McGILL UNIVERSITY
FACULTY OF SCIENCE

DEPARTMENT OF
MATHEMATICS AND STATISTICS

MATH 223 2006 01
LINEAR ALGEBRA

Information for Students
(Winter Term, 2005/2006)

Pages 1 - 9 of these notes may be considered the Course Outline for this course.

W. G. Brown

May 4, 2006

Contents

List of Tables

1 General Information

Distribution Date: This *preliminary* version as of May 4, 2006
(all information is subject to change)
Pages 1 - 9 of these notes may be considered the Course Outline for this course.

These notes may undergo minor corrections or updates during the term: the *definitive* version will be the version accessible at

http://www.math.mcgill.ca/brown/math223b.html

or on WebCT, at

http://www.mcgill.ca/webct or **http://webct.mcgill.ca**

1.1 Instructor and Times

INSTRUCTOR:	Prof. W. G. Brown
CRN:	691
OFFICE:	BURN 1224
OFFICE HOURS:	M15:00→16:00; F09:15→10:10
	(or by appointment) subject to change
TELEPHONE:	398–3836
E-MAIL:	BROWN@MATH.MCGILL.CA
CLASSROOM:	ARTS W 120
CLASS HOURS:	MWF 10:35–11:25

Table 1: Instructor and Times

1.2 Calendar Description

MATH 223 LINEAR ALGEBRA. (3 credits. 3 hours lecture. Prerequisite: **MATH 133 or equivalent.**[1] Restriction: Not open to students in Mathematics programs, nor to students who have taken or are taking MATH 236, MATH 247, or MATH 251. It is open to students in faculty Programs.) Review of matrix algebra, determinants

[1] **This prerequisite is rigid: If you do not have a background at least at the level of MATH 133, you are in violation of McGill regulations if you remain in this course.**

and systems of linear equations. Vector spaces, linear operators, and their matrix representations, orthogonality. Eigenvalues and eigenvectors, diagonalization of Hermitian matrices. Applications.

1.3 Evaluation of Your Progress

1.3.1 Your final grade

(See Table 2, p. 6) Your grade in this course will be a *letter grade*, based on a percentage grade computed from the following components:

1. Approximately 12 **WeBWorK** homework assignments (cf. §1.3.3) — counting together for 10%. All **WeBWorK** assignments must be completed by their posted expiration dates and times. The Assignments will be numbered WW_1, $\ldots WW_{10}$.[2]

2. Two or Three Written homework assignments WR_i $(i = 1, 2, \ldots)$ — counting together for 5%. Written assignments must be completed by their posted expiration dates and times.

3. Two Class Tests, tentatively scheduled for Wednesdays, February 01st, and March 01st[3], 2006, at the regular class time. A single Test Grade will be based on the tests: the first test will count for 5 marks, and the second for 10 of the 15 marks. The Test Grade will count for either 15% or 0% of your final grade, whichever is to your advantage.[4]

4. The final examination — counting for either 70% or 85% of the final grade.

Where a student's performance on the final examination is superior to her performance on the Class Tests, the final examination grade will replace *the test grades* in the calculations. *It is not planned to permit the examination grade to replace the grades on* **WeBWorK** *assignments.*

1.3.2 WeBWorK

We use the **WeBWorK** system, developed at the University of Rochester — which is designed to expose you to a large number of drill problems, and where plagiarism is discouraged. **WeBWorK** *is accessible only over the Web.* **WeBWorK** assignments carry a due date and time; only answers submitted by this date and time will count.

[2]Assignment WW_0 is intended to introduce you to the **WeBWorK** system; it does not count.

[3]Date finalized as March 08th, 2006

[4]This option of replacement by the final examination is available only for the entire Test Grade, not separately for the portion which derives from one or other of the contributing tests.

Some of the assignments have limits to the numbers of times a student may attempt a problem.

1.3.3 WeBWorK Assignments

There will be 6 pairs of **WeBWorK** assignments, accessible at the URL

`http://msr05.math.mcgill.ca/webwork/m223w06`

Your user name will be your 9-digit student number, and your first password will be this same 9-digit student number. The 12 assignments will be paired — each even numbered assignment (WW_2, WW_4, WW_6, WW_8, WW_{10}, WW_{12}) will contain the same types of problems as on the preceding odd numbered assignment (WW_1, WW_3, WW_5, WW_7, WW_9, WW_{11}). THE ODD-NUMBERED ASSIGNMENTS DO NOT COUNT! The intention is that you should use each odd-numbered assignment to thoroughly learn how to solve the problems, and then attempt the following even-numbered assignment. The data on the even numbered assignment will probably be different — different numbers and/or functions, but the same concepts. It is expected that the due date for assignments will be on specified Fridays, at midnight. If you leave your **WeBWorK** assignment until the hours close to the due time on the due date, you should not be surprised if the system is slow to respond. This is not a malfunction, but is simply a reflection of the fact that other students have also been procrastinating! To benefit from the speed that the system can deliver under normal conditions, do not delay your **WeBWorK** until the last possible day! If a systems failure interferes with the due date of an assignment, arrangements may be made to change that date, and an e-mail message may be broadcast to all users (to the e-mail addresses on record), or a note posted in the course announcements on WebCT; but slowness in the system just before the due time will not normally be considered a systems failure.[5]

1.3.4 Final Examination

A 3-hour-long final examination will be scheduled during the regular examination period for the winter term (April 11th, 2006 through April 28th, 2006). You are advised not to make any travel arrangements that would prevent you from being present on campus at any time during this period.

[5]Should you find that the system is responding slowly, *do not* submit your solutions more than once; you may deplete the number of attempts that have been allowed to you for a problem: this will not be considered a systems failure.

1.3.5 Supplemental Assessments

Supplemental Examination. There will be a supplemental examination in this course. (For information about Supplemental Examinations, see

http://www.mcgill.ca/artscisao/departmental/examination/supplemental/.

Note, in particular, that a Supplemental Examination may be written only by a student who has obtained a grade of F or D as a grade in the course, and that the grade on the Supplemental Examination *counts in your average as though you have taken the course again — without a term work component!*)

There is No Additional Work Option. "Will students with marks of D, F, or J have the option of doing additional work to upgrade their mark?" No. ("Additional Work" refers to an option available in certain Arts and Science courses, but not available in MATH 223 2006 01.)

1.3.6 Machine Scoring

"Will the final examination be machine scored?" The final examination will not be machine scored.

1.3.7 Plagiarism

While students are not discouraged from discussing methods for solving **WeBWorK** assignment problems with their colleagues, all the work that you submit — whether through **WeBWorK** or written assignments, or the final examination must be your own. The Senate of the University requires the following message in all course outlines:

> "McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures. (See http://www.mcgill.ca/integrity for more information).
>
> "L'université McGill attache une haute importance à l'honnêteté académique. Il incombe par conséquent à tous les étudiants de comprendre ce que l'on entend par tricherie, plagiat et autres infractions académiques, ainsi que les conséquences que peuvent avoir de telles actions, selon le Code de conduite de l'étudiant et des procédures disciplinaires. (Pour de plus amples renseignements, veuillez consulter le site http://www.mcgill.ca/integrity)."

It is a violation of University regulations to permit others to solve your WeBWorK problems, or to extend such assistance to

others; you could be asked to sign a statement attesting to the originality of your work. The Handbook on Student Rights and Responsibilities[6] states in ¶A.I.15(a) that

> "No student shall, with intent to deceive, represent the work of another person as his or her own in any academic writing, essay, thesis, research report, project or assignment submitted in a course or program of study or represent as his or her own an entire essay or work of another, whether the material so represented constitutes a part or the entirety of the work submitted."

You are also referred to the following URL:

http://www.mcgill.ca/integrity/studentguide/

1.4 Published Materials

1.4.1 Required Text-Book

I am required to continue in this semester the use of the textbook used in MATH 223 2005 09; that book is [1] Schaum's Outline of Theory and Problems of LINEAR ALGEBRA, Third Edition, by S. Lipschutz and M. Lipson; McGraw-Hill (2001). ISBN 0-07-136200-2. Earlier editions of this book are not recommended, as the order of topics has changed.

1.4.2 Website

These notes, and other materials for students in this course, will be accessible at through WebCT, or directly at the following URL:

http://www.math.mcgill.ca/brown/math223b.html

The notes will be in "pdf" (.pdf) form, and can be read using the Adobe Acrobat reader, which many users have on their computers. This free software may be downloaded from the following URL:

http://www.adobe.com/prodindex/acrobat/readstep.html [7]

[8] It is expected that most computers in campus labs should have the necessary software to read the posted materials.

---

[6] http://upload.mcgill.ca/secretariat/greenbookenglish.pdf

[7] At the time of this writing the current version appears to be 7.0.

[8] There is no reason to expect the distribution of problems on the Class Tests or in assignments and examinations from previous years be related to the frequencies of any types of problems on the examination that you will be writing at the end of the term.

Item	#	Due Date	Details
	WW_0		DOES NOT COUNT: introduces **WeBWorK**
WeBWorK	WW_1	13 Jan 06	Does not count in your grade.
Assignments (cf. §1.3.3)	WW_2	20 Jan 06	Same scope as WW_1, but with limited numbers of attempts
	WW_3	27 Jan 06	Does not count in your grade
	WW_4	27 Jan 06	Same scope as WW_3, but with limited attempts
$\boxed{10\%}$	$\vdots$		
	WW_{11}	31 Mar 06	Does not count in your grade
	WW_{12}	07 Apr 06	Same scope as WW_{11}, but with limited attempts
Class Tests	T_1	01 Feb 06	Counts for one-third of TEST GRADE (tentative date)
$\boxed{15\% \text{ or } 0\%}$	T_2	08 Mar 06	Counts for two-thirds of TEST GRADE (revised date)
Final Exam $\boxed{70\% \text{ or } 85\%}$		11–28 Apr 06	Date of exam to be announced by Faculty
Supplemental Exam		23–24 Aug 06	Date of exam to be announced by Faculty; only for students who do not obtain standing at the final. Supplemental exams count in your average like taking the course *again*; the exam counts for 100%.

Table 2: Summary of Course Requirements, as of May 4, 2006 (all dates are subject to change)

Where revisions are made to distributed printed materials — for example these information sheets — I expect that the last version will usually be posted on the Web.

The notes and **WeBWorK** will also be available via a link from the WebCT URL:

http://webct.mcgill.ca

but some other features of WebCT have not yet been implemented.

1.5 Syllabus

In the following list section numbers refer to the text-book [1]. The syllabus may include most if not all of of Chapters 4–13, with omissions as listed below.[9][10]

1. Chapter 1: **Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$, Spatial Vectors.** This chapter is considered review of material in MATH 133, although your exposure may have been largely for small values of n.

2. Chapter 2: **Algebra of Matrices.** This chapter also is considered review.

3. Chapter 3: **Systems of Linear Equations.** I consider §§3.1–3.7 to be review; here again, your familiarity could be from cases of small numbers of equations in small numbers of variables, but you should understand the geometric interpretations of those cases. You should also know about parametric equations for lines in 2 and 3 dimensions, and should be able to solve a small system of linear equations by some sort of row reduction of matrices.

4. Chapter 4: **Vector Spaces.** This material is *not* considered to be review. At this point we will shift gears from high speed to normal discussion of topics considered to be new. In the course of discussion of this chapter, I will also consider the sections of Chapter 3 which were not considered review in item 3 above.

5. Chapter 5: **Linear Mappings.** DESCRIPTION TO BE COMPLETED

6. Chapter 6: **Linear Mappings and Matrices.** DESCRIPTION TO BE COMPLETED

7. Chapter 7: **Inner Product Spaces; Orthogonality.** DESCRIPTION TO BE COMPLETED

8. Chapter 8: **Determinants.** Discussion of this topic will be minimized.

---

[9] If a textbook section is listed below, you should assume that all material in that section is examination material *even if the not every topic has been discussed in the lectures*; however, I may give you information during the term concerning topics that may be considered subsidiary.

Do not assume that a topic is omitted from the syllabus if it has not been tested in a WeBWorK assignment or the Class Tests, or if it has not appeared on any of the old examinations in the course! Some topics to not lend themselves to these types of testing; others may have been omitted simply because of lack of space, or oversight. By the same token, you need not expect every topic in the course to be examined on the final examination.

[10] The details of the specific sections included and excluded may not be available during the first week of course; as these facts become known, the details will be posted in this location in the updated Information document on WebCT.

9. Chapter 9: **Diagonalization: Eigenvalues and Eigenvectors.** DESCRIPTION TO BE COMPLETED
10. Chapter 10: **Canonical Forms.** Omit this Chapter.
11. Chapter 11: **Linear Functionals and the Dual Space.** Omit this Chapter.
12. Chapter 12: **Bilinear, Quadratic, and Hermitian Forms.** Omit this Chapter.
13. Chapter 13: **Linear Operators on Inner Product Spaces.** Omit this Chapter.
14. Chapter 14: **Multilinear Products.** Omit this Chapter.

1.6 Preparation and Workload

1.6.1 Calculators

The use of calculators is not permitted in either the Class Tests or the examination in this course. Students whose previous mathematics courses have been calculator-oriented would be advised to make particular efforts to avoid the use of a calculator in solving problems in this course, in order to develop a minimal facility in manual calculation. This means that *you are urged to do all arithmetic by hand.*

1.6.2 Self-Supervision

The real uses of WeBWorK and the Class Tests. Students often misunderstand the true significance of **WeBWorK** assignments and the Class Tests. While both contribute to your grade, they help you estimate the quality of your progress in the course. Take proper remedial action if you obtain low grades on the test, or if you require many attempts before being able to solve a problem on **WeBWorK**. However, while both **WeBWorK** and the Class Tests have a role to play in learning the subject matter of the course, *neither is as important as reading your textbook, working problems yourself, and attending and listening at lectures and tutorials.*

Does a high grade on WeBWorK indicate the likelihood of a high grade on the final examination? NO! The primary purpose of the **WeBWorK** assignments is as an aid to learning; but, as your work is not being done under examination conditions, and as **WeBWorK** tests only the correctness of answers, not knowledge of methods, you should not use the **WeBWorK** grades as indicators of your likely examination grade. The grades on the Class Tests are somewhat more useful for that purpose.

1.6.3 Escape Routes

At any time, even after the last date for dropping the course, students who are experiencing medical or personal difficulties should not hesitate to consult their advisors or the Student Affairs office of their faculty. Don't allow yourself to be overwhelmed by such problems; the University has resource persons who may be able to help you.

1.6.4 Keep your e-mail address up to date

Both WebCT and **WeBWorK** contain an e-mail address where we may assume you can be reached. If you prefer to use another e-mail address, the most convenient way is to *forward* your mail from your student mailbox, leaving the recorded addresses in these two systems unchanged. You can enter or change a forwarding e-mail address by going to http://webmail.mcgill.ca, and logging in to your student mailbox at po-box.mcgill.ca.[11]

1.6.5 Use of Calculators and Computer Algebra Systems

Insofar as course *content* is concerned, we will emphasize the lowest possible use of technology: we will avoid the use of calculators and computers. You are urged to do all calculations manually, and to avoid the use of computer algebra systems. You should not use a calculator or computer in the solution of **WeBWorK** problems, as it prevents you from developing skill for detecting errors in manual calculations — a skill that you will need for the Class Tests and final examination. At this time there is no plan to permit the use of calculators at test or examination.

[11] It is possible that the University may prevent students from forwarding mail addressed to their "official" address.

> "E-mail is one of the official means of communication between McGill University and its students. As with all official University communications, it is the student's responsibility to ensure that time-critical e-mail is accessed, read, and acted upon in a timely fashion. If a student chooses to forward University e-mail to another e-mail mailbox, it is that student's responsibility to ensure that the alternate account is viable."(cf., http://www.mcgill.ca/email-policy

2 First Class Tests

2.1 Version 1

FIRST CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown

DATE: Wednesday, 01st February, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

INSTRUCTIONS

1. All your writing — even rough work — must be handed in.
2. Calculators are not permitted.
3. *Your neighbour's version of this test may not be the same as yours.*
4. This examination booklet consists of this cover, Pages 11 through 16 containing questions; and Pages 17, 18, which are blank.
5. Show all your work. All solutions are to be written in the space provided on the page where the question is printed. When that spaces is exhausted, you may write *on the facing page*, on one of the blank pages, or on the back cover of the booklet, but you must indicate any continuation clearly on the page where the question is printed! (Please inform the invigilator if you find that your booklet is defective.)

PLEASE DO NOT WRITE INSIDE THIS BOX

1 /2	2 /4	3 /4	4 /2	5 /4	6 /8	7 /6	8 /8
9 /7							
				RAW /45		/	

1. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ -3 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ __________.

2. [4 MARKS] If A, B, and C are 2×2, 2×5, and 5×8 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: ________

   BA: ________

   BC: ________

   AB: ________

3. [4 MARKS] If $A = \begin{pmatrix} 8 & -2 \\ 5 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 6 \\ 0 & 3 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ \end{pmatrix}$

4. [2 MARKS] Find a 3×3 matrix A such that

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ 4 \end{pmatrix}, \; A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ -4 \end{pmatrix}, \text{ and } A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}.$$

$$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(2-3i)(-4-4i)(1-5i) = \_\_\_\_ + \_\_\_\_ \, i,$

   (b) $((5-5i)^2+5)i = \_\_\_\_ + \_\_\_\_ \, i.$

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} -3 & 1 & 1 & -16 \\ 2 & -3 & -1 & 18 \\ -2 & -2 & -3 & 1 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & -5 & -2 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 16 & -20 & -8 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix}$

(b) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 16 & -20 & -8 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 4 & -5 & -2 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix}$

(c) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ -5 & -3 & -1 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 2 & -5 & 1 \\ 2 & 1 & 4 \\ -5 & -3 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ 2 & 1 & 4 \\ -5 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -5 & 1 \\ -5 & -3 & -1 \\ 2 & 1 & 4 \end{bmatrix}$

(e) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ -2 & 4 & -1 \end{bmatrix} = \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ 2 & -4 & -13 \end{bmatrix}$

(f) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ 2 & -4 & -13 \end{bmatrix} = \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ -2 & 4 & -1 \end{bmatrix}$

8. [8 MARKS] Let

$$A = \begin{bmatrix} 0 & -3 \\ -8 & -6 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:

$$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

(b) Write A^{-1} as a product of 4 elementary matrices:

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

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(x, y, z) \mid x, y, z > 0\}$

   (b) $\{(x, y, z) \mid 6x + 2y = 0, -5x - 7z = 0\}$

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

CONTINUATION PAGE FOR PROBLEM NUMBER

You *must* refer to this continuation page on the page where the problem is printed!

2.2 Version 2

FIRST CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 01st February, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. All your writing — even rough work — must be handed in.
2. Calculators are not permitted.
3. *Your neighbour's version of this test may not be the same as yours.*
4. This examination booklet consists of this cover, Pages 20 through 25 containing questions; and Pages 26, 27, which are blank.
5. Show all your work. All solutions are to be written in the space provided on the page where the question is printed. When that spaces is exhausted, you may write *on the facing page*, on one of the blank pages, or on the back cover of the booklet, but you must indicate any continuation clearly on the page where the question is printed! (Please inform the invigilator if you find that your booklet is defective.)

PLEASE DO NOT WRITE INSIDE THIS BOX

1 /2	2 /4	3 /4	4 /2	5 /4	6 /8	7 /6	8 /8
9 /7							
				Raw /45		/	

1. [2 MARKS] Find a 3×3 matrix A such that

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix},\ A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix},\ \text{and}\ A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 4 \end{pmatrix}.$$

$$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

2. [4 MARKS] If A, B, and C are 5×5, 5×8, and 8×6 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   A^2: ________

   AC: ________

   CB: ________

   BA: ________

3. [4 MARKS] If $A = \begin{pmatrix} 4 & 1 \\ 6 & 8 \end{pmatrix}$ and $B = \begin{pmatrix} -3 & 2 \\ 3 & 3 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ \end{pmatrix}$

4. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} -3 \\ 4 \\ 4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ __________.

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(5 - i)(3 - 3i)(1 - 3i) \ =$ _____ $+$ _____ i,

   (b) $((-4 + 5i)^2 - 3)i \ =$ _____ $+$ _____ i.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} -1 & -3 & 1 & 14 \\ -2 & -3 & 3 & 22 \\ 2 & -1 & 1 & 2 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 2 & -4 & 2 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 4 & -8 & 4 \end{bmatrix}$

(b) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 4 & -8 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 2 & -4 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -5 & 2 \\ -3 & 1 & 4 \\ 5 & -4 & 3 \end{bmatrix} = \begin{bmatrix} -3 & -5 & 2 \\ 5 & -4 & 3 \\ -3 & 1 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -5 & 2 \\ 5 & -4 & 3 \\ -3 & 1 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -5 & 2 \\ -3 & 1 & 4 \\ 5 & -4 & 3 \end{bmatrix}$

(e) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ -5 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ 11 & 17 & 7 \end{bmatrix}$

(f) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ 11 & 17 & 7 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ -5 & -3 & -1 \end{bmatrix}$

8. [8 MARKS] Let

$$A = \begin{bmatrix} 0 & 5 \\ 9 & -4 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:

$$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

(b) Write A^{-1} as a product of 4 elementary matrices:

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

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(6x + 3y, -8x + 4y, -2x - 4y) \mid x, y$ arbitrary numbers $\}$

   (b) $\{(x, x - 9, x + 5) \mid x$ arbitrary number $\}$

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

2.3 Version 3

FIRST CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 01st February, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. All your writing — even rough work — must be handed in.
2. Calculators are not permitted.
3. *Your neighbour's version of this test may not be the same as yours.*
4. This examination booklet consists of this cover, Pages 29 through 34 containing questions; and Pages 35, 36, which are blank.
5. Show all your work. All solutions are to be written in the space provided on the page where the question is printed. When that spaces is exhausted, you may write *on the facing page*, on one of the blank pages, or on the back cover of the booklet, but you must indicate any continuation clearly on the page where the question is printed! (Please inform the invigilator if you find that your booklet is defective.)

PLEASE DO NOT WRITE INSIDE THIS BOX

1	2	3	4	5	6	7	8
/4	/2	/2	/4	/4	/8	/6	/8
9 /7							
				RAW /45		/	

1. [4 MARKS] If A, B, and C are 4×4, 4×3, and 3×6 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: ______

   AC: ______

   AB: ______

   BA: ______

2. [2 MARKS] Find a 3×3 matrix A such that

   $$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \\ -5 \end{pmatrix}, \ A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \text{ and } A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}.$$

   $$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

3. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} -2 \\ 1 \\ -4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ ________.

4. [4 MARKS] If $A = \begin{pmatrix} -10 & 7 \\ -1 & -3 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 3 \\ -4 & 6 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_\_ & \_\_\_\_ \\ \_\_\_\_ & \_\_\_\_ \end{pmatrix}$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(3 + 2i)(-4 - i)(5 - 2i) \ = \_\_\_\_ + \_\_\_\_ \ i$,

   (b) $((4 + 2i)^2 + 1)i \ = \_\_\_\_ + \_\_\_\_ \ i$.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} -1 & -3 & 1 & 3 \\ -2 & 1 & 3 & -8 \\ -2 & -1 & 1 & -4 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 4 & 3 & -3 \end{bmatrix} = \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 12 & 9 & -9 \end{bmatrix}$$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 12 & 9 & -9 \end{bmatrix} = \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 4 & 3 & -3 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & 1 & -3 \\ -4 & -1 & 1 \\ -5 & -2 & -5 \end{bmatrix} = \begin{bmatrix} -4 & -1 & 1 \\ -2 & 1 & -3 \\ -5 & -2 & -5 \end{bmatrix}$$

(d) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -1 & 1 \\ -2 & 1 & -3 \\ -5 & -2 & -5 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -3 \\ -4 & -1 & 1 \\ -5 & -2 & -5 \end{bmatrix}$$

(e) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ 3 & -2 & 5 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ -9 & 1 & -1 \end{bmatrix}$$

(f) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ -9 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ 3 & -2 & 5 \end{bmatrix}$$

8. [8 MARKS] Let

$$A = \begin{bmatrix} 0 & -9 \\ 2 & 5 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:

$$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

(b) Write A^{-1} as a product of 4 elementary matrices:

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

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(x, y, z) \mid x + y + z = 8\}$

   (b) $\{(-8x, 2x, -9x) \mid x$ arbitrary number $\}$

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

2.4 Version 4

FIRST CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown

DATE: Wednesday, 01st February, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. All your writing — even rough work — must be handed in.
2. Calculators are not permitted.
3. *Your neighbour's version of this test may not be the same as yours.*
4. This examination booklet consists of this cover, Pages 38 through 43 containing questions; and Pages 44, 45, which are blank.
5. Show all your work. All solutions are to be written in the space provided on the page where the question is printed. When that spaces is exhausted, you may write *on the facing page*, on one of the blank pages, or on the back cover of the booklet, but you must indicate any continuation clearly on the page where the question is printed! (Please inform the invigilator if you find that your booklet is defective.)

PLEASE DO NOT WRITE INSIDE THIS BOX

1	2	3	4	5	6	7	8
/4	/2	/4	/2	/4	/8	/6	/8
9 /7							
				Raw /45		/	

1. [4 MARKS] If $A = \begin{pmatrix} -3 & 7 \\ 4 & 8 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & -3 \\ -3 & -2 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_\_ & \_\_\_\_ \\ \_\_\_\_ & \_\_\_\_ \end{pmatrix}$

2. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} -5 \\ -3 \\ 0 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ _________.

3. [4 MARKS] If A, B, and C are 3×3, 3×7, and 7×9 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: _______

   BC: _______

   BA: _______

   AC: _______

4. [2 MARKS] Find a 3×3 matrix A such that
$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 4 \end{pmatrix},\ A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \\ -1 \end{pmatrix},\ \text{and } A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 5 \\ -1 \end{pmatrix}.$$
$$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

(a) $(3 - 2i)(-1 + 2i)(-4 + i) = \_\_\_\_ + \_\_\_\_\ i$,

(b) $((5 - 5i)^2 - 5)i = \_\_\_\_ + \_\_\_\_\ i$.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} 3 & 3 & -3 & 18 \\ 2 & -3 & 3 & -28 \\ 2 & 0 & -1 & 0 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & 3 & 1 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix} = \begin{bmatrix} -8 & 12 & 4 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix}$$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -8 & 12 & 4 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 1 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -1 & 5 & -1 \\ 1 & -2 & -2 \\ 3 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -1 \\ 1 & -2 & -2 \\ -1 & 5 & -1 \end{bmatrix}$$

(d) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 3 & -2 & -1 \\ 1 & -2 & -2 \\ -1 & 5 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 5 & -1 \\ 1 & -2 & -2 \\ 3 & -2 & -1 \end{bmatrix}$$

(e) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -1 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -9 & 3 & 3 \end{bmatrix}$$

(f) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -9 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

8. [8 MARKS] Let

$$A = \begin{bmatrix} 0 & -7 \\ -5 & -8 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:

$$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

(b) Write A^{-1} as a product of 4 elementary matrices:

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

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(2x - 8y, -4x - 9y, 5x + 9y) \mid x, y$ arbitrary numbers $\}$

   (b) $\{(-6, y, z) \mid y, z$ arbitrary numbers $\}$

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

CONTINUATION PAGE FOR PROBLEM NUMBER $\boxed{}$

You *must* refer to this continuation page on the page where the problem is printed!

3 Solutions to Problems on the First Class Tests

The 45-minute test was administered on February 1st, 2006.
The solutions were mounted on the Web on Sunday, 5 February, 2006

This test was intended to evaluate your preparation in areas covered mainly by the first 2 pairs of **WeBWorK** assignments. The grades are — like the grades on **WeBWorK** assignments, grossly inflated. So you should read them as indicating likely performance on a real test or examination in this course, but to detect areas where you are not adequately prepared. An exception is the last question, which was on recent course material, and was included so that students would not miss lectures to prepare for an insignificant test.

(In order to "equalize" the difficulty level of the 4 versions, the grades on versions ##1,3 were scaled upwards by approximately 15%.[12])

3.1 Version 1

1. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ -3 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ ________.

   Solution: $2(-5) + (-3)2 + (-3)(-3) = -7$.

2. [4 MARKS] If A, B, and C are 2×2, 2×5, and 5×8 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: ______

   BA: ______

   BC: ______

   AB: ______

   Solution: CB is undefined since C has 8 columns, but B has a different number – 2 – of rows. BA is undefined, since B has 5 columns, but A has a different number — 2 — of rows. BC is 2×8. AB is 2×5.

[12] On WebCT the grades will be shown both as the raw grade/45 and the recorded grade/5.0. Any corrections will not be posted immediately, but will await the next uploading of data, which might not be until after **WeBWorK** assignment WW6 has been graded.

3. [4 MARKS] If $A = \begin{pmatrix} 8 & \text{-2} \\ 5 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 6 \\ 0 & 3 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_\_ & \_\_\_\_ \\ \_\_\_\_ & \_\_\_\_ \end{pmatrix}$

   Solution: $AB = \begin{pmatrix} 8 & 42 \\ 5 & 51 \end{pmatrix}$

4. [2 MARKS] Find a 3×3 matrix A such that

   $A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \text{-1} \\ \text{-4} \\ 4 \end{pmatrix}$, $A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ \text{-2} \\ \text{-4} \end{pmatrix}$, and $A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}$.

   $A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$

   Solution:
$$\begin{pmatrix} -1 & 3 & 1 \\ -4 & -2 & 4 \\ 4 & -4 & 1 \end{pmatrix}$$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) (a) $(2-3i)(-4-4i)(1-5i) = \_\_\_\_ + \_\_\_\_\, i$,
   (b) $((5-5i)^2+5)i = \_\_\_\_ + \_\_\_\_\, i$.

   Solution:

   (a) $(2-3i)(-4-4i)(1-5i) = (-20+4i)(1-5i) = 0 + 104i$ or simply $104i$.
   (b) $((5-5i)^2+5)i = (-50i+5)i = 50+5i$.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} \text{-3} & 1 & 1 & \text{-16} \\ 2 & \text{-3} & \text{-1} & 18 \\ \text{-2} & \text{-2} & \text{-3} & 1 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

   Solution: This was a version of Problem 3 on WeBWorK assignments 3 and 4. This quiz was intended to expose difficulties in solving simple matrix or MATH 133 problems, and this question certainly did that. To row reduce this matrix you needed to plan your operations so as to minimize the computational difficulties. If you have been solving such problems before the quiz — as on the WeBWorK assignment — you know that you can plan your row reduction steps so as to keep the numbers relatively simple; and it doesn't require the use of a calculator. (You should not be using a calculator on WeBWorK.) If you permit complicated fractions to enter into your calculations, then you are inviting problems when you are doing hand calculations. Of course, this problem does not show whether you understand the principles of row reduction; it show whether you know how to manage small problems manually without assistance.

The only practical way to grade this problem was to base the grade on what you finally achieved, and to check the work only where your final matrix was close enough to a row reduced matrix that that was feasible. In the solution I have not indicated the row operations over the arrows; it would be better to always show them, but I have had to economize on my time as I had to prepare 4 versions. I have not tried to find the shortest solution: in some cases it would have been faster to scale a row first and then add a multiple of another row.

$$\begin{pmatrix} -3 & 1 & 1 & -16 \\ 2 & -3 & -1 & 18 \\ -2 & -2 & -3 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & -2 & 0 & 2 \\ 2 & -3 & -1 & 18 \\ -2 & -2 & -3 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & -2 & 0 & 2 \\ 0 & -7 & -1 & 22 \\ 0 & 2 & -3 & -3 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & -7 & -1 & 22 \\ 0 & 2 & -3 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 1 & -13 & 10 \\ 0 & 2 & -3 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 26 & -22 \\ 0 & 1 & -13 & 10 \\ 0 & 0 & 23 & -23 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 26 & -22 \\ 0 & 1 & -13 & 10 \\ 0 & 0 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 26 & -22 \\ 0 & 1 & -13 & 10 \\ 0 & 0 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & -5 & -2 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 16 & -20 & -8 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix}$$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 16 & -20 & -8 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 4 & -5 & -2 \\ -3 & 4 & 4 \\ -3 & -4 & -3 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ -5 & -3 & -1 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 2 & -5 & 1 \\ 2 & 1 & 4 \\ -5 & -3 & -1 \end{bmatrix}$$

(d) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ 2 & 1 & 4 \\ -5 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -5 & 1 \\ -5 & -3 & -1 \\ 2 & 1 & 4 \end{bmatrix}$

(e) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ -2 & 4 & -1 \end{bmatrix} = \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ 2 & -4 & -13 \end{bmatrix}$

(f) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ 2 & -4 & -13 \end{bmatrix} = \begin{bmatrix} -2 & -5 & -3 \\ 1 & -2 & -3 \\ -2 & 4 & -1 \end{bmatrix}$

Solution:

(a) $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix}$

8. [8 MARKS] Let

$$A = \begin{bmatrix} 0 & -3 \\ -8 & -6 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:

$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$

(b) Write A^{-1} as a product of 4 elementary matrices:

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

Solution: This was a problem where you could verify the correctness of your solution by multiplying out your matrices. For that reason I have graded it mainly as all or nothing. Many students included in their produce a matrix which was not elementary; in such cases the product would be correct even if the decomposition was not.

(a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -8 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & \frac{3}{4} \\ 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} -\frac{1}{8} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

To see an explicit solution of a problem of this type, see Version 4.

