^2 + 1 = 0 \Rightarrow \lambda^2 = -1 \Rightarrow \lambda = \pm i$

The only roots are complex: $\lambda = i$ and $\lambda = -i$. However, if we permit A to act on $\mathbb{C}^2$ then

$$\begin{aligned} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} &= \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix} \\ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} &= \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix} \end{aligned}$$

Thus i and $-i$ are the eigenvalues, with $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as corresponding eigenvectors.

Example 2

Let $A = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}$. Find the eigenvalues of A , and find a basis for each eigen-space.

Solution

The characteristic equation of A is

$$\begin{aligned} 0 &= \det \begin{bmatrix} 0.5 - \lambda & -0.6 \\ 0.75 & 1.1 - \lambda \end{bmatrix} = (0.5 - \lambda)(1.1 - \lambda) - (-0.6)(0.75) \\ &= \lambda^2 - 1.6\lambda + 1 \end{aligned}$$

From the quadratic formula, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = 0.8 \pm 0.6i$. For the eigenvalue $\lambda = 0.8 - 0.6i$, we study

$$\mathbf{A} - (0.8 - 0.6i) = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} - \begin{bmatrix} 0.8 - 0.6i & 0 \\ 0 & 0.8 - 0.6i \end{bmatrix} = \begin{bmatrix} -0.3 + 0.6i & -0.6 \\ 0.75 & 0.3 + 0.6i \end{bmatrix} \quad (1)$$

Since $0.8 - 0.6i$ is an eigenvalue, we know that the system

$$\begin{aligned} (-0.3 + 0.6i)x_1 - 0.6x_2 &= 0 \\ 0.75x_1 + (0.3 + 0.6i)x_2 &= 0 \end{aligned} \quad (2)$$

has a nontrivial solution (with x_1 and x_2 possibly complex numbers).

$$\begin{aligned} 0.75x_1 &= (-0.3 - 0.6i)x_2 \\ \Rightarrow x_1 &= (-0.4 - 0.8i)x_2 \end{aligned}$$

Taking $x_2 = 5$ to eliminate the decimals, we have $x_1 = -2 - 4i$. A basis for the eigenspace

corresponding to $\lambda = 0.8 - 0.6i$ is $\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$

Analogous calculations for $\lambda = 0.8 + 0.6i$ produce the eigenvector $\mathbf{v}_2 = \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix}$

Check

Compute

$$\mathbf{A}\mathbf{v}_2 = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4 + 2i \\ 4 + 3i \end{bmatrix} = (0.8 + 0.6i)\mathbf{v}_2$$

Example 3

One way to see how multiplication by the $\mathbf{A}$ in Example 2 affects points is to plot an arbitrary initial point-say, $\mathbf{x}_0 = (2, 0)$ and then to plot successive images of this point under repeated multiplications by $\mathbf{A}$. That is, plot

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix}$$

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -0.4 \\ 2.4 \end{bmatrix}$$

$$\mathbf{x}_3 = \mathbf{A}\mathbf{x}_2, \dots$$

Figure 1 shows $\mathbf{x}_0, \dots, \mathbf{x}_8$ as heavy dots. The smaller dots are the locations of $\mathbf{x}_9, \dots, \mathbf{x}_{100}$. The sequence lies along an elliptical orbit.

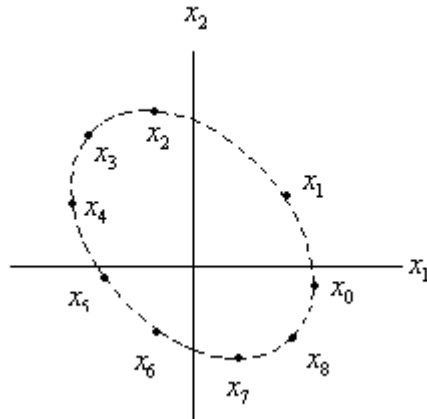


Figure 1 Iterates of a point x_0 under the action of a matrix with a complex eigenvalue

Of course Figure 1 does not explain why the rotation occurs. The secret to the rotation is hidden in the real and imaginary parts of a complex eigenvector.

Real and Imaginary Parts of Vectors

MCQS

The complex conjugate of a complex vector x in C^n is the vector $\bar{x}$ in C^n whose entries are the complex conjugates of the entries in x . The real and imaginary parts of a complex vector x are the vectors $\text{Re } x$ and $\text{Im } x$ formed from the real and imaginary parts of the entries of x .

Example 4

$$\text{If } x = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \text{ then}$$

$$\text{Re } x = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \text{Im } x = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \text{ and } \bar{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix}$$

If B is an $m \times n$ matrix with possibly complex entries, then $\bar{B}$ denotes the matrix whose entries are the complex conjugates of the entries in B . Properties of conjugates for complex numbers carry over to complex matrix algebra:

$$\overline{rx} = \bar{r} \bar{x}, \quad \overline{Bx} = \bar{B} \bar{x}, \text{ and } \overline{rB} = \bar{r} \bar{B}$$

Eigenvalues and Eigenvectors of a Real Matrix that Acts on C^n

Let A be $n \times n$ matrix whose entries are real. Then $\overline{A\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = A\overline{\mathbf{x}}$. If λ is an eigenvalue of A with $\mathbf{x}$ a corresponding eigenvector in C^n , then $A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. Hence $\overline{\lambda}$ is also an eigenvalue of A , with $\overline{\mathbf{x}}$ a corresponding eigenvector. This shows that when A is real, its complex eigenvalues occur in conjugate pairs. (Here and elsewhere, we use the term complex eigenvalue to refer to an eigenvalue $\lambda = a + bi$ with $b \neq 0$.)

Example 5

The eigenvalues of the real matrix in example 2 are complex conjugates namely, $0.8 - 0.6i$ and $0.8 + 0.6i$. The corresponding eigenvectors found in example 2 are also

conjugates: $\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{\mathbf{v}_1}$

Example 6

If $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real and not both zero, then the eigenvalues of C are $\lambda = a \pm bi$. Also, if $r = |\lambda| = \sqrt{a^2 + b^2}$, then

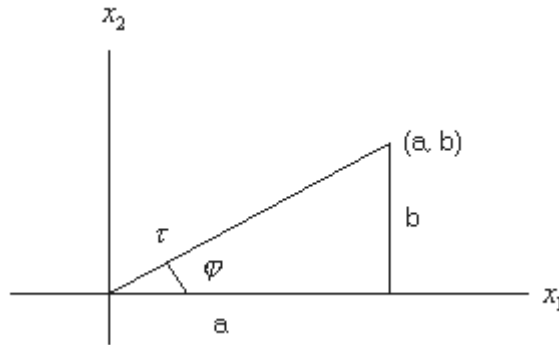


Figure 2

$$C = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

where φ is the angle between the positive x -axis and the ray from $(0, 0)$ through (a, b) . See Fig. 2. The angle φ is called the argument of $\lambda = a + bi$.

Thus, the transformation $\mathbf{x} \rightarrow \mathbf{C}\mathbf{x}$ may be viewed as the composition of a rotation through the angle φ and a scaling by $|\lambda|$.

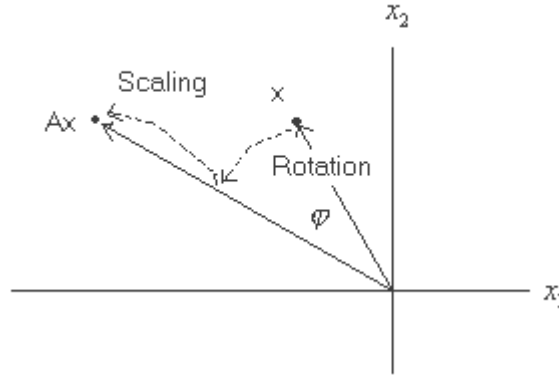


Figure 3 A rotation followed by a scaling

Finally, we are ready to uncover the rotation that is hidden within a real matrix having a complex eigenvalue.

Example 7

Let $\mathbf{A} = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}$, $\lambda = 0.8 - 0.6i$, and $\mathbf{v}_I = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$, as in example 2. Also, let $\mathbf{P}$

be the 2×2 real matrix $\mathbf{P} = [\text{Re } \mathbf{v}_I \quad \text{Im } \mathbf{v}_I] = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$

and let $\mathbf{C} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \frac{1}{20} \begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix} \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$

By Example 6, $\mathbf{C}$ is a pure rotation because $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$.

From $\mathbf{C} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, we obtain

$$\mathbf{A} = \mathbf{P}\mathbf{C}\mathbf{P}^{-1} = \mathbf{P} \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix} \mathbf{P}^{-1}$$

Here is the rotation “inside” $\mathbf{A}$! The matrix $\mathbf{P}$ provides a change of variable, say, $\mathbf{x} = \mathbf{P}\mathbf{u}$. The action of $\mathbf{A}$ amounts to a change of variable from $\mathbf{x}$ to $\mathbf{u}$, followed by a rotation, and then return to the original variable. See Figure 4.

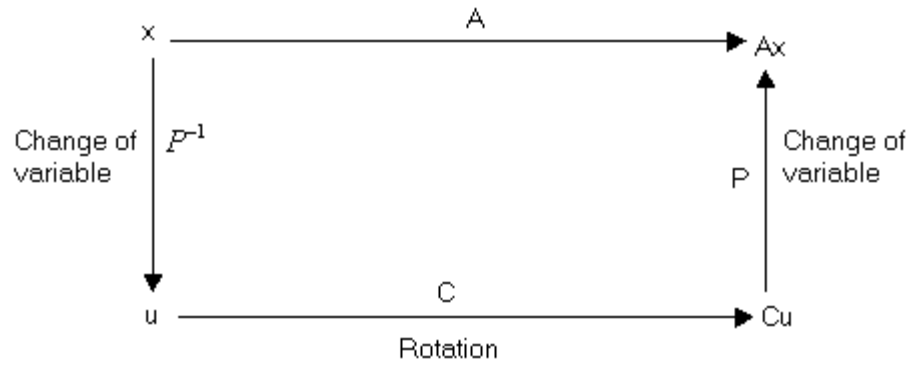


Figure 4 Rotation due to a complex eigenvalue

Theorem

Let A be a real 2×2 matrix with complex Eigen values $\lambda = a - bi$ ($b \neq 0$) and associated eigenvectors v in $\mathbb{C}^2$, then

$$A = PCP^{-1}, \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

$$P = [\text{Re } v \quad \text{Im } v]$$

Lecture 33

Discrete Dynamical Systems

For a dynamical system described by the difference equation $x_{k+1} = Ax_k$, eigenvalues and eigenvectors provide the key to understand the long term behavior and evolution of the system. Here the vector x_k gives information about the system as time (which is denoted by k) passes. Discrete dynamical system has a great application in many scientific fields. For example, modern state space design method of standard undergraduate courses in control system and steady state response of a control system relies heavily on dynamical system.

In this section, we will suppose that a matrix A is diagonalizable, with n linearly independent eigenvectors, $v_1, \dots, v_n$, and corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.

For convenience, assume that the eigenvectors are arranged so that

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

Since $\{v_1, \dots, v_n\}$ is a basis for R^n , any initial vector x_0 can be written uniquely as

$$x_0 = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$$

Since the v_i are eigenvectors,

$$\begin{aligned} x_1 &= Ax_0 \\ &= c_1 Av_1 + c_2 Av_2 + \dots + c_n Av_n \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + c_n \lambda_n v_n \end{aligned}$$

In general,

$$x_k = c_1 (\lambda_1)^k v_1 + \dots + c_n (\lambda_n)^k v_n \quad (k = 0, 1, 2, \dots)$$

We will discuss in the following examples what can happen to x_k if $k \rightarrow \infty$.

Example 1

Denote the owl and wood rat populations at time k by $x_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, where k is the time in months, O_k is the number of owls in the region studied, and R_k is the number of rats (measured in thousands). Suppose that

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned}$$

where p is a positive parameter to be specified. The $(.5)O_k$ in the first equation says that with no wood rats for food only half of the owls will survive each month, while the $(1.1)R_k$ in the second equation says that with no owls as predators, the rat population will grow by 10% per month. If rats are plentiful, then $(.4)R_k$ will tend to make the owl population rise, while the negative term $-p \cdot O_k$ measures the deaths of rats due to

predation by owls. (In fact, $1000p$ is the average number of rats eaten by one owl in one month.) Determine the evolution of this system when the predation parameter p is 0.104.

Solution

The coefficient matrix A for the given equations is

$$A = \begin{bmatrix} 0.5 & 0.4 \\ -p & 1.1 \end{bmatrix}$$

When $p = .104$,

The matrix becomes

$$A = \begin{bmatrix} 0.5 & 0.4 \\ -0.104 & 1.1 \end{bmatrix}$$

For eigenvalues of the coefficient matrix put

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

The eigenvectors corresponding to the eigenvalues $\lambda_1 = 1.02$ and $\lambda_2 = .58$ are

$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

An initial $\mathbf{x}_0$ can be written as $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ then, for $k \geq 0$,

$$\mathbf{x}_k = c_1(1.02)^k \mathbf{v}_1 + c_2(.58)^k \mathbf{v}_2 = c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} + c_2(.58)^k \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

As $k \rightarrow \infty$, $(.58)^k$ rapidly approaches zero. Assume $c_1 \neq 0$. Then, for all sufficiently large k , $\mathbf{x}_k$ is approximately the same as $c_1(1.02)^k \mathbf{v}_1$, and we write

$$\mathbf{x}_k \approx c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \dots\dots\dots(1)$$

The approximation in (1) improves as k increases, and so for large k ,

$$\mathbf{x}_{k+1} \approx c_1(1.02)^{k+1} \begin{bmatrix} 10 \\ 13 \end{bmatrix} = (1.02)c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \approx 1.02\mathbf{x}_k$$

This approximation says that eventually both $\mathbf{x}_k$ entries i.e owls and rats grow by a factor of almost 1.02 each month. Also equation (1) shows that $\mathbf{x}_k$ is approximately a multiple of (10,13). So we can say that the entries in $\mathbf{x}_k$ are almost in the same ratio as 10 to 13. It means that for every 10 owls there are about 13 thousand rats.

Trajectory of a dynamical system

In a 2×2 matrix, geometric description of a system's evolution can enhance the algebraic calculations. We can view that what happens to an initial point $\mathbf{x}_0$ in R^2 , when it is transformed repeatedly by the mapping $\mathbf{x} \rightarrow A\mathbf{x}$. The graph of $\mathbf{x}_0, \mathbf{x}_1, \dots$ is called a trajectory of the dynamical system.

Example 2

Plot several trajectories of the dynamical system $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$, when

$$\mathbf{A} = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

Solution

In a diagonal matrix the eigen values are the diagonal entries and as $\mathbf{A}$ is a diagonal matrix with diagonal entries 0.8 and 0.64, Therefore the eigenvalues of $\mathbf{A}$ are .8 and .64,

with eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then $\mathbf{x}_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Of course, $\mathbf{x}_k$ tends to 0 because $(.8)^k$ and $(.64)^k$ both approach to 0 as $k \rightarrow \infty$. But the way $\mathbf{x}_k$ goes towards 0 is interesting. Figure 1 shows the first few terms of several trajectories that begin at points on the boundary of the box with corners at $(\pm 3, \pm 3)$. The points on a trajectory are connected by a thin curve, to make the trajectory easier to see.

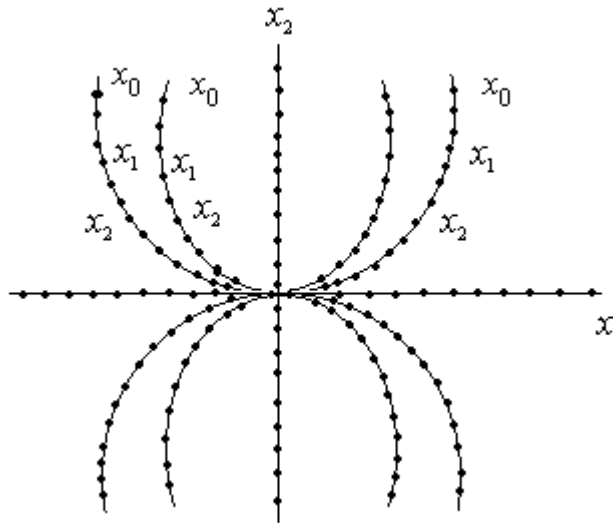


Figure 1 The origin as an attractor

In this example, the origin is called an **attractor of the dynamical system** because all trajectories tend towards **O**. This occurs whenever both eigenvalues are less than 1 in magnitude. The direction of greatest attraction is along the line through **O** and the eigenvector $\mathbf{v}_2$ as it has smaller eigenvalue.

Example 3

Plot several typical solution of the equation $x_{k+1} = Ax_k$, when

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

Solution

In a diagonal matrix the eigen values are the diagonal entries and as A is a diagonal matrix with diagonal entries 1.44 and 1.2, Therefore the eigenvalues of A are 1.44 and

1.2, with eigenvectors $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

As

$$x_0 = c_1 v_1 + c_2 v_2$$

$$= c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} c_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c_2 \end{bmatrix}$$

$$= \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \text{ then}$$

$$x_k = c_1 (1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 (1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Both terms grow in size as $k \rightarrow \infty$, but the first term grows faster because it has larger eigenvalue. Figure 2 shows several trajectories that begin at points quite close to 0.

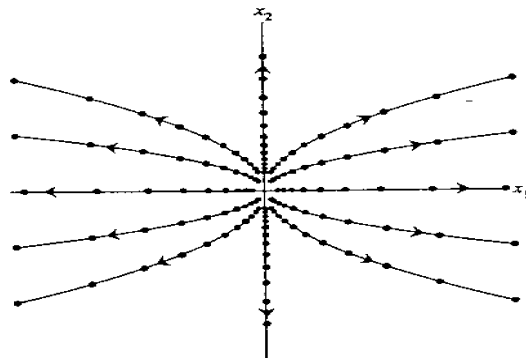


Figure 2 The origin as a repellor

In this example, the origin is called a **repellor** of the dynamical system because all trajectories tend away from **O**. This occurs whenever both eigenvalues are greater than 1

in magnitude. The direction of the greatest repulsion is the line through 0 and that eigen vector which has a larger eigenvalue.

Remark

From the above two examples it may be noticed that when eigenvalues are less than 1 in magnitude, the origin behaves as an attractor and when the eigenvalues are greater than 1 in magnitude, the origin behaves as a repellor.

Example 4

Plot several typical solutions of the equation

$y_{k+1} = Dy_k$, where

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

Show that a solution $\{y_k\}$ is unbounded if its initial point is not on the x_2 -axis.

Solution

Mistakes

The eigenvalues of D are 2.0 and 0.5, with eigenvectors $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

If

$$y_0 = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2$$

$$y_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \text{ then}$$

$$y_k = c_1 (2.0)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 (0.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

If y_0 is on the x_2 -axis, then $c_1 = 0$ and $y_k \rightarrow 0$ as $k \rightarrow \infty$. But if y_0 is not on the x_2 -axis, then the first term in the sum for y_k becomes arbitrarily large, and so $\{y_k\}$ is unbounded. Figure 3 shows ten trajectories that begin near or on the axis.

