

**A FIRST COURSE
IN
LINEAR ALGEBRA**

A FIRST COURSE
IN
LINEAR ALGEBRA
MAT2040 Notebook

Prof. Tom Luo

The Chinese University of Hong Kong, Shenzhen

Prof. Ruoyu Sun

University of Illinois Urbana-Champaign



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

Copyright ©2004 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department with the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic format.

Library of Congress Cataloging-in-Publication Data:

Survey Methodology / Robert M. Groves . . . [et al.].

p. cm.—(Wiley series in survey methodology)

“Wiley-Interscience.”

Includes bibliographical references and index.

ISBN 0-471-48348-6 (pbk.)

1. Surveys—Methodology. 2. Social

sciences—Research—Statistical methods. I. Groves, Robert M. II. Series.

HA31.2.S873 2004

001.4'33—dc22

2004044064

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Contents

Contributors	v
Foreword	vii
Preface	ix
Acknowledgments	xi
Acronyms	xiii
1 Week1	1
1.1 Tuesday	1
1.1.1 Introduction	1
1.1.2 Gaussian Elimination	3
1.1.3 Complexity Analysis	11
1.1.4 Brief Summary	12
1.2 Thursday	14
1.2.1 Row-Echelon Form	14
1.2.2 Matrix Multiplication	16
1.2.3 Special Matrices	19
1.3 Friday	21
1.3.1 Matrix Multiplication	21
1.3.2 Elementary Matrix	22
1.3.3 Properties of Matrix	24
1.3.4 Permutation Matrix	26
1.3.5 LU decomposition	29
1.3.6 LDU decomposition	33
1.3.7 LU Decomposition with row exchanges	35
1.4 Assignment One	36

2	Week2	39
2.1	Tuesday	39
2.1.1	Review	39
2.1.2	Special matrix multiplication case	41
2.1.3	Inverse	44
2.2	Wednesday	49
2.2.1	Remarks on Gaussian Elimination	49
2.2.2	Properties of matrix	50
2.2.3	matrix transpose	53
2.3	Assignment Two	55
2.4	Friday	56
2.4.1	symmetric matrix	56
2.4.2	Interaction of inverse and transpose	57
2.4.3	Vector Space	58
2.5	Assignment Three	68
3	Week3	71
3.1	Tuesday	71
3.1.1	Introduction	71
3.1.2	Review of 2 weeks	72
3.1.3	Examples of solving equations	73
3.1.4	How to solve a general rectangular	78
3.2	Thursday	83
3.2.1	Review	83
3.2.2	Remarks on solving linear system equations	86
3.2.3	Linear dependence	88
3.2.4	Basis and dimension	90
3.3	Friday	96
3.3.1	Review	96

3.3.2	More on basis and dimension	97
3.3.3	What is rank?	99
3.4	Assignment Four	110
4	Midterm	113
4.1	Sample Exam	113
4.2	Midterm Exam	120
5	Week4	127
5.1	Friday	127
5.1.1	Linear Transformation	127
5.1.2	Example: differentiation	129
5.1.3	Basis Change	134
5.1.4	Determinant	136
5.2	Assignment Five	144
6	Week5	147
6.1	Tuesday	147
6.1.1	Formulas for Determinant	147
6.1.2	Determinant by Cofactors	152
6.1.3	Determinant Applications	153
6.1.4	Orthogonality	156
6.2	Thursday	160
6.2.1	Orthogonality	160
6.2.2	Least Squares Approximations	166
6.2.3	Projections	169
6.3	Friday	172
6.3.1	Orthonormal basis	172
6.3.2	Gram-Schmidt Process	176