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(x, y, z) \mid x, y, z > 0\}$
   (b) $\{(x, y, z) \mid 6x + 2y = 0, -5x - 7z = 0\}$

   Solution:

   (a) [3 MARKS] This is not a subspace. One proof is to observe that the vector $\mathbf{0}$ is not present. We can also show that one of the 2 closure rules fails. Here closure under addition does, indeed, hold; it is closure under scalar multiplication that fails. For example, the vector $\mathbf{v} = (1, 1, 1)$ has positive entries, so it is present in the subset. But, with the scalar 0, the product $0\mathbf{v}$ is the vector $\mathbf{0}$, which is not on the subset, as at least one of its entries is not positive.
   (b) [4 MARKS]
      i. We have to prove that the set is not empty, or, equivalently, that it contains the vector $\mathbf{0} = (0, 0, 0)$. Since $6(0) = 2(0) = 0$ and $-5(0)-7(0) = 0$, $(0, 0, 0)$ is in the subset.
      ii. We can prove closure under addition. If $6x_1 + 2y_1 = 0 = 6x_2 + 2y_2$ and $-5x_1 - 7z_1 = 0 = -5x_2 - 7z_2$, then $6(x_1 + x_2) + 2(y_1 + y_2) = (6x_1 + 2y_1) + (6x_2 + 2y_2) = 0 + 0 = 0$, and $-5(x_1 + x_2) - 7(z_1 + z_2) = (-5x_1 - 7z_1) + (-5x_2 - 7z_2) = 0 + 0 = 0$. Thus, if both $\mathbf{u} = (x_1, y_1, z_1)$ and $\mathbf{v} = (x_2, y_2, z_2)$ are in the set, so also is $\mathbf{u}+\mathbf{v} = ((x_1+x_2), (y_1+y_2), (z_1, z_2))$, and the set is closed under vector addition.

iii. Let $\mathbf{u} = (x_1, y_1, z_1)$ be a vector in the set, i.e., suppose that $6x_1 + 2y_1 = 0$ and $-5x_1 - 7z_1 = 0$; and let k be any scalar. Then $6(kx_1) + 2(ky_1) = k(6x_1 + 2y_1) = k(0) = 0$; in the same way we can prove that $-5(kx_1) - 7(kz_1) = 0$, so $k\mathbf{v}$ is also in the set. Thus the set is also closed under scalar multiplication.

3.2 Version 2

1. [2 MARKS] Find a 3×3 matrix A such that

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix}, A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \text{ and } A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \text{-4} \\ 0 \\ 4 \end{pmatrix}.$$

$$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

<u>Solution</u>:

$$\begin{pmatrix} 5 & 3 & -4 \\ 2 & 1 & 0 \\ 2 & 1 & 4 \end{pmatrix}$$

2. [4 MARKS] If A, B, and C are 5×5, 5×8, and 8×6 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

A^2: _______

AC: _______

CB: _______

BA: _______

<u>Solution</u>: A^2 is 5×5. AC is undefined, since A has 5 columns, but C has a different number — 8 — of rows. CB is undefined, since C has 6 columns, but B has a different number — 5 — of rows. BA is undefined, since B has 8 columns, but A has a different number — 5 — of rows.

3. [4 MARKS] If $A = \begin{pmatrix} 4 & 1 \\ 6 & 8 \end{pmatrix}$ and $B = \begin{pmatrix} \text{-3} & 2 \\ 3 & 3 \end{pmatrix}$

Then $AB = \begin{pmatrix} \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ \end{pmatrix}$

Solution: $AB = \begin{pmatrix} -9 & 11 \\ 6 & 36 \end{pmatrix}$

4. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} \text{-3} \\ 4 \\ 4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} \text{-2} \\ 0 \\ 5 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ ________.

   Solution: $(-3)(-2) + 4(0) + 4(5) = 26$.

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(5 - i)(3 - 3i)(1 - 3i) =$ ____ $+$ ____ i,
   (b) $((-4 + 5i)^2 - 3)i =$ ____ $+$ ____ i.

   Solution:

   (a) $(5 - i)(3 - 3i)(1 - 3i) = (12 - 18i)(1 - 3i) = -42 - 54i$.
   (b) $((-4 + 5i)^2 - 3)i = (-12 - 40i)i = 40 - 12i$.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} \text{-1} & \text{-3} & 1 & 14 \\ \text{-2} & \text{-3} & 3 & 22 \\ 2 & \text{-1} & 1 & 2 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

   Solution: This was a version of Problem 3 on WeBWorK assignments 3 and 4. This quiz was intended to expose difficulties in solving simple matrix or MATH 133 problems, and this question certainly did that. To row reduce this matrix you needed to plan your operations so as to minimize the computational difficulties. If you have been solving such problems before the quiz — as on the WeBWorK assignment — you know that you can plan your row reduction steps so as to keep the numbers relatively simple; and it doesn't require the use of a calculator. (You should not be using a calculator on WeBWorK.) If you permit complicated fractions to enter into your calculations, then you are inviting problems when you are doing hand calculations. Of course, this problem does not show whether you understand the principles of row reduction; it show whether you know how to manage small problems manually without assistance.

   The only practical way to grade this problem was to base the grade on what you finally achieved, and to check the work only where your final matrix was close enough to a row reduced matrix that that was feasible. In the solution I have not indicated the row operations over the arrows; it would be better to always show them, but I have had to economize on my time as I had to prepare 4 versions. I have not tried to find the shortest solution: in some cases it would have been faster to scale a row first and then add a multiple of another row.

$$\begin{pmatrix} -1 & -3 & 1 & 14 \\ -2 & -3 & 3 & 22 \\ 2 & -1 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -14 \\ -2 & -3 & 3 & 22 \\ 2 & -1 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -14 \\ 0 & 3 & 1 & -6 \\ 0 & -7 & 3 & 30 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 3 & -1 & -14 \\ 0 & 3 & 1 & -6 \\ 0 & -1 & 5 & 18 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -14 \\ 0 & 3 & 1 & -6 \\ 0 & 1 & -5 & 18 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -14 \\ 0 & 1 & -5 & -18 \\ 0 & 3 & 1 & -6 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 14 & 40 \\ 0 & 1 & -5 & -18 \\ 0 & 0 & 16 & 48 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 14 & 40 \\ 0 & 1 & -5 & -18 \\ 0 & 0 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 2 & -4 & 2 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 4 & -8 & 4 \end{bmatrix}$$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 4 & -8 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 2 \\ -3 & -3 & -5 \\ 2 & -4 & 2 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -5 & 2 \\ -3 & 1 & 4 \\ 5 & -4 & 3 \end{bmatrix} = \begin{bmatrix} -3 & -5 & 2 \\ 5 & -4 & 3 \\ -3 & 1 & 4 \end{bmatrix}$$

(d) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -3 & -5 & 2 \\ 5 & -4 & 3 \\ -3 & 1 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -5 & 2 \\ -3 & 1 & 4 \\ 5 & -4 & 3 \end{bmatrix}$$

(e) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ -5 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ 11 & 17 & 7 \end{bmatrix}$$

(f) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ 11 & 17 & 7 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ -5 & 5 & -3 \\ -5 & -3 & -1 \end{bmatrix}$$

Solution:

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$

8. [8 MARKS] Let
$$A = \begin{bmatrix} 0 & 5 \\ 9 & -4 \end{bmatrix}$$

(a) Showing all your work, write A as a product of 4 elementary matrices:
$A = \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix}$

(b) Write A^{-1} as a product of 4 elementary matrices:
$A^{-1} = \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix} \begin{bmatrix} \_ & \_ \\ \_ & \_ \end{bmatrix}$

Solution: This was a problem where you could verify the correctness of your solution by multiplying out your matrices. For that reason I have graded it mainly as all or nothing. Many students included in their produce a matrix which was not elementary; in such cases the product would be correct even if the decomposition was not. (The solutions given below are not unique.)

(a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & -\frac{4}{9} \\ 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & \frac{4}{9} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{9} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

To see an explicit solution of a problem of this type, look at the solutions to Version 4.

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(6x + 3y, -8x + 4y, -2x - 4y) \mid x, y$ arbitrary numbers $\}$

   (b) $\{(x, x - 9, x + 5) \mid x$ arbitrary number $\}$

   Solution:

   (a) [4 MARKS] Solutions to this question can be constructed analogously to the solution to Problem 9 of Test Version 1, found in these notes on page 50.

   (b) [3 MARKS] This set is not a subspace, as it is closed under neither addition nor scalar multiplication. One could also observe that the vector $\mathbf{0} = (0, 0, 0)$ is not present. For, if $(x, x - 9, x + 5) = (0, 0, 0)$, then the following equations hold *simultaneously*:

   $$x = 0, \quad x - 9 = 0, \quad x + 5 = 0\,,$$

   and each of these equations contradicts the other two!

   This example could be reformulated as a counterexample to closure under scalar multiplication. If the subset is not empty, then there must exist a real number x such that $\mathbf{v} = (x, x - 9, x + 5)$ is present. With the scalar 0, the product $0\mathbf{v} = (0x, 0(x - 9), 0(x + 5)) = (0, 0, 0)$, which we have already shown is not present in the set.

   Closure under addition is also not present. If we sum two vectors $\mathbf{v}_1 = (x_1, x_1 - 9, x_1 + 5)\mathbf{v}_2 = (x_2, x_2 - 9, x_2 + 5)$, we obtain

   $$\mathbf{v}_1 + \mathbf{v}_2 = (x_1 + x_2, (x_1 + x_2) - 18, (x_1 + x_2) + 10),.$$

   The second entry of this vector does not have the proper relationship to the first — it must be equal to the first -9, and it is not. Thus this set is not closed under vector addition either.

3.3 Version 3

1. [4 MARKS] If A, B, and C are 4×4, 4×3, and 3×6 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: ________

   AC: ________

   AB: ________

   BA: ________

   Solution: CB is undefined, since C has 6 columns, but B has a different number — 4 — of rows. AC is undefined, since A has 4 columns, but C has a different number — 3 — of rows. AB is 4×3. BA is undefined since B has 3 columns, but A has a different number — 4 — of rows.

2. [2 MARKS] Find a 3×3 matrix A such that

   $$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \text{-1} \\ 5 \\ \text{-5} \end{pmatrix},\ A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix},\ \text{and } A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}.$$

   $$A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$$

   Solution:

   $$\begin{pmatrix} -1 & 1 & 2 \\ 5 & 0 & 2 \\ -5 & 2 & 4 \end{pmatrix}$$

3. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} \text{-2} \\ 1 \\ \text{-4} \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 4 \\ \text{-4} \\ 4 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} =$ __________.

   Solution: $(-2)4 + 1(-4) + (-4)4 = -28$.

4. [4 MARKS] If $A = \begin{pmatrix} \text{-10} & 7 \\ \text{-1} & \text{-3} \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 3 \\ \text{-4} & 6 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_\_ & \_\_\_\_ \\ \_\_\_\_ & \_\_\_\_ \end{pmatrix}$

Solution: $AB = \begin{pmatrix} -58 & 12 \\ 9 & -21 \end{pmatrix}$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(3+2i)(-4-i)(5-2i) = \_\_\_\_ + \_\_\_\_\, i,$

   (b) $((4+2i)^2+1)i = \_\_\_\_ + \_\_\_\_\, i.$

   Solution:

   (a) $(3+2i)(-4-i)(5-2i) = (-10-11i)(5-2i) = -72-35i$

   (b) $((4+2i)^2+1)i = (13+16i)i = -16+13i.$

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} \text{-1} & \text{-3} & 1 & 3 \\ \text{-2} & 1 & 3 & \text{-8} \\ \text{-2} & \text{-1} & 1 & \text{-4} \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

   Solution: This was a version of Problem 3 on WeBWorK assignments 3 and 4. This quiz was intended to expose difficulties in solving simple matrix or MATH 133 problems, and this question certainly did that. To row reduce this matrix you needed to plan your operations so as to minimize the computational difficulties. If you have been solving such problems before the quiz — as on the WeBWorK assignment — you know that you can plan your row reduction steps so as to keep the numbers relatively simple; and it doesn't require the use of a calculator. (You should not be using a calculator on WeBWorK.) If you permit complicated fractions to enter into your calculations, then you are inviting problems when you are doing hand calculations. Of course, this problem does not show whether you understand the principles of row reduction; it show whether you know how to manage small problems manually without assistance.

   The only practical way to grade this problem was to base the grade on what you finally achieved, and to check the work only where your final matrix was close enough to a row reduced matrix that that was feasible. In the solution I have not indicated the row operations over the arrows; it would be better to always show them, but I have had to economize on my time as I had to prepare 4 versions. I have not tried to find the shortest solution: in some cases it would have been faster to scale a row first and then add a multiple of another row.

$$\begin{pmatrix} -1 & -3 & 1 & 3 \\ -2 & 1 & 3 & -8 \\ -2 & -1 & 1 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -3 \\ -2 & 1 & 3 & -8 \\ -2 & -1 & 1 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -1 & -3 \\ 0 & 7 & 1 & -14 \\ 0 & 5 & -1 & -10 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 3 & -1 & -3 \\ 0 & 7 & 1 & -14 \\ 0 & -2 & -2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -1 & -3 \\ 0 & 7 & 1 & -14 \\ 0 & 1 & 1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -1 & -3 \\ 0 & 1 & 1 & -2 \\ 0 & 7 & 1 & -14 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & -4 & 3 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & -6 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -4 & 3 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

(a) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 4 & 3 & -3 \end{bmatrix} = \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 12 & 9 & -9 \end{bmatrix}$$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 12 & 9 & -9 \end{bmatrix} = \begin{bmatrix} -4 & -2 & -4 \\ -5 & 2 & 4 \\ 4 & 3 & -3 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -2 & 1 & -3 \\ -4 & -1 & 1 \\ -5 & -2 & -5 \end{bmatrix} = \begin{bmatrix} -4 & -1 & 1 \\ -2 & 1 & -3 \\ -5 & -2 & -5 \end{bmatrix}$$

(d) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & -1 & 1 \\ -2 & 1 & -3 \\ -5 & -2 & -5 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -3 \\ -4 & -1 & 1 \\ -5 & -2 & -5 \end{bmatrix}$$

(e) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ 3 & -2 & 5 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ -9 & 1 & -1 \end{bmatrix}$$

(f) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ -9 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -2 \\ 4 & 5 & -4 \\ 3 & -2 & 5 \end{bmatrix}$$

<ins>Solution</ins>:

(a) $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

(b) $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

(c) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$

8. [8 MARKS] Let
$$A = \begin{bmatrix} 0 & -9 \\ 2 & 5 \end{bmatrix}$$

   (a) Showing all your work, write A as a product of 4 elementary matrices:
   $A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$

   (b) Write A^{-1} as a product of 4 elementary matrices:
   $A^{-1} = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$

   Solution: This was a problem where you could verify the correctness of your solution by multiplying out your matrices. For that reason I have graded it mainly as all or nothing. Many students included in their produce a matrix which was not elementary; in such cases the product would be correct even if the decomposition was not. (The solutions given below are not unique.)

   (a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -9 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{2} \\ 0 & 1 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & -\frac{5}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{9} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

   To see an explicit solution of a problem of this type, look at the solutions to Version 4.

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

(a) $\{(x, y, z) \mid x + y + z = 8\}$

(b) $\{(-8x, 2x, -9x) \mid x \text{ arbitrary number }\}$

Solution:

(a) [3 MARKS] This subset is not a subspace. It is not closed under either addition or multiplication by a scalar, and it also does not contain the vector $\mathbf{0} = (0, 0, 0)$.
Suppose that $x_1 + y_1 + z_1 = 8 = x_2 + y_2 + z_2$. Then

$$(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = 8 + 8 = 16 \neq 8$$

so the set is not closed under addition; and, if k is any real number, and $x+y+z = 8$, then $(kx)+(ky)+(kz) = k(x+y+z) = 8k$, so the set is closed only under multiplication by the scalar 1. That the vector $\mathbf{0}$ is not present is obvious, since the sum of its coordinates, $0 + 0 + 0 = 0 \neq 8$.

(b) [4 MARKS] The set consists of all scalar multiples of the vector $(-8, 2, -9)$. Thus it is clearly not empty, and is closed under scalar multiplication. Closure under vector addition follows from the fact that $(-8x_1, 2x_1, -9x_1) + (-8x_2, 2x_2, -9x_2) = (-8x_1-8x_2, 2x_2+2x_2, -9x_1-9x_2) = (-8(x_1+x_2), 2(x_1+x_2), -9(x_1 + x_2))$.

3.4 Version 4

1. [4 MARKS] If $A = \begin{pmatrix} -3 & 7 \\ 4 & 8 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & -3 \\ -3 & -2 \end{pmatrix}$

   Then $AB = \begin{pmatrix} \_\_\_\_ & \_\_\_\_ \\ \_\_\_\_ & \_\_\_\_ \end{pmatrix}$

   Solution: $AB = \begin{pmatrix} -18 & -5 \\ -28 & -28 \end{pmatrix}$

2. [2 MARKS] Let $\mathbf{x} = \begin{pmatrix} -5 \\ -3 \\ 0 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix}$.

   Find the dot product of x and y.

   $\mathbf{x} \bullet \mathbf{y} = $ ________.

   Solution: $(-5)2 + (-3)4 + 0(-4) = -22$.

3. [4 MARKS] If A, B, and C are 3×3, 3×7, and 7×9 matrices respectively, determine which of the following products are defined. For those defined, enter the size of the resulting matrix (e.g., "3 x 4"). For those undefined, give an explicit reason why the product is undefined.

   CB: ______

   BC: ______

   BA: ______

   AC: ______

   Solution: CB is undefined, since C has 9 columns, but B has a different number — 3 — of rows. BC is 3×9. BA is undefined, since B has 7 columns, but A has a different number — 3 — of rows. AC is undefined, since A has 3 columns, but C has a different number — 7 — of rows.

4. [2 MARKS] Find a 3×3 matrix A such that

   $A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \text{-1} \\ \text{-2} \\ 4 \end{pmatrix}$, $A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \text{-3} \\ \text{-2} \\ \text{-1} \end{pmatrix}$, and $A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \text{-2} \\ 5 \\ \text{-1} \end{pmatrix}$.

   $A = \begin{pmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{pmatrix}$

   Solution:

$$\begin{pmatrix} -1 & -3 & -2 \\ -2 & -2 & 5 \\ 4 & -1 & -1 \end{pmatrix}$$

5. [4 MARKS] Showing your detailed work, write the following complex numbers in $a + bi$ form:

   (a) $(3 - 2i)(-1 + 2i)(-4 + i) = \_\_\_\_ + \_\_\_\_\ i$,

   (b) $((5 - 5i)^2 - 5)i = \_\_\_\_ + \_\_\_\_\ i$.

   Solution:

   (a) $(3 - 2i)(-1 + 2i)(-4 + i) = (1 + 8i)(-4 + i) = -12 - 31i$.

   (b) $((5 - 5i)^2 - 5)i = (-50i - 5)i = 50 - 5i$.

6. [8 MARKS] Showing all your work, reduce the matrix $\begin{pmatrix} 3 & 3 & \text{-3} & 18 \\ 2 & \text{-3} & 3 & \text{-28} \\ 2 & 0 & \text{-1} & 0 \end{pmatrix}$ to reduced row-echelon form (called *row canonical form* in your textbook).

   Solution: This was a version of Problem 3 on WeBWorK assignments 3 and 4. This quiz was intended to expose difficulties in solving simple matrix or MATH 133 problems, and this question certainly did that. To row reduce this matrix you needed to plan your operations so as to minimize the computational difficulties. If you have been solving such problems before the quiz — as on the WeBWorK assignment — you know that you can plan your row reduction steps so as to keep the numbers relatively simple; and it doesn't require the use of a calculator. (You should not be using a calculator on WeBWorK.) If you permit complicated fractions to enter into your calculations, then you are inviting problems when you are doing hand calculations. Of course, this problem does not show whether you understand the principles of row reduction; it show whether you know how to manage small problems manually without assistance.

   The only practical way to grade this problem was to base the grade on what you finally achieved, and to check the work only where your final matrix was close enough to a row reduced matrix that that was feasible. In the solution I have not indicated the row operations over the arrows; it would be better to always show them, but I have had to economize on my time as I had to prepare 4 versions. I have not tried to find the shortest solution: in some cases it would have been faster to scale a row first and then add a multiple of another row.

$$\begin{pmatrix} 3 & 3 & -3 & 18 \\ 2 & -3 & 3 & -28 \\ 2 & 0 & -1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 6 \\ 2 & -3 & 3 & -28 \\ 2 & 0 & -1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 6 \\ 0 & -5 & 5 & -40 \\ 0 & -2 & 1 & -12 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1 & -1 & 6 \\ 0 & 1 & -1 & 8 \\ 0 & -2 & 1 & -12 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -1 & 8 \\ 0 & 0 & -1 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -1 & 8 \\ 0 & 0 & 1- & 4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -4 \end{pmatrix}$$

7. [6 MARKS] In the following problems find elementary matrices such that the respective matrix equations hold.

   (a) $\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \text{-2} & 3 & 1 \\ \text{-5} & \text{-2} & 3 \\ \text{-4} & \text{-2} & \text{-2} \end{bmatrix} = \begin{bmatrix} \text{-8} & 12 & 4 \\ \text{-5} & \text{-2} & 3 \\ \text{-4} & \text{-2} & \text{-2} \end{bmatrix}$

(b) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -8 & 12 & 4 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 1 \\ -5 & -2 & 3 \\ -4 & -2 & -2 \end{bmatrix}$$

(c) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} -1 & 5 & -1 \\ 1 & -2 & -2 \\ 3 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -1 \\ 1 & -2 & -2 \\ -1 & 5 & -1 \end{bmatrix}$$

(d) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 3 & -2 & -1 \\ 1 & -2 & -2 \\ -1 & 5 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 5 & -1 \\ 1 & -2 & -2 \\ 3 & -2 & -1 \end{bmatrix}$$

(e) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -1 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -9 & 3 & 3 \end{bmatrix}$$

(f) $$\begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -9 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -4 & 2 \\ -4 & 2 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

<u>Solution</u>:

(a) $$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(b) $$\begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(c) $$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(d) $$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(e) $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

(f) $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

8. [8 MARKS] Let
$$A = \begin{bmatrix} 0 & -7 \\ -5 & -8 \end{bmatrix}$$

   (a) Showing all your work, write A as a product of 4 elementary matrices:
   $$A = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

   (b) Write A^{-1} as a product of 4 elementary matrices:
   $$A^{-1} = \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix}$$

   Solution: This was a problem where you could verify the correctness of your solution by multiplying out your matrices. For that reason I have graded it mainly as all or nothing. Many students included in their produce a matrix which was not elementary; in such cases the product would be correct even if the decomposition was not. (The solutions given below are not unique.)

   (a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -7 \end{pmatrix} \begin{pmatrix} 1 & \frac{8}{5} \\ 0 & 1 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & -\frac{8}{5} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{7} \end{pmatrix} \begin{pmatrix} -\frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

   Here is one way to solve this kind of problem. Row reduce the matrix, keeping a record of the operations; then, to the left of the matrix, write down the elementary matrices that accomplish the row operations, in the order in which they have been applied. The product of those matrices will be the inverse. The decomposition of the original matrix can be obtained by writing the inverses of those matrices in the opposite order. Here

$$\begin{aligned}
A = & \begin{pmatrix} 0 & -7 \\ -5 & -8 \end{pmatrix} \\
R_1 \leftrightarrow R_2 : & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -7 \\ -5 & -8 \end{pmatrix} = \begin{pmatrix} -5 & -8 \\ 0 & -7 \end{pmatrix} \\
\frac{-1}{5}R_1 : & \begin{pmatrix} -\frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -7 \\ -5 & -8 \end{pmatrix} = \begin{pmatrix} 1 & \frac{8}{5} \\ 0 & -7 \end{pmatrix} \\
\frac{-1}{7}R_2 : & \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{7} \end{pmatrix}\begin{pmatrix} -\frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -7 \\ -5 & -8 \end{pmatrix} = \begin{pmatrix} 1 & \frac{8}{5} \\ 0 & 1 \end{pmatrix} \\
R_1 \rightarrow R_1 - \frac{8}{5}R_2 : & \begin{pmatrix} 1 & -\frac{8}{5} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{7} \end{pmatrix}\begin{pmatrix} -\frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -7 \\ -5 & -8 \end{pmatrix} \\
= & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

9. [7 MARKS] Consider the following subsets of $\mathbb{R}^3$. If the set is a subspace, state that fact and prove it. If it is not a subspace, provide an explicit counterexample to demonstrate that fact.

   (a) $\{(2x - 8y, -4x - 9y, 5x + 9y) \mid x, y \text{ arbitrary numbers }\}$

   (b) $\{(-6, y, z) \mid y, z \text{ arbitrary numbers }\}$

   Solution:

   (a) [4 MARKS] Solutions to this question can be constructed analogously to the solution to Problem 9 of Test Version 1, found in these notes on page 50.

   (b) [3 MARKS] This subset is not closed under either vector addition or scalar multiplication. The sum of two vectors in the set will have, as its first entry, $(-6) + (-6) = -12$, which is not of the desired form (-6). Multiplication of a vector in the set by an arbitrary scalar k will not yield a vector in the set unless the $k = 1$, since we would require — looking only at the first entry — that $k(-6) = -6$, implying that $k = 1$. Thus any scalar except 1 yields a counterexample.

   It is not good enough to say that, when you set $y = z = 0$ you obtain a vector $(-6, 0, 0)$, which is not $(0, 0, 0)$. You need to observe that you can't find any values of y, z that will make the triple $(0, 0, 0)$ and that, in attempting to solve the system of equations

$$\begin{aligned} -6 &= 0 \\ y &= 0 \\ z &= 0 \end{aligned}$$

you can't find any values of y or z which will satisfy the first equation; it isn't enough to consider the last 2 equations, since whatever you do there still does not achieve a solution to the first equation.

4 First Written Assignment

Distribution Date: Mounted on the Web on Friday, February 10th, 2006
Solutions are to be submitted by Wednesday, March 1st, 2006

1. A function $F : \mathcal{U} \to \mathcal{V}$ from one real vector space to another is defined to be *odd* if, for any $\mathbf{u} \in \mathcal{U}$, $F(-\mathbf{u}) = -F(\mathbf{u})$.
   (a) Prove carefully that every linear transformation from $\mathcal{U}$ to $\mathcal{V}$ is odd.
   (b) For real vector spaces $\mathcal{U} = \mathcal{V} = \mathbb{R}^1$ construct an example of an odd function $F : \mathcal{U} \to \mathcal{V}$ which is not a linear transformations.

2. (cf. [1, Exercise 4.131, p. 168]) Let $\mathcal{V} = \mathbb{R}_{2,2}$.
   (a) Prove that the following matrices constitute a basis for $\mathcal{V}$: $M_1 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, $M_2 = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$, $M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.
   (b) Find the coordinate vector of the matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ referred to the coordinate system M_1, M_2, M_3, M_4 (in that order).
   (c) Show that the function $F : \mathcal{V} \to \mathcal{V}$ which maps any matrix on to its transpose, i.e., such that $F(A) = A^T$, is a linear transformation. (Hint: look at [1, Theorem 2.3, p. 33], which you may accept without proof.)
   (d) Use your earlier results to determine the matrix of F referred to the given coordinate system .
   (e) Determine the matrix of F
      - referred to the coordinate system M_1, M_2, M_3, M_4 in the domain; and
      - referred to the coordinate system M_4, M_3, M_2, M_1 in the target.
   (f) Determine the matrix of F
      - referred to the coordinate system M_3, M_1, M_4, M_2 in the domain; and
      - referred to the coordinate system M_1, M_2, M_3, M_4 in the target.

3. Consider linear transformations

$$H : \mathbb{R}^4 \to \mathbb{R}^3, \quad K : \mathbb{R}^3 \to \mathbb{R}^4$$

given by

$$H\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 2x+y+4z+3t \\ -3x+2y-13z-t \\ 4x+12z+4t \end{pmatrix}$$

$$K\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ u \\ v \\ w \end{pmatrix}$$

(a) Show that none of the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is in the kernel of H.

(b) Determine a basis for the kernel of H.

(c) Give the matrix of H referred to the standard bases of $\mathbb{R}^4$ and $\mathbb{R}^3$ respectively in the domain and the target.

(d) Find a coordinate system for the domain of H consisting of your full basis for the kernel of H and some vectors from the standard basis. Give the matrix of H referred to your new coordinate system in the domain, and the standard coordinate system in the target.

(e) Determine the matrix of $H \circ K$, referred to the standard bases of its domain and target.

(f) Determine the matrix of $K \circ H$, referred to the standard bases of its domain and target.

(g) Determine a basis for the image of K.

4. Consider the real vector space $\mathcal{V} = \mathbb{R}_3[t]$ of polynomials of degree not exceeding 3. You may assume that is is known that the vectors $\mathbf{f}_i = t^{i-1}$ $(i = 1, 2, 3, 4)$ constitute a basis.

   (a) Show that the operation of differentiation of polynomials in $\mathcal{V}$ — which we shall denote by D — is a linear transformation.

   (b) Find the matrix of D referred to the coordinate system $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$.

   (c) Find a basis for the subspace of $\mathcal{V}$ consisting of polynomials which constitute the kernel of $D^2 + 3D + 2I$.

5 Solutions, First Written Assignment

The assignment was due by 5 p.m. on Wednesday, March 1st, 2006. These draft solutions, which are subject to correction, were mounted on the Web about 09:10 on March 2nd.

1. A function $F : \mathcal{U} \to \mathcal{V}$ from one real vector space to another is defined to be *odd* if, for any $\mathbf{u} \in \mathcal{U}$, $F(-\mathbf{u}) = -F(\mathbf{u})$.

   (a) Prove carefully that every linear transformation from $\mathcal{U}$ to $\mathcal{V}$ is odd.

   (b) For real vector spaces $\mathcal{U} = \mathcal{V} = \mathbb{R}^1$ construct an example of an odd function $F : \mathcal{U} \to \mathcal{V}$ which is not a linear transformations.

   Solution:

   (a) Let $S : \mathcal{U} \to \mathcal{V}$ be a linear transformation. Then, taking $\mathbf{u}$ to be any vector in $\mathcal{U}$, and the scalar -1, we have

   $$S(-\mathbf{u}) = -S(\mathbf{u})$$

   by virtue of one of the properties of linearity. Thus S is an odd function.

   (b) One example would be where $F(1) = 2$, $F(-1) = -2$, $F(x) = 0$ for $x \neq \pm 1$. This function is odd, but it's not linear, since

   $$F(1+1) = F(2) = 0 \neq 2 + 2 = F(1) + F(1)$$

2. (cf. [1, Exercise 4.131, p. 168]) Let $\mathcal{V} = \mathbb{R}_{2,2}$.

   (a) Prove that the following matrices constitute a basis for $\mathcal{V}$: $M_1 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, $M_2 = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$, $M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

   (b) Find the coordinate vector of the matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ referred to the coordinate system M_1, M_2, M_3, M_4 (in that order).

   (c) Show that the function $F : \mathcal{V} \to \mathcal{V}$ which maps any matrix on to its transpose, i.e., such that $F(A) = A^T$, is a linear transformation. (Hint: look at [1, Theorem 2.3, p. 33], which you may accept without proof.)

   (d) Use your earlier results to determine the matrix of F referred to the given coordinate system .

   (e) Determine the matrix of F

- referred to the coordinate system M_1, M_2, M_3, M_4 in the domain; and
- referred to the coordinate system M_4, M_3, M_2, M_1 in the target.

(f) Determine the matrix of F

- referred to the coordinate system M_3, M_1, M_4, M_2 in the domain; and
- referred to the coordinate system M_1, M_2, M_3, M_4 in the target.

<u>Solution</u>:

(a) Consider the matrix equation

$$a_1M_1 + a_2M_2 + a_3M_3 + a_4M_4 = M$$

where $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, i.e.,

$$\begin{pmatrix} a_1 - a_2 & a_1 + a_2 \\ -a_1 + a_3 & -a_1 + a_2 + a_3 + a_4 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

equivalently the 4 scalar equations

$$\begin{aligned} +1a_1 - 1a_2 + 0a_3 + 0a_4 &= a \\ +1a_1 + 1a_2 + 0a_3 + 0a_4 &= b \\ -1a_1 + 0a_2 + 1a_3 + 0a_4 &= c \\ -1a_1 + 1a_2 + 1a_3 + 1a_4 &= d \end{aligned}$$

whose augmented matrix is

$$\left(\begin{array}{rrrr|c} 1 & -1 & 0 & 0 & a \\ 1 & 1 & 0 & 0 & b \\ -1 & 0 & 1 & 0 & c \\ -1 & 1 & 1 & 1 & d \end{array}\right)$$

and which may be row reduced to

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{a+b}{2} \\ 0 & 1 & 0 & 0 & \frac{-a+b}{2} \\ 0 & 0 & 1 & 0 & \frac{a+b+2c}{2} \\ 0 & 0 & 0 & 1 & \frac{a-b-2c+2d}{2} \end{array}\right).$$

When $a = b = c = d$, this shows that the only linear combination of these matrices which equals the matrix 0 is the trivial linear combination: i.e., it shows that the 4 given matrices are linearly independent. For general a, b, c, d

these equations show how every matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be expressed (uniquely) as a linear combination; thus it shows that the given matrices span the space of all 2×2 matrices. (We didn't need this existence part of the proof, as we knew from the fact that the 4 matrices are linearly independent members of a space of dimension 4, that they had to span the space.)