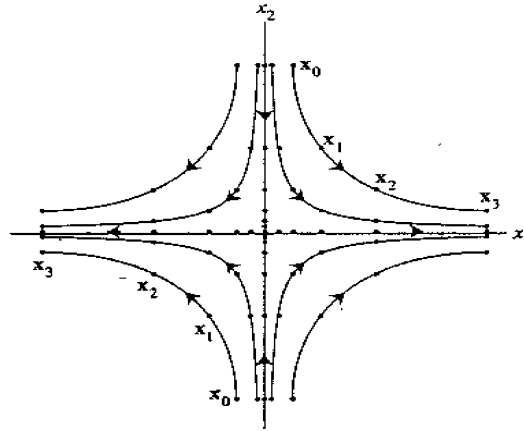


FIGURE 3 The origin as a saddle point.

In this example 'O' is called the saddle point because one eigenvalue is greater than '1' in magnitude and one is less than '1' in magnitude. The origin attracts solution from some directions and repels them in other directions.

Change of Variable

The preceding three examples involved diagonal matrices. To handle the nondiagonal case, we return for a moment to the $n \times n$ case in which eigenvectors of A form a basis

$\{v_1, \dots, v_n\}$ for R^n . Let $P = [v_1 \ \dots \ v_n]$, and let D be the diagonal matrix with the

corresponding eigenvalues on the diagonal. Given a sequence $\{x_k\}$ satisfying $x_{k+1} = Ax_k$,

define a new sequence $\{y_k\}$ by $y_k = P^{-1}x_k$, or equivalently, $x_k = Py_k$.

Substituting these relations into the equation $x_{k+1} = Ax_k$, and using the fact that

$A = PDP^{-1}$, we find that $Py_{k+1} = APy_k = (PDP^{-1})Py_k = PDy_k$

Left-multiplying both sides by P^{-1} , we obtain $y_{k+1} = Dy_k$

If we write y_k as $y(k)$ and denote the entries in $y(k)$ by $y_1(k), \dots, y_n(k)$, then

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

The change of variable from x_k to y_k has decoupled the system of difference equations. The evolution of $y_1(k)$, for example, is unaffected by what happens to $y_2(k), \dots, y_n(k)$, because $y_1(k+1) = \lambda_1 y_1(k)$ for each k .

Example 5

Show that the origin is a saddle point for solutions of $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$, where

$$\mathbf{A} = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

find the directions of greatest attraction and greatest repulsion.

Solution

For eigenvalues of the given matrix put

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

we find that $\mathbf{A}$ has eigenvalues 2 and .5, with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and

$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively. Since $|2| > 1$ and $|.5| < 1$, the origin is a saddle point of the dynamical system. If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then

$$\mathbf{x}_k = c_1(2)^k \mathbf{v}_1 + c_2(.5)^k \mathbf{v}_2$$

In figure 4, the direction of the greatest repulsion is the line through $\mathbf{v}_1$, the eigenvector whose eigenvalue is greater than 1.

The direction of the greatest attraction is determined by the eigenvector $\mathbf{v}_2$ whose eigenvalue is less than 1.

A number of trajectories are shown in figure 4. When this graph is viewed in terms of eigenvector axes, the diagram looks same as figure 4.

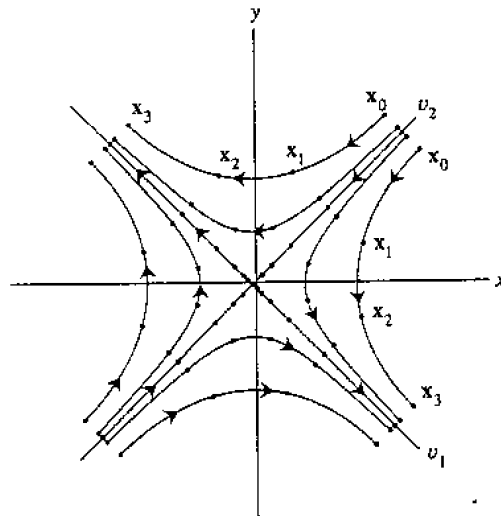


FIGURE 4 The origin as a saddle point.

Complex Eigenvalues

Since the characteristic equation of $n \times n$ matrix involves a polynomial of degree n , the equation always has exactly n roots, in which complex roots are possibly included. If we have **complex eigenvalues then we get complex eigen vectors which** may be divided into two parts: which is real part of the vector and imaginary part of the vector.

Note

If a matrix has two complex eigenvalues in which absolute value of one eigenvalue is greater than 1, then origin behaves as a repeller and iterates of x_0 will spiral outwards around the origin. If the absolute value of an eigenvalue is less than 1, then the origin behaves as an attractor and iterates of x_0 will spiral inward towards the origin.

Example 6

Let

$$A = \begin{bmatrix} 0.8 & 0.5 \\ -0.1 & 1.0 \end{bmatrix} \text{ has eigenvalue } 0.9 \pm 0.2i \text{ with eigenvectors } \begin{bmatrix} 1 \pm 2i \\ 1 \end{bmatrix}.$$

Find the trajectories of the system $x_{k+1} = Ax_k$ with initial vectors $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$,

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ -2.5 \end{bmatrix}.$$

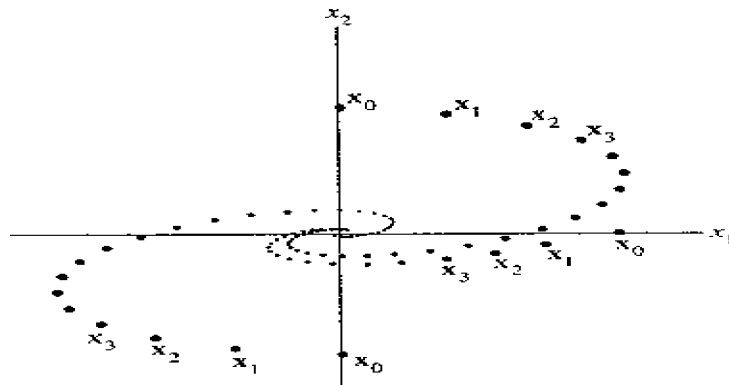
Solution

Given a matrix

$$A = \begin{bmatrix} 0.8 & 0.5 \\ -0.1 & 1.0 \end{bmatrix}$$

The eigenvalues of the matrix are $0.9 \pm 0.2i$ and corresponding eigenvectors are $\begin{bmatrix} 1 \pm 2i \\ 1 \end{bmatrix}$

also the initial vectors are given as $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$.



Example 7

Suppose the search survival rate of a young bird is 50%, and the stage matrix A is

$$A = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.3 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix}$$

What does the stage matrix model predict about this bird?

Solution

Given a matrix $A = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.3 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix}$

for eigenvalues of A we will put $\det(A - \lambda I) = 0$

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -\lambda & 0 & 0.33 \\ 0.3 & -\lambda & 0 \\ 0 & 0.71 & 0.94 - \lambda \end{vmatrix} \\ &= -\lambda(-0.94 + \lambda^2) + 0.33(0.3 \times 0.71) \\ &= 0.94\lambda - \lambda^3 + 0.07029 \end{aligned}$$

for eigenvalues put $\det(A - \lambda I) = 0$

$$0.94\lambda - \lambda^3 + 0.07029 = 0$$

$$\lambda^3 - 0.94\lambda - 0.07029 = 0$$

it gives 3 values of λ which are approximately 1.01, $-0.03 + 0.26i$, $-0.03 - 0.26i$

eigen vector for first value of λ will be $v_1 = (10, 3, 31)$ and for the next two values of λ we will get complex eigen vectors denoted by v_2 and v_3 .

Now

$$x_k = c_1(1.01)^k v_1 + c_2(-0.03 + 0.26i)^k v_2 + c_3(-0.03 - 0.26i)^k v_3$$

As $k \rightarrow \infty$, the second two vectors tend to 0. So x_k becomes more and more like the (real) vector $c_1(1.01)^k v_1$. Thus the long term growth rate of the owl population will be 1.01, and the population will grow slowly. The eigenvector describes the eventual distribution of the owls by life stages: For every 31 adults, there will be about 10 juveniles and three subadults.

Exercise

Suppose that a 3×3 matrix A has eigenvalues $2, 2/5, 3/5$ and corresponding eigenvectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} \quad v_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \quad v_3 = \begin{bmatrix} -3 \\ 3 \\ 7 \end{bmatrix}$$

$$\text{also } x_0 = \begin{bmatrix} -2 \\ -5 \\ -3 \end{bmatrix}$$

Find a general solution of the equation $x_{k+1} = Ax_k$ also find what happens to x_k when $k \rightarrow \infty$.

Lecture 34

Applications to Differential Equations

Differential equation

A differential equation is any equation which contains derivatives, either ordinary derivatives or partial derivatives. Differential equations play an extremely important and useful role in applied mathematics, engineering, and physics. Moreover, a lot of mathematical and numerical machinery has been developed for the solution of differential equations.

System of linear differential equations

A system of linear differential equations can be expressed as:

$$\begin{aligned}x_1' &= a_{11}x_1 + \dots + a_{1n}x_n \\x_2' &= a_{21}x_1 + \dots + a_{2n}x_n \\&\cdot \\&\cdot \\&\cdot \\x_n' &= a_{n1}x_1 + \dots + a_{nn}x_n\end{aligned}$$

where $x_i(t)$ is a function of time, $i = 1, \dots, n$, and the matrix of constant coefficient is

$$A = [a_{ij}]$$

In many applied problems, several quantities are continuously varying in time, and they are related by a system of differential equations:

$$\begin{aligned}x_1' &= a_{11}x_1 + \dots + a_{1n}x_n \\x_2' &= a_{21}x_1 + \dots + a_{2n}x_n \\&\cdot \\&\cdot \\&\cdot \\x_n' &= a_{n1}x_1 + \dots + a_{nn}x_n\end{aligned}$$

Here $x_1, x_2, \dots, x_n$ are differentiable functions of t , with derivatives $x_1', x_2', \dots, x_n'$, and the a_{ij} are constants. Write the system as a matrix differential equation

$$x' = Ax$$

Where

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, x'(t) = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix}, \text{ and } A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & & a_{nn} \end{bmatrix}$$

Superposition of solutions

If u and v are two solutions of $x' = Ax$, then $cu + dv$ is also a solution.

Proof:

If u and v are two solutions of $x' = Ax$ then $u' = Au$ and $v' = Av$. Therefore,

$$\begin{aligned} (cu + dv)' &= cu' + dv' \\ &= cAu + dAv \\ &= A(cu + dv) \end{aligned}$$

Linear combination of any set of solutions is also a solution. This property is called **superposition** of solutions.

Fundamental set of Solutions

If A is an $n \times n$ matrix, then there are n linearly independent functions in a fundamental set, and each solution of $x' = Ax$ is a unique linear combination of these 'n' functions. So we can say that fundamental set of solutions is a basis for the set of all solutions of $x' = Ax$, and the solution set is n -dimensional vector space of functions.

Initial value problem

If a vector x_0 is specified, then the initial value problem is to construct the unique function x such that $x' = Ax$ and $x(0) = x_0$

Example 1

For the system $x' = Ax$. What will be the solution when A is a diagonal matrix and is given by

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

Solution

The solution of the given system can be obtained by elementary calculus. Consider

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

from here we have

$$x_1'(t) = 3x_1(t) \quad \text{and} \quad x_2'(t) = -5x_2(t)$$

first of all we will find solution of $x_1'(t) = 3x_1(t)$ by using calculus.

$$x_1'(t) = 3x_1(t)$$

$$\frac{dx_1}{dt} = 3x_1$$

$$\frac{dx_1}{x_1} = 3dt$$

Integrating both sides

$$\int \frac{dx_1}{x_1} = 3 \int dt$$

$$\ln x_1 = 3t + \ln c_1$$

$$\ln x_1 - \ln c_1 = 3t$$

$$\ln\left(\frac{x_1}{c_1}\right) = 3t$$

taking antilog on both sides

$$\frac{x_1}{c_1} = e^{3t}$$

$$x_1(t) = c_1 e^{3t}$$

Similarly Solution of

$$x_2'(t) = -5x_2(t)$$

will be

$$x_2(t) = c_2 e^{-5t}$$

And it can be written in the form

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1 e^{3t} \\ c_2 e^{-5t} \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t}$$

The given system is said to be decoupled because each derivative of the function depends only on the function itself, not on some combination or coupling of $x_1(t)$ and $x_2(t)$.

Observation

From the above example we observe that a solution of the general equation $x' = Ax$ might be a linear combination of functions of the form

$$x(t) = ve^{\lambda t}$$

Where λ is some scalar and v is some fixed non-zero vector.

$$x(t) = ve^{\lambda t} \dots\dots\dots(1)$$

differentiating w.r.t 't'

$$x'(t) = \lambda ve^{\lambda t}$$

Multiplying both sides of (1) by A

$$Ax(t) = Ave^{\lambda t}$$

Since $e^{\lambda t}$ is never zero, $x'(t)$ will equal to $Ax(t)$ iff $\lambda v = Av$.

Eigenfunctions

Each pair of eigenvalue and its corresponding eigenvector provides a solution of the equation $x' = Ax$ which is called **eigenfunctions** of the differential equation.

Example 2

The circuit in figure 1 can be described by

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -(R_1 + R_2)/C_1 & R_2/C_1 \\ R_2/C_2 & -R_2/C_2 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

Where $v_1(t)$ and $v_2(t)$ are the voltages across the two capacitors at time t .

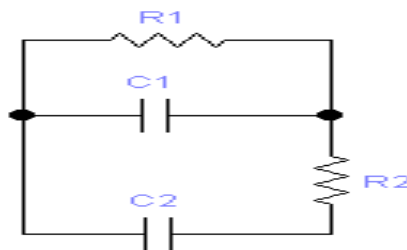


Figure 1

Suppose that resistor R_1 is 2 ohms, R_2 is 1 ohms, capacitor C_1 is 2 farads, and capacitor C_2 is 1 farad and suppose that there is an initial charge of 5 volts on capacitor C_1 and 4 volts on capacitor C_2 . Find formulas for $v_1(t)$ and $v_2(t)$ that describe how the voltages change over time.

Solution

For the data given, we can form A as

$$A = \begin{bmatrix} -(R_1 + R_2)/C_1 & R_2/C_1 \\ R_2/C_2 & -R_2/C_2 \end{bmatrix}$$

$$R_1 = 2$$

$$R_2 = 1$$

$$C_1 = 2$$

$$C_2 = 1 \text{ so}$$

$$-(R_1 + R_2)/C_1 = -(2+1)/2 \Rightarrow -3/2 \Rightarrow -1.5$$

$$R_2/C_1 = 1/2 \Rightarrow 0.5$$

$$R_2/C_2 = 1/1 \Rightarrow 1$$

$$-R_2/C_2 = -1/1 \Rightarrow -1$$

$$A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \text{ and } \mathbf{x}_0 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}. \text{ The vector } \mathbf{x}_0 \text{ lists the initial values of } \mathbf{x}. \text{ From}$$

A , we obtain eigenvalues $\lambda_1 = -.5$ and $\lambda_2 = -2$, by using $\det(A - \lambda I) = 0$ with corresponding

$$\text{eigenvectors } \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

The eigenfunctions $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda_1 t}$ and $\mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda_2 t}$ both satisfy $\mathbf{x}' = A\mathbf{x}$, and so does any linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$. Set

$$\begin{aligned} \mathbf{x}(t) &= c_1 \mathbf{v}_1 e^{\lambda_1 t} + c_2 \mathbf{v}_2 e^{\lambda_2 t} \\ &= c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t} \dots\dots\dots(1) \end{aligned}$$

and note that $\mathbf{x}(0) = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2$. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are obviously linearly independent and hence span $\mathbf{R}^2$, c_1 and c_2 can be found to make $\mathbf{x}(0)$ equal to $\mathbf{x}_0$.

We can find c_1 and c_2 as

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{x}_0] \\ &= \begin{bmatrix} 1 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \text{ row reduction of this matrix gives} \\ &= \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{bmatrix} \end{aligned}$$

$$\text{hence } c_1 = 3 \text{ and } c_2 = -2$$

Put the value of c_1 and c_2 in (1)

$$\mathbf{x}(t) = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

Thus the desired solution of the differential equation $\mathbf{x}' = A\mathbf{x}$ is

$$x(t) = \begin{bmatrix} 3e^{-0.5t} \\ 6e^{-0.5t} \end{bmatrix} - \begin{bmatrix} -2e^{-2t} \\ 2e^{-2t} \end{bmatrix}$$

or
$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 3e^{-0.5t} + 2e^{-2t} \\ 6e^{-0.5t} - 2e^{-2t} \end{bmatrix}$$

As both eigenvalues are negative so origin in this case behaves as an attractor and the direction of greatest attraction will be along the more negative eigenvalue $\lambda = -2$.

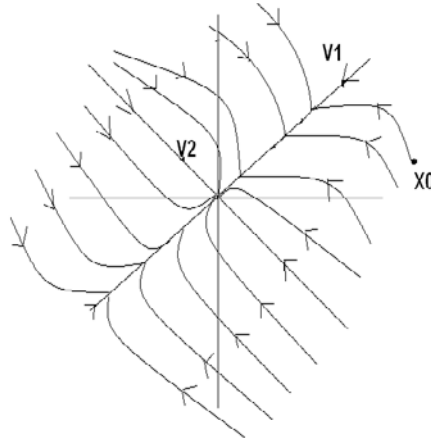


Figure 2 The Origin as an attractor

Example 3

Suppose a particle is moving in a planar force field and its position vector x satisfies $\dot{x} = Ax$ and $x(0) = x_0$, where

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Solve this initial value problem for $t \geq 0$, and sketch the trajectory of the particle.

Solution

The eigenvalues of the given matrix can be found by using $\det(A - \lambda I) = 0$ which are turned out to be $\lambda_1 = 6$ and $\lambda_2 = -1$ with corresponding eigen vectors $v_1 = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

For any constants c_1 and c_2 , the function

$$\begin{aligned} x(t) &= c_1 v_1 e^{\lambda_1 t} + c_2 v_2 e^{\lambda_2 t} \\ &= c_1 \begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t} \end{aligned}$$

We have to find c_1 and c_2 to satisfy $x(0) = x_0$, so

$$x_0 = c_1 \begin{bmatrix} -5 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix} = \begin{bmatrix} -5c_1 \\ 2c_1 \end{bmatrix} + \begin{bmatrix} c_2 \\ c_2 \end{bmatrix}$$

$$= \begin{bmatrix} -5c_1 + c_2 \\ 2c_1 + c_2 \end{bmatrix}$$

$$2.9 = -5c_1 + c_2 \dots \dots \dots (1)$$

$$2.6 = 2c_1 + c_2 \dots \dots \dots (2)$$

Subtracting 1 from 2 we get

$$c_1 = -\frac{3}{70} \text{ and substituting this value of } c_1 \text{ in (1) we get } c_2 = \frac{188}{70}$$

$$x(t) = -\frac{3}{70} \begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + \frac{188}{70} \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

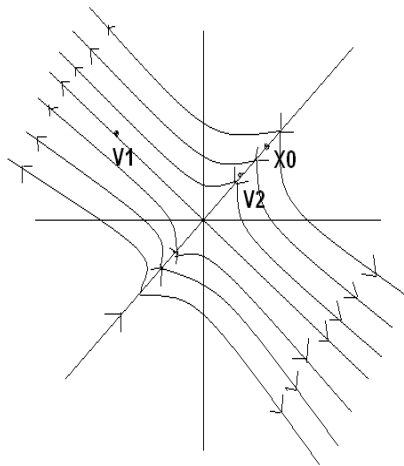


Figure 3 The Origin as a saddle point

As the matrix A has both positive and negative eigenvalues so in this case origin behaves as a saddle point. The direction of greatest repulsion is the line through v_1 and 0 corresponding to the positive eigenvalue. The direction of greatest attraction is the line through v_2 and 0, corresponding to the negative eigenvalue.