6.3.3	The Factorization $A = QR$.	181
6.3.4	Function Space	183
6.3.5	Fourier Series	185
6.4	Assignment Six	186
7	Week6	187
7.1	Tuesday	187
7.1.1	Summary of previous weeks	187
7.1.2	Eigenvalues and eigenvectors	191
7.1.3	Products and Sums of Eigenvalue	195
7.1.4	Application: Page Rank and Web Search	196
7.2	Thursday	199
7.2.1	Review	199
7.2.2	Similarity	199
7.2.3	Diagonalization	201
7.2.4	Powers of A	206
7.2.5	Nondiagonalizable Matrices	207
7.3	Friday	210
7.3.1	Review	210
7.3.2	Fibonacci Numbers	210
7.3.3	Imaginary Eigenvalues	212
7.3.4	Complex Numbers	214
7.3.5	Complex Vectors	214
7.3.6	Spectral Theorem	220
7.3.7	Hermitian matrix	221
7.4	Assignment Seven	223
8	Week7	227
8.1	Tuesday	227
8.1.1	Quadratic form	227

8.1.2	Positive Definite Matrices	232
8.2	Thursday	241
8.2.1	SVD: Singular Value Decomposition	241
8.2.2	Remark on SVD decomposition	245
8.2.3	Best Low-Rank Approximation	253
8.3	Assignment Eight	255
9	Final Exam	257
9.1	Sample Exam	257
9.2	Final Exam	264
10	Solution	271
10.1	Assignment Solutions	271
10.1.1	Solution to Assignment One	271
10.1.2	Solution to Assignment Two	277
10.1.3	Solution to Assignment Three	280
10.1.4	Solution to Assignment Four	286
10.1.5	Solution to Assignment Five	297
10.1.6	Solution to Assignment Six	303
10.1.7	Solution to Assignment Seven	311
10.1.8	Solution to Assignment Eight	321
10.2	Midterm Exam Solutions	328
10.2.1	Sample Exam Solution	328
10.2.2	Midterm Exam Solution	338
10.3	Final Exam Solutions	346
10.3.1	Sample Exam Solution	346
10.3.2	Final Exam Solution	357

A	This is Appendix Title	371
A.1	This is First Level Heading	371
A.1.1	This is Second Level Heading	372

Contributors

ZHI-QUAN LUO, Shenzhen Research Institute of Big Data, Lecturer

RUOYU SUN, Industrial and Enterprise Systems Engineering, Lecturer

JIE WANG, The Chinese University of Hongkong, Shenzhen, Typer

Foreword

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Preface

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

place

date

Acknowledgments

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

I. R. S.

Acronyms

ASTA	Arrivals See Time Averages
BHCA	Busy Hour Call Attempts
BR	Bandwidth Reservation
b.u.	bandwidth unit(s)
CAC	Call / Connection Admission Control
CBP	Call Blocking Probability(-ies)
CCS	Centum Call Seconds
CDTM	Connection Dependent Threshold Model
CS	Complete Sharing
DiffServ	Differentiated Services
EMLM	Erlang Multirate Loss Model
erl	The Erlang unit of traffic-load
FIFO	First in - First out
GB	Global balance
GoS	Grade of Service
ICT	Information and Communication Technology
IntServ	Integrated Services
IP	Internet Protocol
ITU-T	International Telecommunication Unit – Standardization sector
LB	Local balance
LHS	Left hand side

LIFO	Last in - First out
MMPP	Markov Modulated Poisson Process
MPLS	Multiple Protocol Labeling Switching
MRM	Multi-Retry Model
MTM	Multi-Threshold Model
PASTA	Poisson Arrivals See Time Averages
PDF	Probability Distribution Function
pdf	probability density function
PFS	Product Form Solution
QoS	Quality of Service
r.v.	random variable(s)
RED	random early detection
RHS	Right hand side
RLA	Reduced Load Approximation
SIRO	service in random order
SRM	Single-Retry Model
STM	Single-Threshold Model
TCP	Transport Control Protocol
TH	Threshold(s)
UDP	User Datagram Protocol

7.2. Thursday

7.2.1. Review

- **Eigenvalue and eigenvectors:** If for **square** matrix \mathbf{A} we have

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

where $\mathbf{x} \neq \mathbf{0}$, then we say λ is the *eigenvalue*, \mathbf{x} is the *eigenvector* associated with λ .