(b) The last part of the preceding computation shows that

$$\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]_{\{M_1,M_2,M_3,M_4\}} = \frac{1}{2}\begin{pmatrix} a+b \\ -a+b \\ a+b+2c \\ a-b-2c+2d \end{pmatrix}.$$

(c) For any two matrices, A and B, and any scalar k

$$\begin{aligned} F(A+B) &= (A+B)^T && \text{definition of } F \\ &= A^T + B^T && [1, \text{Theorem 2.3(i), p. 33}] \\ &= F(A) + F(B) && \text{definition of } F \\ F(kA) &= (kA)^T && \text{definition of } F \\ &= kA^T && [1, \text{Theorem 2.3(iii), p. 33}] \\ &= kF(A) && \text{definition of } F \end{aligned}$$

showing that both properties defining linearity hold for this transformation.

(d) Applying the formula determined above for the coordinates of any matrix when referred to the given coordinate system, we have

$$\left[M_1^T\right]_{\{M_1,M_2,M_3,M_4\}} = \left[\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\right]_{\{M_1,M_2,M_3,M_4\}} = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$\left[M_2^T\right]_{\{M_1,M_2,M_3,M_4\}} = \left[\begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}\right]_{\{M_1,M_2,M_3,M_4\}} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

$$\left[M_3^T\right]_{\{M_1,M_2,M_3,M_4\}} = \left[\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\right]_{\{M_1,M_2,M_3,M_4\}} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\left[M_4^T\right]_{\{M_1,M_2,M_3,M_4\}} = \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]_{\{M_1,M_2,M_3,M_4\}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

It follows that the matrix of the transpose function is

$$[F]^{\{M_1,M_2,M_3,M_4\}}_{\{M_1,M_2,M_3,M_4\}} = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

(e) The only change from the preceding is in the order of the basis vectors in the coordinate system for the target. This means that we have to change the order of the scalar weights attached to the basis vectors. From the information just determined, we have

$$\left[M_1^T\right]_{\{M_4,M_3,M_2,M_1\}} = \left[\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\right]_{\{M_4,M_3,M_2,M_1\}} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\left[M_2^T\right]_{\{M_4,M_3,M_2,M_1\}} = \left[\begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}\right]_{\{M_4,M_3,M_2,M_1\}} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

$$\left[M_3^T\right]_{\{M_4,M_3,M_2,M_1\}} = \left[\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\right]_{\{M_4,M_3,M_2,M_1\}} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\left[M_4^T\right]_{\{M_4,M_3,M_2,M_1\}} = \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]_{\{M_4,M_3,M_2,M_1\}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

It follows that the matrix of the transpose function is

$$[F]_{\{M_4,M_3,M_2,M_1\}}^{\{M_1,M_2,M_3,M_4\}} = \begin{pmatrix} -1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ -1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

The correspondence of the columns of the matrix with the basis vectors for the domain has not changed, but the rows of the matrix have been rearranged.

(f) We have changed the order of the basis vectors for the domain, but not those of the target. This means that the order of the column vectors in the matrix is the only change:

$$[F]_{\{M_1,M_2,M_3,M_4\}}^{\{M_3,M_1,M_4,M_2\}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ \frac{1}{2} & -1 & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & \frac{1}{2} \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{pmatrix}$$

3. Consider linear transformations

$$H : \mathbb{R}^4 \to \mathbb{R}^3, \quad K : \mathbb{R}^3 \to \mathbb{R}^4$$

given by

$$H\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 2x+y+4z+3t \\ -3x+2y-13z-t \\ 4x+12z+4t \end{pmatrix}$$

$$K\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ u \\ v \\ w \end{pmatrix}$$

(a) Show that none of the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is in the kernel of H.

(b) Determine a basis for the kernel of H.

(c) Give the matrix of H referred to the standard bases of $\mathbb{R}^4$ and $\mathbb{R}^3$ respectively in the domain and the target.

(d) Find a coordinate system for the domain of H consisting of your full basis for the kernel of H and some vectors from the standard basis. Give the matrix of H referred to your new coordinate system in the domain, and the standard coordinate system in the target.

(e) Determine the matrix of $H \circ K$, referred to the standard bases of its domain and target.

(f) Determine the matrix of $K \circ H$, referred to the standard bases of its domain and target.

(g) Determine a basis for the image of K.

Solution:

(a) $$H\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}, \quad H\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad H\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -13 \\ 12 \end{pmatrix},$$

$H\begin{pmatrix}0\\0\\0\\1\end{pmatrix}=\begin{pmatrix}3\\-1\\4\end{pmatrix}$, $\quad H\begin{pmatrix}1\\1\\1\\1\end{pmatrix}=\begin{pmatrix}10\\-15\\20\end{pmatrix}$. None of the five given vectors is mapped on to $\mathbf{0}$, so none of them is in the kernel of H.

(b) To determine the kernel we solve the equation

$$\begin{pmatrix}2x+y+4z+3t\\-3x+2y-13z-t\\4x+12z+4t\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

i.e., the system of homogeneous linear equations whose coefficient matrix is

$$\begin{pmatrix}2&1&4&3\\-3&2&-13&-1\\4&0&12&4\end{pmatrix}.$$

Row reduction yields the equivalent system

$$\begin{aligned}x+3z+t&=0\\y-2z+t&=0\end{aligned}$$

or

$$\begin{aligned}x&=-3z-t\\y&=2z-t\,.\end{aligned}$$

Thus the general solution is

$$\begin{pmatrix}x\\y\\z\\t\end{pmatrix}=\begin{pmatrix}-3z-t\\2z-t\\z\\t\end{pmatrix}=z\begin{pmatrix}-3\\2\\1\\0\end{pmatrix}+t\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}$$

so *one* basis for the kernel is the 2 vectors

$$\begin{pmatrix}-3\\2\\1\\0\end{pmatrix},\quad\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix},$$

obtained by setting all but one of the parameters z, t equal to zero, and that one equal to 1 — in all possible ways.

(c) We have earlier determined the images of the standard basis vectors of the domain, expressed in terms of the standard coordinate system of the target; these vectors form the columns of the matrix, and we obtain

$$[H] = \begin{pmatrix} 2 & 1 & 4 & 3 \\ -3 & 2 & -13 & -1 \\ 4 & 0 & 12 & 4 \end{pmatrix}$$

which matrix we have already worked with above.

(d) Let's take the two vectors we have chosen to be basis vectors of the kernel as the first 2 vectors of our coordinate system; they are both mapped on to $\mathbf{0}$, whose coordinate vector will always be $\mathbf{0}$. The remaining vectors of our coordinate system are to be selected from the standard basis. We have already determined the action of H on them. The matrix of H referred to this new coordinate system will have 2 zero columns, followed by 2 columns from the 3×4 matrix found above; there are many different possible solutions. We have already shown that none of the vectors in the standard basis, so we can certainly select any *one* of them to increase the size of our independent set from 2 to 3. However, it might happen that *two* of these vectors could, together with the basis vectors from the kernel, not form a linearly independent set of 4 vectors. So you should still check that the 4 vectors you propose to use are linearly independent. I am going to work with the 3rd and 4th vectors of the standard basis, but other selections are possible — in fact all other selections are possible. I begin by checking that I have a linearly independent set; I do this by column reduction, observing that $\begin{pmatrix} -3 & -1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ column reduces to $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Now I may determine the matrix of the transformation:

$$[H]^{\left\{\begin{pmatrix} -3 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}}_{\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}} = \begin{pmatrix} 0 & 0 & 4 & 3 \\ 0 & 0 & -13 & -1 \\ 0 & 0 & 12 & 4 \end{pmatrix}.$$

The matrix has two zero columns, corresponding to the basis vectors from the kernel, and the other columns were obtained earlier.

(e) The matrix of K (referred to the standard coordinate systems) is

$$[K] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The matrix of $H \circ K$ (referred to the standard coordinate system in $\mathbb{R}^3$) will be the product

$$[H][K] = \begin{pmatrix} 2 & 1 & 4 & 3 \\ -3 & 2 & -13 & -1 \\ 4 & 0 & 12 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 3 \\ 2 & -13 & -1 \\ 0 & 12 & 4 \end{pmatrix} .$$

(f) The matrix of $K \circ H$ (referred to the standard coordinate system in $\mathbb{R}^4$) will be the product

$$[K][H] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 4 & 3 \\ -3 & 2 & -13 & -1 \\ 4 & 0 & 12 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 1 & 4 & 3 \\ -3 & 2 & -13 & -1 \\ 4 & 0 & 12 & 4 \end{pmatrix} .$$

(g) To find a basis for the image of K, we can *column*-reduce $[K]$. But this matrix is already column-reduced! *One* basis is thus

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

4. Consider the real vector space $\mathcal{V} = \mathbb{R}_3[t]$ of polynomials of degree not exceeding 3. You may assume that is is known that the vectors $\mathbf{f}_i = t^{i-1}$ $(i = 1, 2, 3, 4)$ constitute a basis.

(a) Show that the operation of differentiation of polynomials in $\mathcal{V}$ — which we shall denote by D — is a linear transformation.

(b) Find the matrix of D referred to the coordinate system $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$.

(c) Find a basis for the subspace of $\mathcal{V}$ consisting of polynomials which constitute the kernel of $D^2 + 3D + 2I$.

Solution:

(a) Let

$$\begin{aligned}\mathbf{a} &= a_0t^0 + a_1t^1 + a_2t^2 + a_3t^3\\ \mathbf{b} &= b_0t^0 + b_1t^1 + b_2t^2 + b_3t^3\end{aligned}$$

be any two polynomials in $\mathcal{V}$, and let k be any scalar. Then

$$\begin{aligned}D(\mathbf{a}+\mathbf{b}) &= D\left(\left(a_0t^0 + a_1t^1 + a_2t^2 + a_3t^3\right) + \left(b_0t^0 + b_1t^1 + b_2t^2 + b_3t^3\right)\right)\\ &= D\left(\left(a_0+b_0\right)t^0 + \left(a_1+b_1\right)t^1 + \left(a_2+b_2\right)t^2 + \left(a_3+b_3\right)t^3\right)\\ &= \left(a_1+b_1\right)t^0 + 2\left(a_2+b_2\right)t^1 + 3\left(a_3+b_3\right)t^2\\ &= \left(a_1t^0 + 2a_2t^1 + 3a_3t^2\right) + \left(b_1t^0 + 2b_2t^1 + 3b_3t^2\right)\\ &= D\mathbf{a} + D\mathbf{b}\\ \text{and}\qquad D(k\mathbf{a}) &= D\left(ka_0t^0 + ka_1t^1 + ka_2t^2 + ka_3t^3\right)\\ &= ka_1t^0 + 2ka_2t^1 + 3ka_3t^2\\ &= k\left(a_1t^0 + 2a_2t^1 + 3a_3t^2\right)\\ &= kD\mathbf{a}\end{aligned}$$

(b) Since

$$\begin{aligned}D\mathbf{f}_1 = 0 &= 0\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3 + 0\mathbf{f}_4\\ D\mathbf{f}_2 = 1 &= 1\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3 + 0\mathbf{f}_4\\ D\mathbf{f}_3 = 2t &= 0\mathbf{f}_1 + 2\mathbf{f}_2 + 0\mathbf{f}_3 + 0\mathbf{f}_4\\ D\mathbf{f}_4 = 3t^2 &= 0\mathbf{f}_1 + 0\mathbf{f}_2 + 3\mathbf{f}_3 + 0\mathbf{f}_4\,,\end{aligned}$$

$$[D]^{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4}_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4} = \left(\begin{array}{cccc}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{array}\right).$$

(c) The matrix of $D^2 + 3D + 2I$ is the sum of the corresponding multiples of powers of the matrix $[D]$, i.e.,

$$\left[D^2 + 3D + 2I\right]^{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4}_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4} = \left(\begin{array}{cccc}2&3&2&0\\0&2&6&6\\0&0&2&9\\0&0&0&2\end{array}\right),$$

whose null space can be shown (by row reduction) to be only the vector $\mathbf{0}$, i.e., only the zero polynomial. This space has an empty basis.

6 Second Class Tests

6.1 Version 1

SECOND CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 08th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a)	1(b)	1(c)	1(d)	2	3		
/2	/8	/6	/4	/10	/10		
				RAW /40		/	

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).
   (a) [2 MARKS] Define what is meant by a *basis* of $\mathbb{R}_2[x]$.
   (b) [8 MARKS] Prove that $C = \{1, x-1, (x-1)^2\}$ is a basis of $\mathbb{R}_2[x]$.
   (c) [6 MARKS] You may assume that differentiation is a linear transformation D of $\mathbb{R}_2[x]$. Find its matrix, referred to coordinate system $1, x-1, (x-1)^2$ in both domain and target.
   (d) [4 MARKS] Determine the coordinates of the polynomial $3 + 6x + x^2$ referred to the coordinate system $1, x-1, (x-1)^2$.
2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2, -1, 8, 7, -1)$ and $\mathbf{u}_2 = (1, -2, 4, -1, -2)$.
3. [10 MARKS] Working carefully, and showing all your work, find the values of k so that the following is an inner product on $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$:
$$f(\mathbf{u}, \mathbf{v}) = x_1 y_1 - 4x_1 y_2 - 4x_2 y_1 + k^2 x_2 y_2\,.$$

6.2 Version 2

SECOND CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 08th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a)	1(b)	1(c)	1(d)	2	3(a)	3(b)	
/2	/8	/6	/4	/10	/5	/5	
				Raw /40		/	

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).

   (a) [2 MARKS] Define what is meant by a linear transformation F from $\mathbb{R}_2[x]$ to $\mathbb{R}_2[x]$.

   (b) [8 MARKS] Prove that the function F given by
   $$F(a + bx + cx^2) = c - ax + bx^2$$
   is a linear transformation of $\mathbb{R}_2[x]$.

   (c) [6 MARKS] You may assume that $C = \{1, x+1, (x+1)^2\}$ is a basis of $\mathbb{R}_2[x]$. Find the matrix of F, referred to coordinate system $1, x+1, (x+1)^2$ in the domain and, and to the standard coordinate system, $1, x, x^2$ in the target.

   (d) [4 MARKS] Determine a formula for $(F \circ F \circ F)(a + bx + cx^2)$

2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2, -3, -19, 6, 8)$ and $\mathbf{u}_2 = (-1, 2, 11, -4, -4)$.

3. Show that neither of these forms fully satisfies the conditions for an inner product:

   (a) [5 MARKS] On $\mathbb{R}^3$, where $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = x_1y_1 + x_3y_3$.

   (b) [5 MARKS] On $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = 4x_1y_1 - 7x_1y_2 - 7x_2y_1 + 9x_2y_2$.

6.3 Version 3

SECOND CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 08th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a) /2	1(b) /8	1(c) /6	1(d) /4	2 /10	3 /10		
				Raw /40		/	

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).

   (a) [2 MARKS] Define what is meant by a *basis* of $\mathbb{R}_2[x]$.

   (b) [8 MARKS] Prove that $C = \{1, x-1, (x-1)^2\}$ is a basis of $\mathbb{R}_2[x]$.

   (c) [6 MARKS] You may assume that differentiation is a linear transformation D of $\mathbb{R}_2[x]$. Find its matrix, referred to coordinate system $1, x-1, (x-1)^2$ in both domain and target.

   (d) [4 MARKS] Determine the coordinates of the polynomial $3 + 6x + x^2$ referred to the coordinate system $1, x-1, (x-1)^2$.

2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2, -1, 8, 7, -1)$ and $\mathbf{u}_2 = (1, -2, 4, -1, -2)$.

3. [10 MARKS] Working carefully, and showing all your work, find the values of k so that the following is an inner product on $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$:

$$f(\mathbf{u}, \mathbf{v}) = x_1y_1 - 4x_1y_2 - 4x_2y_1 + k^2x_2y_2 \,.$$

6.4 Version 4

SECOND CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 08th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a)	1(b)	1(c)	1(d)	2	3(a)	3(b)	
/2	/8	/6	/4	/10	/5	/5	
				Raw /40		/	

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).
   (a) [2 MARKS] Define what is meant by a linear transformation F from $\mathbb{R}_2[x]$ to $\mathbb{R}_2[x]$.
   (b) [8 MARKS] Prove that the function F given by
   $$F(a + bx + cx^2) = c - ax + bx^2$$
   is a linear transformation of $\mathbb{R}_2[x]$.
   (c) [6 MARKS] You may assume that $C = \{1, x+1, (x+1)^2\}$ is a basis of $\mathbb{R}_2[x]$. Find the matrix of F, referred to coordinate system $1, x+1, (x+1)^2$ in the domain and, and to the standard coordinate system, $1, x, x^2$ in the target.
   (d) [4 MARKS] Determine a formula for $(F \circ F \circ F)(a + bx + cx^2)$
2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2, -3, -19, 6, 8)$ and $\mathbf{u}_2 = (-1, 2, 11, -4, -4)$.
3. Show that neither of these forms fully satisfies the conditions for an inner product:
   (a) [5 MARKS] On $\mathbb{R}^3$, where $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = x_1y_1 + x_3y_3$.
   (b) [5 MARKS] On $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = 4x_1y_1 - 7x_1y_2 - 7x_2y_1 + 9x_2y_2$.

7 Third Class Tests, (only for students who did not write one of the Second Class Tests)

7.1 Version 1

THIRD CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 15th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a)	1(b)	2(a)	2(b)	2(c)	2(d)		
/7	/3	/6	/10	/4	/10		
				RAW /40		/	

1. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f, g\rangle = \int_0^1 f(x)\, g(x)\, dx$, let $\mathbf{u}_1 = 3x$, $\mathbf{u}_2 = 3 + 2x$, $\mathbf{u}_3 = 4x^2$.
   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for $\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.
   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for $\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

2. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.
   (b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is a basis for $\mathcal{V}$.
   (c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.
   (d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

7.2 Version 2

THIRD CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 15th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

INSTRUCTIONS

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a) /6	1(b) /10	1(c) /4	1(d) /10	2(a) /7	2(b) /3		
				RAW /40		/	

1. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.

   (b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a basis for $\mathcal{V}$.

   (c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.

   (d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

2. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f, g\rangle = \int_0^1 f(x)\, g(x)\, dx$, let $\mathbf{u}_1 = -6x$, $\mathbf{u}_2 = 1 + 3x$, $\mathbf{u}_3 = 2x^2$.

   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.

   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.

7.3 Version 3

THIRD CLASS TEST: MATH 223 2006 01
LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 15th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

INSTRUCTIONS

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a) /7	1(b) /3	2(a) /6	2(b) /10	2(c) /4	2(d) /10		
				RAW /40		/	

1. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f, g\rangle = \int_0^1 f(x)\,g(x)\,dx$, let $\mathbf{u}_1 = 3x$, $\mathbf{u}_2 = 3 + 2x$, $\mathbf{u}_3 = 4x^2$.
   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for $\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.
   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for $\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

2. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22}\end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.
   (b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1\end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}$ is a basis for $\mathcal{V}$.
   (c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22}\end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.
   (d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0\end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1\end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

7.4 Version 4

THIRD CLASS TEST: MATH 223 2006 01

LINEAR ALGEBRA

EXAMINER: Professor W. G. Brown
DATE: Wednesday, 15th March, 2006. TIME: 10:40 – 11:25

FAMILY NAME:

GIVEN NAMES:

STUDENT NUMBER:

Instructions

1. This sheet must be handed in with your solutions.
2. All your writing — even rough work — must be handed in.
3. Calculators are not permitted.
4. *Your neighbour's version of this test may not be the same as yours.*
5. All of the test questions are printed on the back of this sheet.
6. Show all your work.

PLEASE DO NOT WRITE INSIDE THIS BOX

1(a) /6	1(b) /10	1(c) /4	1(d) /10	2(a) /7	2(b) /3		
				Raw /40		/	

1. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.

   (b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a basis for $\mathcal{V}$.

   (c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.

   (d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

2. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f, g\rangle = \int_0^1 f(x)\, g(x)\, dx$, let $\mathbf{u}_1 = -6x$, $\mathbf{u}_2 = 1 + 3x$, $\mathbf{u}_3 = 2x^2$.

   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.

   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.

8 Second Written Assignment

Distribution Date: Friday, March 17th, 2006; corrected/updated March 25th, 2006 and March 31st, 2006
Solutions are to be submitted by Wednesday, April 5th.
Note that Problems ##2,3 are based on Chapter 9, which we will begin on Monday, March 20th.

1. (Problem 8 on the Final Examination of MATH 223 2005 01)
   Let $A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & -2 & 3 & -1 \end{pmatrix}$. Find orthonormal bases of:
   (a) the null space of A
   (b) the row space of A
   (c) the image of the linear mapping given by A.

2. (Problem 1 on the Final Examination of MATH 223 2005 01) Let $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$.
   (a) Find all eigenvalues and corresponding eigenvectors.
   (b) Find a non-singular matrix P such that $D = P^{-1}AP$ is diagonal.
   (c) Find a matrix B such that $B^2 = A$.
   (d) Find $f(A)$, where $f(t) = t^4 - 3t^3 - 6t^2 + 7t + 3$.

3. (Problem 2 on the Final Examination of MATH 223 2004 09) For each of the following matrices A, find the characteristic polynomial and the minimum polynomial. Find the eigenvalues, and a basis for each eigenspace. Decide in each case whether the matrix is diagonalizable over the reals.

$$
\begin{aligned}
A &= \begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{pmatrix} \\
B &= \begin{pmatrix} 3 & 1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\
C &= \begin{pmatrix} 6 & 0 & 12 \\ 0 & 4 & 0 \\ -3 & 0 & -6 \end{pmatrix}
\end{aligned}
$$

9 Solutions to Problems on the Second Class Tests

Release Date: Sunday, 19 March, 2006
subject to correction

The test was written on Wednesday, March 8th, 2006.

9.1 Versions 1 and 3

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).

   (a) [2 MARKS] Define what is meant by a *basis* of $\mathbb{R}_2[x]$.

   (b) [8 MARKS] Prove that $C = \{1, x-1, (x-1)^2\}$ is a basis of $\mathbb{R}_2[x]$.

   (c) [6 MARKS] You may assume that differentiation is a linear transformation D of $\mathbb{R}_2[x]$. Find its matrix, referred to coordinate system $1, x-1, (x-1)^2$ in both domain and target.

   (d) [4 MARKS] Determine the coordinates of the polynomial $3+6x+x^2$ referred to the coordinate system $1, x-1, (x-1)^2$.

<u>Solution</u>: (This problem was modelled, in part, on Problem 3 of the Final Examination in MATH 231 2005 01.)

(a) A *basis* is a set of vectors which
   - are linearly independent; and
   - span $\mathbb{R}_2[x]$.

(b) i. **Proof of linear independence**: Let a, b, c be scalars. Then

$$
\begin{aligned}
& a + b(x-1) + c(x-1)^2 = 0 \\
\Rightarrow\ & (a-b+c)x^0 + (b-2c)x + cx^2 = 0x^0 + 0x^1 + 0x^2 \\
\Rightarrow\ & \left\{ \begin{array}{rrrcl} +a & -b & +c & = & 0 \\ & b & -2c & = & 0 \\ & & c & = & 0 \end{array} \right. \\
\Rightarrow\ & a = b = c = 0\,.
\end{aligned}
$$

ii. **Proof that the vectors span** $\mathbb{R}_2[x]$: Let $a+bx+cx^2$ be any polynomial in $\mathbb{R}_2[x]$. Then

$$\begin{array}{rcl} a+b+cx^2 & = & A+B(x-1)+C(x-1)^2 \\ \Leftrightarrow a+bx+cx^2 & = & (A-B+C)+(B-2C)x+Cx^2 \\ \Leftrightarrow & & \left\{\begin{array}{rrrcl} A & -B & +C & = & a \\ & B & -2C & = & b \\ & & C & = & c \end{array}\right. \\ \Leftrightarrow & & \left\{\begin{array}{rcc} A & = & a+b+c \\ B & = & b+2c \\ C & = & c \end{array}\right. . \end{array}$$

We have thus shown how every polynomial $a+bx+cx^2$ is expressible as a linear combination of the 3 linearly independent vectors $1, x-1, (x-1)^2$, i.e., that the three vectors *span* the space.

(c) We differentiate the three basis vectors, and express each of them as a linear combination of the basis vectors:

$$\begin{array}{rcl} D1 & = & 0 = 0\cdot 1 + 0\cdot(x-1) + 0\cdot(x-1)^2 \\ D(x-1) & = & 1 = 1\cdot 1 + 0\cdot(x-1) + 0\cdot(x-1)^2 \\ D\left((x-1)^2\right) & = & D(x^2-2x+1) \\ & = & 2x-2 = 2(x-1) = 0\cdot 1 + 2\cdot(x-1) + 0\cdot(x-1)^2 \end{array}$$

so the matrix is

$$[D]_{1,x-1,(x-1)^2}^{1,x-1,(x-1)^2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} .$$

(The reason why I differentiated $(x-1)^2$ naively — rather than applying the Chain Rule — is that we haven't proved that the Chain Rule is valid; no marks were deducted if you did use the Chain Rule.)

(d) I will apply the result of Part 1(b)ii above, with $a=3$, $b=6$, $c=1$:

$$[3+6x+x^2]_{1,x-1,(x-1)^2} = \begin{pmatrix} 3+6+1 \\ 6+2 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \\ 1 \end{pmatrix} .$$

Check: $10+8(x-1)+1(x-1)^2 = 10+8x-8+x^2-2x+1 = 3+6x+x^2$.

2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2,-1,8,7,-1)$ and $\mathbf{u}_2 = (1,-2,4,-1,-2)$.

Solution: (This problem is similar to [1, Exercise 7.66, p. 272].)

Let's denote the general vector of $\mathcal{W}$ by $\mathbf{v} = (x_1, x_2, x_3, x_4, x_5)$. The conditions that we wish to impose are

$$\mathbf{u}_1 \bullet \mathbf{v} = 0 = \mathbf{u}_2 \bullet \mathbf{v},$$

which are equivalent to the matrix equation

$$\begin{pmatrix} 2 & -1 & 8 & 7 & -1 \\ 1 & -2 & 4 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Row reduction of the coefficient matrix yields

$$\begin{pmatrix} 1 & 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 & 1 \end{pmatrix}.$$

We see that the general solution is, therefore,

$$\begin{pmatrix} -4x_3 - 5x_4 \\ -3x_4 - x_5 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -5 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus one basis of $\mathcal{W}$ is

$$\mathbf{u}_1 = \begin{pmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -5 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

3. [10 MARKS] Working carefully, and showing all your work, find the values of k so that the following is an inner product on $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$:

$$f(\mathbf{u}, \mathbf{v}) = x_1 y_1 - 4x_1 y_2 - 4x_2 y_1 + k^2 x_2 y_2\,.$$

Solution: (Problem related to [1, Exercise 7.58, p. 271].)

I_1: [3 MARKS] We have to prove linearity in the first argument. Let $\mathbf{w} = (z_1, z_2)$, and let a and c be real numbers. Then

$$a\mathbf{u} + c\mathbf{w} = a(x_1, x_2) + b(z_1, z_2) = (ax_1 + cz_1, ax_2 + cz_2)\,.$$

$$\begin{aligned}
& f(a\mathbf{u} + c\mathbf{w}, \mathbf{v}) \\
= \;& (ax_1 + cz_1)y_1 - 4(ax_1 + cz_1)y_2 - 4(ax_2 + cz_2)y_1 + k^2(ax_2 + cz_2)y_2 \\
= \;& \big((ax_1)y_1 - 4(ax_1)y_2 - 4(ax_2)y_1 + k^2(ax_2)y_2\big) \\
& \quad + \big((cz_1)y_1 - 4(cz_1)y_2 - 4(cz_2)y_1 + k^2(cz_2)y_2\big) \\
= \;& a\left(x_1y_1 - 4x_1y_2 - 4x_2y_1 + k^2x_2y_2\right) \\
& \quad + c\left(z_1y_1 - 4z_1y_2 - 4z_2y_1 + k^2z_2y_2\right) \\
= \;& af(\mathbf{u}, \mathbf{v}) + cf(\mathbf{w}, \mathbf{v})
\end{aligned}$$

so property I_1 holds for *all* values of k.

I_2: [2 MARKS]

$$\begin{aligned}
f(\mathbf{u}, \mathbf{v}) &= x_1y_1 - 4x_1y_2 - 4x_2y_1 + k^2x_2y_2 \\
&= y_1x_1 - 4y_1x_2 - 4y_2x_1 + k^2y_2x_2 \\
&= f(\mathbf{v}, \mathbf{u})
\end{aligned}$$

so property I_2 holds for *all* values of k.

I_3: [5 MARKS]

$$\begin{aligned}
f(\mathbf{u}, \mathbf{u}) &= x_1x_1 - 4x_1x_2 - 4x_2x_1 + k^2x_2x_2 \\
&= x_1^2 - 8x_1x_2 + kx_2^2 \\
&= (x_1 - 4x_2)^2 + (k^2 - 16)x_2^2\,.
\end{aligned}$$

If $k^2 < 16$, it is possible to find values of x_1, x_2 which will make this sum negative: e.g., take $x_2 = 1$, $x_1 = 4$. Thus it is certainly necessary that $|k| \geq 4$. Even if $k = 4$, there are non-zero vectors for which the proposed inner product with itself will be zero, e.g., the vector $(-4, -1)$. However, if we take $|k| > 4$, then the inner product of a vector with itself will be a sum of squares that cannot be negative; and it will be 0 precisely when both $x_2 = 0$ and $x_1 - 4x_2 = 0$, i.e., for the vector $\mathbf{0}$ alone.

9.2 Versions 2 and 4

1. Let $B = \{1, x, x^2\}$ be the "standard" basis of the vector space $\mathbb{R}_2[x]$ (or $\mathbf{P}_2(x)$).

(a) [2 MARKS] Define what is meant by a linear transformation F from $\mathbb{R}_2[x]$ to $\mathbb{R}_2[x]$.

(b) [8 MARKS] Prove that the function F given by

$$F(a+bx+cx^2) = c - ax + bx^2$$

is a linear transformation of $\mathbb{R}_2[x]$.

(c) [6 MARKS] You may assume that $C = \{1, x+1, (x+1)^2\}$ is a basis of $\mathbb{R}_2[x]$. Find the matrix of F, referred to coordinate system $1, x+1, (x+1)^2$ in the domain and, and to the standard coordinate system, $1, x, x^2$ in the target.

(d) [4 MARKS] Determine a formula for $(F \circ F \circ F)(a+bx+cx^2)$

Solution: (This problem was modelled, in part, on Problem 3 of the Final Examination in MATH 231 2005 01.)

(a) F is defined to be linear if it has the following two properties:

Commutes with +: Suppose that $a+bx+cx^2$ and $d+ex+fx^2$ are any vectors in $\mathbb{R}_2[x]$. Then

$$F((a+bx+cx^2)+(d+ex+fx^2)) = F(a+bx+cx^2)+F(d+ex+fx^2)\,.$$

Commutes with scalar multiplication: Suppose that $a+bx+cx^2$ is any polynomial in $\mathbb{R}_2[x]$, and k is any scalar. Then

$$F(k(a+bx+cx^2)) = kF(a+bx+cx^2)\,.$$

(b) i. **Commutes with** +: Suppose that $a+bx+cx^2$ and $d+ex+fx^2$ are any vectors in $\mathbb{R}_2[x]$. Then

$$\begin{aligned}
& F((a+bx+cx^2)+(d+ex+fx^2)) \\
= \; & F((a+d)+(b+e)x+(c+f)x^2) \\
= \; & (c+f)-(a+d)x+(b+e)x^2 \quad \text{by definition of } F \\
= \; & (c-ax+bx^2)+(f-dx+ex^2) \\
& \text{by definition of + in } \mathbb{R}_2[x] \\
= \; & F(a+bx+cx^2)+F(d+ex+fx^2)
\end{aligned}$$

□

ii. **Commutes with scalar multiplication:**: Suppose that $a + bx + cx^2$ is any polynomial in $\mathbb{R}_2[x]$, and k is any scalar. Then

$$\begin{aligned} F(k(a+bx+cx^2)) &= F((ka)+(kb)x+(kc)x^2) \\ &= (kc)-(ka)x+(kb)x^2 \quad \text{by definition of } F \\ &= k(c-ax+bx^2) \\ &\quad \text{by definition of scalar multiplication in } \mathbb{R}_2[x] \\ &= kF(a+bx+cx^2) \end{aligned}$$

□

(c)

$$\begin{aligned} F(1) &= F(1+0x+0x^2) = 0-1x+0x^2 \\ F(x+1) &= F(1+1x+0x^2) = 0-1x+1x^2 \\ F\left((x+1)^2\right) &= F(1+2x+x^2) \\ &= 1-1x+2x^2\,. \end{aligned}$$

The matrix is

$$[F]_{1,x,x^2}^{1,x+1,(x+1)^2} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

(d)

$$\begin{aligned} F(a+bx+cx^2) &= c-ax+bx^2 \\ F\left(F(a+bx+cx^2)\right) &= F(c-ax+bx^2) \\ &= b-cx-ax^2 \\ F\left(F\left(F(a+bx+cx^2)\right)\right) &= F\left(F(c-ax+bx^2)\right) \\ &= F(b-cx-ax^2) \\ &= -a-bx-c^2 \\ \text{i.e.,}\quad (F\circ F\circ F)(a+bx+cx^2) &= -a-bx-cx^2 \end{aligned}$$

2. [10 MARKS] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal (using the standard inner product) to both vectors $\mathbf{u}_1 = (2,-3,-19,6,8)$ and $\mathbf{u}_2 = (-1,2,11,-4,-4)$.

