

**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](http://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

<b>Contributors</b>	<b>v</b>
<b>Foreword</b>	<b>vii</b>
<b>Preface</b>	<b>ix</b>
<b>Acknowledgments</b>	<b>xi</b>
<b>Acronyms</b>	<b>xiii</b>
<b>1 Week1 .....</b>	<b>1</b>
<b>1.1 Tuesday</b>	<b>1</b>
1.1.1 Introduction .....	1
1.1.2 Gaussian Elimination .....	3
1.1.3 Complexity Analysis .....	11
1.1.4 Brief Summary .....	12
<b>1.2 Thursday</b>	<b>14</b>
1.2.1 Row-Echelon Form .....	14
1.2.2 Matrix Multiplication .....	16
1.2.3 Special Matrices .....	19
<b>1.3 Friday</b>	<b>21</b>
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
<b>1.4 Assignment One</b>	<b>36</b>

<b>2</b>	<b>Week2</b>	<b>39</b>
<b>2.1</b>	<b>Tuesday</b>	<b>39</b>
2.1.1	Review	39
2.1.2	Special matrix multiplication case	41
2.1.3	Inverse	44
<b>2.2</b>	<b>Wednesday</b>	<b>49</b>
2.2.1	Remarks on Gaussian Elimination	49
2.2.2	Properties of matrix	50
2.2.3	matrix transpose	53
<b>2.3</b>	<b>Assignment Two</b>	<b>55</b>
<b>2.4</b>	<b>Friday</b>	<b>56</b>
2.4.1	symmetric matrix	56
2.4.2	Interaction of inverse and transpose	57
2.4.3	Vector Space	58
<b>2.5</b>	<b>Assignment Three</b>	<b>68</b>
<b>3</b>	<b>Week3</b>	<b>71</b>
<b>3.1</b>	<b>Tuesday</b>	<b>71</b>
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
<b>3.2</b>	<b>Thursday</b>	<b>83</b>
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
<b>3.3</b>	<b>Friday</b>	<b>96</b>
3.3.1	Review	96

3.3.2	More on basis and dimension	97
3.3.3	What is rank?	99
<b>3.4</b>	<b>Assignment Four</b>	<b>110</b>
<b>4</b>	<b>Midterm</b>	<b>113</b>
<b>4.1</b>	<b>Sample Exam</b>	<b>113</b>
<b>4.2</b>	<b>Midterm Exam</b>	<b>120</b>
<b>5</b>	<b>Week4</b>	<b>127</b>
<b>5.1</b>	<b>Friday</b>	<b>127</b>
5.1.1	Linear Transformation	127
5.1.2	Example: differentiation	129
5.1.3	Basis Change	134
5.1.4	Determinant	136
<b>5.2</b>	<b>Assignment Five</b>	<b>144</b>
<b>6</b>	<b>Week5</b>	<b>147</b>
<b>6.1</b>	<b>Tuesday</b>	<b>147</b>
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
<b>6.2</b>	<b>Thursday</b>	<b>160</b>
6.2.1	Orthogonality	160
6.2.2	Least Squares Approximations	166
6.2.3	Projections	169
<b>6.3</b>	<b>Friday</b>	<b>172</b>
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
<b>6.4</b>	<b>Assignment Six</b>	<b>186</b>
<b>7</b>	<b>Week6</b>	<b>187</b>
<b>7.1</b>	<b>Tuesday</b>	<b>187</b>
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
<b>7.2</b>	<b>Thursday</b>	<b>200</b>
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
<b>7.3</b>	<b>Friday</b>	<b>210</b>
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
<b>7.4</b>	<b>Assignment Seven</b>	<b>223</b>
<b>8</b>	<b>Week7</b>	<b>227</b>
<b>8.1</b>	<b>Tuesday</b>	<b>227</b>
8.1.1	Quadratic form	227



