

**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 and Projection	156
6.2	Thursday	160
6.2.1	Orthogonality and Projection	160
6.2.2	Least Squares Approximations	165
6.2.3	Projections	168
6.3	Friday	171
6.3.1	Orthonormal basis	171
6.3.2	Gram-Schmidt Process	174

6.3.3	The Factorization $A = QR$.	180
6.3.4	Function Space	183
6.3.5	Fourier Series	184
6.4	Assignment Six	186
7	Week6	187
7.1	Tuesday	187
7.1.1	Summary of last two weeks	187
7.1.2	Eigenvalues and eigenvectors	191
7.1.3	Products and Sums of Eigenvalue	196
7.1.4	Application: Page Rank and Web Search	197
7.2	Thursday	200
7.2.1	Review	200
7.2.2	Similarity and eigenvalues	200
7.2.3	Diagonalization	203
7.2.4	Powers of A	208
7.2.5	Nondiagonalizable Matrices	209
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

Chapter 5

Week4

5.1. Friday

5.1.1. Linear Transformation

We start with a matrix \mathbf{A} . When multiplying \mathbf{A} with a vector \mathbf{v} , it essentially transforms \mathbf{v} to another vector \mathbf{Av} . Matrix multiplication $L(\mathbf{v}) = \mathbf{Av}$ gives a **linear transformation**:

Definition 5.1 [linear transformation] A transformation L assigns an output $T(\mathbf{v})$ to each input vector \mathbf{v} in V .

The transformation $L(\cdot)$ is said to be a **linear transformation** if it satisfies

$$L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$$

for all vectors $\mathbf{v}_1, \mathbf{v}_2$ and scalars α, β . ■

Key Observation: If the input is $\mathbf{v} = \mathbf{0}$, the output must be $L(\mathbf{v}) = \mathbf{0}$.

5.1.1.1. The idea of linear transformation

Given the linear transformation $L : \mathbb{R}^n \mapsto \mathbb{R}^m$, let's show that in order to study the output, it suffices to start from the **basis** of our output:

Assume the basis of \mathbb{R}^n is $\{e_1, e_2, \dots, e_n\}$, where $L(e_i) = a_i \in \mathbb{R}^m$ for $i = 1, \dots, n$. **The linearity of transformation extends to the combinations of n vectors.**

Hence given any vector $\mathbf{x} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \in \mathbb{R}^n$, we can express its trans-

formation in matrix multiplication form:

$$\begin{aligned}
L(\mathbf{x}) &= L(x_1e_1 + x_2e_2 + \cdots + x_ne_n) \\
&= x_1L(e_1) + x_2L(e_2) + \cdots + x_nL(e_n) \\
&= x_1a_1 + x_2a_2 + \cdots + x_na_n = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \mathbf{Ax}
\end{aligned}$$

where $a_i := L(e_i)$, and \mathbf{A} is a $m \times n$ matrix with columns a_1, \dots, a_n .

5.1.1.2. Matrix defines linear transformation

Conversely, given $m \times n$ matrix \mathbf{A} , $L(\mathbf{x}) = \mathbf{Ax}$ defines a linear mapping. This is because matrix multiplication is also a linear operator.

- R** Transformations have a new “language”. For example, for *nonlinear* transformation, if there is **no matrix**, we cannot talk about **column space**. But this idea could be rescued. We know the *column space* consists of all outputs \mathbf{Av} , the *null space* consists of all inputs for which $\mathbf{Av} = \mathbf{0}$. We could generalize those terms into “range” and “kernel”:

Definition 5.2 [range] For a linear transformation $L : V \mapsto W$, the range (or image) of L refers to the set of all outputs $L(\mathbf{v})$, which is denoted as:

$$\text{Range}(L) = \{L(\mathbf{x}) : \mathbf{x} \in \mathbf{V}\}$$

Sometimes we also use notation $\text{Im}(L)$ to express the same thing. ■

The range corresponds to the column space. If $L(\mathbf{x}) = \mathbf{Ax}$, we have $\text{Range}(L) = \mathcal{C}(\mathbf{A})$.