   Solution: (This problem is similar to [1, Exercise 7.66, p. 272].)

   Let's denote the general vector of $\mathcal{W}$ by $\mathbf{v} = (x_1,x_2,x_3,x_4,x_5)$. The conditions that we wish to impose are

$$\mathbf{u}_1 \bullet \mathbf{v} = 0 = \mathbf{u}_2 \bullet \mathbf{v},$$

which are equivalent to the matrix equation

$$\begin{pmatrix} 2 & -3 & -19 & 6 & 8 \\ -1 & 2 & 11 & -4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

Row reduction of the coefficient matrix yields

$$\begin{pmatrix} 1 & 0 & -5 & 0 & -4 \\ 0 & 1 & 3 & -2 & 0 \end{pmatrix} .$$

We see that the general solution is, therefore,

$$\begin{pmatrix} 5x_3 - 4x_5 \\ -3x_3 + 2x_4 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} 5 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

Thus one basis of $\mathcal{W}$ is

$$\mathbf{u}_1 = \begin{pmatrix} 5 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

3. Show that neither of these forms fully satisfies the conditions for an inner product:

   (a) [5 MARKS] On $\mathbb{R}^3$, where $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + x_3 y_3$.

   (b) [5 MARKS] On $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$, the function $\langle \mathbf{u}, \mathbf{v} \rangle = 4x_1 y_1 - 7x_1 y_2 - 7x_2 y_1 + 9x_2 y_2$.

   Solution: (Problem related to [1, Exercise 7.60, p. 271].)

   (a) There are non-$\mathbf{0}$ vectors $\mathbf{u}$ such that $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, for example, the vector $(0, 1, 0)$. This contravenes condition I_3.

   (b) Computing $\langle \mathbf{u}, \mathbf{u} \rangle = 4x_1^2 - 14x_1 x_2 + 9x_2^2 = \left(2x_1 - \frac{7}{2}x_2\right)^2 - \frac{13}{4}x_2^2$, we find there are vectors whose inner product with themselves will be negative. For example, take $x_1 = \frac{7}{2}$, $x_2 = 2$. This contravenes I_3.

10 Solutions to Problems on the Third Class Tests

Release Date: Sunday, 19 March, 2006
subject to correction

The test was written on Wednesday, March 15th, 2006, only by students who did not write the 2nd class test a week earlier.

10.1 Versions 1 and 3

1. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f,g\rangle = \int_0^1 f(x)\,g(x)\,dx$, let $\mathbf{u}_1 = 3x$, $\mathbf{u}_2 = 3+2x$, $\mathbf{u}_3 = 4x^2$.

   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for span$\{\mathbf{u}_1,\mathbf{u}_2\}$.

   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for span$\{\mathbf{u}_1,\mathbf{u}_2\}$.

   Solution:

   (a)

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 = 3x \\
\langle \mathbf{v}_1,\mathbf{v}_1\rangle &= \int_0^1 (3x)(3x)\,dx = \left[3x^3\right]_0^1 = 3 \\
\langle \mathbf{u}_2,\mathbf{v}_1\rangle &= \int_0^1 (3+2x)(3x)\,dx = \left[\frac{9x^2}{2}+2x^3\right]_0^1 = \frac{13}{2} \\
\mathbf{v_2} &= \mathbf{u}_2 - \frac{\langle \mathbf{u}_2,\mathbf{v}_1\rangle}{\langle \mathbf{v}_1,\mathbf{v}_1\rangle}\mathbf{v}_1 \\
&= 3+2x-\frac{\frac{13}{2}}{3}(3x) = 3-\frac{9}{2}x
\end{aligned}$$

   (b) We saw earlier that $\|\mathbf{v}_1\|^2 = 3$; hence we may normalize this vector, to obtain a unit vector, $\mathbf{w}_1 = \frac{1}{\sqrt{3}}3x = \sqrt{3}\,x$; we could equally well have taken $\mathbf{w}_1 =$

$-\sqrt{3}\,x$. The vector $\mathbf{v}_2$ may be normalized as follows:

$$\begin{aligned}
\langle \mathbf{v}_2, \mathbf{v}_2 \rangle &= \int_0^1 \left(3 - \frac{9}{2}x\right)^2 dx \\
&= -\frac{2}{27}\left[\left(3 - \frac{9}{2}x\right)^3\right]_0^1 = \frac{9}{4} \\
\Rightarrow \|\mathbf{v}_2\| &= \frac{3}{2} \\
\Rightarrow \mathbf{w}_2 &= 2 - 3x
\end{aligned}$$

2. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.

   (b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is a basis for $\mathcal{V}$.

   (c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.

   (d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

Solution:

(a) i. **Closure under vector addition**: Given two matrices in $\mathcal{V}$, $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, and $\begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix}$, their matrix sum is

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ 0 & a_{22}+b_{22} \end{pmatrix},$$

which is again a matrix in $\mathcal{V}$.

ii. **Closure under scalar multiplication**: Given a matrix in $\mathcal{V}$, and any scalar k, the product is

$$k\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} \\ 0 & ka_{22} \end{pmatrix},$$

which is again a matrix in $\mathcal{V}$.

iii. **Contains 0**: The zero matrix is upper triangular, since the entry in position $(2,1)$ is 0 (as are the other 3 entries). Hence $\mathbf{0} \in \mathcal{V}$.

(b) We have to prove that the 3 vectors are linearly independent, and that they span $\mathcal{V}$.

i. **Linear independence**: Let k_1, k_2, k_3 be scalars. Then

$$\begin{aligned}
& k_1\mathbf{f}_1 + k_2\mathbf{f}_2 + k_3\mathbf{f}_3 = \mathbf{0} \\
\Rightarrow\ & k_1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + k_2\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + k_3\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} k_1 - k_2 & k_3 \\ 0 & k_1 + k_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

ii. **Spanning set**: We have to show that any matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ is expressible as a linear combination of the alleged basis vectors. That is, we need to solve the equation

$$k_1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + k_2\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + k_3\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}.$$

This equation is equivalent to the matrix equation

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{22} \end{pmatrix},$$

which we solve to yield

$$\begin{aligned} k_1 &= \frac{a_{11}+a_{22}}{2} \\ k_2 &= \frac{-a_{11}+a_{22}}{2} \\ k_3 &= a_{12} \end{aligned}$$

The existence of such a solution implies that the vectors span $\mathcal{V}$.

(c) The last calculation shows that

$$\left[\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} = \begin{pmatrix} \frac{a_{11}+a_{22}}{2} \\ \frac{-a_{11}+a_{22}}{2} \\ a_{12} \end{pmatrix}.$$

(d)

$$\begin{aligned} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{1+0}{2} \\ \frac{-1+0}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \\ \left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{0+0}{2} \\ \frac{-0+0}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{0+1}{2} \\ \frac{-0+1}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \\ [1_{\mathcal{V}}]^{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3}_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

10.2 Versions 2 and 4

1. (a) [6 MARKS] $\mathcal{V}$ is defined to be the subset of the space $\mathbb{R}_{2,2}$ of 2×2 real matrices consisting of "upper triangular" matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, where a_{11}, a_{12}, and a_{22} are real numbers. Carefully prove that $\mathcal{V}$ is a subspace of $\mathbb{R}_{2,2}$.

(b) [10 MARKS] Prove carefully that $\mathbf{f}_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a basis for $\mathcal{V}$.

(c) [4 MARKS] Determine the coordinates of the matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ in the coordinate system $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$.

(d) [10 MARKS] You may assume that the matrices $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ form a basis for $\mathcal{V}$. Determine the matrix of the linear transformation which is the identity function $1_{\mathcal{V}}$ if the coordinate system for the domain is $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and and the coordinate system for the target is $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

Solution:

(a) i. **Closure under vector addition**: Given two matrices in $\mathcal{V}$, $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, and $\begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix}$, their matrix sum is

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ 0 & a_{22}+b_{22} \end{pmatrix},$$

which is again a matrix in $\mathcal{V}$.

ii. **Closure under scalar multiplication**: Given a matrix in $\mathcal{V}$, and any scalar k, the product is

$$k\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} \\ 0 & ka_{22} \end{pmatrix},$$

which is again a matrix in $\mathcal{V}$.

iii. **Contains 0**: The zero matrix is upper triangular, since the entry in position $(2, 1)$ is 0 (as are the other 3 entries). Hence $\mathbf{0} \in \mathcal{V}$.

(b) We have to prove that the 3 vectors are linearly independent, and that they span $\mathcal{V}$.

i. **Linear independence**: Let k_1, k_2, k_3 be scalars. Then

$$\begin{aligned}
& k_1\mathbf{f}_1 + k_2\mathbf{f}_2 + k_3\mathbf{f}_3 = \mathbf{0} \\
\Rightarrow\ & k_1\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + k_2\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + k_3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} k_3 - k_1 & k_2 \\ 0 & k_3 + k_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\Rightarrow\ & \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

ii. **Spanning set**: We have to show that any matrix $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$ is expressible as a linear combination of the alleged basis vectors. That is, we need to solve the equation

$$k_1\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + k_2\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + k_3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}.$$

This equation is equivalent to the matrix equation

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{22} \end{pmatrix},$$

which we solve to yield

$$\begin{aligned}
k_1 &= \frac{-a_{11} + a_{22}}{2} \\
k_2 &= a_{12} \\
k_3 &= \frac{a_{11} + a_{22}}{2}
\end{aligned}$$

The existence of such a solution implies that the vectors span $\mathcal{V}$.

(c) The last calculation shows that

$$\left[\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} = \begin{pmatrix} \frac{-a_{11}+a_{22}}{2} \\ a_{12} \\ \frac{a_{11}+a_{22}}{2} \end{pmatrix}.$$

(d)

$$\begin{aligned}
\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{-1+0}{2} \\ 0 \\ \frac{1+0}{2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \\
\left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{-0+0}{2} \\ 1 \\ \frac{0+0}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
\left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} \frac{-0+1}{2} \\ 0 \\ \frac{0+1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \\
[1_{\mathcal{V}}]^{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3}_{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3} &= \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} .
\end{aligned}$$

2. In the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2, with inner product given by $\langle f, g\rangle = \int_0^1 f(x)\,g(x)\,dx$, let $\mathbf{u}_1 = -6x$, $\mathbf{u}_2 = 1+3x$, $\mathbf{u}_3 = 2x^2$.
   (a) [7 MARKS] Showing all your work in detail, apply the Gram-Schmidt process to the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, **in the given order**, to determine an orthogonal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.
   (b) [3 MARKS] From your orthogonal basis just calculated, determine an orthonormal basis, $\mathbf{w}_1$, $\mathbf{w}_2$ for span$\{\mathbf{u}_1, \mathbf{u}_2\}$.

Solution:

(a)

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 = -6x \\
\langle \mathbf{v}_1, \mathbf{v}_1\rangle &= \int_0^1 (-6x)(-6x)\,dx = \left[12x^3\right]_0^1 = 12 \\
\langle \mathbf{u}_2, \mathbf{v}_1\rangle &= \int_0^1 (1+3x)(-6x)\,dx = \left[-3x^2 - 6x^3\right]_0^1 = -9 \\
\mathbf{v_2} &= \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1 \\
&= 1 + 3x - \frac{-9}{12}(-6x) = 1 - \frac{3}{2}x
\end{aligned}$$

(b) We saw earlier that $\|\mathbf{v}_1\|^2 = 12$; hence we may normalize this vector, to obtain a unit vector, $\mathbf{w}_1 = \frac{1}{2\sqrt{3}}(-6)x = -\sqrt{3}\,x$; we could equally well have taken $\mathbf{w}_1 = +\sqrt{3}\,x$. The vector $\mathbf{v}_2$ may be normalized as follows:

$$\begin{aligned}
\langle \mathbf{v}_2, \mathbf{v}_2 \rangle &= \int_0^1 \left(1 - \frac{3}{2}x\right)^2 dx \\
&= -\frac{2}{9}\left[\left(1 - \frac{3}{2}x\right)^3\right]_0^1 = \frac{1}{4} \\
\Rightarrow \|\mathbf{v}_2\| &= \frac{1}{2} \\
\Rightarrow \mathbf{w}_2 &= 2 - 3x
\end{aligned}$$

11 Solutions, Second Written Assignment

Distribution Date: Mounted on the Web on Tuesday, April 11th, 2006
subject to correction

1. (Problem 8 on the Final Examination of MATH 223 2005 01)

   Let $A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & -2 & 3 & -1 \end{pmatrix}$. Find orthonormal bases of:

   (a) the null space of A

   (b) the row space of A

   (c) the image of the linear mapping given by A.

   <u>Solution</u>:

   (a) Under successive row operations the matrix reduces as follows:

$$A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & -2 & 3 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -2 & 1 \end{pmatrix}.$$

   This last is the matrix of coefficients of the system of equations

$$\begin{aligned} +1x - 1y + 1z + 0w &= 0 \\ +0x + 1y - 2z + 1w &= 0 \end{aligned}$$

   which is equivalent to

$$\begin{aligned} x &= z - w \\ y &= 2z - w\,, \end{aligned}$$

   so the solutions of the system are — i.e., the general vector in the null space is

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} z - w \\ 2z - w \\ z \\ w \end{pmatrix} = \begin{pmatrix} z - w \\ 2z - w \\ z \\ w \end{pmatrix} = z \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + w \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Thus a basis for the null space is the vectors

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u_2} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

We will apply the Gram-Schmidt process to them in the given order. In that process $\mathbf{u}_1$ is not changed. $\mathbf{u}_2$ is replaced by

$$\mathbf{u}_2 - \frac{-3}{6}\mathbf{u}_1 = \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

These vectors have been orthogonalized, but not normalized. The length of the first is $\sqrt{1+4+1}$, while the length of the second is $\frac{1}{2} \cdot \sqrt{1+0+1+4} = \frac{\sqrt{6}}{2}$. Hence one orthonormal basis for the null space of A is

$$\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \qquad \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

(b) We have shown that the dimension of the row space of A is 2; hence one basis is the original rows of A; another is the rows of the row reduced matrix we found. Applying the Gram-Schmidt process to the rows of A, i.e., to $\mathbf{v}_1 = (1, -1, 1, 0)$ and $\mathbf{v}_2 = (1, -2, 3, -1)$ in that order, we have to replace $\mathbf{v}_2$ by $\mathbf{v}_2 - \frac{1+2+3+0}{1+1+1+0}\mathbf{v}_1 = (1, -2, 3, -1) - 2(1, -1, 1, 0) = (-1, 0, 1, -1)$. Thus one orthogonal basis for the row space is $(1, -1, 1, 0)$, $(-1, 0, 1, -1)$. Both of these vectors have length $\sqrt{3}$. Hence one orthonormal basis is

$$\left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right), \left(-\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right).$$

(c) The image is generated by the columns of A. Under *column* reduction we obtain $A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & -2 & 3 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ whose first two columns span the image and are already orthonormal: hence they are an orthonormal basis.

2. (Problem 1 on the Final Examination of MATH 223 2005 01) Let $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$.
   (a) Find all eigenvalues and corresponding eigenvectors.
   (b) Find a non-singular matrix P such that $D = P^{-1}AP$ is diagonal.
   (c) Find a matrix B such that $B^2 = A$.
   (d) Find $f(A)$, where $f(t) = t^4 - 3t^3 - 6t^2 + 7t + 3$.

Solution:

(a) The characteristic polynomial is $(t-2)(t-3) - (-1)(-2) = t^2 - 5t + 4 = (t-1)(t-4)$. Thus the eigenvalues are 1 and 4.

Eigenvalue $t = 1$: We solve the system of homogeneous equations with coefficient matrix $\begin{pmatrix} -1 & -2 \\ -1 & -2 \end{pmatrix}$ to obtain one basis vector for its solution space: $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Eigenvalue $t = 4$: We solve the system of homogeneous equations with coefficient matrix $\begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix}$ to obtain one basis vector for its solution space: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Thus, for the eigenvalues $1, 4$ corresponding eigenvectors are $(2, -1)$ and $(1, 1)$ (or any non-zero scalar multiples of each of them).

(b) We form a matrix with the eigenvectors as columns, and observe that

$$A \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Hence

$$\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

so

$$P = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$$

is one matrix which will diagonalize A.

(c) The eigenvalues of A are positive;

$$A = P \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} P^{-1}.$$

Take

$$B = P\begin{pmatrix} \sqrt{1} & 0 \\ 0 & \sqrt{4} \end{pmatrix} P^{-1} = \frac{1}{3}\begin{pmatrix} 4 & 2 \\ 1 & 5 \end{pmatrix}.$$

Verify that $B^2 = P\begin{pmatrix} \sqrt{1} & 0 \\ 0 & \sqrt{4} \end{pmatrix}^2 P^{-1} = P\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} P^{-1} = A$.

(d) Of course, we could just compute the powers and combine them with the given scalar weights. But that is not what the author intended. Since $D = P^{-1}AP$, $A = PDP^{-1}$; and $A^n = PD^nP^{-1}$ after cancellation of matrices with their inverses. Hence

$$\begin{aligned} f(A) &= A^4 - 3A^3 - 6A^2 + 7A + 3I \\ &= P(D^4 - 3D^3 - 6D^2 + 7D + 3I)P^{-1} \\ &= P\begin{pmatrix} 1 - 3 - 6 + 7 + 3 & 0 \\ 0 & 4^4 - 3 \cdot 4^3 - 6 \cdot 4^2 + 7 \cdot 4 + 3 \end{pmatrix} P^{-1} \\ &= P\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}. \end{aligned}$$

Further simplifications are possible. If we divide the characteristic polynomial into $f(t)$, we obtain

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

Instead of evaluating the polynomial at an eigenvalue, it suffices to evaluate the remainder, $-t + 3$ there.

$$\begin{aligned} & \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}^{-1} \\ = & \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \\ = & \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}. \end{aligned}$$

3. (Problem 2 on the Final Examination of MATH 223 2004 09) For each of the following matrices A, find the characteristic polynomial and the minimum polynomial. Find the eigenvalues, and a basis for each eigenspace. Decide in each case whether

the matrix is diagonalizable over the reals.

$$\begin{aligned} A &= \begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{pmatrix} \\ B &= \begin{pmatrix} 3 & 1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\ C &= \begin{pmatrix} 6 & 0 & 12 \\ 0 & 4 & 0 \\ -3 & 0 & -6 \end{pmatrix} \end{aligned}$$

<u>Solution</u>:

(a) The characteristic polynomial is $\det(tI_3 - A)$, which factors as $(t-1)^2(t-5)$, so the eigenvalues are 1, of multiplicity 2, and 5, of multiplicity 1.

Eigenvalue $t = -1$: Setting $t = -1$ we find that the matrix row reduces to

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

corresponding to the single equation, $x = -2y - z$. Hence the solutions to the corresponding homogeneous system are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2y - z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

The eigenspace has dimension 2: two linearly independent eigenvectors for this eigenvalue are $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

Eigenvalue $t = 5$: Setting $t = 5$ we find that the matrix row reduces to

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix};$$

the general solution of the system of homogeneous equations of which this is the coefficient matrix is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ 2z \\ z \end{pmatrix} = z \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

so one eigenvector for this eigenvalue is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. If we apply the matrix A to a matrix

$$P = \begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix},$$

whose columns are these three eigenvectors, we obtain a matrix whose columns are obtained from P by multiplying respectively by the corresponding eigenvalues, and this is the product of a diagonal matrix with P:

$$\begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix},$$

implying that

$$\begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} -2 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix},$$

and so A is diagonalizable over $\mathbb{R}$. By definition, the minimum polynomial is that monic polynomial of lowest degree satisfied by the given matrix. We know that it must have the same distinct roots as the characteristic polynomial, although possibly not to the same multiplicities. In the present case that implies that either the minimum polynomial coincides with the characteristic polynomial, or that it is the polynomial $(t+1)(t-5) = t^2 - 4t - 5$. There is additional information available in this case which could help us, but I will first attack the problem on a naive level: we can simply investigate whether A satisfies this polynomial:

$$\begin{aligned} A^2 - 4A - 5I &= \begin{pmatrix} 5 & 8 & 4 \\ 8 & 17 & 8 \\ 4 & 8 & 5 \end{pmatrix} - 4\begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Students who were writing the examination may have known more about this matrix: it is a symmetric matrix, and, as such, is known to be diagonalizable, which you were able to prove "by brute force". Being diagonalizable, also implies that the simple product of linear factors would, in this case, be the minimal polynomial. But you are not expected to know that for your examination.

(b) In this case $\det(tI - B)$ factorizes as $(t-2)(t^2 - 6t + 10)$, i.e., $(t-2)((t-3)^2 + 1)$, where the quadratic factor is irreducible over $\mathbb{R}$. There is only one real eigenvalue. $t = 2$. The matrix $2I - B$ row reduces to $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, the matrix of coefficients of a homogeneous system whose solutions are generated by $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. This matrix cannot be diagonalized over the reals. (Since the complex roots of the quadratic are distinct, this matrix can be diagonalized over $\mathbb{C}$.)

In this case there are no repeated factors of the characteristic polynomial: the minimum polynomial coincides with the characteristic polynomial.

(c) In this case $\det(tI - C)$ factorizes into $(t-4)t^2$, so the eigenvalues are 0, of multiplicity 2, and 4, of multiplicity 1. Corresponding to eigenvalue 4 there can be only one eigenvector, and we find one to be $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Corresponding to eigenvalue 0 we find the matrix $0I - C$ row reduces to $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and the corresponding homogeneous system has a basis consisting of the single vector $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$. This matrix is not diagonalizable.

In this case one could again consider the polynomial $t(t-4)$ — the product of the linear factors of the characteristic polynomial, each reduced to multiplicity 1 — as a candidate for minimum polynomial. We have not studied the theory of minimum polynomials, so you would not know that the polynomial $t(t-4)$ will *not* be satisfied by the matrix. Let's do the calculations:

$$C(C - 4I) = \begin{pmatrix} 6 & 0 & 12 \\ 0 & 4 & 0 \\ -3 & 0 & -6 \end{pmatrix} \begin{pmatrix} 2 & 0 & 12 \\ 0 & 0 & 0 \\ -3 & 0 & 10 \end{pmatrix} = \begin{pmatrix} -24 & 0 & 48 \\ 0 & 0 & 0 \\ 12 & 0 & 24 \end{pmatrix}$$

which is not the zero matrix. Hence the minimum polynomial coincides with the characteristic polynomial.

12 Timetable for MATH 223 2006 01

Distribution Date: updated Sunday, March 19th, 2006
(Subject to correction and change.)
Section numbers refer to the text-book.

The following schedule is currently under revision; once a stable version is available, it will be mounted in this location in the Information document, and will be visible on WebCT.

MONDAY	WEDNESDAY	FRIDAY
JANUARY		
	04 **Review of §§1.1-1.3**	06 **Review of §§1.4-1.7, 2.1-2.5**
09 **Review of §§2.6-2.10, etc.**	11 **Review of §§3.1-3.2**	13 **§§3.3, 3.4** WW_1
Course changes must be completed on MINERVA by Jan. 17, 2006		
16 **§§3.5-3.7, 3.12**	18 **§§3.12**	20 **§§, 4.1-4.3** WW_2
Deadline for withdrawal with fee refund = Jan. 22, 2006		
23 **§§4.3, 4.5**	25 **§§4.4, 4.6**	27 **§§4.7, 4.8** WW_3
Verification Period: January 30 – February 03, 2006		
30 **§§4.9, 3.8, 3.11**		
FEBRUARY		
	01 **FIRST TEST**	03 **X** WW_4
Deadline for withdrawal (with W) from course via MINERVA = Feb. 12, 2006		
06 **§4.11**	08	10 **§§5.1 – 5.3** WW_5
13 **§§5.4, 5.5**	15 **§§6.1, 6.2**	17 **§§6.1–6.5** WW_6
Study Break: February 20–24, 2006 No lectures, no office hours!		
20 **NO LECTURE**	22 **NO LECTURE**	24 **NO LECTURE**
27 **§§7.1-7.3**		

Notation:

WW_n = Regular **WeBWorK** Assignment WW_n due at midnight on Friday this week (Assignment WW_0 introduces **WeBWorK**, and does not count in your grade.)

Ⓡ = Read Only

X = reserved for eXpansion or review

Section numbers refer to the text-book.

MONDAY		WEDNESDAY		FRIDAY	
MARCH					
		01	**§7.3** *Written Assignment Due*	03	**§§7.4-7.6** WW_7
06	**§7.6**	08	**SECOND TEST**	10	**§§7.6-7.8**
13	**Chapter 8**	15	**THIRD TEST (replaces 2nd)**	17	**Chapter 8, §§6.4, 9.1** WW_8, WW_9
20	**Chapter 9**	22	**Chapter 9**	24	**Chapter 9** WW_{10}
27		29	§	31	WW_{11}
APRIL					
03	§	05	§	07	§ WW_{12}
10		13			

Notation:

WW_n = **WeBWorK** Assignment WW_n due at midnight on Friday this week (Assignment WW_0 introduces **WeBWorK**, and does not count in your grade.)
Ⓡ = Read Only
X = reserved for eXpansion or review
Section numbers refer to the text-book.

13 References

[1] S. Lipschutz and M. Lipson, Schaum's Outline of Theory and Problems of LINEAR ALGEBRA, Third Edition, McGraw-Hill (2001). ISBN 0-07-136200-2.

[2] J. Stewart, Single Varialbe Calculus, Early Transcendentals, 5th Edition. Thomson*Brooks/Cole (2003) ISBN0-534-39330-6.

[3] B. Noble and J. W. Daniel, *Applied Linear Algebra, Third Edition (1988)*. Prentice-Hall, Englewood Cliffs, NJ 07632. ISBN 0–13–041260–0.

A Supplementary Lecture Notes

A.1 Supplementary Notes for the Lecture of Wednesday, January 4th, 2006

Release Date: Wednesday, January 4th, 2006
Subject to further revision

A.1.1 These notes

1. You can expect these notes to be less detailed once we complete the review of prerequisite material.

2. These notes will contain some of the material that I propose to discuss in the lectures, and also some material that will not make it to the lectures. I will be following the order of topics in the textbook closely, sometimes explaining statements that I find require some elaboration. Much of the class time will be spent in discussing problems from the textbook, for which I will be including sketches of solutions wherever possible.

3. I am preparing these notes freshly this year: this means that there will likely be some misprints or other errors that I will not discover before posting the notes. If you find statements that appear troubling, please ask about them or send me an e-mail message. The notes will be progressively corrected as required.

4. The intention is that students read the textbook carefully. When I list definitions and theorems is it intended as a sketch prepared in advance of part of my lectures. I will not always list the full definition: sometimes, particularly when I use the symbol ¡¿ at the beginning of the item, I will only indicate that a definition is needed, without actually stating a definition. There the intention is that either the definition is obvious, or that you are expected to look it up in the textbook, or both.

5. Unlike your textbook, I usually use large parentheses (=round brackets ()) rather than brackets (=square brackets []) in denoting matrices.

6. I try to denote vectors — both row and column vectors — by **boldface** type; matrices will usually be denoted by non-boldface type.

7. I will denote the real numbers, rational numbers, complex numbers, respectively by $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, and general fields usually by $\mathbb{K}$ or $\mathbb{F}$; the textbook uses **boldface** type for these symbols — another commonly used notation.

CHAPTER 1 — Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$. Spatial Vectors

Students are expected to be thoroughly familiar with the material in this chapter, by virtue of the prerequisite of "MATH 133 or equivalent". In some cases your familiarity will be only for dimensions $n = 2, 3$; you may be asked to generalize known results to higher dimensions.

A.1.2 §1.1 Introduction

Definition A.1 (This is a preliminary definition, and will be modified in Chapter 4.)

1. Later in the course I will be more specific about the structure called a *field*. For the present you may think of a *field* as being one of $\mathbb{R}$ (the field of real numbers) or $\mathbb{C}$ (the field of complex numbers).

2. Individual elements of whichever field is under consideration, usually $\mathbb{R}$ or $\mathbb{C}$, are called *scalars*.

3. A vector is a *linear array* consisting of a number of scalars written in either a row or a column; i.e., either as
$$\mathbf{w} = (w_1, w_2, \ldots, w_n) \quad \text{or} \quad \mathbf{w} = \begin{pmatrix} w_1 & w_2 & \ldots & w_n \end{pmatrix}$$
(a *row* vector); or
$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$
(a *column* vector). I will not distinguish between a column vector and a matrix having just 1 column; I will write row vectors in two ways, either with its elements separated by commas, as shown here, or as a matrix having 1 row.

4. If $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ or
$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

are two given vectors, we will write $\mathbf{a} = \mathbf{b}$ if and only if

$$\begin{aligned} m &= n\,, \\ \text{and } a_i &= b_i \quad \text{for } i = 1, 2, \ldots, n. \end{aligned}$$

A.1.3 §1.2 Vectors in $\mathbb{R}^n$

Definition A.2

1. The set of all ordered[13] n-tuples of real numbers constitutes the members, called *points*, of $\mathbb{R}^n$, *real n-space*.
2. The vector whose *components* (or *entries*) are all 0 is called the *zero-vector*, and denoted by $\mathbf{0}$.
3. A vector is said to be *non-zero* if not all of its components are 0.

A.1.4 Default Notation for Points in $\mathbb{R}^n$

While we do need to be able to talk about row- and column vectors when discussing matrices, the formalism for mappings to be introduced in connection with [1, Chapters 5,6] is much cleaner if we insist that points in $\mathbb{R}^n$ should always be represented by *column* vectors. This will particularly be the case when we are interpreting points in $\mathbb{R}^n$ as *coordinate vectors* for points in some real vector space, referred to a fixed coordinate system. Occasionally I will use row vectors, either from forgetfulness or to save space; but, when we come to do our calculations, we will always take such vectors to be column vectors, i.e., matrices with one column.

A.1.5 §1.3 Vector addition and scalar multiplication

In this section and the next vectors are usually written in terms of row vectors; the intention is that definitions and theorems should apply also when all vectors are written as column vectors.

Definition A.3 Let

$$\begin{aligned} \mathbf{a} &= (a_1, a_2, \ldots, a_n) \\ \mathbf{b} &= (b_1, b_2, \ldots, b_n) \end{aligned}$$

and let $k \in \mathbb{R}$ be an arbitrary scalar.

[13] I may eventually suppress the word *ordered*; however, it should be understood that an *unordered* n-tuple is just a set of real numbers. Worse than that, while we permit some of the entries in the ordered n-tuple to be the same, we don't permit repetitions when we list elements of a set. Thus, for correctness, we should always say *ordered*; when we omit it, it is often because of laziness or boredom.

1. We define

$$\begin{aligned}\mathbf{a}+\mathbf{b} &= (a_1+b_1, a_2+b_2, \ldots, a_n+b_n)\\ -\mathbf{b} &= (-b_1, -b_2, \ldots, -b_n)\\ \mathbf{a}-\mathbf{b} &= \mathbf{a}+(-\mathbf{b})\\ k\mathbf{b} &= (kb_1, kb_2, \ldots, kb_n)\end{aligned}$$

   Note that we have not given any meaning to the juxtaposition of a vector *followed* by a scalar.

2. We call $-\mathbf{u}$ the *reversal* of $\mathbf{u}$, $\mathbf{a}+\mathbf{b}$ the *sum*[14] of $\mathbf{a}$, $\mathbf{b}$.

3. If $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}$ are vectors, and $k_1, k_2, \ldots, k_n$ are scalars, then

$$k_1\mathbf{u_1} + k_2\mathbf{u_2} + \cdots + k_n\mathbf{u_n}$$

   is said to be a *linear combination* of the given vectors.

Theorem A.1 *Let*

$$\begin{aligned}\mathbf{a} &= (a_1, a_2, \ldots, a_n)\\ \mathbf{b} &= (b_1, b_2, \ldots, b_n)\\ \mathbf{c} &= (c_1, c_2, \ldots, c_n),\end{aligned}$$

be arbitrary vectors, and let $k, \ell \in \mathbb{R}$ be arbitrary scalars. Then

1. $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$
2. $\mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}$
3. $\mathbf{a}+\mathbf{0} = \mathbf{a}$
4. $\mathbf{a}+(-\mathbf{a}) = \mathbf{0}$
5. $(k+\ell)\mathbf{a} = k\mathbf{a}+\ell\mathbf{a}$
6. $k(\mathbf{a}+\mathbf{b}) = k\mathbf{a}+k\mathbf{b}$
7. $(k\ell)\mathbf{a} = k(\ell\mathbf{a})$
8. $1\mathbf{a} = \mathbf{a}$
9. $-(-\mathbf{a}) = \mathbf{a}$

[14]Note that, while we use the + sign to denote addition here, it is a different operation from addition of real numbers, since it is defined on a different set. Had it been practical, I would have used the type $\boldsymbol{+}$ for addition of vectors, to distinguish the operation from addition of scalars. The distinction is broader than this. We really should use a symbol like $\boldsymbol{+}_{\mathbb{R}^n}$ to indicate the context where the vector addition is taking place.