- **How to compute eigenvalues and eigenvectors?** To solve the eigenvalue problem for matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, you should follow these steps:
 - Compute the characteristic polynomial of $\lambda\mathbf{I} - \mathbf{A}$. The determinant is a polynomial in λ of degree n .
 - Find the roots of this polynomial, by solving $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$. The n roots are the n eigenvalues of \mathbf{A} . They make $\mathbf{A} - \lambda\mathbf{I}$ singular.
 - For each eigenvalue λ , solve $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ to find a corresponding eigenvector \mathbf{x} .

7.2.2. Similarity

The similar matrices have the same eigenvalues:

Definition 7.3 [Similar] If there exists a *nonsingular* matrix \mathbf{S} such that

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S},$$

then we say \mathbf{A} is **similar** to \mathbf{B} . ■

Proposition 7.3 Let \mathbf{A} and \mathbf{B} be $n \times n$ matrices. If \mathbf{B} is *similar* to \mathbf{A} , then \mathbf{A} and \mathbf{B} have the same eigenvalues.

Proof idea. Since eigenvalues are the roots of the *characteristic polynomial*, so it suffices to prove these two polynomials are the same.


Proof. The *characteristic polynomial* for \mathbf{B} is given by

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

Since $\det(\mathbf{S}^{-1}) \det(\mathbf{S}) = 1$, we obtain:

$$\begin{aligned} P_{\mathbf{B}}(\lambda) &= \det(\lambda \mathbf{I} - \mathbf{A}) \\ &= P_{\mathbf{A}}(\lambda). \end{aligned}$$

Since they have the same *characteristic polynomial*, the roots for *characteristic polynomials* of \mathbf{A} and \mathbf{B} must be same. Therefore they have the same eigenvalues. ■

 What is invariant? In other words, what is not changed during matrix transformation?

- **Rank** is invariant under *row transformation*.
- **Eigenvalues** is invariant under *similar transformation*.
- Unluckily, similar matrices usually don't have the same eigenvectors.
It's easy to raise a counterexample.

By using eigenvalues, we have a new proof for $\det(\mathbf{S}^{-1}) = \frac{1}{\det(\mathbf{S})}$:

Proof. Suppose $\det(\mathbf{S}) = \lambda_1 \lambda_2 \dots \lambda_n$, where λ_i 's are eigenvalues of \mathbf{S} . Then there exists \mathbf{x}_i such that

$$\mathbf{S} \mathbf{x}_i = \lambda_i \mathbf{x}_i$$

for $i = 1, \dots, n$.

Since \mathbf{S} is invertible and all λ_i 's are nonzero, we imply that:

$$\mathbf{S}\mathbf{x}_i = \lambda_i\mathbf{x}_i \implies \mathbf{x}_i = \lambda_i\mathbf{S}^{-1}\mathbf{x}_i \implies \mathbf{S}^{-1}\mathbf{x}_i = \frac{1}{\lambda_i}\mathbf{x}_i$$

Hence, $\frac{1}{\lambda_i}$'s are eigenvalues of \mathbf{S}^{-1} . Since $\mathbf{S}^{-1} \in \mathbb{R}^{n \times n}$, $\frac{1}{\lambda_i}$'s ($i = 1, \dots, n$) are the only eigenvalues of \mathbf{S}^{-1} .

Hence the determinant of \mathbf{S}^{-1} is the product of its eigenvalues:

$$\det(\mathbf{S}^{-1}) = \frac{1}{\lambda_1} \frac{1}{\lambda_2} \cdots \frac{1}{\lambda_n} = \frac{1}{\det(\mathbf{S})}.$$

■

We can also use eigenvalue to proof the statement shown below:

Proposition 7.4 \mathbf{A} is singular if and only if $\det(\mathbf{A}) = 0$.