8.1.2	Positive Definite Matrices . . . . .	232
<b>8.2</b>	<b>Thursday</b>	<b>241</b>
8.2.1	SVD: Singular Value Decomposition . . . . .	241
8.2.2	Remark on SVD decomposition . . . . .	245
8.2.3	Best Low-Rank Approximation . . . . .	253
<b>8.3</b>	<b>Assignment Eight</b>	<b>255</b>
<b>9</b>	<b>Final Exam . . . . .</b>	<b>257</b>
<b>9.1</b>	<b>Sample Exam</b>	<b>257</b>
<b>9.2</b>	<b>Final Exam</b>	<b>264</b>
<b>10</b>	<b>Solution . . . . .</b>	<b>271</b>
<b>10.1</b>	<b>Assignment Solutions</b>	<b>271</b>
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
<b>10.2</b>	<b>Midterm Exam Solutions</b>	<b>328</b>
10.2.1	Sample Exam Solution . . . . .	328
10.2.2	Midterm Exam Solution . . . . .	338
<b>10.3</b>	<b>Final Exam Solutions</b>	<b>346</b>
10.3.1	Sample Exam Solution . . . . .	346
10.3.2	Final Exam Solution . . . . .	357

<b>A</b>	<b>This is Appendix Title .....</b>	<b>371</b>
<b>A.1</b>	<b>This is First Level Heading</b>	<b>371</b>
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

## 6.3. Friday

This lecture has two goals. The first is to see **how orthogonality makes it easy to find the projection matrix  $P$  and the projection  $\text{Proj}_{\mathcal{C}(A)} \mathbf{b}$** . The key idea is that *Orthogonality makes the product  $A^T A$  a diagonal matrix*. The second goal is to **show how to construct orthogonal basis of  $\mathcal{C}(A)$** . For matrix  $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ , the columns may not be orthogonal. We intend to convert  $a_1, \dots, a_n$  to orthogonal vectors, which will be the columns of a new matrix  $Q$ .

### 6.3.1. Orthonormal basis

The vectors  $\mathbf{q}_1, \dots, \mathbf{q}_n$  are **orthogonal** when their inner product  $\langle \mathbf{q}_i, \mathbf{q}_j \rangle$  are zero. ( $i \neq j$ .) With one more step—each vector is just divided by its length, then the collection of vectors become **orthogonal unit vectors**. Their lengths are all 1. Then this basis is called **orthonormal**.

**Definition 6.15** [orthonormal] The collection of vectors  $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{R}^m$  is said to be:

- **orthogonal** if  $\langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0$  for all  $i, j$  with  $i \neq j$
- **orthonormal** if  $\|\mathbf{q}_i\|_2 = 1$  for all  $i$  and  $\langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0$  for all  $i, j$  with  $i \neq j$ , or equivalently,

$$\langle \mathbf{q}_i, \mathbf{q}_j \rangle = \begin{cases} 0 & \text{when } i \neq j & (\text{orthogonal vectors}), \\ 1 & \text{when } i = j & (\text{unit vectors: } \|\mathbf{q}_i\| = 1). \end{cases}$$

Moreover, if  $\mathbf{q}_1, \dots, \mathbf{q}_n$  are **orthonormal**, then the basis  $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$  is called **orthonormal basis**. ■

■ **Example 6.6** Given a collection of unit vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

then  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  forms an *orthonormal basis* for  $\mathbb{R}^n$ . ■

If we want to express vector  $\mathbf{b}$  as the linear combination of arbitrary basis (may not be orthogonal)  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ , what should we do?

**Answer:** Solve the system  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where  $\mathbf{A} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$

What if  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  is an **orthogonal** basis? How to find solution  $\mathbf{x}$  s.t.

$$\mathbf{b} = x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + \cdots + x_n\mathbf{q}_n? \quad (6.11)$$

**Answer:** We just do the inner product of each  $\mathbf{q}_i$  with  $\mathbf{b}$  to get the coefficient  $x_i$ :

$$\begin{aligned} \langle \mathbf{q}_i, \mathbf{b} \rangle &= x_1 \langle \mathbf{q}_i, \mathbf{q}_1 \rangle + x_2 \langle \mathbf{q}_i, \mathbf{q}_2 \rangle + \cdots + x_n \langle \mathbf{q}_i, \mathbf{q}_n \rangle \\ &= x_i \langle \mathbf{q}_i, \mathbf{q}_i \rangle = x_i \end{aligned} \quad (6.12)$$

By substituting Eq.(6.12) into Eq.(6.11), we could express  $\mathbf{b}$  as:

$$\mathbf{b} = \sum_{i=1}^n \langle \mathbf{q}_i, \mathbf{b} \rangle \mathbf{q}_i.$$