A.2 Supplementary Notes for the Lecture of Friday, January 6th, 2006

Release Date: Friday, January 6th, 2006
Subject to revision

A.2.1 §1.4 Dot (Inner) Product

This is a temporary definition that we will eventually be generalizing, in [1, Chapter 7]. However we need it in its present, temporary form.

Definition A.4 Let

$$\begin{aligned} \mathbf{a} &= (a_1, a_2, \ldots, a_n)\,, \\ \mathbf{b} &= (b_1, b_2, \ldots, b_n) \end{aligned}$$

be two given vectors in $\mathbb{R}^n$.

1. We define the *dot* or *inner* product of these vectors to be
$$\mathbf{a} \bullet \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n\,.$$
2. The *norm* or *length* of $\mathbf{a}$, denoted by $\|\mathbf{a}\|$ is defined to be[15]
$$\sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\,.$$
3. Where $\|\mathbf{a}\| \neq 0$ and $\|\mathbf{b}\| \neq 0$, the *angle* between $\mathbf{a}$ and $\mathbf{b}$ is defined to be $\arccos\left(\dfrac{\mathbf{a} \bullet \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}\right)$.
4. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be *orthogonal* or *perpendicular* if any one of the following conditions holds:
    (a) $\mathbf{a} = \mathbf{0}$
    (b) $\mathbf{b} = \mathbf{0}$
    (c) The angle between $\mathbf{a}$ and $\mathbf{b}$ is $\dfrac{\pi}{2}$.
5. A vector is a *unit* vector if its length is 1.

[15]Note the symbols: a tiny dot $\cdot$ means "multiplied by"; a big dot $\bullet$ represents the dot product.

6. The *distance* between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, denoted by $d(\mathbf{u}, \mathbf{v})$, is defined to be the length of $\mathbf{a} - \mathbf{b}$; thus
$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}\,.$$

Theorem A.2 *Let* $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ *be any three vectors, and* k, ℓ *any scalars. Then*

1. $(\mathbf{u} + \mathbf{v}) \bullet \mathbf{w} = \mathbf{u} \bullet \mathbf{w} + \mathbf{v} \bullet \mathbf{w}$
2. $\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}$
3. $(k\mathbf{u}) \bullet \mathbf{v} = \mathbf{u} \bullet (k\mathbf{v}) = k(\mathbf{u} \bullet \mathbf{v})$
4. $\|\mathbf{u}\| \geq 0$, *with equality if and only if* $\mathbf{u} = \mathbf{0}$.
5. *If* $\mathbf{u} \neq \mathbf{0}$, $\dfrac{1}{\|\mathbf{u}\|}\mathbf{u}$ *is a unit vector.*[16]
6. $\mathbf{a}$ *and* $\mathbf{b}$ *are orthogonal if and only if* $\mathbf{a} \bullet \mathbf{b} = 0$.[17]

Definition A.5 1. The textbook follows common usage by placing a circumflex over a vector that has unit length, as $\hat{\mathbf{u}}$, when $\|\mathbf{u}\| = 1$.

2. The textbook follows a less universal usage of denoting $\dfrac{1}{\|\mathbf{u}\|}\mathbf{u}$ by $\hat{\mathbf{u}}$; that is, the circumflex is used to denote a vector obtained from the given (non-zero) vector by *normalizing*, i.e., by multiplying by the reciprocal of its length, and thereby creating a unit vector.

3. If $\mathbf{v} \neq \mathbf{0}$, the *(scalar) projection* of $\mathbf{u}$ on $\mathbf{v}$ is defined to be $\mathbf{u} \bullet \hat{\mathbf{v}}$.

4. If $\mathbf{v} \neq \mathbf{0}$, the *(vector) projection* of $\mathbf{u}$ on $\mathbf{v}$, denoted by $\mathrm{proj}(\mathbf{u}, \mathbf{v})$ is defined to be $(\mathbf{u} \bullet \hat{\mathbf{v}})\,\hat{\mathbf{v}}$.[18]

Theorem A.3 (Cauchy, Schwartz, Buniakovski, et al.) *For any vectors* $\mathbf{u}$, $\mathbf{v}$ *in* $\mathbb{R}^n$
$$|\mathbf{u} \bullet \mathbf{v}| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|\,.$$

[16] Note that we have not defined what it means to *divide* a vector by a scalar; that is why I write $\frac{1}{\|\mathbf{u}\|}\mathbf{u}$ rather than $\frac{\mathbf{u}}{\|\mathbf{u}\|}$.

[17] This condition may now be taken as the *definition* of orthogonality: this is the way in which the textbook defines the concept.

[18] Your textbook uses the word *projection* for the *vector* projection.

Theorem A.4 (Minkovski) *For any vectors* $\mathbf{u}$, $\mathbf{v}$ *in* $\mathbb{R}^n$,

$$\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \mathbf{v}\|$$

Theorem A.5

$$\mathrm{proj}(\mathbf{u},\mathbf{v}) = \frac{\mathbf{u} \bullet \mathbf{v}}{\|\mathbf{v}\|}\mathbf{v}$$

A.2.2 §1.5 Located Vectors, Hyperplanes, Lines, Curves in $\mathbb{R}^n$

The theory of vectors is linked to the geometry of Euclidean spaces. Eventually we will be introducing abstract definitions, which will be independent of any coordinates in the spaces. For the present, think of a vector $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ as being associated with the point $(a_1, a_2, \ldots, a_n)$ in $\mathbb{R}^n$. The usual interpretation of this association is through a *directed line segment* originating at the origin, and terminating at the point. The textbook uses slightly different notations: denoting the "vector" by $[a_1, a_2, \ldots, a_n]$, but retaining parentheses (=round brackets) for the coordinates of the point, as $(a_1, a_2, \ldots, a_n)$; the authors may abbreviate the coordinates of the point $A = (a_1, a_2, \ldots, a_n)$, as $A(a_1, a_2, \ldots, a_n)$, or even as $A(a_i)$.

Definition A.6 If $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$ are points of $\mathbb{R}^n$, the line segment (i.e., a piece of a doubly infinite line) directed from A to B is denoted by $\vec{AB}$; if this segment is moved from its original position in a parallel way, preserving the sense of direction from A to B, we obtain a segment directed from the origin to the point with coordinates $(b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n)$; we may "identify" the vector with the vector with these entries, written either this way, or as

$$\begin{pmatrix} b_1 - a_1 & b_2 - a_2 & \ldots & b_n - a_n \end{pmatrix} \qquad \text{or} \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ \ldots \\ b_n - a_n \end{pmatrix}.$$

The definition we have given for vector addition can be shown to be equivalent to either the "triangle rule" or the "parallelogram rule" of adding geometric vectors, as sketched in the textbook in [1, Figure 1-1, p. 2]. The definition of scalar multiplication can be motivated by interpreting multiplication by a positive scalar as lengthening a vector by that factor; and multiplication by a negative scalar can be interpreted as first reversing the vector, and then lengthening by the magnitude of the given scalar.

Definition A.7 1. Let $\mathbf{a} \in \mathbb{R}^n$ be a non-zero vector, and $b \in \mathbb{R}$ a scalar. The set of points $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ such that

$$\mathbf{a} \bullet \mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

is called a *hyperplane with normal direction* $\mathbf{a}$.

2. Let $\mathbf{u} \in \mathbb{R}^n$ be a non-zero vector, and $P = \mathbf{a} = (a_1, a_2, \ldots, a_n)$ be any point in $\mathbb{R}^n$. The set of points

$$\mathbf{a} + t\mathbf{u} = (a_1 + tu_1, a_2 + tu_2, \ldots, a_n + tu_n)$$

where t ranges over $\mathbb{R}$, is said to form a *line* in $\mathbb{R}^n$ through the point $\mathbf{a}$ with direction $\mathbf{u}$; t is the *parameter*.

A.2.3 §1.6 Vectors in $\mathbb{R}^3$ (Spatial Vectors), ijk notation

(This theory will be generalized later in the course.) It is customary to denote

$$\begin{aligned} \mathbf{i} &= (1,0,0) \\ \mathbf{j} &= (0,1,0) \\ \mathbf{k} &= (0,0,1) \end{aligned}$$

so that any vector (x, y, z) may be expressed as $x\mathbf{i} + y\mathbf{j} + k\mathbf{k}$. The vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ have such familiar properties as

$$\|\mathbf{i}\| = \|\mathbf{j}\| = \|\mathbf{k}\| = 1$$

$$\mathbf{i} \bullet \mathbf{j} = \mathbf{j} \bullet \mathbf{k} = \mathbf{k} \bullet \mathbf{i} = 0$$

Notes:

1. It is customary to use the two vectors $\mathbf{i}$ and $\mathbf{j}$ in an analogous way in $\mathbb{R}^2$, and the single vector $\mathbf{i}$ in $\mathbb{R}^1$.

2. While these are unit vectors, the textbook does not normally write them with the circumflex.

Definition A.8 Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be arbitrary vectors in $\mathbb{R}^3$. The *cross-product* $\mathbf{u} \times \mathbf{v}$ is defined to be the vector

$$(u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$$

Theorem A.6 *1. Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be arbitrary vectors in $\mathbb{R}^3$. Then*

(a) $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

(b) $\|\mathbf{u} \times \mathbf{v}\|$ is equal to the area of the parallelogram determined by the two given vectors.

2.

$$\begin{aligned}\mathbf{i}\times\mathbf{j} &= -\mathbf{j}\times\mathbf{i}=\mathbf{k}\\ \mathbf{j}\times\mathbf{k} &= -\mathbf{k}\times\mathbf{j}=\mathbf{i}\\ \mathbf{k}\times\mathbf{i} &= -\mathbf{i}\times\mathbf{k}=\mathbf{j}\\ \mathbf{i}\times\mathbf{i} &= \mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=\mathbf{0}\end{aligned}$$

A.2.4 §1.7 Complex Numbers

The *complex numbers* are simply the points of $\mathbb{R}^2$, endowed with a multiplication rule that associates with any 2 points a third. The multiplication rule is

$$(x_1, y_1)\cdot(x_2, y_2) = (x_1x_2 - y_1y_1, x_1y_2 + x_2y_1)$$

It is customary to denote complex numbers (x, y) by a single symbol z, and to call x the *real* part and y the *imaginary* part. The rationale for these names is that the complex numbers $(x, 0)$ lie on the x-axis, called the "real" axis, associated with the real number system $\mathbb{R}$: multiplication of these "real" numbers is induced by the multiplication of the real parts:

$$(x_1, 0)\cdot(x_2, 0) = (x_1x_2 - 0, 0\times 0) = (x_1x_2, 0)\,.$$

The numbers $(0, y)$ lie on the y-axis, and there the product

$$(0, 1)\cdot(0, 1) = (-1, 0)$$

shows that the unit point on the y-axis is one of the square roots in $\mathbb{C}$ of the real number -1; since no real numbers have this property, the custom arose to call these numbers on the y-axis or "imaginary", and to call the entire axis the "imaginary" axis. The complex number $(0, 1)$ is often denoted by i (not to be confused with $\mathbf{i}$); thus

$$(x, y) = (x, 0) + yi = x + yi$$

if we identify the "real" complex number $(x, 0)$ with the real number x. The mirror image of the complex number $z = (x, y)$ in the real axis, $(x, -y)$, is called its *complex conjugate*, and denoted by $\overline{z}$. The product $z\cdot\overline{z}$ is always real; its non-negative square root is called the *absolute value* or *modulus* of z, and denoted by $|z|$.

A.2.5 §1.8 Vectors in $\mathbb{C}^n$

We will return to this topic later in the course.

A.2.6 Chapter 1: Supplementary Problems

You should be able to solve *all* the problems in (the sections selected from) this chapter. The few problems that I solve here are purely as a sampling, and not intended to suggest that those I select are more important than others.

[**1, Exercise 1.49**] Write $\mathbf{v} = \begin{pmatrix} 9 \\ -3 \\ 16 \end{pmatrix}$ as a linear combination of

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix}.$$

Solution: (I chose this problem at random; what is interesting is that there is a misprint, either in the problem in the textbook, or in the alleged answer: the authors state that

$$\mathbf{v} = 3\mathbf{u_1} - \mathbf{u}_2 + 2\mathbf{u}_3 .$$

Assuming their vectors $\mathbf{u}_i (i = 1, 2, 3)$ are printed correctly, this would indicate that the vector $\mathbf{v}$ should be $\begin{pmatrix} 9 \\ 0 \\ 16 \end{pmatrix}$. I will solve the problem with the given data; but, as the numbers were not "cooked up" to yield pretty solutions, the constants will be complicated fractions.)

We need to find scalars x_1, x_2, x_3 such that

$$\mathbf{v} = k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + k_3\mathbf{u}_3 ,$$

i.e.,

$$\begin{pmatrix} 9 \\ -3 \\ 16 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix};$$

this vector equation can be written as a simultaneous system of linear equations in x_1, x_2, x_3:

$$\begin{aligned} 1x_1 + 2x_2 + 4x_3 &= 9 \\ 3x_2 + 5x_2 - 2x_3 &= -3 \\ 3x_1 - 1x_2 + 3x_3 &= 16 \end{aligned}$$

By standard methods of row reduction we find that

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{88}\begin{pmatrix} 288 \\ -146 \\ 199 \end{pmatrix} ;$$

you may verify the correctness of this solution by computing the linear combination with the given scalar weights and demonstrating that it is equal to $\mathbf{v}$.

[1, Exercise 4.83(c), p. 163] Find the value(s) of k so that $\mathbf{w} = (1, k, 4)$ is a linear combination of the vectors $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (2, 3, 1)$.

Solution: We need to solve the vector equation

$$(1, k, 4) = a(1, 2, 3) + b(2, 3, 1)$$

For scalars k, a, b. By definition of scalar multiplication, this equation is equivalent to

$$(1, k, 4) = (a, 2a, 3a) + (2b, 3b, b)$$

and, by definition of vector addition, this last equation is equivalent to

$$(1, k, 4) = (a + 2b, 2a + 3b, 3a + b) .$$

Two vectors of 3-tuples are equal when corresponding entries are equal:

$$\begin{aligned} 1 &= a + 2b \\ k &= 2a + 3b \\ 4 &= 3a + b \end{aligned}$$

Let's rewrite these equations as an inhomogeneous system in 3 variables:

$$\begin{aligned} 1a + 2b + 0k &= 1 \\ 2a + 3b + (-1)k &= 0 \\ 3a + 1b + 0k &= 4 \end{aligned}$$

By row reduction of the augmented matrix, we find that there is exactly one solution, $(a, b, k) = \left(\frac{7}{5}, -\frac{1}{5}, \frac{11}{5}\right)$, so k is uniquely equal to $\frac{11}{5}$. (Geometrically, the two given vectors span a plane in $\mathbb{R}^3$, and this plane intersects the line $(x, y, z) = (1, k, 4)$ in just one point: $\left(1, \frac{11}{5}, 4\right)$.

[1, Exercise 1.67] Simplify the following complex numbers, by writing each in the form $x + iy$, where $x, y \in \mathbb{R}$:

1. $\frac{1}{2i}$
2. $\frac{2+3i}{7-3i}$
3. i^{15}, i^{25}, i^{34}
4. $\left(\frac{1}{3-i}\right)^2$

Solution:

1. $\frac{1}{2i} = \frac{1}{2i} \cdot \frac{i}{i} = \frac{i}{-2} = 0 - \frac{1}{2}i$
2. $\frac{2+3i}{7-3i} = \frac{2+3i}{7-3i} \cdot \frac{7+3i}{7+3i} = \frac{(2+3i)(7+3i)}{49+9} = \frac{14-9+27i}{58} = \frac{5}{58} + \frac{27}{58}i$
3. $i^{15} = i^{3\cdot 4} \cdot i^2 \cdot i = 1^3 \cdot (-1) \cdot i = -i$ (error in textbook answer), $i^{25} = i^{6\cdot 4} \cdot i = 1^6 \cdot i = i$, $i^{34} = i^{8\cdot 4} \cdot i^2 = 1^8 \cdot (-1) = -1$
4. $\left(\frac{1}{3-i}\right)^2 = \left(\frac{1}{3-i}\right)^2 \cdot \left(\frac{3+i}{3+i}\right)^2 = \frac{(3+i)^2}{(9+1)^2} = \frac{(9-1)+6i}{10^2} = \frac{2}{25} + \frac{3}{50}i$

A.2.7 Some "standard" types of problems that do not appear to be found in the textbook

There are a number of "standard" types of problems that you should be able to solve, by virtue of your background in MATH 133. Here are some:

1. Find the plane in $\mathbb{R}^3$ through 3 given (non-collinear) points.

   Solution: There are a number of ways of solving a problem like this. One method is to assume undetermined coefficients in the equation, and then impose the condition that the plane pass through the given points. This gives 3 equations for the 4 undetermined coefficients, so that the ratios of the coefficients can be found. Since an equation can be multiplied by a non-zero scalar, this is sufficient information to find an equation for the plane.

   Another method is to find vectors joining the points in pairs. The cross product of two of these vectors will give the normal direction to the plane.

2. Find the distance from a point P to a given plane $ax + by + cz = d$.

Solution: There are many methods for solving this problem. The easiest may be to find the magnitude of the scalar projection on to the normal to the plane of the vector joining P to any convenient point in the given plane.

CHAPTER 2 — Algebra of Matrices

As for Chapter 1, students in MATH 223 are expected to be thoroughly familiar with the material in this chapter, by virtue of the prerequisite of "MATH 133 or equivalent".

A.2.8 §2.1 Introduction

A.2.9 §2.2 Matrices

A matrix *over* a field $\mathbb{F}$ — for this review we begin by considering the field to be $\mathbb{R}$: the intention is that you carry over the definitions to general fields without any significant issues. A matrix normally has a finite number of "rows", which we often denoted by the symbol m, and a finite number of "columns", here denoted by n. Thus a typical $m \times n$ matrix is

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} .$$

The horizontal subarrays are the rows, and the vertical subarrays are the columns.

The *zero matrix* 0_{mn} is an $m \times n$ matrix, all of whose entries are 0; where the size of the matrix is clear from the context, the subscript mn may be suppressed: you will know from the context whether we are looking at a scalar or a matrix.

A.2.10 §2.3 Matrix Addition and Scalar Multiplication

An $m \times n$-matrix may be thought of as a vector in $\mathbb{R}^{mn}$ which has been "folded" into m rows, each containing n entries. The definitions of addition and multiplication by a scalar are those that would hold if the matrix had been written linearly as a string of mn scalars; I will not reproduce those definitions here, assuming that all students are familiar with them and with their properties.

A.2.11 §2.4 Summation Symbol

Students are reminded of the use of the symbol Σ to denote summation. When we have a function f of one integer variable i, and where r and s are integers where f is defined, and $r \leq s$, we write $\sum_{i=r}^{s} f(i)$ to denote the sum

$$\sum_{i=r}^{s} f(i) = f(r) + f(r+1) + f(r+2) + \cdots + f(s-1) + f(s)\,.$$

In this notation the symbol i is a *bound* or "dummy" variable: we may replace it throughout the formula by any other symbol not already in use; thus

$$\sum_{i=r}^{s} f(i) = \sum_{j=r}^{s} f(j) = \sum_{A=r}^{s} f(A)\,.$$

A.2.12 §2.5 Matrix Multiplication

Matrix multiplication differs from the multiplication rules we have defined hitherto, in that it is not defined for all pairs of matrices. If A and B are matrices, we define the product AB *only if the number of columns of*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

is equal to the number of rows of

$$B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1p} \\ b_{21} & b_{22} & \ldots & b_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{np} \end{pmatrix}\,.$$

Then we have the following definition

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1p} \\ b_{21} & b_{22} & \ldots & b_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1i}b_{i1} & \sum_{i=1}^{n} a_{1i}b_{i2} & \ldots & \sum_{i=1}^{n} a_{1i}b_{ip} \\ \sum_{i=1}^{n} a_{2i}b_{i1} & \sum_{i=1}^{n} a_{2i}b_{i2} & \ldots & \sum_{i=1}^{n} a_{2i}b_{ip} \\ \ldots & \ldots & \ldots & \ldots \\ \sum_{i=1}^{n} a_{mi}b_{i1} & \sum_{i=1}^{n} a_{mi}b_{i2} & \ldots & \sum_{i=1}^{n} a_{mi}b_{ip} \end{pmatrix}$$

Theorem A.7

1. *Let A, B, C be respectively $m \times n$, $n \times p$ and $p \times q$ matrices. Then the* associative law *holds for matrix multiplication:*

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

2. *Let A, B, C be respectively $m \times n$, $n \times p$ and $n \times p$ matrices. Then the* left distributive law *holds for matrix multiplication:*

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

3. *Let A, B, C be respectively $m \times n$, $m \times n$ and $n \times p$ matrices. Then the* right distributive law *holds for matrix multiplication:*

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

4. *Let A, B be respectively $m \times n$ and $n \times p$ matrices, and let k be any scalar. Then*

$$k(AB) = (kA)B$$

A.3 Supplementary Notes for the Lecture of Monday, January 9th, 2006

Release Date: Monday, January 9th, 2006
Subject to further revision

A.3.1 Summary of Last Week's Lectures

1. I discussed the prerequisites to this course — a first course in linear algebra and related geometry, like McGill's MATH 131, or CEGEP 201-105. Students should have seen things like geometric vectors and vectors of ordered pairs and triples of real numbers, small dimensional matrices, and solutions of small systems of linear equations (possibly interpreted only geometrically). [1, Chapters 1 and 2] are review material, except that you may not have met the complex field $\mathbb{C}$ before; you may also not have seen some of the concepts considered for general dimension n before.

2. I discussed only isolated topics in [1, Chapter 1], mentioning that complex numbers would be reviewed in **WeBWorK** assignments WW_1 and WW_2.

3. I discussed [1, §2.1–2.5]

A.3.2 WeBWorK WW_1

This assignment, which does not count, has been released. You should have access to it, unless you joined the course very recently, where Dr. Hundemer of my Department may not yet have created your account. You may need to review [1, Chapter 1]. For some reason one of the problems in WW_1 on complex numbers appears prior to the introduction to the subject; skip ahead to the introduction, and then go back to the problem. You may also wish to read my notes A.2.4 on page 1009.

A.3.3 §2.6 The Transpose of a Matrix

The *transpose* of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

is the matrix

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}$$

Theorem A.8 *1. Let A and B be matrices with the same numbers of rows and the same numbers of columns. Then*

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

2. Let A be any matrix. Then $\left(A^T\right)^T = A$.

3. Let k be any scalar, and A any matrix. Then

$$(kA)^T = kA^T .$$

4. Let A and B be respectively an $m \times n$ matrix and an $n \times p$ matrix. Then

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

A.3.4 §2.7 Square Matrices

Definition A.9 1. A *square* matrix has the same number of rows and columns.

2. The *(main) diagonal* of an $n \times n$ matrix A consists of the entries a_{11}, a_{22}, ..., a_{nn}.

3. The *trace* of an $n \times n$ matrix A, denoted by $\mathrm{tr}(A)$, is the sum of the *(main) diagonal* entries, i.e.,
$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} .$$

4. The $n \times n$ *identity* matrix I_n is the matrix whose entries are all 0, except for the main diagonal entries, which are 1. Where the size of the matrix is clear from the context, the subscript n may be suppressed.

   The entry in the kth row and ℓth column of I_n is denoted by $\delta_{k\ell}$, called the *Kronecker delta function*.

5. A *scalar* matrix is the product of a scalar with an identity matrix.

Theorem A.9 *1. Let A, B be $n \times n$ matrices, and let k be any scalar. Then*

(a) $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

(b) $\text{tr}(kA) = k \cdot \text{tr}(A)$

(c) $\text{tr}\left(A^T\right) = \text{tr}(A)$

2. *Let A be an $m \times n$ matrix, and B an $n \times m$ matrix. Then* $\text{tr}(AB) = \text{tr}(BA)$.

3. *Let C be an $m \times n$ matrix. Then*

$$CI_n = C = I_m C$$

A.3.5 §2.8 Powers of Matrices, Polynomials in Matrices

While we cannot multiply two arbitrary matrices, we can multiply matrices obtained from a given *square* $n \times n$ matrix by applying the various operations of matrix addition, matrix multiplication, and scalar multiplication, as all of these yield $n \times n$ matrices. We can define the non-negative *powers recursively*:

Definition A.10 Let A be an $n \times n$ matrix. Define

$$A^0 = I_n$$

If A^n has been defined, where n is a non-negative integer, we define

$$A^{n+1} = (A^n)\, A$$

$(n = 0, 1, \ldots)$. If $f(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n$ is a polynomial (with coefficients from the same field as the matrix A, then we can define the *matric polynomial* $f(A)$ by

$$f(A) = b_0 I_n + b_1 A + b_2 A^2 + \cdots + b_n A^n\,.$$

When $f(A)$ is equal to the zero matrix 0_{nn}, A may be called a *zero* or *root* of the polynomial f.

A.3.6 §2.9 Invertible (Nonsingular) Matrices

A matrix A is said to be *invertible* or *nonsingular* if there exists a matrix B such that

$$AB = I \qquad \text{and} \qquad BA = I\,. \tag{1}$$

B is called an *inverse* of A.

Theorem A.10 1. *If A is invertible, then A is square.*

2. *If, for a matrix A, there exists a matrix B such that equation (1) holds, then B is unique with that property. In other words, if A has an inverse, then that inverse is unique. We will denote the inverse of A, where it exists, by A^{-1}.*

3. *If* $B = A^{-1}$*, then* $A = B^{-1}$.

4. *A* 1×1 *matrix* $\begin{pmatrix} a \end{pmatrix}$ *is invertible if and only if* $a \neq 0$*, in which case*

$$\begin{pmatrix} a \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{a} \end{pmatrix}.$$

5. *A matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is invertible if and only if* $ad - bc \neq 0$*, in which case*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

<u>Proof of Theorem A.10</u>:

1. This will be proved in connection with [1, §5.4]

2. First observe that, if A is $m \times n$, then, since the product AB is square, B must be $n \times m$. Thus, if A were to have two inverses B_1 and B_2, they would have the same size. The properties assumed are

$$AB_1 = I_m = AB_2$$

$$B_1A = I_n = B_2A$$

We apply the distributive property of matrix multiplications, and the properties of the identity matrices:

$$B_1 = B_1I_m = B_1(AB_2) = (B_1A)B_2 = I_nB_2 = B_2$$

$\square$

This proof actually proves a stronger result. If we separate the two conditions on the inverse, calling the matrices that act on the two sides of A its *left inverse* and *right inverse*, then this proof can be used to show that, if they both exist, then they must be identical!

3.

4.

5.

A.3.7 §2.10 Special Types of Square Matrices

Definition A.11 1. A *diagonal* matrix is a square matrix whose only non-zero entries — if any — are in the main diagonal. If, reading down the main diagonal, those entries are $d_1, d_2, \ldots$, we may denote the matrix by $\mathrm{diag}(d_1, d_2, \ldots)$.

2. An *upper triangular matrix* is a square matrix whose only non-zero entries — if any — are on or above the main diagonal. That is, if $T = (t_{ij})$ is an $n \times n$ upper triangular matrix, then, for $1 \le i, j \le n$,

$$i > j \Rightarrow t_{ij} = 0\,.$$

Your textbook states that the term *upper triangular* may be abbreviated to *triangular*. This terminology would be avoided, since there is also the following concept:

3. A *lower triangular matrix* is a square matrix whose only non-zero entries — if any — are on or below the main diagonal. That is, if $T = (t_{ij})$ is an $n \times n$ lower triangular matrix, then, for $1 \le i, j \le n$,

$$i < j \Rightarrow t_{ij} = 0\,.$$

Theorem A.11 *1. Sums, powers, and scalar multiples of upper triangular matrices are also upper triangular, as are matric polynomials in a given upper triangular matrix.*

2. An upper triangular matrix is invertible if and only if all of its main diagonal elements are non-zero. When such an inverse exists, it is also upper triangular.

Definition A.12 1. A matrix is *symmetric* if it is equal to its transpose.

2. A matrix is *skew-symmetric* if it is equal to the negative of its transpose.

3. A matrix A is orthogonal if its transpose is its inverse; equivalently, if $AA^T = A^TA = I$

4. A matrix A is *normal* if $AA^T = A^TA$.

Theorem A.12 *1. A symmetric matrix is square.*

2. A skew-symmetric matrix is square.

3. The main diagonal elements of a skew-symmetric matrix are all 0.

A.3.8 §2.11 Complex Matrices

This topic is not considered review, but will not be discussed at this time.

A.3.9 §2.12 Block Matrices

Sometimes it is convenient to subdivide a matrix by imaginary horizontal and vertical lines into component submatrices.

A.3.10 Chapter 2: Supplementary Problems

[**1, Exercise 2.56, p. 54**] Find 2×2 invertible matrices A and B such that $A + B \neq 0$ and $A + B$ is not invertible.

Solution: The textbook gives the example of

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}.$$

This example is correct, and does meet the conditions of the problem; the sum $A+B$ does not have an inverse because its second row (column) is a scalar multiple of its first row (column). But the example is otherwise not very interesting. Other examples would be

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

whose sum is a matrix of 1's; or

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A.4 Supplementary Notes for the Lecture of Wednesday, January 11th, 2006

Release Date: Wednesday, January 11th, 2006
Subject to further revision

I spent a few minutes at the beginning completing some review topics from Chapter 2.

CHAPTER 3 — Systems of Linear Equations

This is the third chapter of prerequisite material; [1, §§3.1–3.7, 3.12] will be reviewed very briefly prior to beginning with the course proper, which will begin with [1, Chapter 4], but which will then return to [1, §§3.8–3.11]. There will not be time in the current course for more than this review, and we will need to use these definitions and results. Your previous contact with these concepts may, however, been restricted to systems where the numbers of equations and/or of variables was not more than, say three.

A.4.1 §3.1 Introduction

The textbook observes that it will be assuming all matrices are over $\mathbb{R}$, and states that "there is almost no loss in generality if the reader assumes that all our scalars...come from the real field". You may make this assumption; however, towards the end of the course we may return to this issue, to see how there are some real differences when we consider certain fields other than $\mathbb{R}$.

A.4.2 §3.2 Basic Definitions. Solutions

Definition A.13 1. A ¡¿ *linear equation* in unknowns $x_1, x_2, \ldots, x_n$

2. ¡¿ the *coefficient* of x_k in an equation

3. ¡¿ the *constant term* of an equation

4. ¡¿ When does $\mathbf{u} = (b_1, b_2, \ldots, b_n)$ *satisfy* the given equation

5. ¡¿ a *system* of linear equations in unknowns $x_1, x_2, \ldots, x_n$

6. An $m \times n$ *system* has m equations, each in the same n unknowns. The first parameter, here m, is usually associated with the number of *rows*, i.e., the *horizontal* lists; the second parameter, here n, is usually associated with the number of *columns*, i.e., the *vertical* lists.

7. An equation is *homogeneous* if the constant to the right of the equal sign is 0; a system is *homogeneous* if all of its equations are homogeneous.

8. An equation is *degenerate* if all of its coefficients are 0.

9. A *(particular) solution* is a vector $\mathbf{u} = (b_1, b_2, \ldots, b_n)$ which simultaneously satisfies all of the equations in the system.

10. The *general solution* of a system is the set of all (particular) solutions.

11. A system of equations is *consistent* if there exists at least one solution; and *inconsistent* or *contradictory* if there exists no solution.

Particular Solutions and General Solutions Remember the jargon introduced above in Definition A.13: a *particular* solution is just one solution, possibly selected to satisfy conditions additional to those specified in the system of equations; the *general solution* is a description of the set of *all* particular solutions. For the field $\mathbb{R}$ — indeed, for all fields that are not *finite*, and you possibly have never met a finite field — the total number of solutions is either 0, 1 or infinite. When the number of solutions is infinite, we will need some systematic way of describing them, since we can't list them all. One way in which this is done by a *parametric* description, generalizing what you already know about parametric equations for lines.[19] I am going to delay the details of this parametric description until I have discussed the concept of *basis* in connection with [1, §4.8], see page 1061 approximately below.

Row and Column Vectors It is sometimes convenient to write the vectors which appear as solutions as *column* vectors and sometimes as *row* vectors. In this course I will not always be fussy about transposing vectors, but will write them in whatever way is practical, in order to agree as closely as possible with the notation of the textbook. If I were writing a book for Mathematics Honours students, I would be much more careful.

Row and Column Numbers I have mentioned in the definitions above that we use the word *row* for the horizontal sets if entries in a matrix, and *column* for the vertical. If we wish to label general entries in a *matrix of coefficients*, we usually use 2 subscripts, where the first denotes the *row number* and the second denotes the *column number*. In

[19]Of course, if you have never seen parametric equations of lines in the plane or 3 dimensions, you probably don't have the prerequisites for this course!

a general system of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n} & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n} & = & b_2 \\ \ldots & & \ldots & & \ldots & & \ldots & = & \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn} & = & b_m \end{array} \tag{2}$$

the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{3}$$

is called the *coefficient matrix*, and the matrix

$$M = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_n \end{array}\right) \tag{4}$$

is called the *augmented matrix*. Here I have represented the augmented matrix as a *partitioned* matrix with an imaginary vertical line; this line is not always shown. We can also write it more compactly as $A = \left(A \mid \mathbf{b} \right)$, where $\mathbf{b}$ denotes the (vertical) matrix of constants from the right sides of the equations.