Decoupling a Dynamical System

For any dynamical system described by $\mathbf{x}' = \mathbf{A}\mathbf{x}$ when $\mathbf{A}$ is diagonalizable, the fundamental set of solutions can be obtained by the methods that we have discussed in examples 2 and 3.

Suppose that the eigenfunctions for $\mathbf{A}$ are

$$\mathbf{v}_1 e^{\lambda_1 t}, \dots, \mathbf{v}_n e^{\lambda_n t}$$

with $\mathbf{v}_1, \dots, \mathbf{v}_n$ linearly independent eigenvectors. Let $\mathbf{P} = [\mathbf{v}_1 \dots \mathbf{v}_n]$ and let $\mathbf{D}$ be the diagonal matrix with entries $\lambda_1, \dots, \lambda_n$, so that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$. Now make a change of variable, defining a new function $\mathbf{y}$ by

$$\mathbf{y}(t) = \mathbf{P}^{-1}\mathbf{x}(t), \text{ or equivalently, } \mathbf{x}(t) = \mathbf{P}\mathbf{y}(t)$$

Substituting these relations into $\mathbf{x}' = \mathbf{A}\mathbf{x}$ gives

$$\mathbf{P}\mathbf{y}'(t) = (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})\mathbf{P}\mathbf{y}(t)$$

$$\mathbf{P}\mathbf{y}'(t) = \mathbf{P}\mathbf{D}\mathbf{y}(t)$$

pre multiplying both side by $\mathbf{P}^{-1}$

$$\mathbf{P}^{-1}\mathbf{P}\mathbf{y}'(t) = \mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{y}(t)$$

$$\mathbf{y}'(t) = \mathbf{D}\mathbf{y}(t) \text{ or}$$

$$\mathbf{y}' = \mathbf{D}\mathbf{y}$$

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

The change of variable from $\mathbf{x}$ to $\mathbf{y}$ has decoupled the system of differential equations because the derivative of each scalar function y_k depends only on y_k . Since $y_1' = \lambda_1 y_1$, we have $y_1(t) = c_1 e^{\lambda_1 t}$, with similar formulas for $y_2, \dots, y_n$. Thus

$$\mathbf{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \text{ where } \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{y}(0) = \mathbf{P}^{-1}\mathbf{x}(0) = \mathbf{P}^{-1}\mathbf{x}_0$$

To obtain the general solution $\mathbf{x}$ of the original system, compute

$$\mathbf{x}(t) = \mathbf{P}\mathbf{y}(t) = [\mathbf{v}_1 \dots \mathbf{v}_n] \mathbf{y}(t) = c_1 \mathbf{v}_1 e^{\lambda_1 t} + \dots + c_n \mathbf{v}_n e^{\lambda_n t}$$

This is the eigenfunction expansion as constructed in example 2.

Complex Eigenvalues

As we know that for a real matrix $\mathbf{A}$, complex eigenvalues and associated eigen vectors come in conjugate pairs i.e if λ is an eigenvalue of $\mathbf{A}$ then $\bar{\lambda}$ will be the 2nd

eigenvalue. Similarly if one eigen vector is v then the other will be $\bar{v}$. So there will be two solutions of $x' = Ax$ and they are

$$x_1(t) = ve^{\lambda t}, x_2(t) = \bar{v}e^{\bar{\lambda}t}$$

Here $x_2(t) = \overline{x_1(t)}$

$$\operatorname{Re}(ve^{\lambda t}) = \frac{1}{2} [x_1(t) + \overline{x_1(t)}]$$

$$\operatorname{Im}(ve^{\lambda t}) = \frac{1}{2i} [x_1(t) - \overline{x_1(t)}]$$

From Calculus

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

$$e^{\lambda t} = 1 + \lambda t + \frac{(\lambda t)^2}{2!} + \dots + \frac{(\lambda t)^n}{n!} + \dots$$

If we write $\lambda = a + ib$

$$e^{(a+ib)t} = e^{at} \cdot e^{ibt} = e^{at} (\cos bt + i \sin bt)$$

So

$$ve^{\lambda t} = (\operatorname{Re} v + i \operatorname{Im} v) \cdot e^{at} (\cos bt + i \sin bt)$$

$$ve^{\lambda t} = \operatorname{Re} v \cdot e^{at} \cos bt + \operatorname{Re} v \cdot e^{at} i \sin bt + i \operatorname{Im} v \cdot e^{at} \cos bt - \operatorname{Im} v \cdot e^{at} \sin bt$$

$$ve^{\lambda t} = e^{at} (\operatorname{Re} v \cos bt - \operatorname{Im} v \sin bt) + ie^{at} (\operatorname{Re} v \sin bt + \operatorname{Im} v \cos bt)$$

So two real solutions of $x' = Ax$ are

$$y_1(t) = \operatorname{Re} x_1(t) = [(\operatorname{Re} v) \cos bt - (\operatorname{Im} v) \sin bt] e^{at}$$

$$y_2(t) = \operatorname{Im} x_1(t) = [(\operatorname{Re} v) \sin bt + (\operatorname{Im} v) \cos bt] e^{at}$$

Example 4

The circuit in figure 4 can be described by the equation

$$\begin{bmatrix} i_L' \\ v_C' \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1 C) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

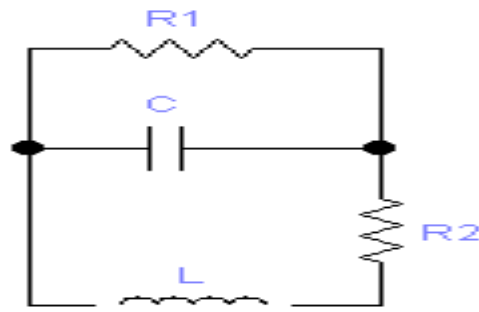


Figure 4

Where i_L is the current passing through the inductor L and v_c is the voltage drop across the capacitor C . Suppose R_1 is 5 ohms, R_2 is .8 ohm, C is .1 farad, and L is .4 henry. Find formulas for i_L and v_c , if the initial current through the inductor is 3 amperes and the initial voltage across the capacitor is 3 volts.

Solution

From the data given, We can form A as

$$A = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1 C) \end{bmatrix}$$

$$R_1 = 5$$

$$R_2 = 0.8$$

$$C = 0.1$$

$$L = 0.4$$

$$-R_2/L = -0.8/0.4 \Rightarrow -8/10 \times 10/4 \Rightarrow -2$$

$$-1/L = -1/0.4 \Rightarrow -10/4 \Rightarrow -2.5$$

$$1/C = 1/0.1 \Rightarrow 10$$

$$-1/(R_1 C) = -1/5 \times 0.1 \Rightarrow -1/0.5 \Rightarrow -2$$

$$A = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$$

$$\begin{aligned}
 \det(A - \lambda I) &= \begin{vmatrix} -2 - \lambda & -2.5 \\ 10 & -2 - \lambda \end{vmatrix} \\
 &= (-2 - \lambda)^2 + 25 \\
 &= 4 + \lambda^2 + 4\lambda + 25 \\
 &= \lambda^2 + 4\lambda + 29
 \end{aligned}$$

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

$$\lambda^2 + 4\lambda + 29 = 0$$

the eigenvalues of A is $-2 \pm 5i$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -i \\ 2 \end{bmatrix}$ The

complex solutions of

$\mathbf{x}' = \mathbf{A}\mathbf{x}$ are complex linear combinations of

$$\begin{aligned}
 \mathbf{x}_1(t) &= \begin{bmatrix} i \\ 2 \end{bmatrix} e^{(-2+5i)t} \\
 &= \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t} \cdot e^{5it} \quad \text{and} \quad \mathbf{x}_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix} e^{(-2-5i)t} \\
 \mathbf{x}_1(t) &= \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t} (\cos 5t + i \sin 5t)
 \end{aligned}$$

By taking real and imaginary parts of $\mathbf{x}_1$, we obtain real solutions:

$$\mathbf{y}_1(t) = \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t}, \mathbf{y}_2(t) = \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t}$$

Since $\mathbf{y}_1$ and $\mathbf{y}_2$ are linearly independent functions, they form a basis for the two-dimensional real vector space of solutions $\mathbf{x}' = \mathbf{A}\mathbf{x}$. Thus, the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t}$$

To satisfy $\mathbf{x}(0) = \mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ we need c_1 and c_2

$$c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$

$$\begin{bmatrix} 0 \\ 2c_1 \end{bmatrix} + \begin{bmatrix} c_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$0 + c_2 = 3$$

Thus

$$2c_1 + 0 = 3$$

$$c_2 = 3$$

$$c_1 = \frac{3}{2} = 1.5$$

$$\begin{aligned}
 \mathbf{x}(t) &= c_1 \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t} \\
 &= 1.5 \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t} \\
 &= \begin{bmatrix} -1.5 \sin 5t e^{-2t} \\ 3 \cos 5t e^{-2t} \end{bmatrix} + \begin{bmatrix} 3 \cos 5t e^{-2t} \\ 6 \sin 5t e^{-2t} \end{bmatrix} \\
 &= \begin{bmatrix} -1.5 \sin 5t e^{-2t} + 3 \cos 5t e^{-2t} \\ 3 \cos 5t e^{-2t} + 6 \sin 5t e^{-2t} \end{bmatrix} \\
 \begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} &= \begin{bmatrix} -1.5 \sin 5t + 3 \cos 5t \\ 3 \cos 5t + 6 \sin 5t \end{bmatrix} e^{-2t}
 \end{aligned}$$

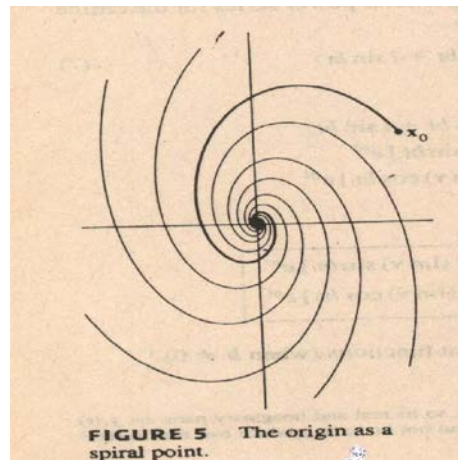


Figure 5 The origin as a spiral point

MCQS

In this figure, due to the complex eigenvalues a rotation is caused by the sine and cosine function and hence the origin is called a spiral point. The trajectories spiral inwards as in this example the real part of the eigenvalue is negative.

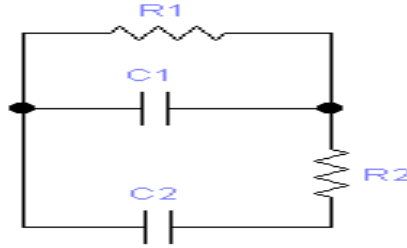
Similarly, it is important to note that when A has a complex eigenvalue with positive real part, the trajectories spiral outwards. And if the real part of the eigenvalue is zero, then the trajectories form ellipses around the origin.

Exercise 1

The circuit in given figure can be described by

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/R_2 C_1 \\ 1/R_2 C_2 & -1/R_2 C_2 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

Where $v_1(t)$ and $v_2(t)$ are the voltages across the two capacitors at time t .



Suppose that resistor R_1 is 1 ohms, R_2 is 2 ohms, capacitor C_1 is 1 farad, and capacitor C_2 is 0.5 farad, and suppose that there is an initial charge of 7 volts on capacitor C_1 and 6 volts on capacitor C_2 . Find formulas for $v_1(t)$ and $v_2(t)$ that describe how the voltages change over time.

Exercise 2

Find formulas for the current i_L and the voltage V_C for the circuit in Example 4, assuming that $R_1 = 10 \text{ ohms}$, $R_2 = 0.2 \text{ ohm}$, $C = 0.4 \text{ farad}$, $L = 0.2 \text{ henry}$ the initial current is 0 amp, and the initial voltage is 16 volts.

Lecture 35

Iterative Estimates for Eigenvalues

Eigenvalues and eigenvectors

Let A be an $n \times n$ square matrix, λ is any scalar value. If there exists a non-zero vector X , such that it satisfies the equation $AX = \lambda X$ then λ is called the eigenvalue and X is called the corresponding eigenvector.

i.e. if

$$AX = \lambda X$$

$$AX - \lambda X = 0$$

$$AX - \lambda IX = 0$$

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

To find the eigenvalues we have to solve the equation

$$|A - \lambda I| = 0$$

Example 1

Find the eigenvalues of $A = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}$.

Solution

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

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

$$|A - \lambda I| = (1-\lambda)(-2-\lambda)$$

Solving $|A - \lambda I| = 0$, we get

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

So, either $(1-\lambda) = 0$ or $(-2-\lambda) = 0$

so, $\lambda = 1$ $\lambda = -2$ are the eigenvalues of A .

Power Method

Let a matrix A is diagonalizable, with n linearly independent eigenvectors, $v_1, v_2, \dots, v_n$ and corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ such that

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

Since $\{v_1, v_2, \dots, v_n\}$ is a basis for R^n , any vector x can be written as

$$x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$$

Multiplying both sides by A, we get

$$\begin{aligned} Ax &= A(c_1 v_1 + c_2 v_2 + \dots + c_n v_n) \\ &= A(c_1 v_1) + A(c_2 v_2) + \dots A(c_n v_n) \\ &= c_1 (Av_1) + c_2 (Av_2) + \dots c_n (Av_n) \\ &= c_1 (\lambda_1 v_1) + c_2 (\lambda_2 v_2) + \dots c_n (\lambda_n v_n) \end{aligned}$$

Again multiplying by A and simplifying as above, we get

$$A^2 x = c_1 (\lambda_1^2 v_1) + c_2 (\lambda_2^2 v_2) + \dots c_n (\lambda_n^2 v_n)$$

Continuing this process we get

$$\begin{aligned} A^k x &= c_1 (\lambda_1^k v_1) + c_2 (\lambda_2^k v_2) + \dots c_n (\lambda_n^k v_n) \\ A^k x &= c_1 (\lambda_1)^k v_1 + c_2 (\lambda_2)^k v_2 + \dots c_n (\lambda_n)^k v_n \quad (k = 1, 2, \dots) \\ \left(\frac{1}{\lambda_1}\right)^k A^k x &= c_1 v_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \dots c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \quad (k = 1, 2, \dots) \\ (\lambda_1)^{-k} A^k x &\rightarrow c_1 v_1 \end{aligned}$$

$$\text{Since } |\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

$$\text{as } k \rightarrow \infty$$

Thus for large k, a scalar multiple of $A^k x$ determines almost the same direction as the eigenvector $c_1 v_1$. Since, positive scalar multiples do not change the direction of the vector, $A^k x$ itself points almost in the same direction as v_1 or $-v_1$, provided $c_1 \neq 0$.

Procedure for finding the Eigenvalues and Eigenvectors **Power Method**

To compute the largest eigenvalue and the corresponding eigenvector of the system

$$A(X) = \lambda(X)$$

Where A is a real, symmetric or un-symmetric matrix, the power method is widely used in practice.

Procedure

Step 1: Choose the initial vector such that the largest element is unity.

Step 2: The normalized vector $v^{(0)}$ is pre-multiplied by the matrix A.

Step 3: The resultant vector is again normalized.

Step 4: This process of iteration is continued and the new normalized vector is repeatedly pre-multiplied by the matrix A until the required accuracy is obtained.

At this point, the result looks like

$$u^{(k)} = [A]v^{(k-1)} = q_k v^{(k)}$$

Here, q_k is the desired largest eigenvalue and $v^{(k)}$ is the corresponding eigenvector.

Example 2

Let $A = \begin{pmatrix} 1.8 & .8 \\ 0.2 & 1.2 \end{pmatrix}$, $v_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ and $x = \begin{pmatrix} -0.5 \\ 1 \end{pmatrix}$. Then A has Eigenvalues 2 and 1, and the eigenspace for $\lambda_1 = 2$ is the line through 0 and v_1 . For $k=0, \dots, 8$, compute $A^k x$ and construct the line through 0 and $A^k x$. What happens as k increases?

Solution

The first three calculations are

$$Ax = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix} \begin{bmatrix} -0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix}$$

$$A^2 x = A(Ax) = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix} \begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix}$$

$$A^3 x = A(A^2 x) = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix} \begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$$

Analogous calculations complete Table 1.

Table 1 Iterates of a Vector

k	0	1	2	3	4	5	6	7	8
$A^k x$	$\begin{bmatrix} -0.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

The vectors $x, Ax, A^2x, \dots, A^k x$ are shown in Fig. 1. The other vectors are growing too long to display. However, line segments are drawn showing the directions of those vectors. In fact, the directions of the vectors are what we really want to see, not the vectors themselves. The lines seem to be approaching the line representing the eigenspaces spanned by v_1 . More precisely, the angle between the line (subspace) determined by $A^k x$ and the line (eigenspaces) determined by v_1 goes to zero as $k \rightarrow \infty$.

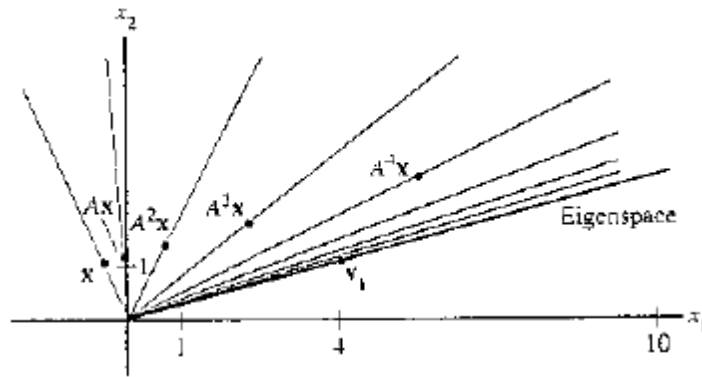


Figure 1 Directions determined by $x, Ax, A^2x, \dots, A^7x$

The vectors $(\lambda_1)^{-k} A^k x$ in (3) are scaled to make them converge to $c_1 v_1$, provided $c_1 \neq 0$. We cannot scale $A^k x$ in this way because we do not know λ_1 but we can scale each $A^k x$ to make its largest entry 1. It turns out that the resulting sequence $\{x_k\}$ will converge to a multiple of v_1 whose largest entry is 1. Figure 2 shows the scaled sequence for Example 1. The eigenvalue λ_1 can be estimated from the sequence $\{x_k\}$ too. When x_k is close to an eigenvector for λ_1 the vector Ax_k is close to $\lambda_1 x_k$ with each entry in Ax_k approximately λ_1 times the corresponding entry in x_k . **Because the largest entry in x_k is 1 the largest entry in Ax_k is close to λ_1 .**

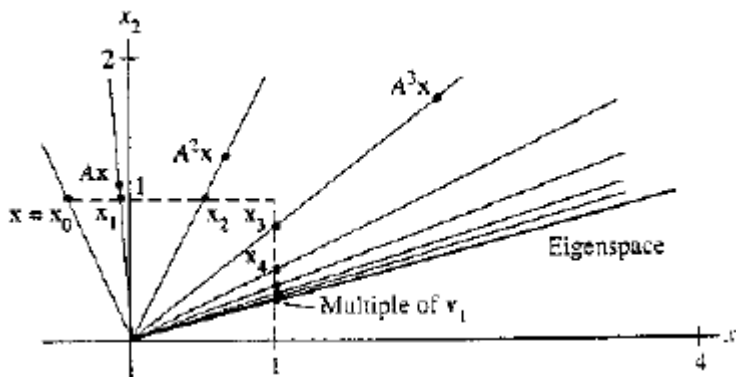


Figure 2 Scaled multiples of $x, Ax, A^2x, \dots, A^7x$.