Proof. Suppose $\det(\mathbf{A}) = \lambda_1\lambda_2\cdots\lambda_n$, where λ_i 's are eigenvalues of \mathbf{A} .

Thus

$$\det(\mathbf{A}) = 0 \iff \exists \lambda_i = 0 \iff \exists \text{ nonzero } \mathbf{x} \text{ s.t. } \mathbf{A}\mathbf{x} = \lambda_i\mathbf{x} = 0\mathbf{x} = \mathbf{0}.$$

Or equivalently, \mathbf{A} is singular.


■

7.2.3. Diagonalization

Proposition (7.3) says if \mathbf{A} is similar to \mathbf{B} , then they have the same eigenvalues.

Question 1. What about the reverse direction?

Question 2. We all approve that the simplest form of a matrix to have eigenvalues $\lambda_1, \dots, \lambda_n$ is the diagonal matrix $\text{diag}(\lambda_1, \dots, \lambda_n)$. Suppose \mathbf{A} has eigenvalues $\lambda_1, \dots, \lambda_n$, is \mathbf{A} similar to the diagonal matrix $\text{diag}(\lambda_1, \dots, \lambda_n)$?

 Why the matrix $\text{diag}(\lambda_1, \dots, \lambda_n)$ has eigenvalues $\lambda_1, \dots, \lambda_n$?

Answer: Let's explain it with $n = 2$:

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} = \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The case for general n is also easy to verify.

The answers to Question 1 and 2 are both **No**! Let's raise a counterexample to explain it:

■ **Example 7.4** We give a counterexample to show that two matrices with the same eigenvalues are not necessarily similar to each other; and \mathbf{A} does not necessarily similar to the corresponding diagonal matrix.

Given $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $P_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda & -1 \\ 0 & \lambda \end{vmatrix}$. Hence its eigenvalues are $\lambda_1 = \lambda_2 = 0$.

Hence, \mathbf{A} and $\mathbf{D} = \text{diag}(0,0)$ have the same eigenvalues. Then we show that \mathbf{A} and \mathbf{D} are not similar:

Assume they are similar, which means there exists invertible matrix \mathbf{S} such that

$$\mathbf{A} = \mathbf{S}^{-1} \mathbf{D} \mathbf{S} = \mathbf{S}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{S} = \mathbf{0} \implies \text{contradiction!}$$

Suppose \mathbf{A} has eigenvalues $\lambda_1, \dots, \lambda_n$, but \mathbf{A} and $\text{diag}(\lambda_1, \dots, \lambda_n)$ may not be similar! We are curious about what kind of matrix can be similar to a diagonal matrix:

Definition 7.4 [Diagonalizable] An $n \times n$ matrix \mathbf{A} is **diagonalizable** if \mathbf{A} is similar to a *diagonal matrix*, that is to say, \exists nonsingular matrix \mathbf{S} and diagonal matrix \mathbf{D} such that

$$\mathbf{S}^{-1} \mathbf{A} \mathbf{S} = \mathbf{D} \tag{7.6}$$

We say \mathbf{S} *diagonalizes* \mathbf{A} .

R Note that Eq.(7.6) can be equivalently written as $\mathbf{AS} = \mathbf{SD}$, or in column-by-column form:

$$\mathbf{A}\mathbf{s}_i = d_i\mathbf{s}_i, \quad i = 1, \dots, n, \quad (7.7)$$

where \mathbf{s}_i denotes the i th column of \mathbf{S} , d_i denotes the (i, i) th entry of \mathbf{D} . The equivalent form Eq.(7.7) also implies that every (\mathbf{s}_i, d_i) must be an eigen-pair of \mathbf{A} . (Proposition (7.5))

Proposition 7.5 Suppose that \mathbf{A} is diagonalizable, then the column vectors of the diagonalizing matrix \mathbf{S} are eigenvectors of \mathbf{A} ; and the diagonal elements of \mathbf{D} are the corresponding eigenvalues of \mathbf{A} .