In this case, from Eq.(6.12) we can see that if columns of  $\mathbf{A}$  are orthogonal, we could easily obtain the solution to  $\mathbf{A}\mathbf{x} = \mathbf{b}$ :

$$x_i = \langle \mathbf{q}_i, \mathbf{b} \rangle, \quad i = 1, 2, \dots, n.$$

**Definition 6.16** [matrix with orthonormal columns] Given a collection of **orthonormal** vectors  $\mathbf{q}_1, \dots, \mathbf{q}_n$ , the matrix

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix}$$

is said to be a matrix with **orthonormal** columns.

Note that a matrix with **orthonormal** columns is often denoted as  $\mathbf{Q}$ . ■

Or equivalently, a matrix  $\mathbf{Q}$  is with **orthonormal** columns if and only if

$$\mathbf{Q}^T \mathbf{Q} = \begin{pmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1^T \mathbf{q}_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \mathbf{q}_n^T \mathbf{q}_n \end{pmatrix} = \mathbf{I}. \quad (6.13)$$

**R** Note that a matrix  $\mathbf{Q}$  with orthonormal columns is *not required to be square*! Moreover,  $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$  in  $\mathbf{Q}$  is *not required to form a basis*.

**Definition 6.17** [orthogonal matrix] A matrix  $\mathbf{Q}$  is said to be **orthogonal** if it is square and its columns are orthonormal. ■

Question: Why we call it an orthogonal matrix, but not an orthonormal matrix?

Answer: Orthogonal matrix usually transform an orthogonal basis into another orthogonal basis by matrix multiplication. This special property requires its column to be **orthonormal**.

■ **Example 6.7** If  $\mathbf{Q}$  is an orthogonal matrix, while  $\hat{\mathbf{Q}}$  is a matrix with orthonormal columns that is **not square**. Do the products  $\mathbf{Q}\mathbf{Q}^T$  and  $\hat{\mathbf{Q}}\hat{\mathbf{Q}}^T$  always be *identity matrix*?

Answer:

- $\mathbf{Q}\mathbf{Q}^T$  is always *identity matrix*. According to equation (6.13), we have  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ .

Hence  $Q^T$  is the left inverse of square matrix  $Q$ , which implies

$$Q^{-1} = Q^T \implies QQ^T = QQ^{-1} = I.$$

Moreover, solving  $Qx = b$  is equivalent to  $x = Q^{-1}b = Q^Tb$ , which is *exactly*

$$x = \begin{bmatrix} \langle q_1, b \rangle \\ \langle q_2, b \rangle \\ \vdots \\ \langle q_n, b \rangle \end{bmatrix}.$$

- Although  $\hat{Q}^T\hat{Q} = I$ , the product  $\hat{Q}\hat{Q}^T$  will never be identity matrix for nonsquare  $\hat{Q}$ . We can verify it by the its rank:

Assume  $\hat{Q} \in \mathbb{R}^{m \times n} (m \neq n)$ . Then it's easy to verify that  $\text{rank}(\hat{Q}\hat{Q}^T) = \text{rank}(\hat{Q})$ .

Since  $\hat{Q}$  has orthonormal columns, the columns of  $\hat{Q}$  are independent, i.e.,  $\text{rank}(\hat{Q}) = n$ . But  $\text{rank}(\hat{Q}\hat{Q}^T) = \text{rank}(\hat{Q}) = n \neq m = \text{rank}(I_m)$ .

Moreover, if  $\hat{Q}$  has only one column  $\hat{q}$ , then  $\hat{Q}\hat{Q}^T = \hat{q}\hat{q}^T = \text{rank}(1) \neq \text{rank}(I_m)$ .

### Proposition 6.2

If  $Q$  has orthonormal columns, then it *leaves lengths unchanged*, in other words,

$$\text{Same length} \quad \|Qx\| = \|x\| \text{ for every vector } x.$$

Also,  $Q$  preserves inner products for vectors, i.e., :

$$\langle Qx, Qy \rangle = \langle x, y \rangle \quad \text{for every vectors } x \text{ and } y.$$

*Proofoutline.*  $\|Qx\|^2 = \|x\|^2$  because

$$\begin{aligned} \langle Qx, Qx \rangle &= x^T Q^T Qx = x^T (Q^T Q)x \\ &= x^T Ix = x^T x \end{aligned}$$



Hence we have  $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ . Just using  $Q^T Q = I$ , we can derive  $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ . ■

Orthogonal matrices are excellent for computations, since the inverse of matrices could usually be converted into transpose.