Steps for finding the eigenvalue and the eigenvector

1. Select an initial vector x_0 whose largest entry is 1.
2. For $k=0,1,\dots$
 - a. Compute Ax_k
 - b. Let μ_k be an entry in Ax_k whose absolute value is as large as possible.
 - c. Compute $x_{k+1} = \left(\frac{1}{\mu_k}\right)Ax_k$

3. For almost all choices of x_0 , the sequence $\{\mu_k\}$ approaches the dominant eigenvalue, and the sequence $\{x_k\}$ approaches a dominant eigenvector.

Example 3

Apply the power method to $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$ with $x_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Stop when $k = 5$, and estimate the dominant eigenvalue and a corresponding eigenvector of A .

Solution

To begin, we calculate Ax_0 and identify the largest entry μ_0 in Ax_0 .

$$Ax_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \mu_0 = 5$$

Scale Ax_0 by $1/\mu_0$ to get x_1 , compute Ax_1 and identify the largest entry in Ax_1

$$x_1 = \frac{1}{\mu_0} Ax_0 = \frac{1}{5} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$Ax_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \mu_1 = 8$$

Scale Ax_1 by $1/\mu_1$ to get x_2 , compute Ax_2 , and identify the largest entry in Ax_2 .

$$x_2 = \frac{1}{\mu_1} Ax_1 = \frac{1}{8} \begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$Ax_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \mu_2 = 7.125$$

Scale Ax_2 by $1/\mu_2$ to get x_3 , and so on. The results of MATLAB calculations for the first five iterations are arranged in the Table 2.

The evidence from the Table 2 strongly suggests that $\{x_k\}$ approaches $(1, .2)$ and $\{\mu_k\}$ approaches 7. If so, then $(1, .2)$ is an eigenvector and 7 is the dominant eigenvalue. This

is easily verified by computing

$$A \begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ .2 \end{bmatrix}$$

TABLE 2 The Power Method for Example 2

k	0	1	2	3	4	5
x_k	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
Ax_k	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

Example 4

Find the first three iterations of the power method applied on the following matrices

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{use } x_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution**1st Iteration**

$$u_1 = Ax_0 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+0+0 \\ 0+0+2 \\ 0+0+2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_1 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = q_1 x_1$$

2nd Iteration

$$u_2 = Ax_1 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+1+0 \\ 0+4+2 \\ 0+1+2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_2 = \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix} = 6 \begin{bmatrix} \frac{1}{6} \\ 1 \\ \frac{1}{2} \end{bmatrix} = q_2 x_2$$

3rd Iteration

$$u_3 = Ax_2 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{6} \\ 1 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{6} + 1 + 0 \\ \frac{1}{3} + 4 + 1 \\ 0 + 1 + 1 \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{16}{3} \\ 2 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_3 = \begin{bmatrix} \frac{7}{6} \\ \frac{16}{3} \\ 2 \end{bmatrix} = \frac{16}{3} \begin{bmatrix} \frac{7}{32} \\ 1 \\ \frac{3}{8} \end{bmatrix} = q_3 x_3$$

Hence the largest eigenvalue after 3 iterations is $\frac{16}{3}$.

The corresponding eigenvector is $\begin{bmatrix} \frac{7}{32} \\ 1 \\ \frac{3}{8} \end{bmatrix}$

The inverse power method

1. Select an initial estimate sufficiently close to λ .
2. Select an initial vector x_0 whose large entry is 1.
3. For $k=0, 1, 2, \dots$

Solve $(A - \alpha I)y_k = x_k$.

Let μ_k be an entry in y_k whose absolute value is as large as possible.

Compute $v_k = \alpha + \left(\frac{1}{\mu_k}\right)$.

Compute $x_{k+1} = \left(\frac{1}{\mu_k}\right)y_k$

4. For almost all the choice of x_0 , the sequence $\{v_k\}$ approaches the eigenvalue λ of A , and the sequence $\{x_k\}$ approaches a corresponding eigenvector.

Example 5

It is not uncommon in some applications to need to know the smallest eigenvalue of a matrix A and to have at hand rough estimates of the eigenvalues. Suppose 21, 3.3, and 1.9 are estimates for the eigenvalues of the matrix A below. Find the smallest eigenvalue, accurate to six decimal places.

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

Solution

The two smallest eigenvalues seem close together, so we use the inverse power method for $A - 1.9I$. Results of a MATLAB calculation are shown in Table 3. Here x_0 was chosen arbitrarily, $y_k = (A - 1.9I)^{-1}x_k$, μ_k is the largest entry in y_k , $v_k = 1.9 + 1/\mu_k$, and $x_{k+1} = (1/\mu_k)y_k$.

Table 3: The Inverse Power Method

K	0	1	2	3	4
x_0	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5736 \\ 0.0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5054 \\ 0.0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5004 \\ 0.0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.50003 \\ 0.00002 \\ 1 \end{bmatrix}$
y_k	$\begin{bmatrix} 4.45 \\ 0.50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ 0.0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ 0.0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ 0.0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ 0.000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
v_k	2.03	2.0008	2.00005	2.000004	2.0000002

Therefore, we can say that eigenvalue is 2 from the matrix

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

Exercise

How can you tell that if a given vector X is a good approximation to an eigenvector of a matrix A : if it is, how would you estimate the corresponding eigenvalue? Experiment with

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \quad \text{And } X = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

Lecture 36

Revision

Revision of Previous Lectures

Lecture 37

Revision

Revision of Previous Lectures

Lecture 38

Inner Product

MCQs

If u and v are vectors in R^n , then we regard u and v as $n \times 1$ matrices. The transpose u^t is a $1 \times n$ matrix, and the matrix product $u^t v$ is a 1×1 matrix which we write as a single real number (a scalar) without brackets.

The number $u^t v$ is called the inner product of u and v . And often is written as $u.v$. This inner product is also referred to as a dot product.

$$\text{If } u = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} \text{ and } v = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$$

Then the inner product of u and v is

$$\begin{bmatrix} u_1 & u_2 & \cdot & \cdot & \cdot & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$$

Example 1

Compute $u.v$ and $v.u$ when $u = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

Solution

$$u = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} \text{ and } v = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$$

$$u^t = [2 \quad -5 \quad -1]$$

$$\begin{aligned} u.v = u^t v &= [2 \quad -5 \quad -1] \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = 2(3) + (-5)(2) + (-1)(-3) \\ &= 6 - 10 + 3 = -1 \end{aligned}$$

$$v^t = [3 \quad 2 \quad -3]$$

$$v \cdot u = v^t u = [3 \quad 2 \quad -3] \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = 3(2) + (2)(-5) + (-3)(-1) \\ = 6 - 10 + 3 = -1$$

Theorem

Let u , v and w be vectors in R^n , and let c be a scalar. Then

- $u \cdot v = v \cdot u$
- $(u + v) \cdot w = u \cdot w + v \cdot w$
- $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$
- $u \cdot u \geq 0$ and $u \cdot u = 0$ if and only if $u = 0$

Observation

$$(c_1 u_1 + c_2 u_2 + \dots + c_p u_p) \cdot w \\ = c_1(u_1 \cdot w) + c_2(u_2 \cdot w) + \dots + c_p(u_p \cdot w)$$

Length or Norm

The length or Norm of v is the nonnegative scalar $\|v\|$ defined by

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \\ \|v\|^2 = v \cdot v$$

Note: For any scalar c , $\|cv\| = |c| \|v\|$

Unit vector

A vector whose length is 1 is called a unit vector. If we divide a non-zero vector v by its length $\|v\|$, we obtain a unit vector u as

$$u = \frac{v}{\|v\|}$$

The length of u is $\|u\| = \frac{1}{\|v\|} \|v\| = 1$

Definition

The process of creating the unit vector u from v is sometimes called normalizing v and we say that u is in the same direction as v . In this case " u " is called the normalized vector.

Example 2

Let $v = (1, 2, 2, 0)$ in R^4 . Find a unit vector u in the same direction as v .

Solution

The length of v is given by

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + v_3^2 + v_4^2}$$

So,

$$\|v\| = \sqrt{1^2 + 2^2 + 2^2 + 0^2} = \sqrt{1+4+4+0} = \sqrt{9} = 3$$

The unit vector u in the direction of v is given as

$$u = \frac{1}{\|v\|} v = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix}$$

To check that $\|u\| = 1$

$$\|u\| = \sqrt{u \cdot u} = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{-2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + (0)^2} = \sqrt{\frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0} = 1$$

Example 3

Let W be the subspace of R^2 spanned by $X = (\frac{2}{3}, 1)$. Find a unit vector Z that is a basis for W .

Solution

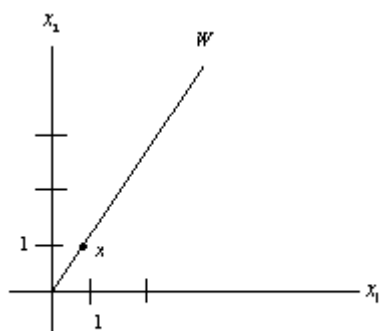
W consists of all multiples of x , as in Fig. 2(a). Any nonzero vector in W is a basis for W . To simplify the calculation, x is scaled to eliminate fractions. That is, multiply x by 3 to get

$$y = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

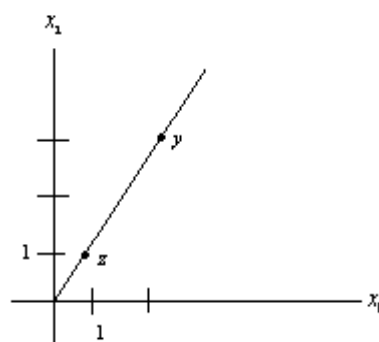
Now compute $\|y\|^2 = 2^2 + 3^2 = 13$, $\|y\| = \sqrt{13}$, and normalize y to get

$$z = \frac{1}{\sqrt{13}} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

See Fig. 2(b). Another unit vector is $(-2/\sqrt{13}, -3/\sqrt{13})$.



(a)



(b)

Figure 2 Normalizing a vector to produce a unit vector.**Definition**

For u and v vectors in R^n , the distance between u and v , written as $\text{dist}(u, v)$, is the length of the vector $u - v$. That is

$$\text{dist}(u, v) = \|u - v\|$$

Example 4

Compute the distance between the vectors $u = (7, 1)$ and $v = (3, 2)$

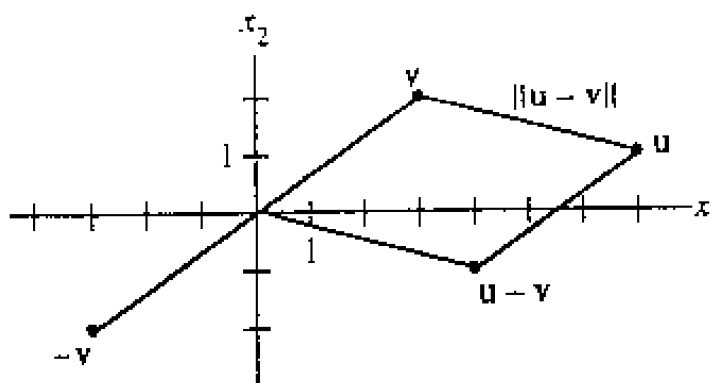
Solution

$$u - v = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7-3 \\ 1-2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\text{dist}(u, v) = \|u - v\| = \sqrt{(4)^2 + (-1)^2} = \sqrt{16+1} = \sqrt{17}$$

Law of Parallelogram of vectors

The vectors, u , v and $u - v$ are shown in the fig. below. When the vector $u - v$ is added to v , the result is u . Notice that the parallelogram in the fig. below shows that the distance from u to v is the same as the distance of $u - v$ to o .

**Example 5**

If $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$, then

$$\begin{aligned} \text{dist}(u, v) = \|u - v\| &= \sqrt{(u - v) \cdot (u - v)} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2} \end{aligned}$$

Definition

Two vectors u and v in R^n are orthogonal (to each other) if $u \cdot v = 0$

MCQS

Note

The zero vector is orthogonal to every vector in R^n because $0^t \cdot v = 0$ for all v in R^n .

The Pythagorean Theorem

Two vectors u and v are orthogonal if and only if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

Orthogonal Complements

The set of all vectors z that are orthogonal to w in W is called the orthogonal complement of W and is denoted by $W^\perp$

Example 6

Let W be a plane through the origin in R^3 , and let L be the line through the origin and perpendicular to W . If z and w are nonzero, z is on L , and w is in W , then the line segment from 0 to z is perpendicular to the line segment from 0 to w ; that is, $z \cdot w = 0$. So each vector on L is orthogonal to every w in W . In fact, L consists of all vectors that are orthogonal to the w 's in W , and W consists of all vectors orthogonal to the z 's in L . That is,

$$L = W^\perp \text{ and } W = L^\perp$$

Remarks

The following two facts about $W^\perp$, with W a subspace of $\mathbf{R}^n$, are needed later in the segment.

- (1) A vector x is in $W^\perp$ if and only if x is orthogonal to every vector in a set that spans W .
- (2) $W^\perp$ is a subspace of $\mathbf{R}^n$.

Theorem 3

Let A be $m \times n$ matrix. Then the orthogonal complement of the row space of A is the null space of A , and the orthogonal complement of the column space of A is the null space of A^T :
 $(\text{Row } A)^\perp = \text{Nul } A$, $(\text{Col } A)^\perp = \text{Nul } A^T$

Proof

The row-column rule for computing Ax shows that if x is in $\text{Nul } A$, then x is orthogonal to each row of A (with the rows treated as vectors in $\mathbf{R}^n$). Since the rows of A span the row space, x is orthogonal to $\text{Row } A$. Conversely, if x is orthogonal to $\text{Row } A$, then x is certainly orthogonal to each row of A , and hence $Ax = \mathbf{0}$. This proves the first statement. The second statement follows from the first by replacing A with A^T and using the fact that $\text{Col } A = \text{Row } A^T$.

Angles in $\mathbf{R}^2$ and $\mathbf{R}^3$

If u and v are nonzero vectors in either $\mathbf{R}^2$ or $\mathbf{R}^3$, then there is a nice connection between their inner product and the angle $\mathcal{G}$ between the two line segments from the origin to the points identified with u and v . The formula is

$$u \cdot v = \|u\| \|v\| \cos \mathcal{G} \quad (2)$$

To verify this formula for vectors in $\mathbf{R}^2$, consider the triangle shown in Fig. 7, with sides of length $\|u\|$, $\|v\|$, and $\|u-v\|$. By the law of cosines,

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\| \|v\| \cos \mathcal{G} \quad \text{MCQS}$$

which can be rearranged to produce

$$\begin{aligned} \|u\| \|v\| \cos \mathcal{G} &= \frac{1}{2} [\|u\|^2 + \|v\|^2 - \|u-v\|^2] \\ &= \frac{1}{2} [u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2] = u_1 v_1 + u_2 v_2 = u \cdot v \end{aligned}$$

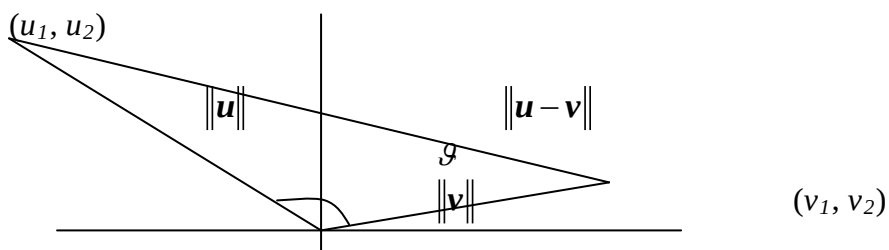


Figure 7 The angle between two vectors.

Example 7

Find the angle between the vectors $u = (1, -1, 2)$, $v = (2, 1, 0)$

Solution

$$u \cdot v = (1)(2) + (-1)(1) + (2)(0) = 2 - 1 + 0 = 1$$

And

$$\|u\| = \sqrt{(1)^2 + (-1)^2 + (2)^2} = \sqrt{1+1+4} = \sqrt{6}$$

$$\|v\| = \sqrt{(2)^2 + (1)^2 + (0)^2} = \sqrt{4+1} = \sqrt{5}$$

Angle between the two vectors is given by

$$\cos \theta = \frac{u \cdot v}{\|u\| \|v\|}$$

Putting the values, we get

$$\cos \theta = \frac{1}{\sqrt{6}\sqrt{5}} = \frac{1}{\sqrt{30}}$$

$$\cos \theta = \frac{1}{\sqrt{30}}$$

$$\theta = \cos^{-1} \frac{1}{\sqrt{30}} = 79.48^\circ$$

Exercises**Q.1**

Compute $u \cdot v$ and $v \cdot u$ when $u = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$ and $v = \begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix}$

Q.2

Let $v = (2, 1, 0, 3)$ in R^4 . Find a unit vector u in the direction opposite to that of v .

Q.3

Let W be the subspace of R^3 spanned by $X = (\frac{1}{2}, \frac{3}{2}, \frac{5}{2})$. Find a unit vector Z that is a basis for W .

Q.4

Compute the distance between the vectors $u = (1, 5, 7)$ and $v = (2, 3, 5)$.

Q.5

Find the angle between the vectors $u = (2, 1, 3)$, $v = (1, 0, 2)$.

Lecture No.39

Orthogonal and Orthonormal sets

Objectives

The objectives of the lecture are to learn about:

- Orthogonal Set.
- Orthogonal Basis.
- Unique representation of a vector as a linear combination of Basis vectors.
- Orthogonal Projection.
- Decomposition of a vector into sum of two vectors.
- Orthonormal Set.
- Orthonormal Basis.
- Some examples to verify the definitions and the statements of the theorems.

Orthogonal Set

Let $S = \{u_1, u_2, \dots, u_p\}$ be the set of non-zero vectors in R^n , is said to be an orthogonal set if all vectors in S are mutually orthogonal. That is $0 \notin S$ and $u_i \cdot u_j = 0 \forall i \neq j, i, j = 1, 2, \dots, p$.

Example

Show that $S = \{u_1, u_2, u_3\}$ is an orthogonal set. Where

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} -1 \\ 2 \\ 7 \\ 2 \end{bmatrix}.$$

Solution

To show that S is orthogonal, we show that each vector in S is orthogonal to other. That is

$$u_i \cdot u_j = 0 \forall i \neq j, i, j = 1, 2, 3.$$

For $i = 1, j = 2$

$$\begin{aligned} u_1 \cdot u_2 &= \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \\ &= -3 + 2 + 1 = 0 \end{aligned}$$

Which implies u_1 is orthogonal to u_2 .

For $i = 1, j = 3$

$$\begin{aligned} u_1 \cdot u_3 &= \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix} \\ &= \frac{-3}{2} - 2 + \frac{7}{2} \\ &= 0. \end{aligned}$$

Which implies u_1 is orthogonal to u_3 .

For $i = 2, j = 3$

$$\begin{aligned} u_2 \cdot u_3 &= \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix} \\ &= \frac{1}{2} - 4 + \frac{7}{2} \\ &= 0. \end{aligned}$$

Which implies u_2 is orthogonal to u_3 .

Thus $S = \{u_1, u_2, u_3\}$ is an orthogonal set.

Theorem

Suppose that $S = \{u_1, u_2, \dots, u_p\}$ is an orthogonal set of non-zero vectors in R^n and $W = \text{Span}\{u_1, u_2, \dots, u_p\}$. Then S is linearly independent set and a basis for W .