Terms that are "absent" Sometimes we will designate the variables in advance, and then write an equation which constrains those variables, but in which some of the variables do not appear explicitly. The intention here usually is that the variables are present with coefficients equal to 0. If the order of the terms in the equation has been prescribed, some of these terms which are equal to 0 may appear at the beginning of the equation; the first variable to appear with a non-zero coefficient is called the *leading* variable of that equation, and its coefficient is the *leading* coefficient for that equation, and also for the corresponding row of the coefficient or augmented matrix.

A.5 Supplementary Notes for the Lecture of Friday, January 13th, 2006

Release Date: Friday, January 13th, 2006
Subject to further revision

A.5.1 Complex numbers again

I began the lecture by discussing some basic properties of complex numbers, so that students could solve problems on **WeBWorK** assignments WW_1 and WW_2. I will return to these topics when we are ready to do non-trivial things with the complex numbers. For now students who are not familiar with the properties of $\mathbb{C}$, and find that [1, §1.7] and my notes and comments in class are not enough, might wish to look in [2, Appendix G]. Specifically today I mentioned in class that, we call the angle made by a vector $\mathbf{z} = (x, y)$ in $\mathbb{R}$ with the positive x-axis the *argument* of $\mathbf{z}$ denote it $\mathrm{Arg}(\mathbf{z})$; and that, with the usual definitions,

$$z = |z|\,(\cos(\mathrm{Arg}(\mathbf{z})), \sin(\mathrm{Arg}(\mathbf{z})))$$

We can show that, when two complex numbers are multiplied, the product has modulus equal to the product of the moduli[20] and argument equal to the sum of the arguments; and the argument of $\frac{1}{z}$ is minus the argument of z.

Definition A.14 The *principal value* of a complex number $\mathbf{z} = x + iy$ is defined to be the particular argument θ that lies in the interval $(-\pi, +\pi]$.

If one insists that the argument always be *principal*, one removes the ambiguity in this coordinate for the point; however, one introduces a discontinuity in the coordinatization of the plane, since points in the 3rd quadrant very close to the x-axis have a very different argument from points in the 2nd quadrant just above the same axis. **It appears that some of the problems on WeBWorK expect that you state an argument only as a principal argument: in particular, Problems ##22, 36 of WW_1, and #22 of WW_2.**

A.5.2 §3.3 Equivalent Systems. Elementary Operations

Our analysis of the solutions of systems of equations is based on the transformation of one system into another which has the same solution set; we call such systems *equivalent*. The transformations we consider will always be *reversible*: the operations we apply can be reversed, and applied in the reverse order to the system we finally obtain, and we can

[20]= plural of "modulus"

recover the original system; and, at every step along the way, the solution set will be the same. The rationale of these transformations is that the system we finally obtain will be one whose solution set is easy to characterize and to determine.

Elementary Operations on a System of Equations We consider 3 Elementary operations:

$[E_1]$: This is an interchange of the positions of two of the equations in the system. Perhaps we should denote it more precisely by $[E_1(i,j)]$, to indicate that the ith and jth equations are being interchanged, where i,j are integers in $\{1,2,\ldots,m\}$.

$[E_2]$: This is the multiplication of an equation by a non-zero constant. Perhaps we should denote it more precisely by $[E_2(i,\alpha)]$, where $1 \le i \le m$, and $\alpha \in \mathbb{R} - \{0\}$.

$[E_3]$: This is the addition of a multiple of one equation to another. Perhaps we should denote it more precisely by $[E_3(i,j,\alpha)]$, where $i,j \in \{1,2,\ldots m\}$, and $\alpha \in \mathbb{R}$. I am following the system of the textbook here; but it should be clear that we could have managed by simply adding one equation to another with multiplier 1, since the previous operation permits us to multiply by a non-zero constant.

Theorem A.13 *Under the application of any one of these Elementary Operations the solution set of a system is not changed.*

What we will be doing in the following sections is to describe an algorithm for applying these operations in such a way as to ultimately obtain a system in a "canonical" form, where the nature of the solutions set is easily determined.

A.5.3 §3.4 Small Square Systems of Linear Equations

I will not discuss this section in the lectures; students should be familiar with its content from MATH 133.

A.6 Supplementary Notes for the Lecture of Monday, January 16th, 2006

Release Date: Monday, January 16th, 2006
Subject to further revision

A.6.1 §3.5 Systems in Triangular and Echelon Forms

The elementary operations discussed earlier can be applied systematically, in order to transform a given system of equations into one whose matrix has a specific "canonical" form. There are a number of such forms that are useful in linear algebra. In all of these "forms" one assumes that the variables appear in the same order in each of the equations; where a variable does not explicitly appear in some equation, we usually interpret it as being present with coefficient 0.

Definition A.15 1. In certain systems of equations the first non-zero term appearing (counting from the left side) may be called the *pivot* term. The associated variable is called the *pivot* variable.

2. A system of equations is said to be in *(row)-echelon* form if the locations of the pivot terms in successive equations move progressively from left to right, with no pivot ever directly below a preceding pivot.

3. The textbook uses the term *triangular* form for the special case of echelon form in which the number of pivots is equal to the number of variables and to the number of equations.

4. Define a linear equation to be *trivial* if it is satisfied by all scalars; equivalently, if it is of the form $0 = 0$; an equation is *nontrivial*, if it is not trivial.

The definition given for *pivot* is ambiguous in 2 ways.

- First, it could happen that all the coefficients in an equation are 0. This situation does not normally affect the solutions: in transforming systems into equivalent systems we normally move such equations to the bottom of the list, which we attain by applying operation $[E_1]$ repeatedly.

- It can happen that all the coefficients of terms to the left of the equal sign are 0. In such a case we look to the right of the equal sign. If the constant term there is also 0, we have the case just discussed. If the term is non-zero, then the equation reads $0 = b$, where $b \neq 0$, which is a contradiction: in such a case the system of equations is inconsistent, as the existence of a solution would imply that this equation could be satisfied.

Theorem A.14 *Suppose a system of r nontrivial equations in n variables is given in echelon form.*

1. *If $r > n$, then at least $r - n$ of the equations will have all coefficients of variables equal to 0. As mentioned above, such equations are either of the form $0 = b$ where $b \neq 0$, or $0 = 0$.*
   (a) *It is often convenient to delete equations of the form $0 = 0$, since this deletion does not change the nature of the solution set of the system.*
   (b) *If there is at least one equation of the form $0 = b$ where $b \neq 0$, then the system is* inconsistent *or* unsolvable, *and has no solutions.*
2. *If $r = n$ the system is in triangular form, and has just one solution.*
3. *If $r < n$, one can arbitrarily assign values to $n-r$ of the variables — called the* free variables *— which are never pivot variables — and then determine the values of the pivot variables from them. This leads to a* parametric *solution to the system, where the non-pivot variables become the* parameters.

The basic idea in solving systems in echelon form is to start with the last equation and move upward, progressively parametrizing non-pivot variables, and expressing pivot variables in terms of them. This procedure is called *back substitution*. One such method will be discussed in the next section.

A.6.2 §3.6 Gaussian Elimination

This *algorithm* proceeds in 2 phases.

- In the first phase the system is transformed through elementary operations applied progressively, moving *forward*, i.e., downward in the system. The textbook calls this phase *Forward Elimination*: whenever a pivot entry is identified, multiples of the equation are subtracted from the following equations, in order to arrange that this variable not appear in any of the following equations.
- In the second phase, called *Backward Elimination*, multiples of the pivot equations are subtracted from equations appearing *earlier* in the system, in order to arrange that the pivot variable does not appear in any of the equations preceding.

  Students in MATH 223 are expected to be familiar with Gaussian elimination or some variant thereof, from their prerequisite course.

Example A.15 (modified from [3, p. 93, Problem 3.2.2]) Use Gaussian elimination for equations to solve

$$\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 5 && (5)\\ 4x_1 + 3x_2 + 7x_3 + 3x_4 &= 8 && (6)\\ -8x_1 - x_2 - x_3 + 3x_4 &= 4 && (7)\\ 6x_1 + x_2 + 2x_3 - x_4 &= 1 && (8)\\ -2x_1 + 0x_2 + x_3 + 2x_4 &= 5 && (9) \end{aligned}$$

Solution: I propose to first subtract multiples of the first equation from the others, in order to eliminate the terms in x_1 appearing in those other equations. If we were proceeding mechanically, we would normally divide the first equation by a scalar to produce coefficient 1 associated with x_1. While that may be necessary eventually, we see that all the terms we need to clear are multiples of $2x_1$, so we delay that step, so that we may continue to work with integer coefficients. The term $2x_1$ that we are subtracting from the other equations may be called the *pivot*. Adding appropriate scalar multiples of the first equation to the others yields:

$$\begin{aligned} 2x_1 + 1x_2 + 2x_3 + 1x_4 &= 5 && (10)\\ 0x_1 + 1x_2 + 3x_3 + 1x_4 &= -2 && (11)\\ 0x_1 + 3x_2 + 7x_3 + 7x_4 &= 24 && (12)\\ 0x_1 - 2x_2 - 4x_3 - 4x_4 &= -14 && (13)\\ 0x_1 + 1x_2 + 3x_3 + 3x_4 &= 10 && (14) \end{aligned}$$

We will now subtract multiples of the second equation from the third, fourth, and fifth to clear the bottom half of the second column:

$$\begin{aligned} 2x_1 + 1x_2 + 2x_3 + 1x_4 &= 5 && (15)\\ 0x_1 + 1x_2 + 3x_3 + 1x_4 &= -2 && (16)\\ 0x_1 + 0x_2 - 2x_3 + 4x_4 &= 30 && (17)\\ 0x_1 + 0x_2 + 2x_3 - 2x_4 &= -18 && (18)\\ 0x_1 + 0x_2 + 0x_3 + 2x_4 &= 12 && (19) \end{aligned}$$

We add the third equation to the fourth:

$$\begin{aligned} 2x_1 + 1x_2 + 2x_3 + 1x_4 &= 5 && (20)\\ 0x_1 + 1x_2 + 3x_3 + 1x_4 &= -2 && (21)\\ 0x_1 + 0x_2 - 2x_3 + 4x_4 &= 30 && (22)\\ 0x_1 + 0x_2 + 0x_3 + 2x_4 &= 12 && (23)\\ 0x_1 + 0x_2 + 0x_3 + 2x_4 &= 12 && (24) \end{aligned}$$

and then subtract the fourth from the fifth:

$$\begin{aligned} 2x_1 + 1x_2 + 2x_3 + 1x_4 &= 5 && (25)\\ 0x_1 + 1x_2 + 3x_3 + 1x_4 &= -2 && (26)\\ 0x_1 + 0x_2 - 2x_3 + 4x_4 &= 30 && (27)\\ 0x_1 + 0x_2 + 0x_3 + 2x_4 &= 12 && (28)\\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0\,. && (29) \end{aligned}$$

This transformed system may now be solved by *backward substitution*: equation (28) yields $x_4 = 6$; substitution of this value in equation (27) yields $x_3 = -3$; substitution of both of these values in equation (26) yields $x_2 = 1$; and a final substitution in equation (26) yields $x_1 = 2$.

A.6.3 §3.7 (Row-)Echelon Matrices, Row Canonical Form, Row Equivalence

Our discussion hitherto has been applied to systems of equations. We now recognize that the operations we discussed could be performed on the coefficients detached from the matrices, and proceed to investigate the properties of matrices that are preserved under those operations.

Definition A.16 In reviewing [1, §3.3] I defined *elementary operations* on systems of equations. Analogous to these we can define *row operations* on matrices. The textbook uses the same symbols, $[E_1]$, $[E_2]$, $[E_3]$ to describe these *row* operations.

The following transformations of a matrix A are called *row operations*:

1. $[E_1]$ Interchange the ith row with the jth row.
2. $[E_2]$ Replace the ith row by k times the ith row, where k is any non-zero scalar.
3. $[E_3]$ Add to the jth row k times the ith row, where $i \neq j$.

Analogous operations on the columns of A are called *column operations*.

Definition A.17 The textbook uses the notation $A \sim B$ to mean that there exists a sequence of elementary row operations that can transform matrix A into matrix B.

Canonical forms We are describing various related kinds of "canonical forms" into which systems of equations and their related matrices can be transformed by operations on the equations or the rows. Some of these "canonical" forms are more convenient to use in the formulation of theorems; others are more efficient to use in practical computations, as they may involve fewer operations and so more efficient algorithms. I will usually not be considering questions of algorithmic efficiency in this course.

Gaussian and Gauss-Jordan reduction of matrices In the preceding section we considered the procedure of *Gaussian elimination*, under which a system of equations was transformed into one which could be solved by back-substitution, starting from the last equation in the transformed system. We now apply the same ideas to matrices. There are several levels of reduction possible.

Definition A.18 1. The first non-zero entry in any row, if any, (counting from the left), is called a *pivot*.

2. A matrix is said to be in *row-echelon form* if it has the properties that

   (a) All rows consisting of 0's alone are together at the bottom of the matrix: if a row has at least one non-zero entry, then all rows above it also have at least one non-zero entry.

   (b) In the rows that follow a non-zero row, under the pivot entry in that non-zero row, there are only zeros.

   (c) *Column-echelon form*, *column canonical form*, *reduced column-echelon form* *column-reduced form* can be defined analogously, in terms of column operations, and on the entries in any *row* containing a pivot entry (which is then the first non-zero entry in a column counting down from the top).

3. A matrix is said to be in *row canonical form*, or *reduced row-echelon form*, or in *row-reduced form* if it is in row-echelon form, and, in addition to being in row-echelon form, it has the following properties:

   3. The pivot in each non-zero row is a 1.

   4. In the rows above any non-zero row, in the column containing the pivot of that non-zero row, there are only zeros.

Some of the results we state will be in terms of matrices which are in row-echelon form, and some in terms of matrices which are further reduced. In general, the more we reduce the matrices, the "prettier" are the results. The downside is that the additional reduction is inefficient. In this course I will not be attempting to produce algorithms that are particularly efficient, so I will not be systematic in trying to streamline algorithms.

Theorem A.16 *1. If matrices B and C are both matrices in row echelon form obtained from a matrix A by row operations, then they have precisely the same numbers of pivot variables.*

2. Every matrix admits just one matrix in row canonical form to which it is row equivalent.

At this point the textbook becomes careless, and the authors define the *rank* of a matrix. While there is nothing wrong with the definition, they give another definition in [1, §4.9, p. 131]. You can only define a concept once! I will delay the definition they give here and, in discussing [1, §4.9, p. 131], show that the concept defined there does have the property desired here.

A.6.4 The row reduction algorithms

The Gaussian algorithm And what is the algorithm that permits the reduction of a matrix? I will not write the algorithm out formally.

1. We begin on the left side of the matrix, moving in until we meet a column which has some non-zero entries. If there is no such column, then the matrix is the zero matrix, and the algorithm terminates.

2. When we meet a column with a non-zero entry, we apply a row operation to move one *row* containing such a non-zero entry to the top of the matrix. Then we subtract multiples of that row from all the rows below in which, in the given column, there are non-zero entries. This clears all entries in the column below the top entry, which becomes our first pivot.

3. We leave the first row alone now, and apply the same algorithm to the rows below the first row: finding, if there is one, a pivot in another column, moving a row to the top, just below the row we have just found, and clearing the entries below the pivot. In this way we obtain a matrix in row echelon form.

Gauss-Jordan Reduction We can continue our reduction in two ways:

1. Scale the rows so that the pivots are all 1.

2. Subtract multiples of pivot rows from the rows *above* in order to make the pivots the unique non-zero entry in the column. (This, perhaps, justifies the term *pivot*.)

I now skip to [1, §3.12] before introducing vector spaces.

A.6.5 §3.12 Elementary Matrices

Definition A.19 Corresponding to the *elementary operations* on systems of linear equations are the *elementary matrices*. These are *square* matrices, which the textbook denotes respectively by E_1, E_2, and E_3. As with the notation for elementary transformations, I suggest a more detailed notation would be better, to indicate precisely what operation is effected *by premultiplying* a given matrix by one of these elementary matrices. The elementary matrix corresponding to row operation $[E_i]$ on an $n \times n$ matrix is obtained from the identity matrix I_n by applying to it that very row operation $(i = 1, 2, 3)$. *Note that the same matrices serve to perform* column *operations when a given matrix is multiplied by them* on the right*; however the specific function of a matrix of type E_3 depends on whether it is used on the right or the left: the matrix which effects an addition of the ith row to the jth when applied on the left, will effect an addition of the jth column to the ith when applied on the right.*

A.7 Supplementary Notes for the Lecture of Wednesday, January 18th, 2006

Release Date: Wednesday, January 18th, 2006
Subject to further revision

A.7.1 §3.12 Elementary Matrices (conclusion)

I now indicate an application of elementary matrices in finding the inverse of an invertible matrix.

Theorem A.17 *Let k be a positive integer, and let A_i be an $m_i \times m_{i+1}$ matrix which is invertible ($i = 1, 2, \ldots, k$). (While we will see eventually that $m_1 = m_2 = \cdots = m_{k+1}$, we do not require that fact now, and it has not yet been proved.) Then the product matrix $A_1A_2 \ldots A_k$ is also invertible, and*

$$(A_1A_2 \ldots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \ldots A_1^{-1} .$$

Sketch of Proof: First prove the theorem for $k = 2$. This can be done by simply computing the product

$$(A_1A_2)(A_2^{-1}A_1^{-1}) = \left((A_1A_2)A_2^{-1}\right) A_1^{-1} = \left(A_1\left(A_2A_2^{-1}\right)\right) A_1^{-1} = A_1IA_1^{-1} = A_1A_1^{-1} = I$$

and the product

$$(A_2^{-1}A_1^{-1})(A_1A_2) = \cdots = I$$

Then the proof is extended by induction to a product of k matrices.

Theorem A.18 *An invertible matrix can have no row or column which consists entirely of 0's.*

Proof: Suppose that the ith row of the $m \times n$ matrix A consists only of 0's. then the i row of any product AB of A with any matrix with n rows will consist entirely of 0's; in particular, this means that AB cannot be equal to the identity matrix, since it has a non-zero entry in every row. To prove the analogous result for columns, consider products CA where C has m columns. □

Theorem A.19 *If A is row equivalent to a matrix B, then either both or neither of A and B are invertible.*

Proof: Any row operation on a matrix A can be achieved by premultiplying[21] A by an elementary matrix. All elementary matrices are invertible. Suppose that $B = F_1F_2 \ldots F_kA$

[21] i.e. multiplying on the left

where F_i is an elementary matrix ($i = 1, \ldots, k$), and B is row reduced. We know that the elementary matrices are invertible.[22] Thus $A = F_k^{-1} F_{k-1}^{-1} \ldots F_1^{-1} B$. Then, if A is invertible, $B^{-1} = A^{-1} F_k^{-1} F_{k-1}^{-1} \ldots F_1^{-1}$. Similarly, if B is invertible,

$$A^{-1} = B^{-1} F_1 F_2 \ldots F_k$$

□

Corollary A.20 (to Theorem A.19) *A square matrix is invertible if and only if it can be row reduced to an identity matrix.*

A.7.2 Supplementary Problems

[1, Exercise 3.67(d), p. 113] Find the inverse of the following matrix, if the matrix is invertible.

$$D = \begin{pmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix}$$

Solution: Note that we still have not seen an explanation why only square matrices can have inverses. That will be done in Theorem A.60, page 1102. Let's first find the inverse BY THE VERY INEFFICIENT METHOD of row reducing the given matrix using elementary matrices applied on the left. These operations can be applied in various orders.

1. We first scale the first row by multiplying by $\frac{1}{2}$. This is equivalent to premultiplying by an elementary matrix which is obtained by applying the operation $[E_1]$ with constant $\frac{1}{2}$ to the 3×3 identity matrix:

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix}$$

2. Next we subtract 5 times the first row from the second:

$$\begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 2 & 1 \end{pmatrix}$$

[22] Can you prove this? For example, any matrix of form E_1 is its own inverse.

3. We multiply the 2nd row by -2 in order to normalize the pivot entry:

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

4. We now add copies of the 2nd row to the first and third, to clear the entries above and below the pivot entry in position $(2,2)$:

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

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

5. We multiply the 3rd row by -1, then subtract multiples of the new 3rd rows from the 1st and 2nd to clear the entries above the pivot in position $(3,3)$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

6. Now let's examine what we have done. We have multiplied the given matrix on the left by a sequence of elementary matrices, eventually causing the matrix to transform into the identity. Since matrix multiplication is associate, we can group the multiplications so that we first multiply the elementary matrices, then multiply their product, in the given order, by the given matrix. Since that yields the identity, the product of the elementary matrices must be the

inverse. That product is

$$
\begin{aligned}
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \\
\times &\begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix}
\begin{pmatrix} 1 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{5}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix}
\begin{pmatrix} -2 & 1 & 0 \\ 5 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
= &\begin{pmatrix} 8 & -3 & -1 \\ -5 & 2 & 1 \\ 10 & -4 & -1 \end{pmatrix}
\end{aligned}
$$

which must be the inverse of the given matrix. Verify by multiplication that it is, indeed, the inverse!

7. The purpose of this exercise, solved in this very inefficient way, is to demonstrate why the procedure discussed in the following theorem is effective.

Determination of the inverse of an invertible matrix A Practice using the following algorithm given to determine an inverse:

Theorem A.21 *Let a square matrix A be given.*

1. *Write A as the left half of a partitioned matrix of the form $M = \left(\, A \mid I \,\right)$ having twice as many columns as rows.*
2. *Perform* row *operations on M in order to transform A into an identity matrix.*
3. *The right half of M will be A^{-1}.*
4. *If, however, A has no inverse, then the procedure of row reduction will not terminate in an identity matrix; but, rather, in a matrix with at least one row of 0's.*

Proof: The operations performed on the right side of the matrix are precisely the multiplications of elementary matrices in the order of their application. Thus this method is simply a systematization of the method applied in the preceding example.

CHAPTER 4 – Vector Spaces

In this chapter we establish the universe in which we will be working. From here on all of our results will be expressed in terms of *vector spaces*, which will be defined abstractly, to generalize the many concrete examples we have already met.

After defining this abstract context, we set about demonstrating how specific examples can be described through the use of matrices. This motif will be continued in [1, Chapter 6], after we first study the abstract concept of Linear Mappings or Linear Transformations in [1, Chapter 5].

After discussion of [1, §§4.1–4.8], I will return to the omitted sections of [1, Chapter 3], which will be discussed in connection with [1, §§4.9, 4.10], and will then complete the discussion of the Chapter with [1, §4.11].

A.7.3 §4.1 Introduction

This short section reminds you of set theoretic notation, which we will be using.

A.8 Supplementary Notes for the Lecture of Friday, January 20th, 2006

Release Date: Friday, January 20th, 2006
Subject to further revision

A.8.1 §4.2 Vector Spaces

In this chapter we start to act like mathematicians. Having observed technically complicated properties of row and column vectors, we now begin to analyze the properties in ways that will produce results that we can apply to other structures that we observe have similar properties. We try to distance ourselves from the specific details that motivated our abstraction, returning only when we are losing our way.

The abstraction that we develop — the *abstract vector space* may appear complicated at first. Don't try to memorize the properties immediately, but remember the general structure. The context is a *field* $\mathbb{K}$ — a set in which we can perform the types of operations we are familiar with in $\mathbb{R}$: addition (and subtraction), multiplication (and division by elements different from 0), where the two operations interact in the distributivity properties; there are two special elements — 0 is the "identity" element for addition, and 1 is the "identity" element for multiplications. In this context we consider a set of objects called *vectors* on which the main operation we define first is *addition*, which is intended to generalize the addition of row and column vectors that we have been discussing. The vectors will have an *additive identity*, which we still denote by $\mathbf{0}$, although it should no longer be thought of as being a string of 0's. Finally, the field $\mathbb{K}$ "operates" or "acts" on the vectors, in a way that we still write as "multiplication" by a scalar: we may write $k\mathbf{v}$ to look like a product, but we are really thinking of k as a function which maps vectors on to vectors. Any theorems we prove for these generalized vector spaces will apply immediately to the row and column vectors of real numbers that we have been studying; but these theorems will also apply to many other structured sets.

Definition A.20 Let $\mathbb{K}$ be a given field. A *vector space* $\mathcal{V}$ *over* $\mathbb{K}$ is a (non-empty)[23] set whose properties we will divide into two lists:

[A]: These are axioms that concern the way in which vectors interact among themselves, under an operation that we call *addition*, and write using the sign $+$.

[23] I have placed the words *non-empty* in parentheses because we don't need to assume that; it will be a consequence of one of the coming conditions that will require the presence of an element called $\mathbf{0}$: since the set needs to contain $\mathbf{0}$, it can't be empty! One of the guiding principles of axiom construction is to try to make the set of axioms as *economical* as possible — try not to assume anything that you can get for free, as a consequence of some of the other axioms.

$[M]$: These are axioms that concern the way in which the field elements "operate" on the set of vectors; we call this operation *scalar multiplication*, and write it the way we did before, writing the scalar immediately to the left of the vector.

While we write the two operations in a way that reminds us of the operations on row and column vectors, this is only a mnemonic tool. The A axioms postulate a structure on the set of vectors, a structure that we call an *Abelian group*, in deference to the Norwegian mathematician, Nils Henrik Abel (1802-1829)[24]. We require that $\mathcal{V}$ have an "addition" property, i.e., a function from $\mathcal{V}$ to $\mathcal{V} \times \mathcal{V}$ given by $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w}$ with the following properties:

$[A_1]$ **(Addition is *associative*):** For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$,

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

$[A_2]$ **(Existence of an *additive identity*):** There exists an element, which we usually denote by $\mathbf{0}$, with the property that, for every vector $\mathbf{v} \in \mathcal{V}$,

$$\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}$$

$[A_3]$ **(Existence of *additive inverses*):** For every vector $\mathbf{v} \in \mathcal{V}$ there exists a vector which we denote by $-\mathbf{v} \in \mathcal{V}$ with the property that

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$$

$[A_4]$ **(*Commutativity* of addition):** For all vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}$,

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

The second part of the structure of a vector space can be viewed as "scalar multiplication"[25]. We assume the following properties:

$[M_1]$ **(left distributivity):** For any scalar $k \in \mathbb{K}$, and any vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}$,

$$k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$$

$[M_2]$ **(right distributivity):** For any scalars $k, \ell \in \mathbb{K}$, and any vector $\mathbf{v} \in \mathcal{V}$,

$$(k + \ell)\mathbf{v} = k\mathbf{v} + \ell\mathbf{v}$$

[24] Note the short lifetime!

[25] although it is more correctly viewed as the scalar elements each acting as a function mapping $\mathcal{V}$ into itself

$[M_3]$ **(associativity):** For any scalars $k, \ell \in \mathbb{K}$, and any vector $\mathbf{v} \in \mathcal{V}$,

$$(k\ell)\mathbf{v} = k(\ell\mathbf{v})$$

$[M_4]$ **(identity for scalar multiplication):** For any vector $\mathbf{v} \in \mathcal{V}$,

$$1\mathbf{v} = \mathbf{v}$$

These axioms have not been presented in the most economical way. We can prove various other simple propositions, some of which are already foreseen in the kind of notation we have been using.

Theorem A.22 *1. For every vector $\mathbf{v}$ in a vector space there is just one element that satisfies axiom $[A_2]$.*

2. *In any vector space there is just one vector $\mathbf{0}$ which satisfies axiom $[A_2]$.*

3. *For any scalar $k \in \mathbb{K}$, $k\mathbf{0} = \mathbf{0}$.*

4. *For any vector $\mathbf{v} \in \mathcal{V}$, $0\mathbf{v} = \mathbf{0}$.*

5. *If $k\mathbf{v} = \mathbf{0}$, then either $k = 0$ or $\mathbf{v} = \mathbf{0}$ (or both).*

6. *For any $k \in \mathbb{K}$ and any $v \in \mathcal{V}$, $(-k)\mathbf{v} = k(-\mathbf{v}) = -k\mathbf{v}$.*

Supplementary Problems

[1, Exercise 4.72, p. 162] Let $\mathcal{V}$ be the set of ordered pairs (a, b) of real numbers, with addition in $\mathcal{V}$ and scalar multiplication on $\mathcal{V}$ defined by

$$\begin{aligned} (a,b) + (c,d) &= (a+c, b+d) \\ k(a,b) &= (ka, 0) \end{aligned}$$

Show that $\mathcal{V}$ satisfies all the axioms of a vector space except $[M_4]$; that is, except $1\mathbf{u} = \mathbf{u}$ for all $\mathbf{u} \in \mathcal{V}$. This shows that this particular axiom is not a consequence of the other axioms.

Solution: I will not give the full proof. To show that the specific axiom fails we need to provide only one instance where it fails — i.e., just one specific ordered pair (a, b). Clearly, we have to take a pair where $b \neq 0$, for otherwise the property would be present. So, for example, I can take $(a, b) = (0, 1)$, and $k = 1$ (since $k = 0$ would not be a good choice to prove the failure of the axiom). With the given definition we have

$$k(a,b) = 1(0,1) = (1(0), 0) = (0,0) \neq (0,1) = (a,b)\,.$$

Mathematicians usually search for the smallest set of assumptions to give the structure they want; this example shows that, unless we assume property $[M_4]$, we won't get it "for free", i.e., as a consequence of the other assumptions, since all of the other axioms can be shown to be satisfied for this particular definition of scalar multiplication.

While I don't expect you to become an expert on this aspect of the algebra of vector spaces, I want you to experience the spirit in which mathematicians polish systems of axioms. I don't expect you to be able to solve [1, Exercise 4.73, p. 163], which is in this spirit.

[1, Exercise 4.74, p. 163] Let $\mathcal{V}$ be the set of ordered pairs (a, b) of real numbers. In each of the following cases show that $\mathcal{V}$ is not a vector space over $\mathbb{R}$, if addition and scalar multiplication are defined in these ways.

1.
$$\begin{aligned} (a,b)+(c,d) &= (a+d, b+c) \\ k(a,b) &= (ka, kb) \end{aligned}$$

2.
$$\begin{aligned} (a,b)+(c,d) &= (a+c, b+d) \\ k(a,b) &= (a, b) \end{aligned}$$

3.
$$\begin{aligned} (a,b)+(c,d) &= (0,0) \\ k(a,b) &= (ka, kb) \end{aligned}$$

4.
$$\begin{aligned} (a,b)+(c,d) &= (ac, bd) \\ k(a,b) &= (ka, kb) \end{aligned}$$

<u>Solution</u>: This problem should be approached by comparing the proposed definitions with the definition of $\mathbb{R}^2$ in Example A.23 on page 1044 of these notes. That should suggest how to go about finding a counterexample to one of the axioms.

1. Here the definition of scalar multiplication is the same as in $\mathbb{R}^2$, so we expect the problem will be in the confused definition of addition.

Here is an undirected attack that doesn't work too well, just to show you how one sometimes has to go back to the drawing board: Taking a general (a, b), and $(c, d) = (0, 0)$, we see that

$$(a, b) + (0, 0) = (a + 0, b + 0) = (a, b)$$

so the $\mathbf{0}$ will have to be $(0, 0)$. But then we find that the vector $-(a, b)$ will have to be $(-b, -a)$. This leads to a problem, since, by [1, Theorem 4.1(iv), p. 117] $-(a, b) = (-a, -b)$. We could have to have $(-a, -b) = (-b, -a)$ for all $a, b \in \mathbb{R}$, and this is contradicted by any choice of $a \neq b$. This is a valid proof, but it is not very elegant — it depends on using a theorem, rather than proving the discrepancy by direct reference to the axioms.

Here is a better proof. I will show that the associativity property of addition, $[A_1]$ fails.

$$((0,0)+(0,1))+(1,0) = (0+1,0+0)+(1,0) = (1,0)+(1,0) = (1+0,0+1) = (1,1)$$

$$(0,0)+((0,1)+(1,0)) = (0,0)+(0+0,1+1) = (0,0)+(0,2) = (0+2,0+0) = (2,0)$$

For associativity to hold we would need to have $1 = 2$ and $1 = 0$, which are both nonsense. We conclude from this specific counterexample that the given definition of vector addition does not give the structure of a vector space.

2. Suppose that $\mathcal{V}$ is a vector space. It follows from the first equation that $(0, 0) + (0, 0) = (0, 0)$, so $(0, 0)$ is the vector $\mathbf{0}$. Setting $k = 1, 2$ in the second equation, we obtain $(2 - 1)(a, b) = (0, 0)$, so $a = 0 = b$ for all a, b, which is a contradiction. We conclude that $\mathcal{V}$ is *not* a vector space.

3. This time the difficulty will again derive from the definition of vector addition. Axiom $[A_2]$ states that, for any vector $(a, b) \in \mathcal{V}$ there must exist a vector called $\mathbf{0}$ such that
$$(a, b) + \mathbf{0} = (a, b)\,.$$
But our definition of vector addition here says that all sums are equal to $(0, 0)$. So, if we take as (a, b) anything but $(0, 0)$, we will get a contradiction. Here is my counterexample. Let $(a, b) = (1, 0)$, and let (c, d) be the vector $\mathbf{0}$ postulated by Axiom $[A_2]$. Then,
$$\begin{aligned}(1,0) &= (a,b) && \text{(the selection we have made)}\\ &= (a,b)+\mathbf{0} && \text{by Axiom } [A_2]\\ &= (a,b)+(c,d) &&\\ &= (0,0) && \text{by the given definition of } +\end{aligned}$$
Comparing the first components, this implies that $1 = 0$. From this contradiction we see that $\mathcal{V}$ is not a vector space, and the given axiom cannot hold for this definition of vector addition.