**When Least Squares Meet Orthogonality.** In particular, if  $Q \in \mathbb{R}^{m \times n}$  has orthonormal columns, the least square problem is easy:

Although  $Q\mathbf{x} = \mathbf{b}$  may not have a solution, but the normal equation

$$Q^T Q \hat{\mathbf{x}} = Q^T \mathbf{b}$$

must have the unique solution  $\hat{\mathbf{x}} = Q^T \mathbf{b}$ . Why? Since  $Q^T Q = I$ , we derive

$$\hat{\mathbf{x}} = (Q^T Q)^{-1} Q^T \mathbf{b} = Q^T \mathbf{b}.$$

### 6.3.1.1. Summary

Hence the **least squares solution** to  $Q\mathbf{x} = \mathbf{b}$  is  $\hat{\mathbf{x}} = Q^T \mathbf{b}$ . In other words,  $QQ^T \mathbf{b} \approx \mathbf{b}$ . The **projection matrix** is  $P = QQ^T$ . Note that the projection  $\text{Proj}_{\mathcal{C}(Q)}(\mathbf{b}) = QQ^T \mathbf{b}$  doesn't equal to  $\mathbf{b}$  in general.

For general matrix  $A$ , the projection matrix is more complicated:

$$P = A(A^T A)^{-1} A^T.$$

### 6.3.2. Gram-Schmidt Process

"Orthogonal is good". So our goal for this section is: *Given a collection of independent vectors, how to make them orthonormal?*

We start with three independent vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  in  $\mathbb{R}^3$ . In order to construct orthonormal vectors, firstly we construct three **orthogonal** vectors  $\mathbf{A}, \mathbf{B}, \mathbf{C}$ . Secondly we divide  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  by their lengths to get three **orthonormal** vectors  $\mathbf{q}_1 = \frac{\mathbf{A}}{\|\mathbf{A}\|}, \mathbf{q}_2 = \frac{\mathbf{B}}{\|\mathbf{B}\|}, \mathbf{q}_3 = \frac{\mathbf{C}}{\|\mathbf{C}\|}$ .

- Firstly we set  $\mathbf{A} = \mathbf{a}$ .

- The next vector  $\mathbf{B}$  must be perpendicular to  $\mathbf{A}$ . Look at the figure (6.4) below, We find that  $\mathbf{B} = \mathbf{b} - \text{Proj}_A(\mathbf{b})$ . Or equivalently,

$$\text{First Gram-Schmidt step} \quad \mathbf{B} = \mathbf{b} - \frac{\langle \mathbf{A}, \mathbf{b} \rangle}{\langle \mathbf{A}, \mathbf{A} \rangle} \mathbf{A}.$$

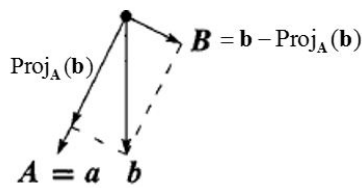


Figure 6.4: Subtract projection to get  $\mathbf{B} = \mathbf{b} - \text{Proj}_A \mathbf{b}$ .

You can take inner product between  $\mathbf{A}$  and  $\mathbf{B}$  to verify that  $\mathbf{A}$  and  $\mathbf{B}$  are orthogonal in Figure (6.4). Note that  $\mathbf{B}$  is not zero (otherwise  $\mathbf{a}$  and  $\mathbf{b}$  would be dependent. We will show it later.)