Proof

Suppose

$$0 = c_1 u_1 + c_2 u_2 + \dots + c_p u_p.$$

Where $c_1, c_2, \dots, c_p$ are scalars.

$$u_1 \cdot 0 = u_1 \cdot (c_1 u_1 + c_2 u_2 + \dots + c_p u_p)$$

$$\begin{aligned} 0 &= u_1 \cdot (c_1 u_1) + u_1 \cdot (c_2 u_2) + \dots + u_1 \cdot (c_p u_p) \\ &= c_1 (u_1 \cdot u_1) + c_2 (u_1 \cdot u_2) + \dots + c_p (u_1 \cdot u_p) \\ &= c_1 (u_1 \cdot u_1) \end{aligned}$$

Since S is orthogonal set, so, $u_1 \cdot u_2 + \dots + u_1 \cdot u_p = 0$ but $u_1 \cdot u_1 > 0$.

Therefore $c_1 = 0$. Similarly, it can be shown that $c_2 = c_3 = \dots = c_p = 0$

Therefore by definition $S = \{u_1, u_2, \dots, u_p\}$ is linearly independent set and by definition of basis is a basis for subspace W .

Example

If $S = \{u_1, u_2\}$ is an orthogonal set of non-zero vector in R^2 . Show that S is linearly independent set. Where

$$u_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } u_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Solution

To show that $S = \{u_1, u_2\}$ is linearly independent set, we show that the following vector equation

$$c_1 u_1 + c_2 u_2 = 0.$$

has only the trivial solution. i.e. $c_1 = c_2 = 0$.

$$c_1 u_1 + c_2 u_2 = 0$$

$$c_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3c_1 \\ c_1 \end{bmatrix} + \begin{bmatrix} -c_2 \\ 3c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$3c_1 - c_2 = 0$$

$$c_1 + 3c_2 = 0$$

Solve them simultaneously, gives

$$c_1 = c_2 = 0.$$

Therefore if S is an orthogonal set then it is linearly independent.

Orthogonal basis

Let $S = \{u_1, u_2, \dots, u_p\}$ be a basis for a subspace W of R^n , is also an orthogonal basis if S is an orthogonal set.

Theorem

If $S = \{u_1, u_2, \dots, u_p\}$ is an orthogonal basis for a subspace W of R^n . Then each y in W can be uniquely expressed as a linear combination of $u_1, u_2, \dots, u_p$. That is

$$y = c_1 u_1 + c_2 u_2 + \dots + c_p u_p.$$

Where

$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

Proof

$$\begin{aligned} y \cdot u_1 &= (c_1 u_1 + c_2 u_2 + \dots + c_p u_p) \cdot u_1 \\ &= (c_1 u_1) \cdot u_1 + (c_2 u_2) \cdot u_1 + \dots + (c_p u_p) \cdot u_1 \\ &= c_1 (u_1 \cdot u_1) + c_2 (u_1 \cdot u_2) + \dots + c_p (u_1 \cdot u_p) \\ &= c_1 (u_1 \cdot u_1). \end{aligned}$$

Since S is orthogonal set, so, $u_1 \cdot u_2 + \dots + u_1 \cdot u_p = 0$ but $u_1 \cdot u_1 > 0$.

Hence

$$c_1 = \frac{y \cdot u_1}{u_1 \cdot u_1} \text{ and similarly } c_2 = \frac{y \cdot u_2}{u_2 \cdot u_2}, \dots, c_p = \frac{y \cdot u_p}{u_p \cdot u_p}.$$

Example

The set $S = \{u_1, u_2, u_3\}$ as in first example is an orthogonal basis for R^3 . Express y as a linear combination of the vectors in S . Where

$$y = \begin{bmatrix} 6 & 1 & -8 \end{bmatrix}^T$$

Solution

We want to write

$$y = c_1 u_1 + c_2 u_2 + c_3 u_3$$

Where c_1, c_2 and c_3 are to be determined.

By the above theorem

$$\begin{aligned} c_1 &= \frac{y \cdot u_1}{u_1 \cdot u_1} \\ &= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}} = \frac{11}{11} = 1 \end{aligned}$$

$$\begin{aligned}
 c_2 &= \frac{y \cdot u_2}{u_2 \cdot u_2} \\
 &= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}} = \frac{-12}{6} = -2
 \end{aligned}$$

And

$$\begin{aligned}
 c_3 &= \frac{y \cdot u_3}{u_3 \cdot u_3} \\
 &= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 7 \\ 2 \end{bmatrix}}{\begin{bmatrix} -1 \\ 2 \\ -2 \\ 7 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ -2 \\ 7 \\ 2 \end{bmatrix}} = \frac{-33}{33/2} = -2
 \end{aligned}$$

Hence

$$y = u_1 - 2u_2 - 2u_3.$$

Example

The set $S = \{u_1, u_2, u_3\}$ is an orthogonal basis for R^3 . Write y as a linear combination of the vectors in S . Where

$$y = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}, u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution

We want to write

$$y = c_1 u_1 + c_2 u_2 + c_3 u_3$$

Where c_1, c_2 and c_3 are to be determined.

By the above theorem

$$c_1 = \frac{y \cdot u_1}{u_1 \cdot u_1} = \frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} = \frac{3-7+0}{1+1+0} = -2$$

$$c_2 = \frac{y \cdot u_2}{u_2 \cdot u_2} = \frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} = \frac{3+7+0}{1+1} = 5$$

And

$$c_3 = \frac{y \cdot u_3}{u_3 \cdot u_3} = \frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} = \frac{4}{1} = 4$$

Hence

$$y = -2u_1 + 5u_2 + 4u_3.$$

Exercise

The set $S = \{u_1, u_2, u_3\}$ is an orthogonal basis for R^3 . Write y as a linear combination of the vectors in S . where

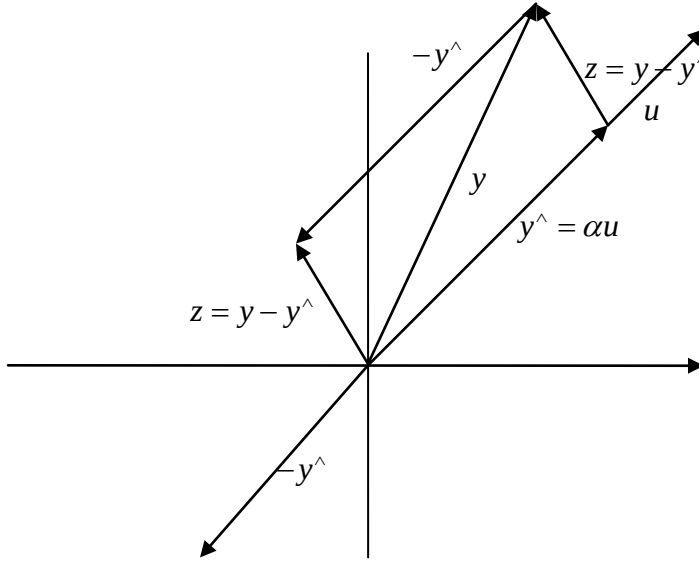
$$y = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, u_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \\ -1 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} \frac{8}{3} \\ \frac{16}{3} \\ \frac{8}{3} \end{bmatrix}$$

An Orthogonal Projection (Decomposition of a vector into the sum of two vectors)

Decomposition of a non- zero vector $y \in R^n$ into the sum of two vectors in such a way, one is multiple of $u \in R^n$ and the other orthogonal to u . That is

$$y = y^\wedge + z$$

Where $y^\wedge = \alpha u$ for some scalar α and z is orthogonal to u .



In the above figure a vector y is decomposed into two vectors $z = y - y^\wedge$ and $y^\wedge = \alpha u$. Clearly it can be seen that $z = y - y^\wedge$ is orthogonal to u and $y^\wedge = \alpha u$ is a multiple of u . Since $z = y - y^\wedge$ is orthogonal to u .

Therefore

$$z \cdot u = 0$$

$$(y - y^\wedge) \cdot u = 0$$

$$(y - \alpha u) \cdot u = 0$$

$$y \cdot u - \alpha(u \cdot u) = 0$$

$$\Rightarrow \alpha = \frac{y \cdot u}{u \cdot u}$$

And

$$z = y - y^\wedge$$

$$= y - \frac{y \cdot u}{u \cdot u} u$$

Hence

$$\hat{y} = \frac{y \cdot u}{u \cdot u} u, \text{ which is an orthogonal projection of } y \text{ onto } u.$$

And

$$\begin{aligned} z &= y - \hat{y} \\ &= y - \frac{y \cdot u}{u \cdot u} u \end{aligned} \quad \text{is a component of } y$$

Example

Let $y = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

Find the orthogonal projection of y onto u . Then write y as a sum of two orthogonal vectors, one in $\text{span}\{u\}$ and one orthogonal to u .

Solution

Compute

$$\begin{aligned} y \cdot u &= \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40 \\ u \cdot u &= \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20 \end{aligned}$$

The orthogonal projection of y onto u is $\hat{y} = \frac{y \cdot u}{u \cdot u} u = \frac{40}{20} u = 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$ and the

component of y orthogonal to u is $y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

The sum of these two vectors is y . That is, $\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$
 $y \quad \hat{y} \quad (y - \hat{y})$

This decomposition of y is illustrated in Fig. 3. **Note:** If the calculations above are correct, then $\{\hat{y}, y - \hat{y}\}$ will be an orthogonal set. As a check, compute

$$\hat{y} \cdot (y - \hat{y}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

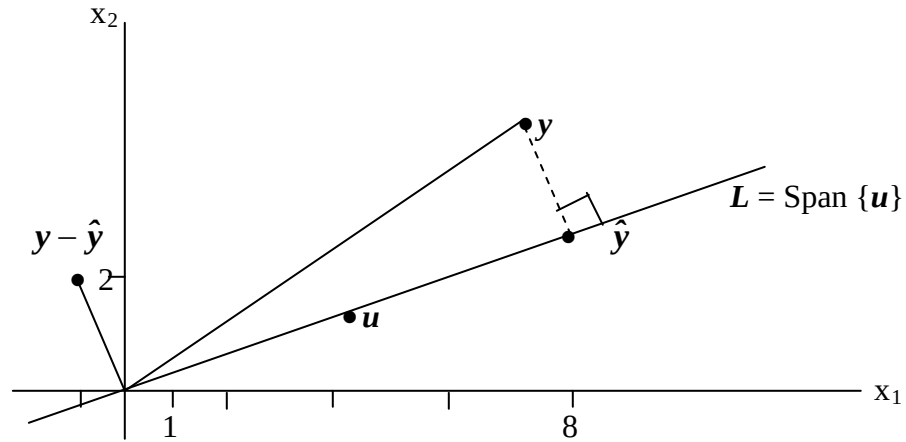
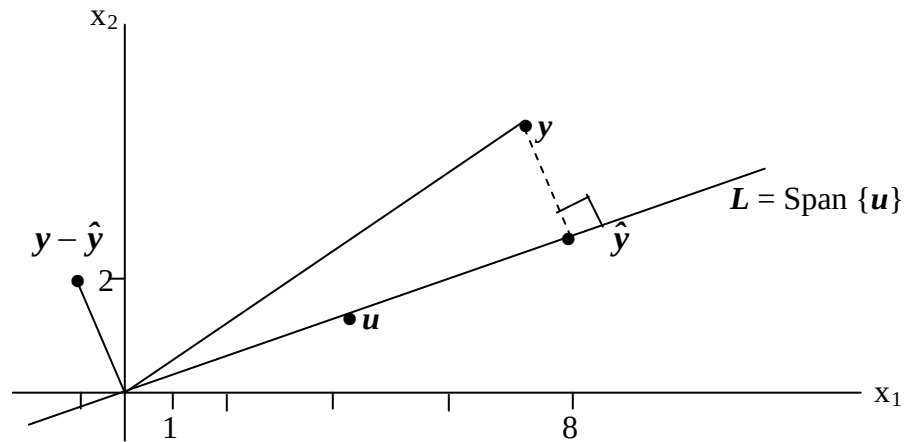


Figure 3 The orthogonal projection of y on to a line through the origin.

Example

Find the distance in figure below from y to L .



Solution

The distance from y to L is the length of the perpendicular line segment from y to the orthogonal projection $\hat{y}$.

The length equals the length of $y - \hat{y}$.

This distance is

$$\|y - \hat{y}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

Example

Decompose $\mathbf{y} = (-3, -4)$ into two vectors $\hat{\mathbf{y}}$ and $\mathbf{z}$, where $\hat{\mathbf{y}}$ is a multiple of $\mathbf{u} = (-3, 1)$ and $\mathbf{z}$ is orthogonal to $\mathbf{u}$. Also prove that $\hat{\mathbf{y}} \cdot \mathbf{z} = 0$

Solution

It is very much clear that $\hat{\mathbf{y}}$ is an orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ and it is calculated by applying the following formula

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

$$= \frac{\begin{bmatrix} -3 \\ -4 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix}}{\begin{bmatrix} -3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \frac{9 - 4}{9 + 1} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -3 \\ -4 \end{bmatrix} - \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -3 + \frac{3}{2} \\ -4 - \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ -\frac{9}{2} \end{bmatrix}$$

So,

$$\hat{\mathbf{y}} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} \text{ and } \mathbf{z} = \begin{bmatrix} -\frac{3}{2} \\ -\frac{9}{2} \end{bmatrix}$$

Now

$$\begin{aligned}
 y^\wedge \cdot z &= \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} -\frac{3}{2} \\ -\frac{9}{2} \end{bmatrix} \\
 &= \frac{9}{4} - \frac{9}{4} \\
 &= 0
 \end{aligned}$$

Therefore, $y^\wedge$ is orthogonal to z

Exercise

Find the orthogonal projection of a vector $y = (-3, 2)$ onto $u = (2, 1)$. Also prove that $y = y^\wedge + z$, where $y^\wedge$ a multiple of u and z is orthogonal to u .

Orthonormal Set

Let $S = \{u_1, u_2, \dots, u_p\}$ be the set of non-zero vectors in R^n , is said to be an orthonormal set if S is an orthogonal set of unit vectors.

Example

Show that $S = \{u_1, u_2, u_3\}$ is an orthonormal set. Where

$$u_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ -\frac{1}{\sqrt{5}} \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Solution

To show that S is an orthonormal set, we show that it is an orthogonal set of unit vectors.

It can be easily prove that S is an orthogonal set because $u_i \cdot u_j = 0 \forall i \neq j, i, j = 1, 2, 3$.

Furthermore

$$u_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, u_3 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

$$\begin{aligned} u_1 \cdot u_1 &= \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix} \cdot \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix} \\ &= \frac{4}{5} + 0 + \frac{1}{5} \\ &= 1 \end{aligned}$$

$$\begin{aligned} u_2 \cdot u_2 &= \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \\ &= 0 + 1 + 0 \\ &= 1 \end{aligned}$$

And

$$\begin{aligned} u_3 \cdot u_3 &= \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix} \\ &= \frac{1}{5} + \frac{4}{5} \\ &= 1 \end{aligned}$$

Hence

$S = \{u_1, u_2, u_3\}$ is an orthonormal set.

Orthonormal basis

Let $S = \{u_1, u_2, \dots, u_p\}$ be a basis for a subspace W of R^n , is also an orthonormal basis if S is an orthonormal set.

Example

Show that $S = \{u_1, u_2, u_3\}$ is an orthonormal basis of R^3 , where

$$u_1 = \begin{bmatrix} \frac{3}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \end{bmatrix}, u_2 = \begin{bmatrix} \frac{-1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} \frac{-1}{\sqrt{66}} \\ \frac{-4}{\sqrt{66}} \\ \frac{7}{\sqrt{66}} \end{bmatrix}.$$

Solution

To show that $S = \{u_1, u_2, u_3\}$ is an orthonormal basis, it is sufficient to show that it is an orthogonal set of unit vectors. That is

$$u_i \cdot u_j = 0 \forall i \neq j, i, j = 1, 2, 3.$$

And

$$u_i \cdot u_i = 1 \forall i = 1, 2, 3.$$

Clearly it can be seen that

$$u_1 \cdot u_2 = 0,$$

$$u_1 \cdot u_3 = 0$$

And

$$u_2 \cdot u_3 = 0.$$

Furthermore

$$u_1 \cdot u_1 = 1,$$

$$u_2 \cdot u_2 = 1$$

And

$$u_3 \cdot u_3 = 1.$$

Hence S is an orthonormal basis of R^3 .

Theorem

A $m \times n$ matrix U has orthonormal columns if and only if $U^t U = I$

Proof

Keep in mind that in an if and only if statement, one part depends on the other, so, each part is proved separately. That is, we consider one part and then prove the other part with the help of that assumed part.

Before proving both sides of the statements, we have to do some extra work which is necessary for the better understanding.

Let $u_1, u_2, \dots, u_m$ be the columns of U . Then U can be written in matrix form as

$$U = [u_1 \ u_2 \ u_3 \ \dots \ u_m]$$

Taking transpose, it becomes

$$U^t = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ \vdots \\ u_m^t \end{bmatrix}$$

$$U^t U = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ \vdots \\ u_m^t \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 & \dots & u_m \end{bmatrix} = \begin{bmatrix} u_1^t u_1 & u_1^t u_2 & u_1^t u_3 & \dots & u_1^t u_m \\ u_2^t u_1 & u_2^t u_2 & u_2^t u_3 & \dots & u_2^t u_m \\ u_3^t u_1 & u_3^t u_2 & u_3^t u_3 & \dots & u_3^t u_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_m^t u_1 & u_m^t u_2 & u_m^t u_3 & \dots & u_m^t u_m \end{bmatrix}$$

As $u \cdot v = v^t u$

Therefore

$$U^t U = \begin{bmatrix} u_1 \cdot u_1 & u_1 \cdot u_2 & u_1 \cdot u_3 & \dots & u_1 \cdot u_m \\ u_2 \cdot u_1 & u_2 \cdot u_2 & u_2 \cdot u_3 & \dots & u_2 \cdot u_m \\ u_3 \cdot u_1 & u_3 \cdot u_2 & u_3 \cdot u_3 & \dots & u_3 \cdot u_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_m \cdot u_1 & u_m \cdot u_2 & u_m \cdot u_3 & \dots & u_m \cdot u_m \end{bmatrix}$$

Now, we come to prove the theorem.

First suppose that $U^t U = I$, and we prove that columns of U are orthonormal.

Since, we assume that

$$\begin{aligned}
 U^t U &= \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \dots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \dots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \dots & u_3.u_m \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \dots & u_m.u_m \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}
 \end{aligned}$$

Clearly, it can be seen that

$$u_i . u_j = 0 \text{ for } i \neq j \quad i, j = 1, 2, \dots, m$$

and

$$u_i . u_j = 1 \text{ for } i = j \quad i, j = 1, 2, \dots, m$$

Therefore, columns of U are orthonormal.

Next suppose that the columns of U are orthonormal and we will show that $U^t U = I$.

Since we assume that columns of U are orthonormal, so, we can write

$$u_i . u_j = 0 \text{ for } i \neq j \quad i, j = 1, 2, \dots, m$$

and

$$u_i . u_j = 1 \text{ for } i = j \quad i, j = 1, 2, \dots, m$$

$$\text{Hence, } U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \dots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \dots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \dots & u_3.u_m \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \dots & u_m.u_m \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

That is

$$U^t U = I.$$

Which is our required result.

Exercise

Prove that the following matrices have orthonormal columns using above theorem.

- (1) $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
- (2) $\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$
- (3) $\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$

Solution (1)

Let

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

$$U^t = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then

$$\begin{aligned} U^t U &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

$$U^t U = I$$

Therefore, by the above theorem, U has orthonormal columns.