Proposition 7.6 The diagonalizing matrix \mathbf{S} is not unique.

Proof. Suppose there exists a diagonalizing matrix \mathbf{S} , verify by yourself that $\alpha\mathbf{S}$ is also a diagonalizing matrix for any $\alpha \neq 0$. ■

R We know that the reverse of proposition (7.3) is not true. However, if we add one more constraint that all eigenvalues of \mathbf{A} are distinct, the reverse is true. We will give a proof of it later.

1. If \mathbf{A} is $n \times n$ and \mathbf{A} has n distinct eigenvalues, then \mathbf{A} is diagonalizable. If the eigenvalues are not distinct, then \mathbf{A} may or may not be diagonalizable depending on whether \mathbf{A} has n linearly independent eigenvectors.

Why is diagonalizable good?

Theorem 7.3 — Diagonalization. A $n \times n$ matrix \mathbf{A} is *diagonalizable* iff \mathbf{A} has n independent eigenvectors.

Proof. Necessity. For n eigen-pairs $(\lambda_i, \mathbf{x}_i)$ of \mathbf{A} , suppose that \mathbf{x}_i 's are independent.

We after-multiply \mathbf{A} with $\mathbf{S} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix}$. The first column of \mathbf{AS} is $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1$. Hence we obtain the result for the product \mathbf{AS} :

$$\mathbf{A} \text{ times } \mathbf{S} \quad \mathbf{AS} = \mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \dots & \lambda_n\mathbf{x}_n \end{bmatrix}. \quad (7.8)$$

Note that the right side of Eq.(7.8) is essentially the product \mathbf{SD} :

$$\mathbf{S} \text{ times } \mathbf{D} \begin{bmatrix} \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \dots & \lambda_n \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \mathbf{SD}.$$

Hence we obtain $\mathbf{AS} = \mathbf{SD}$. Since \mathbf{x}_i 's are independent, there exists the inverse \mathbf{S}^{-1} .

Therefore, $\mathbf{D} = \mathbf{S}^{-1}\mathbf{AS}$.

Sufficiency. If \mathbf{A} is diagonalizable, then there exists \mathbf{S} and \mathbf{D} such that

$$\mathbf{D} = \mathbf{S}^{-1}\mathbf{AS} \quad (7.9)$$

where \mathbf{S} is nonsingular. Suppose $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, and $\mathbf{S} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix}$, where \mathbf{x}_i 's are independent.

The Eq.(7.9) can be equivalently written as $\mathbf{AS} = \mathbf{SD}$, i.e., $\mathbf{Ax}_i = \lambda_i \mathbf{x}_i$ for $i = 1, 2, \dots, n$.

Hence \mathbf{x}_i 's are the independent eigenvectors of \mathbf{A} associated with λ_i 's. ■

Diagonalizable matrix is very useful. For diagonalizable matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, it follows that its eigenvectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ are independent, i.e., form a basis for \mathbb{R}^n . Then for any $\mathbf{y} \in \mathbb{R}^n$, there exists (c_1, c_2, \dots, c_n) such that

$$\mathbf{y} = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_n \mathbf{x}_n$$

If we consider matrix \mathbf{A} as representation of linear transformation, we obtain

$$\begin{aligned} \mathbf{Ay} &= c_1 \mathbf{Ax}_1 + \dots + c_n \mathbf{Ax}_n \\ &= c_1 \lambda_1 \mathbf{x}_1 + \dots + c_n \lambda_n \mathbf{x}_n \end{aligned}$$

Hence, the linear transformation from \mathbf{y} into $\mathbf{A}\mathbf{y}$ is equivalent to transforming the coordinate coefficients from (c_1, \dots, c_n) into $(c_1\lambda_1, \dots, c_n\lambda_n)$:

$$\mathbf{y} \xrightarrow{\mathbf{A}} \mathbf{A}\mathbf{y}$$

$$(c_1, \dots, c_n) \xrightarrow{D=\text{diag}(\lambda_1, \dots, \lambda_n)} (c_1\lambda_1, \dots, c_n\lambda_n) = (c_1, \dots, c_n) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

We are curious about whether there is an useful way to determine whether \mathbf{A} is diagonalizable.