- Then we want to construct another vector  $\mathbf{C}$ . Most likely  $\mathbf{c}$  is **not** perpendicular to  $\mathbf{A}$  and  $\mathbf{B}$ . What we do is to **subtract  $\mathbf{c}$  off its projections onto the column space of  $\mathbf{A}$  and  $\mathbf{B}$  to get  $\mathbf{C}$ :**

$$\begin{aligned} \mathbf{C} &= \mathbf{c} - \text{Proj}_{\text{span}\{\mathbf{A}, \mathbf{B}\}}(\mathbf{c}) \\ \text{Next Gram-Schmidt step} \quad &= \mathbf{c} - \text{Proj}_A(\mathbf{c}) - \text{Proj}_B(\mathbf{c}) \\ &= \mathbf{c} - \frac{\langle \mathbf{A}, \mathbf{c} \rangle}{\langle \mathbf{A}, \mathbf{A} \rangle} \mathbf{A} - \frac{\langle \mathbf{B}, \mathbf{c} \rangle}{\langle \mathbf{B}, \mathbf{B} \rangle} \mathbf{B}. \end{aligned}$$

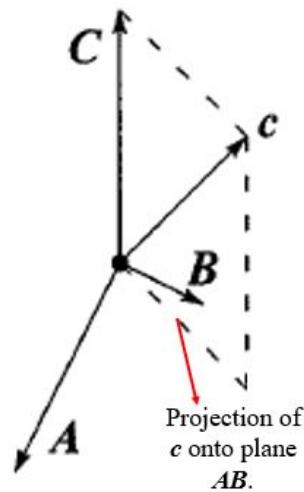


Figure 6.5: Subtract  $\mathbf{c}$  off its projections onto the column space of  $\mathbf{A}$  and  $\mathbf{B}$  to get  $\mathbf{C}$

- Finally we get orthogonal vectors  $\mathbf{A}, \mathbf{B}, \mathbf{C}$ . Orthonormal vectors  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$  are obtained by dividing their lengths (shown in Figure (6.6)):

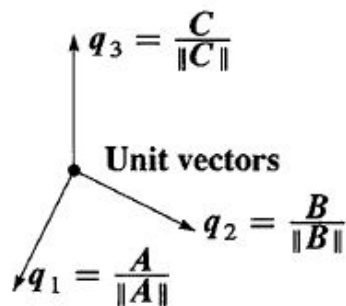


Figure 6.6: Final Gram-Schmidt step

Next we show an example of Gram-Schmidt step:

■ **Example 6.8** How to construct orthonormal vectors from

$$\mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} ?$$

- Firstly we set  $\mathbf{A} = \mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ .

•

$$\begin{aligned} \mathbf{B} &= \mathbf{b} - \text{Proj}_{\mathbf{A}}(\mathbf{b}) = \mathbf{b} - \frac{\langle \mathbf{A}, \mathbf{b} \rangle}{\langle \mathbf{A}, \mathbf{A} \rangle} \mathbf{A} \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} 2^{-1} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} \end{aligned}$$

•

$$\begin{aligned} \mathbf{C} &= \mathbf{c} - \text{Proj}_{\mathbf{A}}(\mathbf{c}) - \text{Proj}_{\mathbf{B}}(\mathbf{c}) = \mathbf{c} - \frac{\langle \mathbf{A}, \mathbf{c} \rangle}{\langle \mathbf{A}, \mathbf{A} \rangle} \mathbf{A} - \frac{\langle \mathbf{B}, \mathbf{c} \rangle}{\langle \mathbf{B}, \mathbf{B} \rangle} \mathbf{B} \\ &= \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} 2^{-1} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}^T \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \left(\frac{1}{2}\right)^{-1} \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

Hence we obtain our orthonormal vectors:

$$\mathbf{q}_1 = \frac{\mathbf{A}}{\|\mathbf{A}\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \mathbf{q}_2 = \frac{\mathbf{B}}{\|\mathbf{B}\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \mathbf{q}_3 = \frac{\mathbf{C}}{\|\mathbf{C}\|} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

And we derive the orthogonal matrix  $Q$ :

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

When will the Gram-Schmidt process “fail”? Let’s describe this process in general case first, then we answer this question.