(2) And (3) are left for reader.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let x and y be in R^n . Then

- a) $\|Ux\| = \|x\|$
- b) $(Ux) \cdot (Uy) = x \cdot y$
- c) $(Ux) \cdot (Uy) = 0$ iff $x \cdot y = 0$

Example

$$\text{Let } U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \text{ and } X = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$$

Verify that $\|Ux\| = \|x\|$

Solution

Notice that U has orthonormal columns and

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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

$$\|Ux\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|x\| = \sqrt{2+9} = \sqrt{11}$$

Lecture No.40 Orthogonal Decomposition

Objectives

The objectives of the lecture are to learn about:

- Orthogonal Decomposition Theorem.
- Best Approximation Theorem.

Orthogonal Projection

The orthogonal projection of a point in R^2 onto a line through the origin has an important analogue in R^n .

That is given a vector Y and a subspace W in R^n , there is a vector $\hat{y}$ in W such that

- 1) $\hat{y}$ is the unique vector in W for which $y - \hat{y}$ is orthogonal to W , and
- 2) $\hat{y}$ is the unique vector in W closest to y .



We observe that whenever a vector y is written as a linear combination of vectors $u_1, u_2, \dots, u_n$ in a basis of R^n , the terms in the sum for y can be grouped into two parts so that y can be written as $y = z_1 + z_2$, where z_1 is a linear combination of some of the u_i 's, and z_2 is a linear combination of the rest of the u_i 's. This idea is particularly useful when $\{u_1, u_2, \dots, u_n\}$ is an orthogonal basis.

Example 1

Let $\{u_1, u_2, \dots, u_5\}$ be an orthogonal basis for R^5 and let $y = c_1 u_1 + c_2 u_2 + \dots + c_5 u_5$.

Consider the subspace $W = \text{Span}\{u_1, u_2\}$ and write y as the sum of a vector z_1 in W and a vector z_2 in $W^\perp$.

Solution

Write

$$y = \underbrace{c_1u_1 + c_2u_2}_{z_1} + \underbrace{c_3u_3 + c_4u_4 + c_5u_5}_{z_2}$$

where $z_1 = c_1u_1 + c_2u_2$ is in Span of $\{u_1, u_2\}$, and $z_2 = c_3u_3 + c_4u_4 + c_5u_5$ is in Span of $\{u_3, u_4, u_5\}$.

To show that z_2 is in $W^\perp$, it suffices to show that z_2 is orthogonal to the vectors in the basis $\{u_1, u_2\}$ for W . Using properties of the inner product, compute

$$z_2 \cdot u_1 = (c_3u_3 + c_4u_4 + c_5u_5) \cdot u_1 = c_3u_3 \cdot u_1 + c_4u_4 \cdot u_1 + c_5u_5 \cdot u_1 = 0$$

since u_1 is orthogonal to u_3, u_4 , and u_5 , a similar calculation shows that $z_2 \cdot u_2 = 0$

Thus, z_2 is in $W^\perp$.

Orthogonal decomposition theorem

Let W be a subspace of R^n , then each y in R^n can be written uniquely in the form

$$y = \hat{y} + z$$

Where $\hat{y} \in W$ and $z \in W^\perp$

Furthermore, if $\{u_1, u_2, \dots, u_p\}$ is any orthogonal basis for W , then

$$\hat{y} = c_1u_1 + c_2u_2 + \dots + c_nu_n, \text{ where } c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

Proof

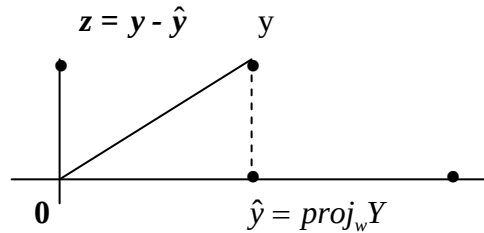


Fig: Orthogonal projection of y on to W.

Firstly, we show that $\hat{y} \in W, z \in W^\perp$. Then we will show that $y = \hat{y} + z$ can be represented in a unique way.

Suppose W is a subspace of R^n and let $\{u_1, u_2, \dots, u_p\}$ be an orthogonal basis for W .

$$\text{As } \hat{y} = c_1u_1 + c_2u_2 + \dots + c_pu_p \text{ where } c_j = \frac{y \cdot u_j}{u_j \cdot u_j}$$

Since $\{u_1, u_2, \dots, u_p\}$ is the basis for W and $y^\wedge$ is written as a linear combination of these basis vectors. Therefore, by definition of basis $y^\wedge \in W$.

Now, we will show that $z = y - y^\wedge \in W^\perp$. For this it is sufficient to show that $z \perp u_j$ for each $j = 1, 2, \dots, p$.

Let $u_1 \in W$ be an arbitrary vector.

$$\begin{aligned}
 z \cdot u_1 &= (y - y^\wedge) \cdot u_1 \\
 &= y \cdot u_1 - y^\wedge \cdot u_1 \\
 &= y \cdot u_1 - (c_1 u_1 + c_2 u_2 + \dots + c_p u_p) \cdot u_1 \\
 &= y \cdot u_1 - c_1 (u_1 \cdot u_1) - c_2 (u_2 \cdot u_1) - \dots - c_p (u_p \cdot u_1) \\
 &= y \cdot u_1 - c_1 (u_1 \cdot u_1) \quad \text{where } u_j \cdot u_1 = 0, j = 2, 3, \dots, p \\
 &= y \cdot u_1 - \frac{y \cdot u_1}{u_1 \cdot u_1} (u_1 \cdot u_1) \\
 &= y \cdot u_1 - y \cdot u_1 \\
 &= 0
 \end{aligned}$$

Therefore, $z \perp u_1$.

Since u_1 is an arbitrary vector, therefore $z \perp u_j$ for $j = 1, 2, \dots, p$.

Hence by definition of $W^\perp$, $z \in W^\perp$.

Now, we must show that $y = y^\wedge + z$ is unique by contradiction.

Let $y = y^\wedge + z$ and $y = y_1^\wedge + z_1$, where $y^\wedge, y_1^\wedge \in W$ and $z, z_1 \in W^\perp$, also $z \neq z_1$ and $y^\wedge \neq y_1^\wedge$. Since above representations for y are equal, that is

$$\begin{aligned}
 y^\wedge + z &= y_1^\wedge + z_1 \\
 \Rightarrow y^\wedge - y_1^\wedge &= z_1 - z
 \end{aligned}$$

Let

$$s = y^\wedge - y_1^\wedge$$

Then

$$s = z_1 - z$$

Since W is a subspace, therefore, by closure property

$$s = y^\wedge - y_1^\wedge \in W$$

Furthermore, $W^\perp$ is also a subspace, therefore by closure property

$$s = z_1 - z \in W^\perp$$

Since

$s \in W$ and $s \in W^\perp$. Therefore by definition $s \perp s$

That is $s \cdot s = 0$

Therefore

$$s = y^\wedge - y_1^\wedge = 0$$

$$\Rightarrow y^\wedge = y_1^\wedge$$

Also

$$z_1 = z$$

This shows that representation is unique.

Example

$$\text{Let } u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \text{ and } y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Observe that $\{u_1, u_2\}$ is an orthogonal basis for $W = \text{span}\{u_1, u_2\}$, write y as the sum of a vector in W and a vector orthogonal to W .

Solution

Since $y^\wedge \in W$, therefore $y^\wedge$ can be written as following:

$$\begin{aligned} y^\wedge &= c_1 u_1 + c_2 u_2 \\ &= \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 \\ &= \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -2 \\ 10 \\ 1 \end{bmatrix} \\ y - y^\wedge &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix} \\ &= 7/5 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \end{aligned}$$

Above theorem ensures that $y - y^\wedge$ is in $W^\perp$.

You can also verify by $(y - y^\wedge) \cdot u_1 = 0$ and $(y - y^\wedge) \cdot u_2 = 0$.

The desired decomposition of y is

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Example

Let $W = \text{span}\{u_1, u_2\}$, where $u_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$

Decompose $y = \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix}$ into two vectors; one in W and one in $W^\perp$. Also verify that these

two vectors are orthogonal.

Solution

Let $y^\wedge \in W$ and $z = y - y^\wedge \in W^\perp$.

Since $y^\wedge \in W$, therefore $y^\wedge$ can be written as following:

$$\begin{aligned} y^\wedge &= c_1 u_1 + c_2 u_2 \\ &= \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 \\ &= \frac{9}{7} \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + \frac{3}{7} \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} \\ y^\wedge &= \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} \end{aligned}$$

Now

$$\begin{aligned} z = y - y^\wedge &= \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} \\ z &= \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \end{aligned}$$

Now we show that $z \perp y^\wedge$, i.e. $z \cdot y^\wedge = 0$

$$z \cdot y^\wedge = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} = 0$$

Therefore $z \perp y^\wedge$.

Exercise

Let $W = \text{span}\{u_1, u_2\}$, where $u_1 = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$

Write $y = \begin{bmatrix} 6 \\ -8 \\ 12 \end{bmatrix}$ as a sum of two vectors; one in W and one in $W^\perp$. Also verify that these

two vectors are orthogonal.

Best Approximation Theorem

Let W is a finite dimensional subspace of an inner product space V and y is any vector in V . The best approximation to y from W is then $\text{Proj}_W y$, i.e for every w (that is not $\text{Proj}_W y$) in W , we have

$$\|y - \text{Proj}_W y\| < \|y - w\|.$$

Example

Let $W = \text{span}\{u_1, u_2\}$, where $u_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$, $u_2 = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix}$. Then using above

theorem, find the distance from y to W .

Solution

Using above theorem the distance from y to W is calculated using the following formula

$$\|y - \text{Proj}_W y\| = \|y - y^\wedge\|$$

Since, we have already calculated

$$y - y^\wedge = \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

$$\text{So } \|y - y^\wedge\| = \sqrt{6}$$

Example

The distance from a point y in R^n to a subspace W is defined as the distance from y to the nearest point in W .

Find the distance from y to $W = \text{span}\{u_1, u_2\}$, where

$$y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

By the Best Approximation Theorem, the distance from y to W is $\|y - \hat{y}\|$, where $\hat{y} = \text{proj}_W y$. Since, $\{u_1, u_2\}$ is an orthogonal basis for W , we have

$$\hat{y} = \frac{15}{30}u_1 + \frac{-21}{6}u_2 = \frac{1}{2} \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$y - \hat{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|y - \hat{y}\|^2 = 3^2 + 6^2 = 45$$

The distance from y to W is $\sqrt{45} = 3\sqrt{5}$.

Theorem

If $\{u_1, u_2, \dots, u_p\}$ is an orthonormal basis for a subspace W of R^n , then

$$\text{Proj}_W y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \dots + (y \cdot u_p)u_p$$

$$\text{If } U = [u_1 \ u_2 \ \dots \ u_p]$$

$$\text{then } \text{Proj}_W y = UU^T y \quad \forall y \text{ in } R^n$$

Example

$$\text{Let } u_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, y = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$$

and $W = \text{span}\{u_1, u_2\}$. Use the fact that u_1 and u_2 are orthogonal to compute $\text{Proj}_W y$.

Solution

$$\begin{aligned} \text{Pr } o j_w y &= \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 \\ &= \frac{88}{66} u_1 + \frac{-2}{6} u_2 \end{aligned}$$

$$= \frac{4}{3} \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = y$$

In this case, y happens to be a linear combination of u_1 and u_2 . So y is in W . The closest point in W to y is y itself.

Lecture # 41**Orthogonal basis, Gram-Schmidt Process, Orthonormal basis****Example 1**

Let $W = \text{Span} \{x_1, x_2\}$, where

$$x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Find the orthogonal basis $\{v_1, v_2\}$ for W .

Solution

Let P be a projection of x_2 on to x_1 . The component of x_2 orthogonal to x_1 is $x_2 - P$, which is in W as it is formed from x_2 and a multiple of x_1 .

Let $v_1 = x_1$ and compute

$$v_2 = x_2 - P = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Thus, $\{v_1, v_2\}$ is an orthogonal set of nonzero vectors in W , $\dim W = 2$ and $\{v_1, v_2\}$ is a basis of W .

Example 2

For the given basis of a subspace $W = \text{Span} \{x_1, x_2\}$,

$$x_1 = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$$

Find the orthogonal basis $\{v_1, v_2\}$ for W .

Solution

Set $v_1 = x_1$ and compute

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix} - \frac{10}{20} \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix}$$

Thus, an orthogonal basis for W is $\left\{ \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix} \right\}$

Theorem

Given a basis $\{x_1, \dots, x_p\}$ for a subspace W of R^n . Define

$$v_1 = x_1 \quad v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \dots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$$

Then $\{v_1, \dots, v_p\}$ is an orthogonal basis for W .

In addition

$$\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{x_1, \dots, x_k\} \text{ for } 1 \leq k \leq p$$

Example 3

The following vectors $\{x_1, x_2, x_3\}$ are linearly independent

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Construct an orthogonal basis for W by Gram-Schmidt Process.

Solution

To construct orthogonal basis we have to perform the following steps.

Step 1 Let $v_1 = x_1$

Step 2

$$\text{Let } v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

Since
$$v_2 = x_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Step 3

$$v_3 = x_3 - \left[\frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \right] = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \left[\frac{2}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right] = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

$$v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Thus, $\{v_1, v_2, v_3\}$ is an orthogonal set of nonzero vectors in W .

Example 4

Find an orthogonal basis for the column space of the following matrix by Gram-Schmidt Process.

$$\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$$

Solution

Name the columns of the above matrix as x_1, x_2, x_3 and perform the Gram-Schmidt Process on these vectors.

$$x_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix}, x_3 = \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix}$$

Set $v_1 = x_1$

$$\begin{aligned} v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\ &= \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix} - (-3) \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\ &= \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} - \frac{5}{2} \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix} \\ &\quad \left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix} \right\} \end{aligned}$$

Thus, orthogonal basis is

Example 5

Find the orthonormal basis of the subspace spanned by the following vectors.

$$x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Solution

Since from example # 1, we have

$$v_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Orthonormal Basis

$$u_1 = \frac{1}{\|v_1\|} v_1 = \frac{1}{\sqrt{45}} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}, u_2 = \frac{1}{\|v_2\|} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example 6

Find the orthonormal basis of the subspace spanned by the following vectors.

$$x_1 = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$$

Solution

Firstly we find v_1 and v_2 by Gram-Schmidt Process as

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

Set $v_1 = x_1$

$$v_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{15}{30} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -5/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$$

$$\text{Now } \|v_2\| = \frac{1}{\sqrt{54}} = \frac{1}{3\sqrt{6}} \text{ and since } \|v_1\| = \sqrt{30}$$

Thus, the orthonormal basis for W is

$$\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|} \right\} = \left\{ \begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \right\}$$

Theorem

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $\text{Col } A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Example 7

Find a QR factorization of matrix A

$$A = \begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$$

Solution

Firstly find the orthonormal basis by applying Gram Schmidt process on the columns of A . We get the following matrix Q .

$$Q = \begin{bmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{bmatrix}$$

$$\text{Now } R = Q^T A = \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

Verify that $A=QR$.

Theorem

If $\{u_1, \dots, u_p\}$ is an orthonormal basis for a subspace W of R^n , then

$$\text{proj}_W y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \dots + (y \cdot u_p)u_p$$

If $U = [u_1 \ u_2 \ \dots \ u_p]$,

then $\text{proj}_W y = UU^T y \ \forall y \text{ in } R^n$

The Orthogonal Decomposition Theorem

Let W be a subspace of R^n . Then each y in R^n can be written uniquely in the form

$$y = \hat{y} + z$$

where $\hat{y}$ is in W and z is in

In fact, if $\{u_1, \dots, u_p\}$ is any orthogonal basis of W , then

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \dots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$$

and $z = y - \hat{y}$. The vector $\hat{y}$ is called the orthogonal projection of y onto W and is often written as $\text{proj}_W y$.

Best Approximation Theorem

Let W be a subspace of R^n , y is any vector in R^n and $\hat{y}$ the orthogonal projection of y onto W . Then $\hat{y}$ is the closest point in W to y , in the sense that

for all v in W distinct from $\hat{y}$.

$$\|y - \hat{y}\| < \|y - v\|$$

The vector $\hat{y}$ in this theorem is called the best approximation to y by elements of W .

Exercise 1

Let $W = \text{Span} \{x_1, x_2\}$, where

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

Construct an orthonormal basis for W .

Exercise 2

Find an orthogonal basis for the column space of the following matrix by Gram-Schmidt Process.

$$\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & -5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$$

Exercise 3

Find a QR factorization of

$$A = \begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$$

Lecture # 42

Least Square Solution

Best Approximation Theorem

Let W be a subspace of R^n , y be any vector in R^n and $\hat{y}$ the orthogonal projection of y onto W . Then $\hat{y}$ is the closest point in W to y , in the sense that $\|y - \hat{y}\| < \|y - v\|$ for all v in W distinct from $\hat{y}$

The vector $\hat{y}$ in this theorem is called the best approximation to y by elements of W .

Least-squares solution

The most important aspect of the least-squares problem is that no matter what “ x ” we select, the **vector Ax will necessarily be in the column space $\text{Col } A$** . So we seek an x that makes Ax the closest point in **$\text{Col } A$ to b** . Of course, if b happens to be in $\text{Col } A$, then b is Ax for some x and such an x is a “**least-squares solution**.”

Solution of the General Least-Squares Problem

Given A and b as above, apply the Best Approximation Theorem stated above to the subspace $\text{Col } A$. Let $\hat{b} = \text{proj}_{\text{Col } A} b$

Since $\hat{b}$ is in the column space of A , the equation $Ax = \hat{b}$ is consistent, and there is an $\hat{x}$ in R^n such that

$$A\hat{x} = \hat{b} \quad (1)$$

Since $\hat{b}$ is the closest point in $\text{Col } A$ to b , a vector $\hat{x}$ is a least-squares solution of $Ax = b$ if and only if $\hat{x}$ satisfies $A\hat{x} = \hat{b}$. Such an $\hat{x}$ in R^n is a list of weights that will build $\hat{b}$ out of the columns of A .

Normal equations for $\hat{x}$

Suppose that $\hat{x}$ satisfies $A\hat{x} = \hat{b}$. By the Orthogonal Decomposition Theorem the projection $\hat{b}$ has the property that $b - \hat{b}$ is orthogonal to $\text{Col } A$, so $b - A\hat{x}$ is orthogonal to each column of A . If a_j is any column of A , then $a_j \cdot (b - A\hat{x}) = 0$, and $a_j^T (b - A\hat{x}) = 0$.

Since each a_j^T is a row of A^T ,

$$A^T (b - A\hat{x}) = 0 \quad (2)$$

$$A^T b - A^T A\hat{x} = 0$$

$$\mathbf{A^T A \hat{x} = A^T b} \quad (3)$$

The matrix equation (3) represents a system of linear equations commonly referred to as the **normal equations for $\hat{x}$** .

Since set of least-squares solutions is nonempty and any such $\hat{x}$ satisfies the normal equations. Conversely, suppose that $\hat{x}$ satisfies $A^T A \hat{x} = A^T b$. Then it satisfies that $b - A \hat{x}$ is orthogonal to the rows of A^T and hence is orthogonal to the columns of A . Since the columns of A span $\text{Col } A$, the vector $b - A \hat{x}$ is orthogonal to all of $\text{Col } A$. Hence the equation $b = A \hat{x} + (b - A \hat{x})$ is a decomposition of b into the sum of a vector in $\text{Col } A$ and a vector orthogonal to $\text{Col } A$. By the uniqueness of the orthogonal decomposition, $A \hat{x}$ must be the orthogonal projection of b onto $\text{Col } A$. That is, $A \hat{x} = \hat{b}$ and $\hat{x}$ is a least-squares solution.