Definition 5.3 [kernel] The kernel of L refers to the set of all inputs for which $L(\mathbf{v}) = \mathbf{0}$, which is denoted as:

$$\ker(L) = \{\mathbf{x} : L(\mathbf{x}) = \mathbf{0}\}$$

Kernel corresponds to the null space. If $L(\mathbf{x}) = \mathbf{Ax}$, we have $\ker(L) = N(\mathbf{A})$.

R For linear transformation $L : \mathbf{V} \mapsto \mathbf{W}$, where $L(\mathbf{x}) = \mathbf{Ax}$. We have two rules:

$$L(\cdot) : \begin{cases} N(\mathbf{A}) \mapsto \{\mathbf{0}\} \\ \mathbf{V} \mapsto \text{col}(\mathbf{A}) \end{cases}$$

5.1.2. Example: differentiation

Key idea of this section:

Suppose we know $L(\mathbf{v}_1), \dots, L(\mathbf{v}_n)$ for the basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then the linearity property produces $L(\mathbf{v})$ for every other input vector \mathbf{v}

Reason: Every \mathbf{v} has a unique combination $c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$ of the basis vector \mathbf{v}_i . Suppose L is a linear transformation, then $L(\mathbf{v})$ must be the **same combination** $c_1L(\mathbf{v}_1) + \dots + c_nL(\mathbf{v}_n)$ of the **known outputs** $L(\mathbf{v}_i)$.

Derivative is a linear transformation. The derivative of the functions $1, x, x^2, x^3$ are $0, 1, 2x, 3x^2$. If we consider “**taking the derivative**” as a transformation, whose inputs and outputs are functions, then we claim that the **derivative transformation** is **linear**:

$$L(\mathbf{v}) = \frac{d\mathbf{v}}{dx} \quad \text{obeys the linearity rule} \quad \frac{d}{dx}(c\mathbf{v} + d\mathbf{w}) = c \frac{d\mathbf{v}}{dx} + d \frac{d\mathbf{w}}{dx}$$

If we consider $1, x, x^2, x^3$ as vectors instead of functions, we notice they form a basis for the space $\mathbf{V} := \{\text{polynomials with degree} \leq 3\}$. Find derivatives of these four basis tells us all derivatives in \mathbf{V} :

■ **Example 5.1** Given any vector \mathbf{v} in \mathbf{V} , it can be expressed as $\mathbf{v} = a + bx + cx^2 + dx^3$.

We want to find the derivative transformation output for \mathbf{v} :

$$\begin{aligned} L(\mathbf{v}) &= aL(1) + bL(x) + cL(x^2) + dL(x^3) \\ &= a \times (0) + b \times (1) + c \times (2x) + d \times (3x^2) \\ &= b + 2cx + 3dx^2 \end{aligned}$$

Can we express this linear transformation L by a matrix \mathbf{A} ? The answer is Yes:

The derivative transforms the space \mathbf{V} of cubics to the space \mathbf{W} of quadratics. The basis for \mathbf{V} is $1, x, x^2, x^3$. The basis for \mathbf{W} is $1, x, x^2$. It follows that *The derivative matrix is 3 by 4*:

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \text{matrix form of derivative } L.$$

Why do we define the derivative matrix? Because **multiplying by \mathbf{A} agrees with transforming by L** . The derivative of $\mathbf{v} = a + bx + cx^2 + dx^3$ is $L(\mathbf{v}) = b + 2cx + 3dx^2$. The same numbers $b, 2c, 3d$ appear when we multiply by matrix \mathbf{A} :

$$\text{Take the derivative} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix}.$$

What does the matrix $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ and $\begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix}$ mean?