### 6.3.2.1. Gram-Schmidt process in general case

**Algorithm: Gram-Schmidt Process**

**Input:** a collection of vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n$ , presumably linear independent.

Firstly construct orthogonal vectors  $\mathbf{A}_1, \dots, \mathbf{A}_n$ .

$$\mathbf{A}_1 = \mathbf{a}_1.$$

To construct  $\mathbf{A}_j$ ,  $j \in \{2, \dots, n\}$ , we compute  $\mathbf{a}_j$  minus its projection in the column space spanned by  $\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{j-1}\}$ :

$$\begin{aligned} \mathbf{A}_j &= \mathbf{a}_j - \text{Proj}_{\text{span}\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{j-1}\}}(\mathbf{a}_j) \\ &= \mathbf{a}_j - \text{Proj}_{\mathbf{A}_1}(\mathbf{a}_j) - \text{Proj}_{\mathbf{A}_2}(\mathbf{a}_j) - \dots - \text{Proj}_{\mathbf{A}_{j-1}}(\mathbf{a}_j) \\ &= \mathbf{a}_j - \frac{\langle \mathbf{A}_1, \mathbf{a}_j \rangle}{\langle \mathbf{A}_1, \mathbf{A}_1 \rangle} \mathbf{A}_1 - \frac{\langle \mathbf{A}_2, \mathbf{a}_j \rangle}{\langle \mathbf{A}_2, \mathbf{A}_2 \rangle} \mathbf{A}_2 - \dots - \frac{\langle \mathbf{A}_{j-1}, \mathbf{a}_j \rangle}{\langle \mathbf{A}_{j-1}, \mathbf{A}_{j-1} \rangle} \mathbf{A}_{j-1} \end{aligned}$$

Secondly, after getting  $\mathbf{A}_1, \dots, \mathbf{A}_n$ , we can construct orthonormal vectors:

$$\mathbf{q}_j = \frac{\mathbf{A}_j}{\|\mathbf{A}_j\|} \quad \text{for } j = 1, 2, \dots, n.$$

So when do this process fail? When  $\exists j$  such that  $\mathbf{A}_j = \mathbf{0}$ , we cannot continue this process anymore:

**Proposition 6.3**  $\mathbf{A}_j \neq \mathbf{0}$  for  $\forall j$  if and only if  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  are independent.

*Proofoutline.*  $\mathbf{A}_j = \mathbf{0} \iff \mathbf{a}_j = \text{Proj}_{\text{span}\{\mathbf{A}_1, \dots, \mathbf{A}_{j-1}\}}(\mathbf{a}_j)$ . It suffices to prove  $\exists j$  s.t.  $\mathbf{A}_j = \mathbf{0}$  if and only if  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  are dependent.

*Sufficiency.* Given  $\mathbf{A}_j = \mathbf{0}$ , then  $\mathbf{a}_j = \text{Proj}_{\text{span}\{\mathbf{A}_1, \dots, \mathbf{A}_{j-1}\}}(\mathbf{a}_j) \in \text{span}\{\mathbf{A}_1, \dots, \mathbf{A}_{j-1}\}$ . It's easy to verify that  $\text{span}\{\mathbf{A}_1, \dots, \mathbf{A}_{j-1}\} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}$ . Hence  $\mathbf{a}_j \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}$ . Hence  $\mathbf{a}_1, \dots, \mathbf{a}_j$  are dependent. Thus  $\mathbf{a}_1, \dots, \mathbf{a}_n$  are dependent.

*Necessity.* Given dependent vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n$ , obviously,  $\mathbf{a}_n \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}\}$ . It's easy to verify that  $\mathbf{a}_n = \text{Proj}_{\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{n-1}\}}(\mathbf{a}_n)$ . Thus  $\mathbf{a}_n = \text{Proj}_{\text{span}\{\mathbf{A}_1, \dots, \mathbf{A}_{n-1}\}}(\mathbf{a}_n) \implies \mathbf{A}_n = \mathbf{0}$ . ■

### 6.3.3. The Factorization $\mathbf{A} = \mathbf{QR}$

We know that Gaussian Elimination leads to *LU decomposition*; in fact, Gram-Schmidt process leads to *QR factorization*. These two decomposition methods are quite important in Linear Algebra, let's discuss QR factorization briefly:

Given a matrix  $\mathbf{A} = \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \end{bmatrix}$ , we finally end with a matrix  $\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix}$ . How are these two matrices related?

*Answer:* Since the linear combination of  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  leads to  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$  (vice versa), there must be a third matrix connecting  $\mathbf{A}$  to  $\mathbf{Q}$ . This third matrix is the triangular  $\mathbf{R}$  such that  $\mathbf{A} = \mathbf{QR}$ .

Let's discuss a specific example to show how to do QR factorization.