Definition

If A is $m \times n$ and b is in R^n , a least-squares solution of $Ax = b$ is an $\hat{x}$ in R^n such that

$$\|b - A\hat{x}\| \leq \|b - Ax\| \quad \forall x \in R^n$$

Theorem

The set of least-squares solutions of $Ax = b$ coincides with the nonempty set of solutions of the normal equations

$$A^T A \hat{x} = A^T b$$

Example 1

Find the least squares solution and its error from the following matrices,

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution

Firstly we find

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \quad \text{and}$$

$$A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^T A \hat{x} = A^T b$ becomes $\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$

Row operations can be used to solve this system, but since $A^T A$ is invertible and 2×2 , it is probably faster to compute $(A^T A)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$

$$\text{Therefore, } \hat{x} = (A^T A)^{-1} A^T b = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{Now again as } A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$\text{Then } A\hat{x} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

$$\text{Hence } \mathbf{b} - A\hat{x} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

$$\text{So } \|\mathbf{b} - A\hat{x}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

The least-squares error is $\sqrt{84}$. For any x in R^2 , the distance between b and the vector Ax is at least $\sqrt{84}$.

Example 2

Find the general least-squares solution of $Ax = b$ in the form of a free variable with

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

Solution

Firstly we find, $A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ and

$$A^T b = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Then augmented matrix for $A^T A \hat{x} = A^T b$ is

$$\left[\begin{array}{cccc|c} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{array} \right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

The general solution is $x_1 = 3 - x_4$, $x_2 = -5 + x_4$, $x_3 = -2 + x_4$, and x_4 is free.

So the general least-squares solution of $Ax = b$ has the form

$$\hat{x} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Theorem

The matrix $A^T A$ is invertible iff the columns of A are linearly independent. In this case, the equation $Ax = b$ has only one least-squares solution $\hat{x}$, and it is given by

$$\hat{x} = (A^T A)^{-1} A^T b$$

Example 3

Find the least squares solution to the following system of equations.

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix}$$

Solution

As

$$A^T = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 55 & 12 & 45 \\ 12 & 26 & 28 \\ 45 & 28 & 77 \end{bmatrix}$$

Now

$$A^T b = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -9 \\ -5 \\ 12 \end{bmatrix}$$

As

$$A^T A \hat{x} = A^T b$$

$$\begin{bmatrix} 55 & 12 & 45 \\ 12 & 26 & 28 \\ 45 & 28 & 77 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -9 \\ -5 \\ 12 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -.676 \\ -.776 \\ .834 \end{bmatrix}$$

Example 4

Compute the least square error for the solution of the following equation

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix}$$

Solution

$$\tilde{A}x = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} -.676 \\ -.776 \\ .834 \end{bmatrix}$$

$$\tilde{A}x = \begin{bmatrix} 0.548 \\ 1.652 \\ -2.172 \\ 3.494 \end{bmatrix}$$

As least square error

$$\|\epsilon\| = \|b - A\bar{x}\|$$

is as small as possible, or in other words is smaller than all other possible choices of x .

As

$$b - A\hat{x} = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix} - \begin{bmatrix} 0.548 \\ 1.652 \\ -2.172 \\ 3.494 \end{bmatrix} = \begin{bmatrix} -0.548 \\ 0.652 \\ 0.172 \\ 1.494 \end{bmatrix}$$

$$\|\epsilon\|^2 = \epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2 + \epsilon_4^2$$

Thus, least square error is

$$\|\epsilon\| = \|b - A\bar{x}\| = \sqrt{(-0.548)^2 + (0.652)^2 + (0.172)^2 + (1.494)^2}$$

$$= 0.3003 + 0.4251 + 0.02958 + 2.23 = 2.987$$

Theorem

Given an $m \times n$ matrix A with linearly independent columns. Let $A = QR$ be a QR factorization of A , then for each b in R^m , the equation $Ax = b$ has a unique least-squares solution, given by

$$\hat{x} = R^{-1}Q^T b$$

Example 1

Find the least square solution for $A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, b = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$

Solution

First of all we find QR factorization of the given matrix A. For this we have to find out orthonormal basis for the column space of A by applying Gram-Schmidt Process, we get the matrix of orthonormal basis Q,

$$Q = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \text{ And}$$

$$R = Q^T A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\text{Then } Q^T b = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

$$\text{The least-squares solution } \hat{x} \text{ satisfies } R\hat{x} = Q^T b; \text{ that is, } \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

$$\text{This equation is solved easily and yields } \hat{x} = \begin{bmatrix} 10 \\ -6 \\ 2 \end{bmatrix}.$$

Example 2

Find the least squares solution $R\hat{x} = Q^T b$ to the given matrices,

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution

First of all we find QR factorization of the given matrix A. Thus, we have to make

Orthonormal basis by applying Gram Schmidt process on the columns of A,

Let $v_1 = x_1$

$$v_2 = x_2 - P = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Thus, the orthonormal basis are

$$\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|} \right\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus

$$Q = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } Q^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Now } R = Q^T A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & 5 \\ 0 & 2 \end{bmatrix} \quad \text{And } Q^T b = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Thus, least squares solution of $R\hat{x} = Q^T b$ is

$$\begin{bmatrix} 15 & 5 \\ 0 & 2 \end{bmatrix} \hat{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} 1 \\ 1.8 \end{bmatrix}$$

Exercise 1

Find a least-squares solution of $Ax = b$ for

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Exercise 2

Find the least-squares solution and its error of $Ax = b$ for

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, b = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Exercise 3

Find the least squares solution $R\hat{x} = Q^T b$ to the given matrices,

$$A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, b = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$$

Lecture # 43

Inner Product Space

Inner Product Space

In mathematics, an inner product space is a vector space with the additional structure called an inner product. This additional structure associates each pair of vectors in the space with a scalar quantity known as the inner product of the vectors. Inner products allow the rigorous introduction of intuitive geometrical notions such as the length of a vector or the angle between two vectors. They also provide the means of defining orthogonality between vectors (zero inner product). Inner product spaces generalize Euclidean spaces (in which the inner product is the dot product, also known as the scalar product) to vector spaces of any (possibly infinite) dimension, and are studied in functional analysis.

Definition

An inner product on a vector space V is a function that to each pair of vectors u and v associates a real number $\langle u, v \rangle$ and satisfies the following axioms,

For all u, v, w in V and all scalars C :

- 1) $\langle u, v \rangle = \langle v, u \rangle$
- 2) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- 3) $\langle cu, v \rangle = c \langle u, v \rangle$
- 4) $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = 0$

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

Example 1

Fix any two positive numbers say 4 & 5 and for vectors $u = \langle u_1, u_2 \rangle$ and $v = \langle v_1, v_2 \rangle$ in R^2 set

$$\langle u, v \rangle = 4u_1v_1 + 5u_2v_2$$

Show that it defines an inner product.

Solution

Certainly Axiom 1 is satisfied, because

$$\langle u, v \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle v, u \rangle.$$

If $w = (w_1, w_2)$, then

$$\begin{aligned} \langle u + v, w \rangle &= 4(u_1 + v_1)w_1 + 5(u_2 + v_2)w_2 \\ &= 4u_1w_1 + 5u_2w_2 + 4v_1w_1 + 5v_2w_2 = \langle u, w \rangle + \langle v, w \rangle \end{aligned}$$

This verifies Axiom 2.

For Axiom 3, we have $\langle cu, v \rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c(4u_1v_1 + 5u_2v_2) = c \langle u, v \rangle$

For Axiom 4, note that $\langle u, u \rangle = 4u_1^2 + 5u_2^2 \geq 0$, and $4u_1^2 + 5u_2^2 = 0$ only if $u_1 = u_2 = 0$, that is, if $u = \mathbf{0}$. Also, $\langle 0, 0 \rangle = 0$.

So $\langle u, v \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2$ defines an inner product on $\mathbf{R}^2$.

Example 2

Let A be symmetric, positive definite $n \times n$ matrix and let u and v be vectors in $\mathfrak{R}^n$. Show that $\langle u, v \rangle = u^t A v$ defines an inner product.

Solution

We check that

$$\begin{aligned}\langle u, v \rangle &= u^t A v = u \cdot A v = A v \cdot u \\ &= A^t v \cdot u = (v^t A)^t u = v^t A u = \langle v, u \rangle\end{aligned}$$

Also

$$\begin{aligned}\langle u, v + w \rangle &= u^t A (v + w) = u^t A v + u^t A w \\ &= \langle u, v \rangle + \langle u, w \rangle\end{aligned}$$

And

$$\langle cu, v \rangle = (cu)^t A v = c(u^t A v) = c \langle u, v \rangle$$

Finally since A is positive definite

$$\langle u, u \rangle = u^t A u > 0 \quad \text{for all } u \neq 0$$

$$\text{So } \langle u, u \rangle = u^t A u = 0 \quad \text{iff } u = 0$$

So $\langle u, v \rangle = u^t A v$ is an inner product space.

Example 3

Let $t_0, \dots, t_n$ be distinct real numbers. For p and q in $\mathbf{P}_n$, define

$$\langle p, q \rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \dots + p(t_n)q(t_n)$$

Show that it defines inner product.

Solution

Certainly Axiom 1 is satisfied, because

$$\begin{aligned}\langle p, q \rangle &= p(t_0)q(t_0) + p(t_1)q(t_1) + \dots + p(t_n)q(t_n) \\ &= q(t_0)p(t_0) + q(t_1)p(t_1) + \dots + q(t_n)p(t_n) = \langle q, p \rangle\end{aligned}$$

If $r = r(t_0) + r(t_1) + \dots + r(t_n)$, then

$$\begin{aligned}\langle p + q, r \rangle &= [p(t_0) + q(t_0)]r(t_0) + [p(t_1) + q(t_1)]r(t_1) + \dots + [p(t_n) + q(t_n)]r(t_n) \\ &= [p(t_0)r(t_0) + p(t_1)r(t_1) + \dots + p(t_n)r(t_n)] + [q(t_0)r(t_0) + q(t_1)r(t_1) + \dots + q(t_n)r(t_n)] \\ &= \langle p, r \rangle + \langle q, r \rangle\end{aligned}$$

This verifies Axiom 2.

For Axiom 3, we have

$$\begin{aligned}\langle cp, q \rangle &= [cp(t_0)]q(t_0) + [cp(t_1)]q(t_1) + \dots + [cp(t_n)]q(t_n) \\ &= c[p(t_0)q(t_0) + p(t_1)q(t_1) + \dots + p(t_n)q(t_n)] = c \langle p, q \rangle\end{aligned}$$

For Axiom 4, note that

$$\langle p, p \rangle = [p(t_0)]^2 + [p(t_1)]^2 + \dots + [p(t_n)]^2 \geq 0$$

Also, $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. (We still use a boldface zero for the zero polynomial, the zero vector in P_n .) If $\langle p, p \rangle = 0$, then p must vanish at $n + 1$ points: $t_0, \dots, t_n$. This is possible only if p is the zero polynomial, because the degree of p is less than $n + 1$. Thus $\langle p, q \rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \dots + p(t_n)q(t_n)$ defines an inner product on P_n .

Example 4

Compute $\langle p, q \rangle$ where $p(t) = 4 + t$ $q(t) = 5 - 4t^2$

Refer to P_2 with the inner product given by evaluation at -1, 0 and 1 in example 2.

Solution

$$P(-1) = 3, P(0) = 4, P(1) = 5$$

$$q(-1) = 1, q(0) = 5, q(1) = 1$$

$$\begin{aligned}\langle p, q \rangle &= P(-1)q(-1) + P(0)q(0) + P(1)q(1) \\ &= (3)(1) + (4)(5) + (5)(1) \\ &= 3 + 20 + 5\end{aligned}$$

Example 5

Compute the orthogonal projection of q onto the subspace spanned by p , for p and q in the above example.

Solution

The orthogonal projection of q onto the subspace spanned by p

$$P(-1) = 3, P(0) = 4, P(1) = 5$$

$$q(-1) = 1, q(0) = 5, q(1) = 1$$

$$q \cdot p = 28 \quad p \cdot p = 50$$

$$\begin{aligned}\hat{q} &= \frac{q \cdot p}{p \cdot p} p = \frac{28}{50}(4 + t) \\ &= \frac{56}{25} + \frac{14}{25}t\end{aligned}$$

Example 6

Let V be P_2 , with the inner product from example 2 where

$$t_0 = 0, t_1 = \frac{1}{2} \text{ and } t_2 = 1$$

Let $p(t) = 12t^2$ and $q(t) = 2t - 1$

Compute $\langle p, q \rangle$ and $\langle q, q \rangle$

Solution

$$\begin{aligned}\langle p, q \rangle &= p(0)q(0) + p\left(\frac{1}{2}\right)q(1) \\ &= (0)(-1) + (3)(0) + (12)(1) = 12\end{aligned}$$

$$\begin{aligned}\langle q, q \rangle &= [q(0)]^2 + \left[q\left(\frac{1}{2}\right)\right]^2 + [q(1)]^2 \\ &= (-1)^2 + (0)^2 + (1)^2 = 2\end{aligned}$$

Norm of a Vector

Let V be an inner product space with the inner product denoted by $\langle u, v \rangle$ just as in R^n , we define the length or norm of a vector V to be the scalar

$$\|v\| = \sqrt{\langle u, v \rangle} \quad \text{or} \quad \|v\|^2 = \langle u, v \rangle$$

- 1) A unit vector is one whose length is 1.
- 2) The distance between u & v is $\|u - v\|$ vectors u & v are orthogonal if $\langle u, v \rangle = 0$

Example 7

Compute the length of the vectors in example 3.

Solution

$$\begin{aligned}\|p\|^2 &= \langle p, p \rangle = [p(0)]^2 + \left[p\left(\frac{1}{2}\right)\right]^2 + [p(1)]^2 \\ &= (0)^2 + (3)^2 + (12)^2 = 153 \\ \|p\| &= \sqrt{153}\end{aligned}$$

In example 3 we found that

$$\begin{aligned}\langle q, q \rangle &= 2 \\ \text{Hence} \quad \|q\| &= \sqrt{2}\end{aligned}$$

Example 8

Let $\mathfrak{R}^2$ have the inner product of example 1 and let $x=(1,1)$ and $y=(5,-1)$

- a) Find $\|x\|$, $\|y\|$ and $|\langle x, y \rangle|^2$
- b) Describe all vectors (z_1, z_2) that are orthogonal to y .

Solution

- a) We have $x=(1,1)$ and $y=(5,-1)$
And $\langle x, y \rangle = 4x_1y_1 + 5x_2y_2$

$$\begin{aligned}
 \|x\| &= \sqrt{\langle x, x \rangle} = \sqrt{4(1)(1) + 5(1)(1)} \\
 &= \sqrt{4+5} = \sqrt{9} = 3 \\
 \|y\| &= \sqrt{\langle y, y \rangle} = \sqrt{4(5)(5) + 5(-1)(-1)} \\
 &= \sqrt{100+5} = \sqrt{105}
 \end{aligned}$$

$$\begin{aligned}
 |\langle x, y \rangle|^2 &= \langle x, y \rangle \langle x, y \rangle \\
 &= [4(1)(5) + 5(1)(-1)]^2 \\
 &= [20 - 5]^2 \\
 &= [15]^2 = 225
 \end{aligned}$$

b) All vectors $z = (z_1, z_2)$ orthogonal to $y = (5, -1)$

$$\langle y, z \rangle = 0$$

$$4(5)(z_1) + 5(-1)(z_2) = 0$$

$$20z_1 - 5z_2 = 0$$

$$4z_1 - z_2 = 0$$

$$\begin{bmatrix} 4 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0$$

So all multiples of $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ are orthogonal to y .

Example 9

Let V be P_4 with the inner product in example 2 involving evaluation of polynomials at $-2, -1, 0, 1, 2$ and view P_2 as a subspace of V . Produce an orthogonal basis for P_2 by applying the Gram Schmidt process to the polynomials $1, t$ & t^2 .

Solution

Given polynomials $1 \quad t \quad t^2$ at $-2, -1, 0, 1$ and 2

$$\begin{array}{lcl}
 \text{Polynomial:} & \mathbf{1} & \mathbf{t} & \mathbf{t^2} \\
 & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} & \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix} \\
 \text{Vector of values:} & 1, & 0, & 0
 \end{array}$$

The inner product of two polynomials in V equals the (standard) inner product of their corresponding vectors in $\mathbf{R}^5$. Observe that t is orthogonal to the constant function 1 . So

take $p_0(t) = 1$ and $p_1(t) = t$. For p_2 , use the vectors in $\mathbf{R}^5$ to compute the projection of t^2 onto $\text{Span}\{p_0, p_1\}$:

$$\langle t^2, p_0 \rangle = \langle t^2, 1 \rangle = 4 + 1 + 0 + 1 + 4 = 10$$

$$\langle p_0, p_0 \rangle = 5$$

$$\langle t^2, p_1 \rangle = \langle t^2, t \rangle = -8 + (-1) + 0 + 1 + 8 = 0$$

The orthogonal projection of t^2 onto $\text{Span}\{1, t\}$ is $\frac{10}{5}p_0 + 0p_1$. Thus

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

An orthogonal basis for the subspace P_2 of V is:

Polynomial:	p_0	p_1	p_2
	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$
Vector of values:	1,	0,	-2

Best Approximation in Inner Product Spaces

A common problem in applied mathematics involves a vector space V whose elements are functions. The problem is to approximate a function f in V by a function g from a specified subspace W of V . The “closeness” of the approximation of f depends on the way $\|f - g\|$ is defined. We will consider only the case in which the distance between f and g is determined by an inner product. In this case the best approximation to f by functions in W is the orthogonal projection of f onto the subspace W .

Example 10

Let V be P_4 with the inner product in example 5 and let p_0, p_1 & p_2 be the orthogonal basis for the subspace P_2 , find the best approximation to $p(t) = 5 - \frac{1}{2}t^4$ by polynomials in P_2 .

Solution:

The values of p_0, p_1 , and p_2 at the numbers $-2, -1, 0, 1$, and 2 are listed in $\mathbf{R}^5$ vectors in

Polynomial:	p_0	p_1	p_2
-------------	-------	-------	-------

$$\text{Vector of values: } \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$$

The corresponding values for $\mathbf{p}$ are: $-3, 9/2, 5, 9/2$, and -3 .

We compute

$$\begin{aligned} \langle p, p_0 \rangle &= 8 & \langle p, p_1 \rangle &= 0 & \langle p, p_2 \rangle &= -31 \\ \langle p_0, p_0 \rangle &= 5, & \langle p_2, p_2 \rangle &= 14 \end{aligned}$$

Then the best approximation in $\mathbf{V}$ to $\mathbf{p}$ by polynomials in $\mathbf{P}_2$ is

$$\begin{aligned} \hat{p} &= \text{proj}_{\mathbf{P}_2} p = \frac{\langle p, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle p, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle p, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 \\ &= \frac{8}{5} p_0 + \frac{-31}{14} p_2 = \frac{8}{5} - \frac{31}{14} (t^2 - 2). \end{aligned}$$

This polynomial is the closest to $\mathbf{P}$ of all polynomials in $\mathbf{P}_2$, when the distance between polynomials is measured only at $-2, -1, 0, 1$, and 2 .