Theorem 7.4 If $\lambda_1, \dots, \lambda_k$ are *distinct* eigenvalues of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ ($n \geq k$) with the corresponding eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_k$, then $\mathbf{x}_1, \dots, \mathbf{x}_k$ are linearly independent.

Proof. • Let's start with the case $k = 2$. Assume that $\lambda_1 \neq \lambda_2$ but $\mathbf{x}_1, \mathbf{x}_2$ are dependent, i.e., $\exists (c_1, c_2) \neq \mathbf{0}$ s.t.

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0}. \quad (7.10)$$

Postmultiplying \mathbf{A} for Eq.(7.10) both sides results in

$$\mathbf{A}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = \mathbf{0} \implies c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 = \mathbf{0}. \quad (7.11)$$

Eq.(7.10) $\times \lambda_2$ - Eq.(7.11) results in:

$$(c_1\lambda_2 - c_1\lambda_1)\mathbf{x} = \mathbf{0}. \implies c_1(\lambda_2 - \lambda_1)\mathbf{x} = \mathbf{0}.$$

Since $\lambda_1 \neq \lambda_2$ and $\mathbf{x} \neq \mathbf{0}$, we derive $c_2 = 0$. Similarly, if we let Eq.(7.10) $\times \lambda_1$ - Eq.(7.11) to cancel c_2 , then we get $c_1 = 0$.

Therefore, $(c_1, c_2) = \mathbf{0}$ leads to a contradiction!

- How to proof this statement for general k ?

Assume there exists $(c_1, \dots, c_k) \neq \mathbf{0}$ s.t.

$$c_1 \mathbf{x}_1 + \dots + c_k \mathbf{x}_k = \mathbf{0} \quad (7.12)$$

Then we obtain two equations from Eq.(7.12):

$$\mathbf{A}(c_1 \mathbf{x}_1 + \dots + c_k \mathbf{x}_k) = c_1 \lambda_1 \mathbf{x}_1 + c_2 \lambda_2 \mathbf{x}_2 + \dots + c_k \lambda_k \mathbf{x}_k = \mathbf{0}. \quad (7.13)$$

$$\lambda_k(c_1 \mathbf{x}_1 + \dots + c_k \mathbf{x}_k) = c_1 \lambda_k \mathbf{x}_1 + c_2 \lambda_k \mathbf{x}_2 + \dots + c_k \lambda_k \mathbf{x}_k = \mathbf{0}. \quad (7.14)$$

We can let Eq.(7.13)–Eq.(7.14) to cancel \mathbf{x}_k :

$$c_1(\lambda_1 - \lambda_k) \mathbf{x}_1 + \dots + c_k(\lambda_{k-1} - \lambda_k) \mathbf{x}_{k-1} = \mathbf{0}. \quad (7.15)$$

By repeatedly applying the trick from (7.12) to (7.15), we can show that

$$c_1(\lambda_1 - \lambda_k) \dots (\lambda_1 - \lambda_2) \mathbf{x}_1 = \mathbf{0} \quad \text{which forces } c_1 = 0.$$

Similarly every $c_i = 0$ for $i = 1, \dots, n$. Here is the contradiction!

■

Corollary 7.1 If all eigenvalues of \mathbf{A} are *distinct*, then \mathbf{A} is *diagonalizable*

7.2.4. Powers of \mathbf{A}

Matrix Powers. If $\mathbf{A} = \mathbf{S}^{-1} \mathbf{D} \mathbf{S}$, then $\mathbf{A}^2 = (\mathbf{S}^{-1} \mathbf{D} \mathbf{S})(\mathbf{S}^{-1} \mathbf{D} \mathbf{S}) = \mathbf{S}^{-1} \mathbf{D}^2 \mathbf{S}$.