■ **Example 6.9** Given  $\mathbf{A} = \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \end{bmatrix}$ , whose columns are independent, then we can use Gram-Schmidt process to obtain the corresponding orthonormal vectors  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$  from  $\mathbf{a}, \mathbf{b}, \mathbf{c}$ . As a result, we can write  $\mathbf{A}$  as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ 0 & \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ 0 & 0 & \mathbf{q}_3^T \mathbf{c} \end{bmatrix}$$

We define  $\mathbf{R} \triangleq \begin{bmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ 0 & \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ 0 & 0 & \mathbf{q}_3^T \mathbf{c} \end{bmatrix}$ ,  $\mathbf{Q} \triangleq \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix}$ .  
Hence  $\mathbf{A}$  could be factorized into:

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

where  $\mathbf{R}$  is upper triangular,  $\mathbf{Q}$  is a matrix with orthonormal columns. ■

QR factorization holds for every matrix with independent columns:

**Theorem 6.6** Every  $m \times n$  matrix  $\mathbf{A}$  with ind. columns can be factorized as

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

where  $\mathbf{Q}$  is a matrix with *orthonormal columns*,  $\mathbf{R}$  is an upper triangular matrix (always square).

We omit the proof of this theorem. Now we show that the inverse of  $\mathbf{R}$  always exists:

*Proof.* suppose  $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$ ,  $\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix}$ . Thus we derive

$$\mathbf{R} = \mathbf{Q}^{-1}\mathbf{A} = \mathbf{Q}^T\mathbf{A} = \begin{bmatrix} \mathbf{q}_1^T \mathbf{a}_1 & \mathbf{q}_1^T \mathbf{a}_2 & \dots & \mathbf{q}_1^T \mathbf{a}_n \\ 0 & \mathbf{q}_2^T \mathbf{a}_2 & \dots & \mathbf{q}_2^T \mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{q}_n^T \mathbf{a}_n \end{bmatrix}$$

For every step  $j$  we have

$$\mathbf{A}_j = \mathbf{a}_j - \text{Proj}_{\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}}(\mathbf{a}_j), \quad \mathbf{q}_j = \frac{\mathbf{A}_j}{\|\mathbf{A}_j\|}.$$

Since  $\langle \mathbf{A}_j, \mathbf{a}_j \rangle = \langle \mathbf{a}_j, \mathbf{a}_j \rangle - \langle \text{Proj}_{\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}}(\mathbf{a}_j), \mathbf{a}_j \rangle = \|\mathbf{a}_j\|^2 - \|\text{Proj}_{\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}}(\mathbf{a}_j)\|^2 > 0$ , we have  $\langle \mathbf{q}_j, \mathbf{a}_j \rangle = \frac{\langle \mathbf{A}_j, \mathbf{a}_j \rangle}{\|\mathbf{A}_j\|} > 0$ . Hence the diagonal of  $\mathbf{R}$  are all positive. Hence this triangular matrix is *invertible*. ■

**Proposition 6.4** If  $A = QR$ , then the least squares solution is given by:

$$\mathbf{x} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q}^T \mathbf{b} = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b}.$$

*Explain:* Since we have

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} \mathbf{x} = \mathbf{R}^T \mathbf{R} \mathbf{x}$$

$$\mathbf{A}^T \mathbf{b} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$$

it's equivalent to solve  $\mathbf{R}^T \mathbf{R} \mathbf{x} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$ .

Since  $\mathbf{R}$  is *invertible*, we solve by back substitution to get

$$\mathbf{x} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q}^T \mathbf{b} = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b}.$$

■

### 6.3.4. Function Space

Sometimes we may also discuss orthonormal basis and Gram-Schmidt process on function space. There is a simple example:

■ **Example 6.10** For subspace  $\text{span}\{1, x, x^2\} \subset \mathcal{C}[-1, 1]$ , firstly, how to define orthogonal for the basis  $\{1, x, x^2\}$ ?