Cauchy – Schwarz Inequality

For all u, v in V

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

Triangle Inequality

For all u, v in V

$$\|u + v\| \leq \|u\| + \|v\|$$

Proof

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ \|u + v\|^2 &= (\|u\| + \|v\|)^2 \end{aligned}$$

$$\Rightarrow \|u + v\| = \|u\| + \|v\|$$

Inner product for $C[a, b]$

Probably the most widely used inner product space for applications is the vector space $C[a, b]$ of all continuous functions on an interval $a \leq t \leq b$, with an inner product that will describe.

Example 11

For f, g in $C[a, b]$, set

$$\langle f, g \rangle = \int_a^b f(t) g(t) dt$$

Show that it defines an inner product on $C[a, b]$.

Solution

Inner product Axioms 1 to 3 follow from elementary properties of definite integrals

1. $\langle f, g \rangle = \langle g, f \rangle$
2. $\langle f + h, g \rangle = \langle f, g \rangle + \langle h, g \rangle$
3. $\langle cf, g \rangle = c \langle f, g \rangle$

For Axiom 4, observe that

$$\langle f, f \rangle = \int_a^b [f(t)]^2 dt \geq 0$$

The function $[f(t)]^2$ is continuous and nonnegative on $[a, b]$. If the definite integral of $[f(t)]^2$ is zero, then $[f(t)]^2$ must be identically zero on $[a, b]$, by a theorem in advanced calculus, in which case f is the zero function. Thus $\langle f, f \rangle = 0$ implies that f is the zero function of $[a, b]$.

So $\langle f, g \rangle = \int_a^b f(t)g(t)dt$ defines an inner product on $C[a, b]$.

Example 12

Compute $\langle f, g \rangle$ where $f(t) = 1 - 3t^2$ and $g(t) = t - t^3$ on $v = C[0, 1]$.

Solution

Let V be the space $C[a, b]$ with the inner product

$$\langle f, g \rangle = \int_a^b f(t) g(t) dt$$

$$f(t) = 1 - 3t^2, \quad g(t) = t - t^3$$

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 (1 - 3t^2)(t - t^3) dt \\ &= \int_0^1 (3t^5 - 4t^3 + t) dt \\ &= \left[\frac{1}{2}t^6 - t^4 + \frac{1}{2}t^2 \right]_0^1 \\ &= 0 \end{aligned}$$

Example 13

Let V be the space $C[a, b]$ with the inner product

$$\langle f, g \rangle = \int_b^a f(t) g(t) dt$$

Let W be the subspace spanned by the polynomials

$$P_1(t) = 1, P_2(t) = 2t - 1 \text{ \& } P_3(t) = 12t^2$$

Use the Gram – Schmidt process to find an orthogonal basis for W .

Solution

Let $q_1 = p_1$, and compute

$$\langle p_2, q_1 \rangle = \int_0^1 (2t - 1)(1) dt = (t^2 - t) \Big|_0^1 = 0$$

So p_2 is already orthogonal to q_1 , and we can take $q_2 = p_2$. For the projection of p_3 onto $W_2 = \text{Span} \{q_1, q_2\}$, we compute

$$\langle p_3, q_1 \rangle = \int_0^1 12t^2 \cdot 1 dt = 4t^3 \Big|_0^1 = 4$$

$$\langle q_1, q_1 \rangle = \int_0^1 1 \cdot 1 dt = t \Big|_0^1 = 1$$

$$\langle p_3, q_2 \rangle = \int_0^1 12t^2 (2t - 1) dt = \int_0^1 (24t^3 - 12t^2) dt = 2$$

$$\langle q_2, q_2 \rangle = \int_0^1 (2t - 1)^2 dt = \frac{1}{6} (2t - 1)^3 \Big|_0^1 = \frac{1}{3}$$

Then
$$\text{proj}_{W_2} p_3 = \frac{\langle p_3, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 + \frac{\langle p_3, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{1/3} q_2 = 4q_1 + 6q_2$$

And
$$q_3 = p_3 - \text{proj}_{W_2} p_3 = p_3 - 4q_1 - 6q_2$$

As a function, $q_3(t) = 12t^2 - 4 - 6(2t - 1) = 12t^2 - 12t + 2$. The orthogonal basis for the subspace W is $\{q_1, q_2, q_3\}$

Exercises

Let $\mathfrak{R}^2$ have the inner product of example 1 and let $x = (1, 1)$ and $y = (5, -1)$

- a) Find $\|x\|$, $\|y\|$ and $|\langle x, y \rangle|^2$ b) Describe all vectors (z_1, z_2) that are orthogonal to y .

- 2) Let $\mathfrak{R}^2$ have the inner product of Example 1. Show that the Cauchy-Schwarz inequality holds for $x = (3, -2)$ and $y = (-2, 1)$

Exercise 3-8 refer to P_2 with the inner product given by evaluation at -1, 0 and 1 in example 2.

- 3) Compute $\langle p, q \rangle$ where $p(t) = 4 + t$ $q(t) = 5 - 4t^2$

- 4) Compute $\langle p, q \rangle$ where $p(t) = 3t - t^2$ $q(t) = 3 + t^2$

- 5) Compute $\|p\|$ and $\|q\|$ for p and q in exercise 3.

- 6) Compute $\|p\|$ and $\|q\|$ for p and q in exercise 4.

- 7) Compute the orthogonal projection of q onto the subspace spanned by p , for p and q in Exercise 3.

8) Compute the orthogonal projection of q onto the subspace spanned by p , for p and q in Exercise 4.

9) Let P^3 have the inner product given by evaluation at $-3, -1, 1$, and 3 . Let

$$p_0(t) = 1, \quad p_1(t) = t, \quad \text{and} \quad p_2(t) = t^2$$

a) Compute the orthogonal projection of P_2 on to the subspace spanned by P_0 and P_1 . b) Find a polynomial q that is orthogonal to P_0 and P_1 such that $\{p_0, p_1, q\}$ is an orthogonal basis for $\text{span}\{p_0, p_1, q\}$. Scale the polynomial q so that its vector of values at $(-3, -1, 1, 3)$ is $(1, -1, -1, 1)$

10) Let P^3 have the inner product given by evaluation at $-3, -1, 1$, and 3 . Let

$$p_0(t) = 1, \quad p_1(t) = t, \quad \text{and} \quad p_2(t) = t^2$$

Find the best approximation to $p(t) = t^3$ by polynomials in $\text{Span}\{p_0, p_1, q\}$.

11) Let p_0, p_1, p_2 be the orthogonal polynomials described in example 5, where the inner product on P_4 is given by evaluation at $-2, -1, 0, 1$, and 2 . Find the orthogonal projection of t^3 onto $\text{Span}\{p_0, p_1, p_2\}$

12) Compute $\langle f, g \rangle$ where $f(t) = 1 - 3t^2$ and $g(t) = t - t^3$ on $v = C[0, 1]$.

13) Compute $\langle f, g \rangle$ where

$$f(t) = 5t - 3 \text{ and } g(t) = t^3 - t^2 \text{ on } v = C[0, 1].$$

14) Compute $\|f\|$ for f in exercise 12.

15) Compute $\|g\|$ for g in exercise 13.

16) Let V be the space $C[-2, 2]$ with the inner product of Example 7. Find an orthogonal basis for the subspace spanned by the polynomials $1, t, t^2$.

17)

Let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be two vectors in R^2 . Show that $\langle u, v \rangle = 2u_1v_1 + 3u_2v_2$

defines an inner product.

Lecture 44

Application of inner product spaces

Definition

An inner product on a vector space V is a function that associates to each pair of vectors u and v in V , a real number $\langle u, v \rangle$ and satisfies the following axioms, for all u, v, w in V and all scalars c :

1. $\langle u, v \rangle = \langle v, u \rangle$
2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
3. $\langle cu, v \rangle = c \langle u, v \rangle$
4. $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = 0$.

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

Least Squares Lines

The simplest relation between two variables x and y is the linear equation $y = \beta_0 + \beta_1 x$. Often experimental data produces points $(x_1, y_1), \dots, (x_n, y_n)$ that when graphed, seem to lie close to a line. Actually we want to determine the parameters β_0 and β_1 that make the line as “close” to the points as possible. There are several ways to measure how close the line is to the data. The usual choice is to add the squares of the residuals. The least squares line is the line $y = \beta_0 + \beta_1 x$ that minimizes the sum of the squares of the residuals.

If the data points are on the line, the parameters β_0 and β_1 would satisfy the equations

<i>predicted</i> value	<i>Observed</i> value
$\beta_0 + \beta_1 x_1 = y_1$	
$\beta_0 + \beta_1 x_2 = y_2$	
.	
.	
.	
$\beta_0 + \beta_1 x_n = y_n$	

We can write this system as

$$X\beta = y$$

Where $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & x_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, $y = \begin{bmatrix} y_0 \\ y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$

Computing the least-squares solution of $X\beta = y$ is equivalent to finding the β that determines the least-squares line.

Example 1

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points (2, 1), (5, 2), (7, 3), (8, 3).

Solution

$$X\beta = y$$

$$\text{Here } X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

For the least-squares solution of $X\beta = y$, obtain the normal equations (with the new notation) :

$$X^T X \hat{\beta} = X^T y$$

i.e, compute

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

The normal equations are

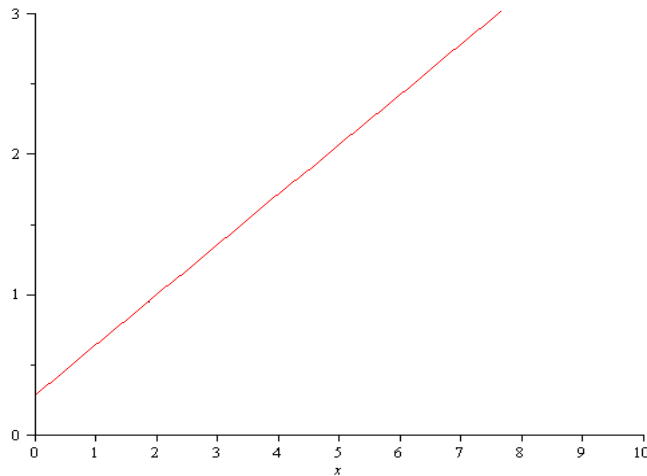
$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Hence,

$$\begin{aligned} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} &= \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} \\ &= \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} \\ &= \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix} \end{aligned}$$

Thus, the least -squares line has the equation

$$y = \frac{2}{7} + \frac{5}{14}x$$



Weighted Least-Squares

Let y be a vector of n observations, $y_1, y_2, \dots, y_n$ and suppose we wish to approximate y by a vector $\hat{y}$ that belongs to some specified subspace of $\mathbb{R}^n$ (as discussed previously that $\hat{y}$ is written as Ax so that $\hat{y}$ was in the column space of A). Now suppose the approximating vector $\hat{y}$ is to be constructed from the columns of matrix A . Then we find an $\hat{x}$ that makes $A\hat{x} = \hat{y}$ as close to y as possible. So that measure of closeness is the weighted error

$$\|Wy - W\hat{y}\|^2 = \|Wy - WA\hat{x}\|^2$$

Where W is the diagonal matrix with (positive) $w_1, \dots, w_n$ on its diagonal, that is

$$W = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & & \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ \cdot & & & & \cdot \\ 0 & & \dots & & w_n \end{bmatrix}$$

Thus, $\hat{x}$ is the ordinary least-squares solution of the equation

$$WAX = Wy$$

The normal equation for the weighted least-squares solution is

$$(WA)^T WAX = (WA)^T Wy$$

Example 2

Find the least squares line $y = \beta_0 + \beta_1 x$ that best fits the data $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$, $(2, 3)$. Suppose that the errors in measuring the y -values of the last two data points are greater than for the other points. Weight this data half as much as the rest of the data.

Solution

Write X, β and y

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, y = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

For a weighting matrix, choose W with diagonal entries 2, 2, 2, 1 and 1.

Left-multiplication by W scales the rows of X and y :

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

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, Wy = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

For normal equation, compute

$$(WX)^T WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix}, \text{ and } (WX)^T Wy = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

And solve

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix}^{-1} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 25 & 9 \\ 9 & 14 \end{bmatrix} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 25 & 9 \\ 9 & 14 \end{bmatrix} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 1169 \\ 55 \end{bmatrix} = \begin{bmatrix} 4.3 \\ 0.2 \end{bmatrix}$$

Therefore, the solution to two significant digits is $\beta_0 = 4.3$ and $\beta_1 = 0.20$.

Hence the required line is $y = 4.3 + 0.2x$

In contrast, the ordinary least-squares line for this data can be found as:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 200 \\ -5 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -0.1 \end{bmatrix}$$

Hence the equation of least-squares line is

$$y = 1.0 - 0.1x$$

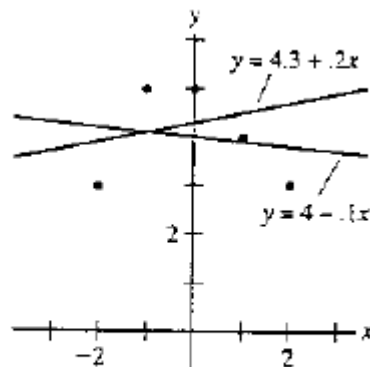


FIGURE 1 Weighted and ordinary least-squares lines.

What Does Trend Analysis Mean?

An aspect of technical analysis that tries to predict the future movement of a stock based on past data. Trend analysis is based on the idea that what has happened in the past gives traders an idea of what will happen in the future.

Linear Trend

A first step in analyzing a time series, to determine whether a linear relationship provides a good approximation to the long-term movement of the series computed by the method of semi averages or by the method of least squares.

Note

The simplest and most common use of trend analysis occurs when the points $t_0, t_1, \dots, t_n$ can be adjusted so that they are evenly spaced and sum to zero.

Example

Fit a quadratic trend function to the data $(-2,3)$, $(-1,5)$, $(0,5)$, $(1,4)$ and $(2,3)$

Solution

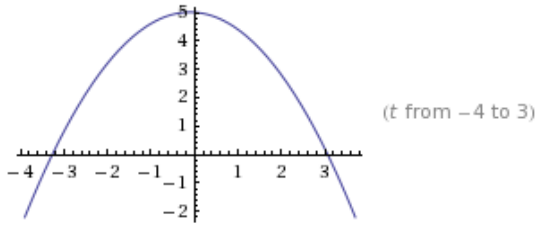
The t-coordinates are suitably scaled to use the orthogonal polynomials found in Example 5 of the last lecture. We have

$$\begin{array}{lcl}
 \text{Polynomial :} & p_0 & p_1 & p_2 & \text{data : } g \\
 & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix} & \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} \\
 \text{Vector of values :} & & & & \\
 \hat{p} = \frac{\langle g, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle g, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle g, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2
 \end{array}$$

$$= \frac{20}{5} p_0 - \frac{1}{10} p_1 - \frac{7}{14} p_2$$

$$\text{and} \quad \hat{p}(t) = 4 - 0.1t - 0.5(t^2 - 2)$$

Since, the coefficient of p_2 is not extremely small, it would be reasonable to conclude that the trend is at least quadratic.



Above figure shows that approximation by a quadratic trend function

Fourier series

If f is a 2π -periodic function then

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mt + b_m \sin mt)$$

is called Fourier series of f where

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos mt \, dt \quad \text{and}$$

$$b_m = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin mt \, dt$$

Example

Let $C[0, 2\pi]$ has the inner product

$$\langle f, g \rangle = \int_0^{2\pi} f(t) g(t) \, dt$$

and let m and n be unequal positive integers. Show that $\cos mt$ and $\cos nt$ are orthogonal.

Solution

When $m \neq n$

$$\begin{aligned} \langle \cos mt, \cos nt \rangle &= \int_0^{2\pi} \cos mt \cos nt \, dt \\ &= \frac{1}{2} \int_0^{2\pi} [\cos(mt + nt) + \cos(mt - nt)] \, dt \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \left[\frac{\sin(mt+nt)}{m+n} + \frac{\sin(mt-nt)}{m-n} \right]_0^{2\pi} \\
&= 0
\end{aligned}$$

Example

Find the n th-order Fourier approximation to the function

$$f(t) = t \text{ on the interval } [0, 2\pi].$$

Solution

We compute

$$\frac{a_0}{2} = \frac{1}{2} \cdot \frac{1}{\pi} \int_0^{2\pi} t \, dt = \frac{1}{2\pi} \left[\frac{1}{2} t^2 \right]_0^{2\pi} = \pi$$

and for $k > 0$, using integration by parts,

$$a_k = \frac{1}{\pi} \int_0^{2\pi} t \cos kt \, dt = \frac{1}{\pi} \left[\frac{1}{k^2} \cos kt + \frac{t}{k} \sin kt \right]_0^{2\pi} = 0$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} t \sin kt \, dt = \frac{1}{\pi} \left[\frac{1}{k^2} \sin kt - \frac{t}{k} \cos kt \right]_0^{2\pi} = -\frac{2}{k}$$

Thus, the n th-order Fourier approximation of $f(t) = t$ is

$$\pi - 2 \sin t - \sin 2t - \frac{2}{3} \sin 3t - \cdots - \frac{2}{n} \sin nt$$

The norm of the difference between f and a Fourier approximation is called the mean square error in the approximation.

It is common to write

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mt + b_m \sin mt)$$

This expression for $f(t)$ is called the Fourier series for f on $[0, 2\pi]$. The term $a_m \cos mt$, for example, is the projection of f onto the one-dimensional subspace spanned by $\cos mt$.

Example

Let $q_1(t) = 1$, $q_2(t) = t$, and $q_3(t) = 3t^2 - 4$. Verify that $\{q_1, q_2, q_3\}$ is an orthogonal set in $C[-2, 2]$ with the inner product

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

Solution:

$$\langle q_1, q_2 \rangle = \int_{-2}^2 1 \cdot t dt = \frac{1}{2} t^2 \Big|_{-2}^2 = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^2 1 \cdot (3t^2 - 4) dt = (t^3 - 4t) \Big|_{-2}^2 = 0$$

$$\langle q_2, q_3 \rangle = \int_{-2}^2 t \cdot (3t^2 - 4) dt = \left(\frac{3}{4} t^4 - 2t^2 \right) \Big|_{-2}^2 = 0$$

Exercise

- Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(0, 1), (1, 1), (2, 2), (3, 2)$.
- Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(-1, 0), (0, 1), (1, 2), (2, 4)$.
- Find the least-squares line $y = \beta_0 + \beta_1 x$ that best fits the data $(-2, 0), (-1, 0), (0, 2), (1, 4), (2, 4)$, assuming that the first and last data points are less reliable. Weight them half as much as the three interior points.
- To make a trend analysis of six evenly spaced data points, one can use orthogonal polynomials with respect to evaluation at the points $t = -5, -3, -1, 1, 3$ and 5

(a). Show that the first three orthogonal polynomials are

$$p_0(t) = 1, \quad p_1(t) = t, \quad \text{and} \quad p_2(t) = \frac{3}{8}t^2 - \frac{35}{8}$$

(b) Fit a quadratic trend function to the data $(-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)$

5: For the space $C[0, 2\pi]$ with the inner product defined by

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt$$

- Show that $\sin mt$ and $\sin nt$ are orthogonal when $m \neq n$
- Find the third-order Fourier approximation to $f(t) = 2\pi - t$
- Find the third order Fourier approximation to $\cos^3 t$, without performing any

integration calculations.

6: Find the first-order and third order Fourier approximations to

$$f(t) = 3 - 2 \sin t + 5 \sin 2t - 6 \cos 2t$$



Revision of Previous Lectures