In general, $\mathbf{A}^k = (\mathbf{S}^{-1} \mathbf{D} \mathbf{S}) \dots (\mathbf{S}^{-1} \mathbf{D} \mathbf{S}) = \mathbf{S}^{-1} \mathbf{D}^k \mathbf{S}$.

Eigenvalues of matrix powers. We may ask if eigenvalues of \mathbf{A} are $\lambda_1, \dots, \lambda_n$, then what is the eigenvalues of \mathbf{A}^k ? The answer is intuitive, the eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$. However, you may use the wrong way to prove this statement:

Proposition 7.7 If eigenvalues of $n \times n$ matrix \mathbf{A} are $\lambda_1, \dots, \lambda_n$, then eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$.

Wrong proof 1: Assume $\mathbf{A} = \mathbf{S}^{-1}\mathbf{D}\mathbf{S}$, then $\mathbf{A}^k = \mathbf{S}^{-1}\mathbf{D}^k\mathbf{S}$. Suppose $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, then $\mathbf{D}^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$. Hence eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$.

This proof is wrong, because \mathbf{A} may not be *diagonalizable*, which means \mathbf{A} may not have the form $\mathbf{A} = \mathbf{S}^{-1}\mathbf{D}\mathbf{S}$. ■

Wrong proof 2: If $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, then $\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda^2\mathbf{x}$. Hence for general k , $\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}$.

This proof only states that if λ is the eigenvalue of \mathbf{A} , then λ^k is the eigenvalues of \mathbf{A}^k . Unfortunately, it still cannot derive this proposition. Because it does not prove that if λ are the eigenvalues with multiplicity m , then λ^k are the eigenvalues of \mathbf{A}^k with multiplicity m .

Let's raise a counterexample: Let eigenvalues of \mathbf{A} be $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 2$; the eigenvalues of \mathbf{A}^2 could be $1^2, 2^2, 2^2$. Hence \mathbf{A} has the eigenvalues 1 with multiplicity 2; while \mathbf{A}^2 has the eigenvalue 1^2 with multiplicity 1. So this \mathbf{A} and \mathbf{A}^2 is a contradiction for this proof. In other words, this proof fails to determine the multiplicity of eigenvalues. ■

R The proposition(7.7) could be proved using **Jordan form**, i.e., for any matrix \mathbf{A} there exists invertible matrix \mathbf{S} such that $\mathbf{A} = \mathbf{S}^{-1}\mathbf{U}\mathbf{S}$, where \mathbf{U} is an upper triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then $\mathbf{A}^k = \mathbf{S}^{-1}\mathbf{U}^k\mathbf{S}$, where \mathbf{U}^k is an upper triangular matrix with diagonal entries $\lambda_1^k, \dots, \lambda_n^k$. Hence the eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$.

7.2.5. Nondiagonalizable Matrices

Sometimes we face some matrices that have too few eigenvalues. (don't count with multiplicity)

For example, given $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, it's easy to verify that its eigenvalue is $\lambda = 0$ and eigenvectors are of the form $\mathbf{x} = \begin{bmatrix} c \\ 0 \end{bmatrix}$.

This 2×2 matrix cannot be diagonalized. Why? Let's introduce a definition first:

Definition 7.5 [Eigenspace] Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ has k distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Then the eigenspace for \mathbf{A} associated with λ_i is the collection of all eigenvectors associated with the eigenvalue λ_i , i.e., the null space $N(\lambda_i \mathbf{I} - \mathbf{A})$. ■

Why does this 2 by 2 matrix \mathbf{A} cannot be diagonalizable? Because the dimension of its eigenspace is too small, i.e., $\text{eigenspace}(\mathbf{A}, \lambda = 0) = 1 < 2$. In general, if the dimension of eigenspace associated with the eigenvalue λ_i is less than the multiplicity of this eigenvalue, then this matrix cannot be diagonalizable. We will discuss it in the next lecture.