*Pre-requisite Knowledge:* Inner product.

$$\langle f, g \rangle = \int_a^b f g \, dx \text{ for } f, g \in \mathcal{C}[a, b]. \quad \|f\|^2 = \int_a^b f^2 \, dx$$

If we have defined inner product, then we can talk about *orthogonality* for  $\{1, x, x^2\}$ . It's easy to verify that

$$\langle 1, x \rangle = 0 \quad \langle x, x^2 \rangle = 0 \quad \langle 1, x^2 \rangle = \frac{2}{3}.$$



If we do the Gram-Schmidt Process similarly, we obtain:

$$\mathbf{A} = 1, \quad \mathbf{B} = x, \quad \mathbf{C} = x^2 - \frac{\langle 1, x^2 \rangle}{\langle 1, 1 \rangle} 1 - \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} x = x^2 - \frac{1}{3}$$

where  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  are *orthogonal*. We can divide their length to obtain orthonormal basis:

$$\begin{aligned} \mathbf{q}_1 &= \frac{\mathbf{A}}{\|\mathbf{A}\|} = \frac{1}{\sqrt{\int_{-1}^1 1^2 dx}} = \frac{1}{2} \\ \mathbf{q}_2 &= \frac{\mathbf{B}}{\|\mathbf{B}\|} = \frac{x}{\sqrt{\int_{-1}^1 x^2 dx}} = \frac{x}{2/3} = \frac{3}{2}x \\ \mathbf{q}_3 &= \frac{\mathbf{C}}{\|\mathbf{C}\|} = \frac{x^2 - \frac{1}{3}}{\sqrt{\int_{-1}^1 (x^2 - \frac{1}{3})^2 dx}} = \frac{x^2 - \frac{1}{3}}{\frac{8}{45}} = \frac{45x^2 - 15}{8} \end{aligned}$$

Hence  $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$  is the orthonormal basis for  $\{1, x, x^2\}$ . ■

■ **Example 6.11** Consider the collection  $\mathcal{F}$  of functions defined on  $[0, 2\pi]$ , where

$$\mathcal{F} := \{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos mx, \sin mx, \dots\}$$

Using various trigonometric identities, we can show that if  $f$  and  $g$  are **distinct**(different) functions in  $\mathcal{F}$ , we have  $\int_0^{2\pi} fg \, dx = 0$ . For example,

$$\langle \sin x, \sin 2x \rangle = \int_0^{2\pi} \sin x \sin 2x \, dx = \int_0^{2\pi} \frac{1}{2} (\cos x - \cos 3x) \, dx = 0.$$

And moreover, if  $f = g$ , we have  $\int_0^{2\pi} f^2 \, dx = \pi$ . For example,

$$\langle \sin 5x, \sin 5x \rangle = \int_0^{2\pi} \sin^2 5x \, dx = \int_0^{2\pi} \frac{1}{2} (1 + \cos 10x) \, dx = \pi.$$

In conclusion, the collection of functions  $\{1, \sin mx, \cos mx\}$  for  $k = 1, 2, \dots$  are *orthogonal* in  $\mathcal{C}[0, 2\pi]$ . Note that this set is **not orthonormal**. ■

This example gives a motivation of the fourier transformation:

### 6.3.5. Fourier Series

Since we have shown the orthogonality of  $\mathcal{F}$  in Example.(6.11), our question is that what kind of function can be written as the linear combination of functions from  $\mathcal{F}$ .

The Fourier series of a function is its expansion into sines and cosines:

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots$$

where  $f(x) \in \mathcal{C}[0, 2\pi]$ . So our question turns into what kind of function could be expressed as fuourier series?

**Theorem 6.7** If a function  $f$  have the finite length in its function space  $\mathcal{C}[a, b]$ , then it could be expressed as *fourier series*.

But how to compute the coefficients  $a_i$ 's and  $b_j$ 's? The key is orthogonality! For example, in order to get  $a_1$ , we just do the inner product between  $f(x)$  and  $\cos x$ :

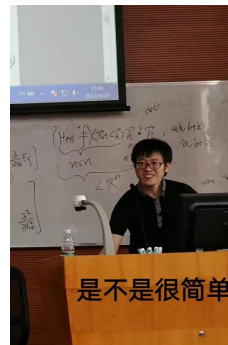


Figure 6.7: Enjoy fourier series!

$$\langle f(x), \cos x \rangle = a_1 \langle \cos x, \cos x \rangle + 0 \implies a_1 = \frac{\langle f(x), \cos x \rangle}{\langle \cos x, \cos x \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos x \, dx$$

Similarly we derive

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos mx \, dx \quad b_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin mx \, dx.$$