It is the **coordinate vector** of \mathbf{v} and $L(\mathbf{v})$. If we consider $a + bx + cx^2 + dx^3$ as a

vector, then it's better for us to study its corresponding coordinate vector

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Hence, taking derivative of \mathbf{v} is the same as multiplying matrix \mathbf{A} by its coordinate vector. ■

5.1.2.1. The inverse of the derivative.

The integral is the inverse of the derivative. . That is from the Fundamental Theorem of Calculus. We review it from the perspective of linear algebra. The integral transformation L^{-1} that *takes the integral from 0 to x* is also linear! Applying L^{-1} to $1, x, x^2$, which are $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$:

$$\text{Integration is } L^{-1} \quad \int_0^x 1 \, dx = x, \quad \int_0^x x \, dx = \frac{1}{2}x^2, \quad \int_0^x x^2 \, dx = \frac{1}{3}x^3.$$

By linearity, the integral of $\mathbf{w} = B + Cx + Dx^2$ is $L^{-1}(\mathbf{w}) = Bx + \frac{1}{2}Cx^2 + \frac{1}{3}Dx^3$. The integral of a quadratic is a cubic. The input space of L^{-1} is the quadratics, the output space is the cubics. **Integration takes \mathbf{W} back to \mathbf{V} .** Integration matrix will be 4 by 3:

$$\text{Take the integral} \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ B \\ \frac{1}{2}C \\ \frac{1}{3}D \end{bmatrix}.$$

If our input is $\mathbf{w} = B + Cx + Dx^2$, our output integral is $0 + Bx + \frac{1}{2}Cx^2 + \frac{1}{3}Dx^3$.

The derivative and the integration are essentially matrix multiplication. We have the corresponding derivative and integration matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

I want to call this matrix \mathbf{A}^{-1} , though rectangular matrices don't have inverses. Note that \mathbf{A}^{-1} is the **right inverse** of matrix \mathbf{A} ! (Do you remember the definition that shown in mid-term?)

$$\mathbf{A}\mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{but} \quad \mathbf{A}^{-1}\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This is reasonable. If you integrate a function and then differentiate, you get back to the start. Hence $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. But if you differentiate before integrating, the constant term is lost.

The integral of the derivative of 1 is zero.

$$L^{-1}L(1) = \text{integral of zero function} = 0.$$

Summary:. In this example, we want to take the derivative. Then we let \mathbf{V} be a vector space of polynomials with degree ≤ 3 . Its basis is given by $E = \{1, x, x^2, x^3\}$. Any $v \in \mathbf{V}$ there is a unique linear combination of the basis vectors that equals to v :

$$v = a + bx + cx^2 + dx^3$$

We write the coordinate vector of v w.r.t. to E :

$$[v]_E = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Then we postmultiply \mathbf{A} by $[v]_E$ to get the corresponding coordinate vector of output space:

$$[L(v)]_F = \mathbf{A}[v]_E$$

where $F = \{1, x, x^2\}$.

Here we give the formal definition for the coordinate vector:

Definition 5.4 [coordinate vector] Let \mathbf{V} be a vector space of dimension n and let $B = \{v_1, v_2, \dots, v_n\}$ be an **ordered** basis for \mathbf{V} . Then for any $v \in \mathbf{V}$ there is a unique linear combination of the basis vectors such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$$

where $\alpha_1, \dots, \alpha_n$ are scalars.

The **coordinate vector** of v w.r.t. to B is defined by

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Hence, vector v could be expressed as: $v = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [v]_B$. ■

More specifically, the linear transformation of vectors is essentially the matrix multiplication of the corresponding coordinate vectors:

Theorem 5.1 Let $E = \{v_1, \dots, v_n\}$ be a basis for \mathbf{V} ; $F = \{w_1, \dots, w_m\}$ be a basis for \mathbf{W} . Given linear transformation $L : \mathbf{V} \mapsto \mathbf{W}$, for any vector $v \in \mathbf{V}$, there exists $m \times n$ matrix \mathbf{A} such that

$$[L(v)]_F = \mathbf{A}[v]_E$$

If we let $\mathbf{W} = \mathbf{V}$, then we obtain a more commonly useful corollary:

Corollary 5.1 Given linear transformation $L : \mathbf{V} \mapsto \mathbf{V}$. We set $E = \{\alpha_1, \dots, \alpha_n\}$ to be the basis of \mathbf{V} . Then given any vector v , there exists $n \times n$ matrix \mathbf{A} such that

$$[L(v)]_E = \mathbf{A}[v]_E$$

5.1.3. Basis Change

Basis Change is essentially matrix multiplication. Suppose $L : \mathbf{V} \mapsto \mathbf{V}$. $E = \{v_1, \dots, v_n\}$ is a basis for \mathbf{V} , $F = \{u_1, \dots, u_n\}$ is another basis for \mathbf{V} . Then vector u_1, \dots, u_n could be expressed by vectors v_1, \dots, v_n . So we set

$$u_1 = s_{11}v_1 + s_{12}v_2 + \dots + s_{1n}v_n,$$

$$u_2 = s_{21}v_1 + s_{22}v_2 + \dots + s_{2n}v_n,$$

$$\dots$$

$$u_n = s_{n1}v_1 + s_{n2}v_2 + \dots + s_{nn}v_n.$$

We could write this system into matrix form:

$$(u_1, \dots, u_n) = (v_1, \dots, v_n) \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \vdots & \vdots & \dots & \vdots \\ s_{n1} & s_{n2} & \dots & s_{nn} \end{pmatrix}.$$

We set $\mathbf{S} = (s_{ij})$. Hence we obtain:

$$(u_1, \dots, u_n) = (v_1, \dots, v_n)\mathbf{S}. \quad (5.1)$$

You should **prove it by yourself** that \mathbf{S} is invertible. Hence we have:

$$(u_1, \dots, u_n)\mathbf{S}^{-1} = (v_1, \dots, v_n). \quad (5.2)$$

We can express linear transformation in terms of different basis. Given any vector $x \in \mathbf{V}$, we want to study the relationship between $L(x)$ and $[x]_F$:

$$\begin{aligned} L(x) &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [L(x)]_E \\ &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times (\mathbf{A}[x]_E) \quad \leftarrow \text{due to corollary (5.1)} \\ &= \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \mathbf{S}^{-1} \times (\mathbf{A}[x]_E) \end{aligned} \quad (5.3)$$

- We claim that $[x]_E = \mathbf{S}[x]_F$:

For any vector $x \in \mathbf{V}$, we obtain:

$$\begin{aligned} x &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [x]_E \\ &= \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \times [x]_F \\ &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times \mathbf{S}[x]_F \end{aligned}$$

Hence $[x]_E = \mathbf{S}[x]_F$.

Substituting $[x]_E = \mathbf{S}[x]_F$ into Eq.(5.3), we obtain:

$$L(x) = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \mathbf{S}^{-1} \mathbf{A} \mathbf{S} [x]_F$$

What do the following process mean? We know that given basis $E = \{v_1, \dots, v_n\}$, per-

forming linear transformation on any vector x is just the same as matrix multiplication:

$$L(x) = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times \mathbf{A}[x]_E$$

In summary,

1. The linear transformation is essentially postmultiplying matrix for the coordinate vector:

$$x = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [x]_E \implies L(x) = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times \mathbf{A}[x]_E$$

2. If we change another basis $F = \{u_1, \dots, u_n\}$, we must change \mathbf{A} into $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$:

$$x = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \times [x]_F \implies L(x) = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \times \mathbf{S}^{-1}\mathbf{A}\mathbf{S}[x]_F$$

It suffices to define $\mathbf{B} := \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, The matrix \mathbf{B} is said to be **similar** to \mathbf{A} .

Definition 5.5 [Similar] Let \mathbf{A}, \mathbf{B} be $n \times n$ matrix. If there exists invertible $n \times n$ matrix \mathbf{S} such that $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, then we say that \mathbf{A} is **similar** to \mathbf{B} . ■

5.1.4. Determinant

The determinant of a **square matrix** is a single number, which contains many amazing amount of information about the matrix. It has four major uses:

The determinant is zero if and only if the matrix has no inverse.