4. Take $(c, d) = (0, 0)$. Then we have, for any a, b,

$$(a, b) + (0, 0) = (a(0), b(0)) = (0, 0)$$

If we add to the extreme members of this equation the vector $-(0, 0)$, we obtain

$$(a, b) + \mathbf{0} = \mathbf{0}$$

which implies, by axiom $[A_2]$, that

$$(a, b) = \mathbf{0}\,.$$

But we did not restrict (a, b): thus all vectors are equal to $\mathbf{0}$, so they are equal to each other. In particular, this implies that $(1, 0)$ and $(2, 0)$ are the same, so $1 = 2$. We again have a contradiction, which tells us that the given definition of $+$ does not yield a vector space.

A.8.2 §4.3 Examples of Vector Spaces

To describe an example we need to specify the field, the set of vectors, the addition law, the vector $\mathbf{0}$, the additive inverses, and the law of scalar multiplication. Often the latter four specifications are "obvious", and may not be explicitly stated.

Example A.23 (Column) n-vectors over a field, $\mathbb{K}^n$

1. For any field $\mathbb{K}$, and any positive integer n, we could let the "underlying set" of a vector space be the "ordered n-tuples" of field elements, written horizontally with parentheses and separated by commas, as $(v_1, v_2, \ldots, v_n)$. For reasons that will become clear later in the course, I find it more convenient to make this standard example be written as "column vectors", and I will represent them by $n\times 1$ matrices. So the typical element of this *set* is

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{30}$$

This defines the "underlying set" of the vector space, but we still need to specify the operations of vector addition and multiplication by a scalar, the vector $\mathbf{0}$, and the additive inverse $-\mathbf{v}$ of any vector $\mathbf{v}$.

2. As you might have expected, I will define the vector $\mathbf{0}$ to be

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and the additive inverse of a vector $\mathbf{v}$ with entries as given above in equation (36) by

$$-\mathbf{v} = \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix} . \tag{31}$$

This looks obvious, but keep in mind that the minus sign that precedes the vector is not the same as the minus sign preceding its entries: the second is describing the additive inverse of an element of $\mathbb{K}$; the first is describing the additive inverse of an element of the vector space we are constructing here: one of them is a function from $\mathbb{K}$ to $\mathbb{K}$ which maps every scalar k on to $-k$; the other is a function from $\mathcal{V}$ to $\mathcal{V}$ which maps every *vector* $\mathbf{v}$ on to $-\mathbf{v}$. These are different functions, as they have different domains.

3. Addition is *defined* by

$$\mathbf{u}+\mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{pmatrix} . \tag{32}$$

Here again, we use the same symbol for addition in two different sets — in $\mathbb{K}$ and in $\mathbb{K}^n$.

4. Multiplication of a vector $\mathbf{v}$ by a scalar k is defined by

$$k\mathbf{v} = k \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{pmatrix} . \tag{33}$$

5. Now we should systematically check each of the 8 axioms for a vector space, to be sure that every one of them is satisfied in the generality specified in each. Only

then are we entitled to call $\mathbb{K}^n$ a vector space. And then we can be assured that every theorem we prove for vector spaces will apply to this important, particular example.

Example A.24 Row n-vectors over a field: Ordered n-tuples can also be written as rows. With the obvious definitions, this can be shown to be a vector space. In this course I don't intend to make a strong distinction between this and the previous example, in spite of the fact that they are different.

A.9 Supplementary Notes for the Lecture of Monday, January 23rd, 2006

Release Date: Monday, January 23rd, 2006
Subject to further revision

A.9.1 §4.3 Examples of Vector Spaces (conclusion)

Not all of these examples were discussed in the lecture.

Example A.25 Diagonal $n \times n$ matrices. Ordered n-tuples can also be written as diagonal matrices. Here we write the ordered n-tuple $(d_1, d_2, \ldots, d_n)$ as

$$\operatorname{diag}(d_1, d_2, \ldots, d_n) = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_n \end{pmatrix}$$

The vector $\mathbf{0}$ will here be $\operatorname{diag}(0, 0, \ldots, 0)$. The inverse, addition, and scalar multiplication will be defined in the obvious way. Again, one needs to verify that all 8 of the axioms are satisfied before being able to call this a vector space.

Example A.26 $\mathbb{K}_{m,n}$, $m \times n$ matrices over the field $\mathbb{K}$: Ordered mn-tuples can also be written as $m \times n$ matrices. Here the definitions of addition and scalar multiplication are those given earlier for matrices, and $\mathbf{0}$ is $0_{m,n}$.

Example A.27 $\mathbb{K}_n[t]$, Polynomials of degree not more than n in an *indeterminate*.[26] ***t***: The 0 element is the 0 polynomial; polynomials are added by adding corresponding coefficients; multiplication by a scalar k is achieved by multiplying each coefficient by k.

Example A.28 $\mathbb{K}[t]$, Polynomials of all degrees in an *indeterminate*.[27] ***t***: $\mathbf{0}$ is the 0 polynomial, i.e., $0t^0$; polynomials are added by adding corresponding coefficients; multiplication by a scalar k is achieved by multiplying each coefficient by k.

Example A.29 $\mathbf{F}(X)$, Functions mapping a given set X into the field $\mathcal{K}$. The following definitions can normally be assumed, unless others are specifically stated:

Definition A.21 Let $f, g \in \mathbf{F}(X)$, and $k \in \mathbb{K}$.

[26] the variable
[27] the variable

1. The function $f+g$ is defined by specifying its action on any point $x \in X$: $(f+g)(x) = f(x) + g(x)$
2. The function kf is defined by specifying its action on any point $x \in X$: $(kf)(x) = kf(x)$. Note that the right side means the product in $\mathbb{K}$ of the scalars k and $f(x)$; on the left side, kf means the function called $kf : X \to \mathbb{K}$.
3. The function $-f$ is defined by $(-f)(x) = -f(x)$, i.e., it maps any point $x \in X$ on to the additive inverse of the field element $f(x)$.
4. The function $0 : X \to \mathbb{K}$ is defined by $x \mapsto 0$ for all $x \in X$.

As with the other examples, we now need to prove that the 8 axioms of vector spaces are satisfied.

Example A.30 Any Subspace of a Given Vector Space. In [1, §4.5] below we will be defining a structure called a *subspace* of a given vector space. A subspace will be, in particular, itself a vector space (using the structure of the vector space in which it is contained).

Supplementary Problems

[1, Exercise 4.75, p. 163] Consider the set of infinite sequences $\mathbf{a} = (a_0, a_1, a_2, \ldots, a_n \ldots)$ of real numbers. If we define sums and scalar products analogously to $\mathbb{R}^n$, i.e., if $\mathbf{a} = (a_0, a_1, a_2, \ldots, a_n \ldots)$ and $\mathbf{b} = (b_0, b_1, b_2, \ldots, b_n \ldots)$ are given sequences, and k is any real number, then by defining

$$\begin{aligned} \mathbf{0} &= (0, 0, 0, \ldots, 0 \ldots) \\ \mathbf{a} + \mathbf{b} &= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n \ldots) \\ k\mathbf{a} &= (ka_0, ka_1, ka_2, \ldots, ka_n \ldots) \\ -\mathbf{a} &= (-a_0, -a_1, -a_2, \ldots, -a_n \ldots) \end{aligned}$$

show that we obtain a vector space.

A.9.2 §4.5 Subspaces

Definition A.22 Suppose that $\mathcal{V}$ is a vector space over a field $\mathbb{K}$, and that W is a sub*set* of $\mathcal{V}$ with the property that, with the same definitions as for $\mathcal{V}$ of $\mathbf{0}$, of vector addition and scalar multiplication, and of $-\mathbf{v}$ for any $v \in \mathcal{V}$, W is a vector space over the field $\mathbb{K}$. Then we call W a *subspace* of $\mathcal{V}$.[28]

[28] In practice I will try to use the same type of symbols for subspaces, so I will use "calligraphic" type rather than simple capitals, as $\mathcal{W}$.

To verify that a set with given definitions constitutes a vector space can be a time-consuming exercise. Fortunately, when we are considering a subset of a known vector space, the following theorem renders this verification much simpler:

Theorem A.31 *A subset W of a vector space $\mathcal{V}$ over a field $\mathbb{K}$ is a subspace of $\mathcal{V}$ if and only if:*

1. $\mathbf{0} \in W$*; and*
2. (a) *W is closed under vector addition: i.e., whenever $\mathbf{u}, \mathbf{v} \in W$, then $\mathbf{u} + \mathbf{v} \in W$; and*
   (b) *W is closed under scalar multiplication: i.e., whenever $\mathbf{u} \in \mathcal{W}$ and $k \in \mathbb{K}$, then $k\mathbf{u} \in W$.*

This theorem can be rendered more satisfying to a mathematician by combining the two closure conditions into one:

Corollary A.32 (to Theorem A.31) *A subset W of a vector space $\mathcal{V}$ over a field $\mathbb{K}$ is a subspace of $\mathcal{V}$ if and only if:*

1. $\mathbf{0} \in W$*; and*
2. *W is closed under linear combinations: i.e., whenever $\mathbf{u}, \mathbf{v} \in \mathcal{W}$ and $k, \ell \in \mathbb{K}$, then $k\mathbf{u} + \ell\mathbf{v} \in W$.*

There are other ways in which the theorem can be made more "economical". For example, the following:

Corollary A.33 (to Theorem A.31) *A subset W of a vector space $\mathcal{V}$ over a field $\mathbb{K}$ is a subspace of $\mathcal{V}$ if and only if:*

1. $W \neq \emptyset$*; and*
2. *Whenever $\mathbf{u}, \mathbf{v} \in W$ and $k \in \mathbb{K}$, then $k\mathbf{u} + \mathbf{v} \in W$.*

We will not spend time in this course investigating these "improved" versions of the theorem: these are activities that are an obsession of mathematicians, and we can't always convince outsiders that they are valuable uses of our time.

A.9.3 Supplementary Problems

[1, Exercise 4.77, p. 163] Determine whether or not W is a subspace of $\mathbb{R}^3$, where W consists of all vectors $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$ such that

1. $a = 3b$
2. $a \leq b \leq c$
3. $ab = 0$
4. $a + b + c = 0$
5. $b = a^2$;
6. $a = 2b = 3c$.

Solution:

1. The condition $a = 3b$ is equivalent to choosing the points on the plane $x = 3y$ through the origin. The set contains the vector $(0, 0, 0)$; it is easy to show that the sum of two vectors with the property again have the property, as does the multiple of a vector with the property by a scalar.[29]
2. The use of inequalities rather than equalities causes this to define a subset which is not a vector space. One counterexample can be seen with the vector $(a, b, c) = (0, 1, 0)$, and the scalar $k = -1$. A vector space must be closed under scalar multiplication; but $k(a, b, c) = (0, -1, 0)$ which fails to have the given property, even though (a, b, c) is in $\mathcal{V}$. Thus $\mathcal{V}$ is not a subspace.
3. If we take the vectors $(1, 0, 0)$ and $(0, 1, 0)$, we see that they both have the property that the product of the first 2 entries is 0. But their sum is $(1, 1, 0)$, which lacks that property. Hence the subset of vectors with this property is not closed under vector addition, and cannot be a subspace of $\mathbb{R}^3$.
4. The property $a + b + c = 0$ does, indeed, define a subspace. Again, I will not provide a full proof, but leave that to the student.
5. The vectors $(1, 1, 0)$ and $(-1, 1, 0)$ both have the property that $b = a^2$; however, their sum lacks that property, so the set of vectors is not closed under addition, and cannot be a vector space. It is not closed under scalar multiplication, either, as is seen from consideration of the vector $(-1)(-1, 1, 0) = (1, -1, 0)$, which is not in W even though $(-1, 1, 0)$ is.

[29] This is not a proof: a valid proof would require that every fact be proved. I could be bluffing!

6. The vectors with the property $a = 2b = 3c$ do form a subspace: these are the points on the line through the origin with parametric equations $(x, y, z) = (0, 0, 0) + t(6, 3, 2)$. This set contains $(0, 0, 0)$, and the student should supply a proof that it is closed under addition and scalar multiplications.

[**1, Exercise 4.81, p. 163**] Show that each of the following subsets is a subspace of $\mathcal{V} = F(\mathbb{R})$:

1. The set W_1 of all *bounded* functions.[30]
2. The set W_2 of all *even* functions.[31]
3. (NOT IN THE TEXTBOOK) The set W_3 of all *odd* functions.[32]

Solution:

1. Suppose that f_1 and f_2 are bounded functions, and that k is a scalar. If $|f_1(x)| \leq M_1$ and $|f_2(x)| \leq M_2$ for all x, then

$$\begin{aligned} |\,(f_1 + f_2)\,(x)| &= |f_1(x) + f_2(x)| \quad \text{by definition of } f_1 + f_2 \\ &\leq |f_1(x)| + |f_2(x)| \quad \text{property of absolute value} \\ &\leq M_1 + M_2 \end{aligned}$$

which implies that $f_1 + f_2$ is also bounded. Similarly,

$$\begin{aligned} |kf_1(x)| &= |k| \cdot |f_1(x)| \\ &\leq |k| \cdot M_1 \end{aligned}$$

implying that kf_1 is also bounded. Thus the set of bounded functions is closed under addition and scalar multiplication. Finally, we observe that the function 0 is bounded, since $|0(x)| \leq 0$ for all x.

2. TO BE COMPLETED

[30]

Definition A.23 A function $f \in F(\mathbb{R})$ is *bounded* if there exists a real number $M \in \mathbb{R}$ such that $(\forall x \in \mathbb{R})|f(x)| \leq M$.

[31]

Definition A.24 A function $f \in F(\mathbb{R})$ is *even* if, for all x, $f(-x) = f(x)$.

[32]

Definition A.25 A function f is *odd* if, $(\forall x \in \mathbb{R})(f(-x) = -f(x))$.

3. TO BE COMPLETED

[1, Exercise 4.82(b), p. 163] The textbook asks you to show that the following is a subspace of the space of sequences defined in [1, Exercise 4.75, p. 163]: the set of all sequences having only a finite number of non-zero elements. The claim is **false**, and this is more serious than a typographical error. We can see that the zero sequence

$$\mathbf{0} = (0, 0, 0, \ldots, 0, \ldots)$$

is not contained in this subset. The problem is that the set is not closed under either addition or multiplication by a scalar! If you add to a vector its additive inverse, you obtain a vector not in the set; and similarly, if you multiply a vector in the set by the scalar 0, you again obtain the vector $\mathbf{0}$, which is not in the set! The error can be "fixed" by adjoining to the set of sequences with only a finite number of non-0 elements, the vector consisting of only 0 elements. In that sense it's not a very serious error, just an oversight.

A.10 Supplementary Notes for the Lecture of Wednesday, January 25th, 2006

Release Date: Wednesday, January 25th, 2006
Subject to further revision

A.10.1 §4.4 Linear Combinations, Spanning Sets

We have met the concept of *linear combination* before, when applied to row and column vectors of scalars. Here we renew the definition in the context of abstract vector spaces; there are no surprises.

Definition A.26 1. A *linear combination* of a finite set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a sum of scalar multiples of the form

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n ,$$

where $k_1, k_2, \ldots, k_n$ are scalars.

2. For any vector space $\mathcal{V}$, a spanning set S is a subset of $\mathcal{V}$ with the property that every vector $v \in \mathcal{V}$ is expressible as a linear combination of vectors in S.

A.10.2 Supplementary Problems

[**1, Exercise 4.83, p. 163**] Consider the vectors

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} .$$

1. Write $\mathbf{w} = \begin{pmatrix} 1 \\ 3 \\ 8 \end{pmatrix}$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$.
2. Write $\mathbf{w} = \begin{pmatrix} 2 \\ 3 \\ 8 \end{pmatrix}$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$.
3. Find k so that $\mathbf{w} = \begin{pmatrix} 1 \\ k \\ 4 \end{pmatrix}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.
4. Find conditions on a, b, c so that $\mathbf{w} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Solution:

1. We have to determine scalars k and ℓ such that $k\mathbf{u}+\ell\mathbf{v}$ is equal to the given vector, i.e., such that the following equations hold

$$\begin{aligned} 1k+2\ell &= 1 \\ 2k+3\ell &= 3 \\ 3k+1\ell &= 8\,. \end{aligned}$$

   The system is equivalent (after Gauss-Jordan reduction) to

$$\begin{aligned} k+0\ell &= 3 \\ 0k+1\ell &= -1 \\ 0k+0\ell &= 0 \end{aligned}$$

   so there is a unique solution,

$$\begin{pmatrix}1\\3\\8\end{pmatrix} = 3\begin{pmatrix}1\\2\\3\end{pmatrix} - 1\begin{pmatrix}2\\3\\1\end{pmatrix}\,.$$

2. Here we have to solve the system

$$\begin{aligned} 1k+2\ell &= 2 \\ 2k+3\ell &= 4 \\ 3k+1\ell &= 5\,, \end{aligned}$$

   which is equivalent (after Gauss-Jordan reduction) to

$$\begin{aligned} k+0\ell &= 2 \\ 0k+1\ell &= 0 \\ 0k+0\ell &= 1 \end{aligned}$$

   which system is inconsistent, because of the contradictory nature of the third equation of the system: the given vector is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.[33]

[33]The wording of the problem is objectionable: the authors should not have suggested that there existed a solution!

3. We have to find scalars a and b such that

$$\begin{aligned} 1a + 2b &= 1 \\ 2a + 3b &= k \\ 3a + 1b &= 4\,. \end{aligned}$$

The system is equivalent (after Gauss-Jordan reduction) to

$$\begin{aligned} a + 0b &= 2k - 3 \\ 0a + 1b &= -k + 2 \\ 0a + 0b &= 5k - 11\,. \end{aligned}$$

There can be a solution only if $5k - 11 = 0$, i.e., $k = \frac{11}{5}$. In that case the system does indeed have a solution — a *unique* solution — $a = 2k - 3 = \frac{7}{3}$, and $b = -k + 2 = \frac{4}{5}$.

4. Here we have to solve for k, ℓ the system

$$\begin{aligned} 1k + 2\ell &= a \\ 2k + 3\ell &= b \\ 3k + 1\ell &= c\,. \end{aligned}$$

The system is equivalent (after Gauss-Jordan reduction) to

$$\begin{aligned} k + 0\ell &= -3a + 2b \\ 0k + 1\ell &= 2a - b \\ 0k + 0\ell &= 7a - 5b + c\,. \end{aligned}$$

There cannot be a solution unless the third equation is not a contradiction, i.e., unless $7a - 5b + c = 0$. For any scalars a, b, c with this property, i.e., such that the point (a, b, c) lies on the plane $7x - 5y + z = 0$ through the origin in $\mathbb{R}^3$, we then take $k = -3a + 2b$ and $\ell = 2a - b$.

A.10.3 §4.5 Subspaces (conclusion)

Unsolved Problems

[1, Exercise 4.80, p. 163] (This result contains a clever trick, and is worth memorizing.) Suppose that $\mathcal{U}$ and $\mathcal{W}$ are subspaces of $\mathcal{V}$ for which $\mathcal{U} \cup \mathcal{W}$ is a subspace. Show that $\mathcal{U} \subseteq \mathcal{W}$ or $\mathcal{W} \subseteq \mathcal{U}$.

Solution: Suppose that

$$\begin{aligned} \mathcal{U} &\not\subseteq \mathcal{W} \qquad \text{and} \\ \mathcal{W} &\not\subseteq \mathcal{U}. \end{aligned}$$

Then there exist vectors $\mathbf{u}, \mathbf{w}$ such that

$$\mathbf{u} \in \mathcal{U} \tag{34}$$
$$\mathbf{u} \notin \mathcal{W} \tag{35}$$
$$\mathbf{w} \in \mathcal{W} \tag{36}$$
$$\mathbf{w} \notin \mathcal{U}. \tag{37}$$

Since $\mathcal{U}$ and $\mathcal{W}$ are subspaces, they are closed under scalar multiplication, in particular under multiplication by the scalar -1, which we know maps a vector on to its negative. It follows that

$$\begin{aligned} -\mathbf{u} &\in \mathcal{U} \\ -\mathbf{w} &\in \mathcal{W}. \end{aligned}$$

Now comes the cleverness. "Where", we ask, "is the vector $\mathbf{u} + \mathbf{w}$"? Since

$$\begin{aligned} \mathbf{u} &\in \mathcal{U} \subseteq \mathcal{U} \cup \mathcal{W} \qquad \text{and} \\ \mathbf{w} &\in \mathcal{W} \subseteq \mathcal{U} \cup \mathcal{W}, \end{aligned}$$

and since $\mathcal{U} \cup \mathcal{W}$ is a subspace, their sum must also be in this union. But the union consists of all points in one or the other of $\mathcal{U}$ or $\mathcal{W}$. If $\mathbf{u} + \mathbf{w} \in \mathcal{U}$, then the closure of $\mathcal{U}$ implies that $(-\mathbf{u}) + (\mathbf{u} + \mathbf{w}) = \mathbf{w}$ is also in this subspace, which is a contradiction; similarly, if $\mathbf{u} + \mathbf{w} \in \mathcal{W}$, then the closure of $\mathcal{W}$ implies that $(\mathbf{u} + \mathbf{w}) + (-\mathbf{w}) = \mathbf{u}$ is also in this subspace. Both of these consequences contradict our earlier assumptions. We conclude, by *reductio ad absurdum* that there was a flaw in one of our earlier hypotheses. We assumed that *both* of the statements (35), (37) should be true — i.e., that neither of the subspaces was contained in the other. Hence at least one of these is false: one of the subspaces must be contained totally in the other.

A.10.4 §4.6 Linear Spans, Row Space of a Matrix

Definition A.27 Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ be vectors in a vector space $\mathcal{V}$. We define the *(linear) span* of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, as

$$\mathrm{span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) = \{k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_n\mathbf{u}_n | k_i \in \mathbb{K}, i = 1, \ldots, n\};$$

that is, the *span* consists of all linear combinations of the given vectors. More generally, if U is a subset of $\mathcal{V}$, then $\text{span}(U)$ denotes the set of all linear combinations of vectors taken from U.

An alternative notation used by some authors is $\langle U \rangle$ instead of $\text{span}(U)$; if U is finite, its elements could also be listed between the wedge brackets.

Theorem A.34 *Let U be a subset of a vector space $\mathcal{V}$.*

1. $\text{span}(U)$ *is a subspace of $\mathcal{V}$. This subspace is sometimes called the subspace of $\mathcal{V}$* generated by *U.*

2. *If $\mathcal{W}$ is a subspace of $\mathcal{V}$ containing the elements of U, then* $\text{span}(U)$ *is a subspace of $\mathcal{W}$.*

3. $\text{span}(U)$ *is the intersection of all subspaces of $\mathcal{V}$ containing the elements of U; we may say that* $\text{span}(U)$ *is the* smallest *subspace of $\mathcal{V}$ which contains the elements of U.*

Proof:

1. **vector 0:** The vector $\mathbf{0}$ is expressible as a linear combination (with all scalars equal to 0) of vectors in U; i.e., $\mathbf{0} \in \text{span}(U)$.

   Closure under addition and scalar multiplication: Let $k_i, \ell_i (i = 1, 2, \ldots, n)$ be scalars, and $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in U$. Then two general vectors in $\text{span}(U)$ have the forms $\sum_{i=1}^{n} k_i \mathbf{u}_i$, $\sum_{i=1}^{n} \ell_i \mathbf{v}_i$; denote them respectively by $\mathbf{u}, \mathbf{v}$. If α is any scalar,

$$
\begin{aligned}
\mathbf{u} + \mathbf{v} &= \sum_{i=1}^{n} k_i \mathbf{u}_i + \sum_{i=1}^{n} \ell_i \mathbf{v}_i \\
&= \sum_{i=1}^{n} (k_i + \ell_i)\, \mathbf{u}_i \\
&\in \text{span}(U) \text{ ; and} \\
\alpha \mathbf{u} &= \alpha \sum_{i=1}^{n} k_i \mathbf{u}_i \\
&= \sum_{i=1}^{n} (\alpha k_i)\, \mathbf{u}_i \\
&\in \text{span}(U)\,.
\end{aligned}
$$

   Thus $\text{span}(U)$ is a subspace of $\mathcal{V}$.

A.11 Supplementary Notes for the Lecture of Friday, January 27th, 2006

Release Date: Friday, January 27th, 2006
Subject to further revision

A.11.1 §4.6 Linear Spans, Row Space of a Matrix (continuation)

I will not have time to prove the remaining parts of Theorem A.34 in the lectures. Remember the last part of the theorem: *the span of a set of vectors is the intersection of all subspaces containing those vectors.*

Definition A.28 Let $A = \begin{pmatrix} a_{ij} \end{pmatrix}_{i=1,\ldots,m;j=1,\ldots,n}$ be a given matrix over $\mathbb{K}$.

1. The *row space* $\mathrm{rowsp}(A)$ is defined by

$$\begin{aligned}\mathrm{rowsp}(A) &= \mathrm{span}(\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{pmatrix}, \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2n} \end{pmatrix}, \\ &\qquad \ldots, \begin{pmatrix} \ldots & \ldots & \ldots & \ldots \end{pmatrix}, \ldots, \begin{pmatrix} a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}).\end{aligned}$$

2. The *column space* is defined to be

$$\mathrm{span}\left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}\right).$$

A.11.2 §4.7 Linear Dependence and Independence

Definition A.29 1. Vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of a vector space $\mathcal{V}$ are said to be *linearly dependent* if there exist scalars $k_1, k_2, \ldots, k_n$, not all 0, such that

$$k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_n\mathbf{u}_n = \mathbf{0}.$$

2. Vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of a vector space $\mathcal{V}$ are said to be *linearly independent* if they are not linearly dependent.

3. The empty set of vectors, $\emptyset$, is defined to be linearly independent.

4. An infinite set U of vectors is said to be *linearly independent* if no finite subset of U is linearly dependent.

Theorem A.35 *1. The vector* $\mathbf{0}$ *cannot be a member of an independent set of vectors.*

2. *A set of vectors are linearly dependent if and only if one of them can be expressed as a linear combination of the others.*
3. *The set* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *are linearly independent if and only if the following implication always holds:*

$$k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_n\mathbf{u}_n = \mathbf{0} \Rightarrow k_i = 0 \quad (i = 1, \ldots, n)\,.$$

4. *The non-zero rows of a matrix in row canonical form, or a matrix in echelon form are linearly independent.*

A.11.3 Supplementary Problems

[1, Exercise 4.91, p. 165] Show that the following functions f, g, h are linearly independent:

$$\begin{aligned} f(t) &= e^t \\ g(t) &= \sin t \\ h(t) &= t^2 \end{aligned}$$

Solution: Later, in [1, §5.2], we will review carefully what we mean by a *function*. These 3 functions are all to be interpreted as having $\mathbb{R}$ as their domain; their codomains (targets) could be taken to be the same set, or possibly a restricted subset; for convenience, let's assume that they all are taken to have the same codomain: otherwise we can't consider them as being members of a vector space; it's simplest to think of the targets as all being $\mathbb{R}$. Remember the definitions of sum and scalar product of functions in $\mathbf{F}(X)$ (cf. Example A.29). We need to show that, if there are 3 scalars a, b, c such that

$$af + bg + ch = 0\,, \tag{38}$$

then $a = b = c = 0$. First let's remember the meaning of equation (38). It says that the function $af + bg + ch$ is identical to the function 0; for that to be true we need to know that the 2 functions act in exactly the same way on every point in their domain (which must be the same). There is not problem with the sameness of the domains, as all 3 functions are defined for all real numbers. Let's consider the equation at certain specific numbers t in the domain:

1. When $t = 0$,

$$\begin{aligned} (af + bg + ch)(0) &= af(0) + bg(0) + ch(0) \\ &= ae^0 + b\sin 0 + c0^0 \\ &= a + 0 + 0 = a \end{aligned}$$

and this must be equal to the value of the constant function 0 at the point 0, i.e., to 0; so we have

$$a = 0\,. \tag{39}$$

2. Equation (38) now reduces to

$$bg + ch = 0\,. \tag{40}$$

Let's take the specific instances of this equation when $t = \pm\frac{\pi}{2}$: adding the two equations gives

$$c\left(\frac{\pi^2}{2}\right) = 0\,,$$

from which we can conclude that $c = 0$.

3. This leaves us with the equation $bg = 0$, i.e., $bg(t) = 0$ for all t. If we take $t = \frac{\pi}{2}$ again, we can conclude that $b = 0$. Thus we have shown that the only linear combination of the functions which is equal to 0 is the trivial linear combination, with all scalars equal to 0: i.e., the 3 functions are linearly independent.

How did I know what values to choose for t? This is a matter of strategy: I had to select values that would produce equations that could be solved to show that the scalars are all 0. Other choices were possible. Sometimes the selection of the values may require some careful planning.

Suppose we tried to solve the same kind of problem with functions $\sin^2(t)$, $\cos^2(t)$, and 5. The solution would break down, because these functions are, in fact, linearly dependent: $1\sin^2(t) + 1\cos^2(t) + \frac{-1}{5}\cdot 5$ is a non-trivial linear combination of the functions which is always equal to 0.

Now that you have seen this solution, try to prove the following sets of functions are linearly independent:

$$\text{(i)}\ \left\{\begin{array}{rcl} f(t) & = & e^t \\ g(t) & = & \sin t \\ h(t) & = & t \end{array}\right\} \qquad \text{(ii)}\ \left\{\begin{array}{rcl} f(t) & = & e^t \\ g(t) & = & e^{2t} \\ h(t) & = & e^{3t} \end{array}\right\} \qquad \text{(iii)}\ \left\{\begin{array}{rcl} f(t) & = & \sin 2t \\ g(t) & = & \sin t \\ h(t) & = & \sin 3t \end{array}\right\}$$

A.12 Supplementary Notes for the Lecture of Monday, January 30th, 2006

Release Date: Monday, January 30th, 2006
Subject to further revision

A.12.1 §4.6 Linear Spans, Row Space of a Matrix (conclusion)

Recall Definition A.28 of the row and column spaces of a matrix.

Theorem A.36 *1. Row operations on a matrix do not change its row space.*

2. *Column operations on a matrix do not change its column space.*

3. *Every matrix is row equivalent to a unique matrix in row canonical form.*

4. *Every matrix is column equivalent to a unique matrix in column canonical form.*

When we row reduce a matrix, we are simply replacing one set of vectors spanning its row space, by a very special spanning set.

A.12.2 §4.8 Basis and Dimension

Definition A.30 A *basis*[34] for a vector space $\mathcal{V}$ is a linearly independent set of vectors whose span is $\mathcal{V}$.

We begin by investigating properties of bases as *sets* of vectors. For most of this course we will be interested in finite bases that have been assigned a specific *order*; the properties of *ordered bases* or *coordinate systems* will be investigated starting on page 1068 below, beginning with §A.32 of these notes.

Theorem A.37 *1. If a subset U of $\mathcal{V}$ is a basis, then every vector in $\mathcal{V}$ can be written in an unique way as a linear combination of vectors in U.*

2. *(Steinitz "Exchange" Lemma) Suppose that $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ are a linearly independent set of vectors in* $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$. *Then $m \leq n$, and there exist vectors $\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_{n-m}}$ such that $1 \leq i_j \leq n (j = 1, 2, \ldots, n-m)$ and*
$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \operatorname{span}\left(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_{n-m}}\right) .$$

3. *If a vector space $\mathcal{V}$ contains a basis with a finite number, n, of elements, then all bases of $\mathcal{V}$ consist of exactly n elements.*

[34]The plural in English of *basis* is *bases*, pronounced "baseez", not to be confused with the English plural of the noun *base*.

Proof:

1. Suppose that $\mathbf{v} \in \mathcal{V}$ has the property that it can be written as a linear combination of two subsets of vectors in U. Without limiting generality, we can assume that both of these subsets are the same, just by including in each of the sums vectors which were previously excluded, but with scalar multipliers equal to 0. Thus we need only to consider the possibility that there exists a set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\} \subseteq U$ with the property that there exist scalars $k_1, k_2, \ldots, k_n$, $\ell_1, \ell_2, \ldots, \ell_n$ such that

$$\sum_{i=1}^{n} k_i \mathbf{u_i} = \sum_{i=0}^{n} \ell_i \mathbf{u_i} .$$

This equation implies that $\sum_{i=1}^{n} (k_i - \ell_i)\, \mathbf{u_i} = \mathbf{0}$. Since the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent, each of the scalars $k_i - \ell_i = 0$ $(i = 1, 2, \ldots, n)$; or, equivalently, $k_i = \ell_i$ $(i = 1, 2, \ldots, n)$ $\square$.

Definition A.31 The common number of basis elements is then called the *dimension* of $\mathcal{V}$, and denoted by $\dim \mathcal{V}$.

Theorem A.38

Let $\mathcal{V}$ be a vector space of finite dimension n.