It can be used to calculate the area or volume of a box. $|\det(\mathbf{A})|$ is the volume of the parallelepiped $\mathcal{P} = \{y = \sum_{i=1}^m \alpha_i \mathbf{a}_i \mid \alpha_i \in [0, 1]\}$:

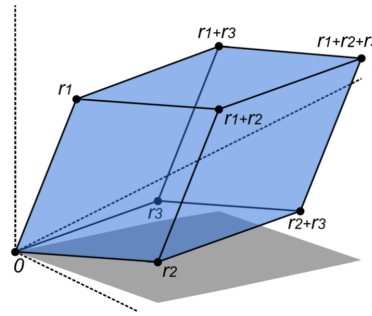


Figure 5.1: The parallelepiped $\mathcal{P} = \{y = \sum_{i=1}^3 \alpha_i \mathbf{a}_i \mid \alpha_i \in [0, 1]\}$, where r_1, r_2, r_3 are $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ on \mathbb{R}^3


The product of all the pivots $= (\pm 1) \times$ the determinant. For a 2 by 2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the pivots are a and $d - (\frac{c}{a})b$. The product of pivots is the determinant:

$$\text{Product of pivots } a(d - \frac{c}{a}b) = ad - bc \quad \text{which is } \det \mathbf{A}$$

Compute determinants to find \mathbf{A}^{-1} and $\mathbf{A}^{-1}\mathbf{b}$. (Cramer's Rule).

5.1.4.1. The properties of the Determinant

We don't intend to define the determinant directly by its formulas. It's better to start with its properties. These properties are simple, but they prepare for the formulas.

 Brackets for the matrix, straight bars for its determinant. For example,

$$\text{The determinant of } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ is } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinant is written in two ways, $\det \mathbf{A}$ or $|\mathbf{A}|$.

We will introduce three basic properties, then we will show how properties 1 – 3 derive other properties.

1. The determinant of the n by n identity matrix is 1.

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \quad \text{and} \quad \begin{vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{vmatrix} = 1.$$

2. The determinant changes sign when two rows are exchanged. (sign reversal)

$$\text{Check: } \begin{vmatrix} c & d \\ a & b \end{vmatrix} = - \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad (\text{both sides equal } bc - ad).$$

3. The determinant is a linear function of each row separately. (all other rows stay fixed).

$$\text{multiply row 1 by any number } t \quad \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\text{Add row 1 of } A \text{ to row 1 of } B: \quad \begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c & d \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ c & d \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c & d \end{vmatrix}$$

Note that this rule **does not** mean $\det(\mathbf{A} + \mathbf{B}) = \det \mathbf{A} + \det \mathbf{B}$.

Note that this rule **does not** mean $\det(t\mathbf{A}) = t \det(\mathbf{A})$.

Actually, $\det(t\mathbf{A}) = t^n \det \mathbf{A}$. This is reasonable. Imagining that expanding a rectangle by 2, its area will increase by 4. Expand an n -dimensional box by t and its volume will increase by t^n .

Pay special attention to property 1 ~ 3. They completely determine the $\det \mathbf{A}$. We could stop here to find a formula for determinants. But before that we prefer to derive other properties that follow directly from the first three:

4. If two rows of A are equal, then $\det A = 0$.

$$\text{Check 2 by 2: } \begin{vmatrix} a & b \\ a & b \end{vmatrix} = 0.$$

Property 4 follows from Property 2.

Proofoutline. Exchange the two equal row. The determinant D is supposed to change sign. But also the matrix is not changed, so we have $-D = D \implies D = 0$. ■

5. Adding a constant multiple of a row to another row doesn't change $\det A$.

$$\begin{vmatrix} a+lc & b+ld \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} lc & ld \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + l \begin{vmatrix} c & d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det A$$

Conclusion: The determinant is not changed by the usual elimination step from A to U . Since every row exchange reverses the sign, we have $\det A = \pm \det U$.

6. If A is triangular, then $\det A = \text{product of diagonal entries}$.

$$\text{Triangular} \quad \begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = ad \quad \text{and also} \quad \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = ad$$

Suppose all diagonal entries of A are nonzero. We do Gaussian Elimination to convert A into diagonal matrix:

$$\det \begin{bmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{bmatrix} = a_{11}a_{22}\dots a_{nn}.$$

Factor a_{11} from the first row by property 3; then factor a_{22} from the second row;..... Finally the determinant is $a_{11} \times a_{22} \times a_{33} \dots \times a_{nn} \times \det I = a_{11} \times a_{22} \times a_{33} \dots \times a_{nn}$.

7. $\det(AB) = \det(A)\det(B)$.

Proof.

- If $|\mathbf{B}|$ is zero, it's easy to verify that \mathbf{B} is singular, then \mathbf{AB} is singular. Thus $\det(\mathbf{AB}) = 0 = \det(\mathbf{A})\det(\mathbf{B})$.
- Suppose $|\mathbf{B}|$ is not zero, and \mathbf{A}, \mathbf{B} is $n \times n$ matrix. Consider the ratio $D(\mathbf{A}) = \frac{|\mathbf{AB}|}{|\mathbf{B}|}$. Check that this ratio has properties 1,2,3. If so, $D(\mathbf{A})$ has to be the determinant, say, $|\mathbf{A}|$. Thus we have $|\mathbf{A}| = \frac{|\mathbf{AB}|}{|\mathbf{B}|}$:

Property 1 (*Determinant of I*) If $\mathbf{A} = \mathbf{I}$, then the ratio becomes $D(\mathbf{A}) = \frac{|\mathbf{B}|}{|\mathbf{B}|} = 1$.

Property 2 (*Sign reversal*) When two rows of \mathbf{A} are exchanged, the same two rows of \mathbf{AB} are also exchanged. Therefore $|\mathbf{AB}|$ changes sign and so does the ratio $\frac{|\mathbf{AB}|}{|\mathbf{B}|}$.

Property 3 (*Linearity*) When row 1 of \mathbf{A} is multiplied by t , so is row 1 of \mathbf{AB} . Thus the ratio is also increased by t . Thus we still have $|\mathbf{A}| = \frac{|\mathbf{AB}|}{|\mathbf{B}|}$. If we Add row 1 of \mathbf{A}_1 to row 1 of \mathbf{A}_2 . Then row 1 of $\mathbf{A}_1\mathbf{B}$ also adds to row 1 of $\mathbf{A}_2\mathbf{B}$. By property three, determinants add. After dividing by $|\mathbf{B}|$, the ratios add. Hence we still have $|\mathbf{A}| = \frac{|\mathbf{AB}|}{|\mathbf{B}|}$.

Conclusion: The ratio $D(\mathbf{A})$ has the same three properties that defines determinant, hence it equals $|\mathbf{A}|$. Hence we obtain the product rule $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

■

Immediately here follows a corollary:

Corollary 5.2

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$$

8. The transpose A^T has the same determinant as A .

$$\text{Transpose} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} \quad \text{Both sides equal } ad - bc.$$

- Proof.*
- When A is singular, A^T is also singular. Hence $|A^T| = |A| = 0$.
 - Otherwise A has LU decomposition $PA = LU$. Transposing both sides gives $A^T P^T = U^T L^T$. By product rule we have

$$\det P \det A = \det L \det U \quad \text{and} \quad \det A^T \det P^T = \det U^T \det L^T.$$

- Firstly, $\det L = \det L^T = 1$. (By property 6, they both have 1's on the diagonal).
- Secondly, $\det U = \det U^T$. (By property 6, they have the same diagonal)
- Thirdly, $\det P = \det P^T$. (Verify by yourself that $P^T P = I$. Hence $|P^T| |P| = 1$. Since permutation matrix is obtained by exchanging rows of I , the only possible value for determinant of permutation matrix is ± 1 . Hence P and P^T must both equal to 1 or both equal to -1).

So L, U, P has the same determinants as L^T, U^T, P^T , Hence we have $\det A = \det A^T$.

■