1. *Any set of $n+1$ or more vectors are linearly dependent.*
2. *Any set of exactly n linearly independent vectors is a basis.*
3. *Any spanning set consisting of exactly n vectors is linearly dependent.*
4. *All subspaces of $\mathcal{V}$ have dimension not exceeding n. The only subspace of dimension n is $\mathcal{V}$ itself. The only subspace of dimension 0 is the set $\{\mathbf{0}\}$.*

Example A.39 In the vector space $\mathbb{R}_n[t]$ of polynomials of degree not exceeding n in an indeterminate t, (cf. A.27) one spanning set is the $n+1$ polynomials

$$1, t, t^2, \ldots, t^n$$

since every vector is evidently a linear combination of these vectors. Moreover,

$$\sum_{i=0}^{n} k_i t^i = \mathbf{0} \Rightarrow k_i = 0 \quad \text{for all } i = 0, 1, \ldots, n\,,$$

so the vectors are linearly independent, and constitute a basis. (This proof is not quite as trivial as it looks, as we are using two different + signs here: there is the + sign that is part of the notation for polynomials, i.e., which separates successive parts of the notation for a polynomial, and there is the + sign which denotes vector addition; they aren't the same operation at all, although we can show that there is no serious harm in confusing them.)

Example A.40 In the vector space $\mathbb{R}[t]$ of polynomials of all degrees in an indeterminate t, (cf. A.27) one spanning set is the infinitely many polynomials

$$1, t, t^2, \ldots, t^n, \ldots$$

since every vector is evidently a linear combination of these vectors. Moreover,

$$\sum_{i=0}^{n} k_i t^i = \mathbf{0} \Rightarrow k_i = 0 \quad \text{for all } i = 0, 1, \ldots, n\,,$$

so the vectors are linearly independent, and constitute a basis.[35]

Example A.41 In the vector space $\mathbb{R}_{m,n}$ of $m \times n$ real matrices, the set of matrices having a 1 as just one entry and 0's as all remaining $mn - 1$ entries evidently constitutes a spanning set which is linearly independent, i.e., a basis.

A.12.3 Supplementary Problems

[1, Exercise 4.100(a), p. 165] Find a homogeneous system of linear equations whose solution space is spanned by the following set of three vectors:

$$\begin{pmatrix} 1 \\ -2 \\ 0 \\ 3 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ -3 \\ 2 \\ 5 \\ -3 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -2 \\ 1 \\ 2 \\ -2 \end{pmatrix}.$$

Solution: The vectors are to be interpreted as being elements of the vector space $\mathbf{R}^5$. Let's denote the general point of that space by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}.$$

[35] In Definition A.29.4 on page 1058 a definition was given of what is meant by the linear independence of an infinite set of vectors: no finite subset is linearly independent. In the present case the linear dependence of a finite subset would entail that some nontrivial linear combination of powers of the indeterminate would be equal to the zero vector, i.e., to the 0 polynomial.

The solutions to this problem will not be unique, unless we ask, for example, that the matrix of coefficients be in row canonical form. We are looking for equations of the form

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = 0\,.$$

If we impose the condition that the coordinates of the three given points satisfy this equation, we obtain the system of equations

$$\begin{aligned} 1a_1 - 2a_2 + 0a_3 + 3a_4 - 1a_5 &= 0 \\ 2a_1 - 3a_2 + 2a_3 + 5a_4 - 3a_5 &= 0 \\ 1a_1 - 2a_2 + 1a_3 + 2a_4 - 2a_5 &= 0 \end{aligned}$$

Row reduction of this system yields the system with augmented matrix

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 0 & 3 & -1 & 0 \\ 0 & 1 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{array}\right).$$

We could work with this matrix in its present form; however, to make the situation more transparent, I will fully reduce the matrix by back substitution, i.e., by subtracting multiples of latter rows from earlier ones, obtaining the following system of equations which is equivalent to the original one:

$$\begin{aligned} 1a_1 + 0a_2 + 0a_3 + 5a_4 + 1a_5 &= 0 \\ 0a_1 + 1a_2 + 0a_3 + 1a_4 + 1a_5 &= 0 \\ 0a_1 + 0a_2 + 1a_3 - 1a_4 - 1a_5 &= 0\,. \end{aligned}$$

Finally, I will transpose to the right sides of the equations all terms involving the non-pivot variables, obtaining

$$\begin{aligned} a_1 &= -5a_4 - 1a_5 \\ a_2 &= -1a_4 - 1a_5 \\ a_3 &= 1a_4 + 1a_5\,. \end{aligned}$$

This system is equivalent to the original system. But it is obvious that any assignment of values to a_4 and a_5, and extended to a_1, a_2, a_3 by the formulæ given by these equations, will yield a solution. Thus we see that the set of solutions to the original system are precisely the vectors

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = a_4 \begin{pmatrix} -5 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + a_5 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Indeed, we can see that the solutions are all linear combinations of the column vectors on the right! Two equations satisfied by the points are, therefore,

$$\begin{aligned} -5x_1 - 1x_2 + 1x_3 + 1x_4 + 0x_5 &= 0 \\ -1x_1 - 1x_2 + 1x_3 + 0x_4 + 1x_5 &= 0 \end{aligned}$$

All other equations will be linear combinations of these 2: these form a *basis* for the vector space of equations which are satisfied by the given points.

A.13 Supplementary Notes for the Lecture of Friday, February 3rd, 2006

Release Date: Friday, February 3rd, 2006
Subject to further revision

<u>Date of Second Class Test</u>

The date of the second class test, previously announced as Wednesday, 01 March, 2006, will be changed tentatively to

Wednesday, 08 March, 2006.

If there is a substantial majority of those present at a class willing to shift the date, we can consider such a further change.

A.13.1 Supplementary Problems

[1, Exercise 4.101, p. 165] Determine whether each of the following is a basis of the vector space $\mathbb{R}_n[t]$:

1. $1,\ 1+t,\ 1+t+t^2,\ 1+t+t^2+t^3,\ \ldots,\ 1+t+t^2+\cdots+t^n$
2. $1+t,\ t+t^2,\ t^2+t^3,\ \ldots,\ t^{n-1}+t^n$

<u>Solution</u>:

1. Define $\mathbf{u}_i = 1+t+\cdots+t^i$ $(i = 0, 1, \ldots, n)$. Suppose that a linear combination

$$\sum_{i=0}^{n} k_i \mathbf{u}_i = \mathbf{0}\,.$$

Collecting powers of t according to the definition of vector addition in this vector space, we find that this equation implies that

$$\left(\sum_{i=0}^{n} k_i\right) t^0 + \left(\sum_{i=1}^{n} k_i\right) t^1 + \left(\sum_{i=2}^{n} k_i\right) t^2 + \cdots + \left(\sum_{i=n}^{n} k_i\right) t^n$$

is the zero polynomial, i.e., that all of its coefficients are 0, so

$$\sum_{i=0}^{n} k_i = \sum_{i=1}^{n} k_i = \sum_{i=2}^{n} k_i = \cdots = \sum_{i=n}^{n} k_i = 0\,.$$

Working from right to left in this sequence of equations, i.e., using back substitution, we find that these $n+1$ equations are equivalent to

$$k_0 = k_1 = \cdots = k_n = 0$$

so the given $n+1$ vectors are linearly independent. Since the dimension of the space is precisely $n+1$, we may conclude that these $n+1$ vectors constitute a basis.

But this last observation is a consequence of the application of a theorem. So let's prove by "first principles" that these vectors form a basis. We have already shown they are linearly independent. A basis is a "linearly independent spanning set", so what remains is to show that these vectors span $\mathbb{R}_n[t]$. To show that they span the space, take a general vector in the space,

$$\mathbf{v} = a_0 t^0 + a_1 t^1 + a_2 t^2 + \cdots + a_n t^n$$

and ask whether there exist scalars k_0, k_1, ..., k_n such that

$$\sum_{i=0}^{n} k_i \mathbf{u}_i = \mathbf{v}\,.$$

This vector equation says that the polynomial on the left is equal to the polynomial $\mathbf{v}$, which *means* that the coefficients of corresponding powers of t are all equal. That entails $n+1$ equations:

$$\begin{array}{rcl} k_0 + k_1 + k_2 + \cdots + k_n &=& a_0 \\ k_1 + k_2 + \cdots + k_n &=& a_1 \\ k_2 + \cdots + k_n &=& a_2 \\ \ldots &=& \ldots \\ k_n &=& a_n \end{array}$$

which is equivalent (under elementary operations) to

$$\begin{array}{rcl} k_0 &=& a_0 - a_1 \\ k_1 &=& a_1 - a_2 \\ k_2 &=& a_2 - a_3 \\ \ldots &=& \ldots \\ k_{n-1} &=& a_{n-1} - a_n \\ k_n &=& a_n. \end{array}$$

What we want from this system is not the explicit solution, but rather the fact that there *exists* a solution. We have shown that every vector in the space is a linear combination of the given vectors, i.e., that they *span* the space. Combined with the earlier proof of linear independence, we have thus proved that the given $n+1$ vectors constitute a basis.

2. This set is certainly not a basis, since it contains only n vectors, and the dimension is $n+1$. Let's observe, however, that the vectors are linearly independent. For, if we had scalars $k_1, k_2, \ldots, k_n$ such that

$$k_1(1+t) + k_2(t+t^2) + k_3(t^2+t^3) + \cdots + k_n(t^{n-1}+t^n) = \mathbf{0}\,,$$

this would imply that

$$k_1 1 + (k_1+k_2)\,t^1 + (k_2+k_3)\,t^2 + (k_3+k_4)\,t^3 + \cdots + (k_{n-1}+k_n)t^n$$

is the zero polynomial, i.e., that all of its coefficients are 0, implying that

$$k_1 = k_2 = \cdots = k_n = 0\,.$$

We could extend this set by one suitable vector to obtain a basis; for example, adjoin t^n; but, for example, the vector $(1+t)^2$ would not work, since it is a linear combination of the first two vectors given.

A.13.2 §4.9 Application to Matrices, Rank of a Matrix

Definition A.32 The *rank* of a matrix A, denoted by $\operatorname{rank}(A)$, is defined to be $\dim(\operatorname{rowsp}(A))$.

Theorem A.42 *Let A be an $m\times n$ matrix, and E an elementary $m\times m$ matrix.*

1. $\operatorname{rank}(EA) = \operatorname{rank}(A)$.

2. *(This strengthens the preceding result.)*

$$\operatorname{rowsp}(EA) = \operatorname{rowsp}(A)$$

3.

$$\operatorname{colsp}(AE) = \operatorname{colsp}(A)$$

4. *Let B be in row-echelon or row canonical form, and $A\sim B$. Then a basis for* $\operatorname{rowsp}(A)$ *is the set of non-zero rows of B.*

5. *Suppose that a specific set of columns of A are linearly dependent. Then, after the application of any row operation to A, the same columns will be linearly dependent.*

6. $\operatorname{rank}(A) = \dim(\operatorname{colsp}(A))$.[36]

36

Proof:

1. E_1**:** This type of elementary matrix effects a reordering of the rows of A. Since the concept of rank is concerned only with the *set* of rows, not with their order, premultiplying by E_1 has no effect on the rank.

   E_2**:** This type of elementary matrix multiplies a row by a non-zero scalar. Since the question of linear independence is concerned only with the question of whether a scalar multiplier is, or is not 0, any scaling of a row can be countered by a compensating scaling of the corresponding row; so this type of operation again has no effect on rank.

   E_3**:** It suffices to consider the case where 1 times one row is added to another; to simplify the exposition, without limiting the generality, I will consider the specific case where the second row is added to the first, i.e., where

$$E_3 = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

   Denote the rows of A by $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$. Suppose that there exist scalars, $k_1, k_2, \dots, k_m$, not all zero, such that

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m = \mathbf{0}\,..$$

   Then

$$k_1(\mathbf{a}_1 + \mathbf{a}_2) + (k_2 - 1)\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m = \mathbf{0}\,,$$

   so a dependence relation holding in A is inherited by E_3A. Conversely, suppose that there exist scalars, $k_1, k_2, \dots, k_m$, not all zero, such that

$$k_1(\mathbf{a}_1 + \mathbf{a}_2) + k_2\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m = \mathbf{0}\,;$$

   then

$$k_1\mathbf{a}_1 + (k_2 + 1)\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m = \mathbf{0}\,,$$

   so a dependence relation holding in E_3A corresponds to one in A. We may conclude that the dependent sets in A always correspond to dependent sets in EA, and that the rank of a matrix is unchanged by premultiplication by an elementary matrix; or, equivalently, by any row operation.

Definition A.33 This latter quantity is sometimes called the *column rank* of A.

2. This is a strengthening of the preceding result, and I continue with the notation for naming the row vectors. The row space is spanned by $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. As observed, operation E_1 does not affect the *set* of vectors, so it cannot affect the space spanned by them. Operation E_2 scales one vector, so any linear combination expressed in terms of the rows of A can equally well be expressed in terms of the rows of E_2A. I leave it to the student to see how the addition or subtraction of one vector to/from another does not affect the subspace spanned.

3. The result can be proved analogously to the preceding.

4. B is obtained from A by premultiplying by elementary matrices. Hence it has the same row space as A. Suppose that the non-zero rows of B are $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$, where
$$\mathbf{w}_i = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \qquad (i = 1, 2, \ldots, r; r \leq m)\,.$$
Since B is in echelon form, there exists for each i an integer j_i such that $a_{i1} = a_{i2} = \cdots = a_{i,j_i-1} = 0$, $a_{ij_i} \neq 0$, $(i = 1, 2, \ldots, r)$, where $j_1 < j_2 < \ldots$. Suppose that there exist scalars $k_1, k_2, \ldots, k_r$ such that $k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_1\mathbf{w}_r = \mathbf{0}$. Then, examining the j_1th entry in each of the vectors, we see that $k_1 a_{1j_1} = 0$, implying that $k_1 = 0$. This serves as the basic case (anchor of induction) for an induction proof that all of the k_i are 0, which shows that the vectors are linearly independent.

5. Suppose that some set of columns $\#\#i_1, i_2, \ldots, i_s$ of A are linearly dependent; i.e., that there exist scalars, $x_1, \ldots, x_n$, not all 0, such that the sum of x_j times the jth column is $\mathbf{0}$; equivalently, suppose that
$$A\mathbf{x} = \mathbf{0}\,, \tag{41}$$
where $\mathbf{x}$ is interpreted as a column vector. We can prove that precisely the same vector $\mathbf{x}$ is a solution of the transformed system of equations. This is easily seen by multiplying both sides of equation (41) on the left by the corresponding elementary matrix — call it E. So we have
$$E(A\mathbf{x}) = E\mathbf{0} = \mathbf{0} \Rightarrow (EA)\mathbf{x} = \mathbf{0}\,,$$
where EA is the matrix of the transformed system, and $\mathbf{x}$ is the *same* non-zero vector giving a dependence relation. So the non-trivial dependence relations correspond to the vectors in the solution space of the corresponding system of homogeneous equations.

6. By the preceding item, the column rank is unchanged under row operations. The pivot columns of a matrix in row-echelon form are evidently linearly independent and constitute a basis for the column space. Hence the column rank is equal to the number of pivots, which, in turn, has been shown to equal the row rank.

We saw in the preceding theorem that we could transform a set of row vectors by row operations in order to obtain a basis of the space they span. But such transformations will usually alter the vectors, so that the basis obtained consists of vectors different from the ones with which we began. We have actually proved more. It is possible, given a set of vectors, to find a basis for their span consisting only of vectors in the given set. This can be done by *column reducing* the matrix of which the given vectors are the rows[37]. When one has obtained a matrix in echelon form, one set of vectors which forms a basis will be those containing the pivots. This is the content of [1, Algorithm 4.2 (Casting-out Algorithm), p. 133].

Unsolved Problems

[1, Exercise 4.107, p. 166] Determine which of the following matrices have the same row space:

$$\begin{aligned} A &= \begin{pmatrix} 1 & -2 & -1 \\ 3 & -4 & 5 \end{pmatrix} \\ B &= \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & -1 \end{pmatrix} \\ C &= \begin{pmatrix} 1 & -1 & 3 \\ 2 & -1 & 10 \\ 3 & -5 & 1 \end{pmatrix} \end{aligned}$$

<u>Solution</u>: Let us row reduce each of the matrices to row reduced form.

$$\begin{aligned} A &= \begin{pmatrix} 1 & -2 & -1 \\ 3 & -4 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 7 \\ 0 & 2 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 4 \end{pmatrix}, \\ B &= \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \\ C &= \begin{pmatrix} 1 & -1 & 3 \\ 2 & -1 & 10 \\ 3 & -5 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & 4 \\ 0 & -2 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Thus the row spaces of both A and C are generated by the same basis of rows $\begin{pmatrix} 1 & 0 & 7 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 & 4 \end{pmatrix}$; so they must be identical row spaces. The basis we have found for the row space of B is different. Does that mean that the *space* is different? One can see that the 3 row spaces all have dimension 2; but how do we know the row space of B is different from the row space of A and C? Look at the

[37]or, alternatively, row reducing a matrix formed from the given vectors as columns

vector $\begin{pmatrix} 1 & 0 & 1 \end{pmatrix}$. If this were in the row space generated by the generators we found for the row spaces of A and C, there would exist constants k, ℓ, such that

$$k\begin{pmatrix} 1 & 0 & 7 \end{pmatrix} + \ell\begin{pmatrix} 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}.$$

This equation is equivalent to the system

$$\begin{aligned} k + \ell &= 1 \\ \ell &= 0 \\ 7k + 4\ell &= 1 \end{aligned}$$

which has no solutions. Since there exists a vector in the row space of B which is not in the row spaces of A, C, the row spaces are different.

We now have the vector space machinery available to simplify the reading of the sections we missed in Chapter 3.

A.13.3 Summary of Chapter 4 (IN PROGRESS)

Following is a brief, simplified list of things you need to know from Chapter 4.

Definitions

- *vector space* (8 axioms)
- *linear combination*
- *span, spanning set*
- *subspace*
- *row space* and *column space* of a matrix
- *linear dependence*
- *linear independence*
- *basis* = linearly independent spanning set
- *dimension* = number of vectors in any basis
- *standard basis of* $\mathbb{K}^n$: n vectors whose entries are all but one of them equal to 0, and the remaining entry is 1.
- *coordinate systems* STILL TO BE DISCUSSED
- *coordinates* STILL TO BE DISCUSSED

Specific Examples

- definitions of $\mathbb{K}^n$, $\mathbb{K}[t]$, $\mathbb{K}_n[t]$, $\mathbb{K}_{m,n}$, $\mathbf{F}(X)$ for all of which you should know the definitions of vector addition, scalar multiplication, the vector $\mathbf{0}$, and a convenient basis. (We have not yet formally considered the *standard basis* of $\mathbb{K}^n$, which consists of the n vectors having all but one entry equal to 0, and the remaining entry equal to 1.

Important Results

- equivalent properties to linear independence, and to linear dependence (e.g., that some vector of the set is a linear combination of the others)
- properties of bases:
  - the original definition (linearly independent spanning set)
  - all bases have the same number of elements (=dimension)
  - a set of linearly independent vectors which is not a basis can be augmented by a vector outside of their span to form a larger linearly independent set
  - any set of vectors whose number equals the dimension spans the entire vector space
  - any set of vectors whose number exceeds the dimension of the space must be linearly dependent
- a linearly independent set cannot contain $\mathbf{0}$
- review of solution of systems of linear equations in the light of Chapter 4 STILL TO BE COMPLETED
- isomorphism of vector spaces over $\mathbb{K}$ with some $\mathbb{K}^n$ STILL TO BE DISCUSSED

Types of Problems you should be able to solve In the Supplementary Problems you should be able to solve the following numbers (all preceded by 4.); some of the problems not shown in the list have already been solved in class: 71, 75, 77, 83, 84, 89, 90, 91, 92, 93, 97, 98, 99, 100, 101, 102, 103, 107, 108, 109, 110, 113; 128 - 133 (not covered yet). For many of these problems there is a worked problem which will suggest how to solve it. If you can't find such, and can't solve a particular kind of problem, ask about it.

A.14 Supplementary Notes for the Lecture of Monday, February 6th, 2006

Release Date: Monday, February 6th, 2006
Subject to further revision

A.14.1 Results of the First Class Test

Solutions to the problems have been mounted on the Web in the present document, beginning on page 46. A discussion of the significance of the grade appears in that location. The graded papers will be returned at today's lecture.

A.14.2 WeBWorK assignments WW_5 and WW_6

Following are some clarifications of issues raised by the notation and terminology of this assignment and its successor. There could be other issues: the systems manager is not able to supply me with copies of all possible versions of all problems.

- The symbol $\mathbb{R}^{m\times n}$ is sometimes used where we would write $\mathbb{R}_{m,n}$.
- There does not appear to be anything wrong with Problem 6: read it carefully, and remember that you are working with polynomials of the form $a_2t^2 + a_1t^1 + a_0t^0$.
- As mentioned earlier, the symbol $C^n(I)$ represents real-valued functions defined on the interval $[0, 1]$ having n derivatives, where the nth derivative is also known to be continuous.
- The symbol $M_n(\mathbb{R})$ appears to mean the vector space of all $n \times n$ matrices with coefficients taken from $\mathbb{R}$.

One problem on some students' versions refers to a real vector space $\mathbb{C}^n$. This requires some explanation. The space $\mathbb{C}^n$, when considered as a vector space over the complex field, has dimension n. However, it is also possible to consider this same set of ordered n-tuples of complex numbers as a vector space over the real numbers, in the obvious way. However, when we do this, vectors like $(i, 0, 0, \ldots, 0)$ and $(1, 0, 0, \ldots, 0)$ are no longer dependent: over $\mathbb{C}$ they are linearly dependent, because i times the first plus 1 times the second is the vector $\mathbf{0}$. But, over the reals, we don't have a scalar i available. Thus, over the reals, the space $\mathbb{C}^n$ has dimension $2n$. *I do not consider this an appropriate problem for a service course in linear algebra, and have told you the answer so that you will not be penalized on the 6th assignment.*

A.14.3 §3.8 Gaussian Elimination. Matrix Formulation

A.14.4 §3.9 Matrix Equation of a System of Linear Equations

Consider a general system of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n} & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n} & = & b_2 \\ \ldots & & \ldots & & \ldots & & \ldots & = & \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn} & = & b_m \end{array} . \tag{42}$$

The *coefficient matrix* is

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} . \tag{43}$$

The equation may be written more compactly as

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{44}$$

of m linear equations in n unknowns. Define

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{45}$$

a solution vector, and

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} . \tag{46}$$

$$M = \left(\begin{array}{c|c} A & \mathbf{b} \end{array} \right) \tag{47}$$

is the *augmented matrix* of the system. We may also write the system as

$$M \begin{pmatrix} X \\ \hline -1 \end{pmatrix} = \mathbf{0} \tag{48}$$

Theorem A.43 *1. A solution* $\mathbf{x}$ *of (44) exists if and only if* $\mathbf{b} \in \mathrm{colsp}(A)$.

2. *A solution* $\mathbf{x}$ *of (44) exists if and only if* $\mathrm{rank}(A) = \mathrm{rank}(M)$.

3. *The solution* $\mathbf{x}$ *to (44) is unique if and only if the common value of the ranks of* A *and* M *is* n.

4. *If the field is infinite, the number of solutions, if it exceeds 1, is infinite.*

Proof:

1. The column vector $\mathbf{x}$, when postmultiplying A, creates a linear combination of the column vectors of A.

2. After row reduction of the augmented matrix it may happen that there is a pivot in the column corresponding to the augmentation column. When that happens, the row containing that pivot corresponds to an equation of the form $0 = 1$, which is a contradiction. Such a system of equations cannot have a solution, since the existence of a solution would entail the truth of this equation. We conclude that the original system had no solution.

   If, however, there is no pivot in the last column of the reduced augmentation matrix, then the variables can be partitioned into two classes: those corresponding to columns containing a pivot, and the others. Rewrite the system of equations so that the variables corresponding to the pivots appear on the left sides, and all other terms are on the right. Let now the non-pivot variables be assigned arbitrary values. Proceeding from the bottom equation to the top, determine the values that such an assignment will induce for the pivot variables, which can be considered the *dependent* variables; the other variables may be considered the *independent* variables. In this way one obtains a solution to the system.

3. The partitioning of the variables described above, into independent and dependent variables, will always produce multiple solutions, provided there is at least one independent variable. If there are no independent variables, i.e., if the number of pivot variables is equal to the total number of variables, then the solution is uniquely determined. Note that the number of equations could, however, have been larger than the number of variables.

Corollary A.44 (to Theorem A.43) *Let* $m = n$ *in (44), so that the number of equations is equal to the number of variables.*

1. *The solution to (44) is unique if and only if the coefficient matrix and augmented matrix have equal rank. In that case the coefficient matrix is invertible.*

2. *Conversely, if the coefficient matrix of a system of n linear equations is invertible, then the system has a unique solution.*

Proof:

1.

2. If A is invertible, then we can multiply both sides of the equation $A\mathbf{x} = \mathbf{b}$ on the left by A^{-1}, obtaining
$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$$
so that the solution is, indeed, unique.[38]

A.14.5 §3.10 Systems of Linear equations and Linear Combinations of Vectors

A.14.6 §3.11 Homogeneous Systems of Linear Equations

Definition A.34 A system (44) is said to be *homogeneous* if $\mathbf{b} = \mathbf{0}$, and the system may be written in the form
$$A\mathbf{x} = \mathbf{0}\,. \tag{49}$$

Theorem A.45 *The solutions of a homogeneous system (49) form a vector space, a subspace of the space $\mathbb{K}^n$ (written as column vectors).*

Proof: We observe that $A\mathbf{0} = \mathbf{0}$, and that, if $A\mathbf{x}_1 = \mathbf{0} = A\mathbf{x}_2$, then $A(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{0}$, and $A(k\mathbf{x}_1) = \mathbf{0}$, where k is any scalar. Thus the solution set satisfies the criteria for being a subspace of the vector space $\mathbb{K}^n$.

Definition A.35 We call the vector space of solutions the *solution space* or *null space* of the given homogeneous system.

Theorem A.46
1. *The solution space of a homogeneous system of equations (49) is not changed under elementary operations on the system, or, equivalently, under elementary operations on the matrix A of coefficients. Hence the matrix may be transformed into row canonical form without altering the set of solutions.*

2. *When the matrix of coefficients is in row canonical form, the non-zero rows — i.e., the rows containing pivots — represent equations which express the pivot variables as linear combinations of the non-pivot variables. Thus the variables are partitioned into two classes: the non-pivot variables, which may be thought of as* independent parameters *whose values may be chosen without restriction from the field, and the pivot variables, which may be thought of as being determined by these parameters.*

[38] We have not yet resolved the question of whether a non-square matrix can be invertible.

3. *One basis for the solution space may be obtained by assigning to all but one of the non-pivot variables the value 0, and to the pivot variables the values determined by the corresponding equations. Thus the dimension of the solution space is equal to the excess of the number of variables over the number of non-zero rows in the row canonical coefficient matrix.*

A.14.7 Inhomogeneous Systems of Linear Equations

Definition A.36 For a general linear system (44) of equations, the system $A\mathbf{x} = \mathbf{0}$ is called the *associated homogeneous system*,

Theorem A.47 *Let $\mathbf{v}_0$ be a particular solution of the system of equations (44).*

1. *Every particular solution $\mathbf{v}$ of (44) is expressible in the form $\mathbf{v} = \mathbf{v}_0 + \mathbf{x}$, where $\mathbf{x}$ is a solution of the associated homogeneous system.*

2. *If $\mathbf{x}$ is any solution of the associated homogeneous system, then $\mathbf{v} = \mathbf{v}_0 + \mathbf{x}$ is a particular solution of (44).*

3. *The solutions of (44) form a vector space if and only if (44) is homogeneous.*

A.14.8 Supplementary Problems

[1, Exercise 3.54(c), p. 112] Solve

$$\begin{aligned} x + 2y + 4z - 5t &= 3 \\ 3x - y + 5z + 2t &= 4 \\ 5x - 4y + 6z + 9t &= 2 \end{aligned}$$

Solution: Let's apply row operations to the augmented matrix,

$$
\begin{aligned}
M &= \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 3 & -1 & 5 & 2 & 4 \\ 5 & -4 & 6 & 9 & 2 \end{array}\right) \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 5 & -4 & 6 & 9 & 2 \end{array}\right) \\
&\qquad \text{under the row operation } R_2 := R_2 - 3R_1 \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 0 & -14 & -14 & 34 & -13 \end{array}\right) \\
&\qquad \text{under the row operation } R_3 := R_3 - 5R_1 \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 0 & 0 & 0 & 0 & -3 \end{array}\right) \\
&\qquad \text{under the row operation } R_3 := R_3 - 2R_2
\end{aligned}
$$

This tells us that the original system has the same solutions as the system of equations corresponding to the last augmented matrix above. But the third equation in that system is $0 = -3$, which is a contradiction. From this contradiction we conclude that this last system has no solutions, and so the original system also had no solutions.

Let's modify the problem: Solve

$$
\begin{aligned}
x + 2y + 4z - 5t &= 3 \\
3x - y + 5z + 2t &= 4 \\
5x - 4y + 6z + 10t &= 2
\end{aligned}
$$

Solution: Let's apply row operations to the augmented matrix,

$$\begin{aligned}
M &= \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 3 & -1 & 5 & 2 & 4 \\ 5 & -4 & 6 & 10 & 2 \end{array}\right) \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 5 & -4 & 6 & 10 & 2 \end{array}\right) \\
&\qquad \text{under the row operation } R_2 := R_2 - 3R_1 \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 0 & -14 & -14 & 35 & -13 \end{array}\right) \\
&\qquad \text{under the row operation } R_3 := R_3 - 5R_1 \\
&\sim \left(\begin{array}{rrrr|r} 1 & 2 & 4 & -5 & 3 \\ 0 & -7 & -7 & 17 & -5 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right) \\
&\qquad \text{under the row operation } R_3 := R_3 - 2R_2
\end{aligned}$$

This system can be solved by *back substitution*: the last equation gives the value of t uniquely as $t = -3$. This value can be substituted into the preceding equation, which can first be simplified by being multiplied by $-\frac{1}{7}$:

$$y + z - \frac{17}{7}t = \frac{5}{7}$$

so

$$y = -z + \frac{17}{7} \cdot (-3) + \frac{5}{7} = -z - \frac{46}{7}$$

where z may be given any real value. Finally these values of y, z, t may be substituted into

$$x = -2y - 4z + 5t + 3$$

to yield

$$x = -2\left(-z - \frac{46}{7}\right) - 4z + 5\,(-3) + 3 = -2z + \frac{8}{7}$$

so we have a the general solution,

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = z \begin{pmatrix} -2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{8}{7} \\ -\frac{46}{7} \\ 0 \\ -3 \end{pmatrix}$$

where z is serving as a parameter: the solutions are points on a line in 4-dimensional space. One point on this line is found by taking $z = 0$: $\left(\frac{8}{7}, -\frac{46}{7}, 0, -3\right)$, which is a *particular* solution of the system. The associated homogeneous system has a 1-dimensional solution space, spanned by the vector $(-2, -1, 1 - 0)$. This line of solutions is the intersection of three hyperplanes. In the original system there were also 3 hyperplanes; but the system was "overdetermined": the three hyperplanes did not intersect in any common point.

[1, Exercise 3.60(a), p. 112] Find the dimension and a basis of the general solution W of (the) homogeneous system

$$\begin{aligned} x + 3y + 2z - s - t &= 0 \\ 2x + 6y + 5z + s - t &= 0 \\ 5x + 15y + 12z + s - 3t &= 0 \end{aligned}$$

Solution: We shall reduce the augmented matrix to Gaussian-reduced form:

$$\left(\begin{array}{rrrrr|r} 1 & 3 & 2 & -1 & -1 & 0 \\ 2 & 6 & 5 & 1 & -1 & 0 \\ 5 & 15 & 12 & 1 & -3 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{rrrrr|r} 1 & 3 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 2 & 6 & 2 & 0 \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{rrrrr|r} 1 & 3 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The system of equations which corresponds to this Gaussian-reduced matrix has the same solution set as the original system, and can be solved by *backward substitution*; we write by moving to the right side of the equal signs all terms except those corresponding to the *pivots*:

$$x = 0 - 3y - 2z + s + t \tag{50}$$
$$z = 0 - 3s - t \tag{51}$$

For arbitrary values of s and t we obtain the value of z from equation (51), and, substituting in equation (50), obtain the value of x for any value of the third 'independent" variable, y:

$$x = -3y - 2(-3s - t) + s + t = -3y + 7s + 3t$$

This is *backward substitution* — working upwards from the equation corresponding to the last non-zero row in the matrix to the first row. This solution could be obtained directly by carrying the row reduction on through *Gauss-Jordan* reduction to

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & -1 & 0 \\ 2 & 6 & 5 & 1 & -1 & 0 \\ 5 & 15 & 12 & 1 & -3 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 2 & 6 & 2 & 0 \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{ccccc|c} 1 & 3 & 0 & -7 & -3 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

From the system

$$\begin{aligned} x &= 0 - 3y + 7s + 3t \\ z &= 0 - 3s - t \end{aligned}$$

we obtain that

$$\begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} = \begin{pmatrix} -3y + 7s + 3t \\ y \\ -3s - t \\ s \\ t \end{pmatrix} = y \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 7 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

We have thus expressed every solution as a *linear combination* of the three vectors

$$\left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Thus these three vectors *generate* or *span* the space of solutions to this system of homogeneous equations. These vectors constitute a *basis* for the space, which has *dimension* 3.

A.14.9 §3.13 *LU* decomposition

Omit this section for the present.

A.14.10 §4.10 Sums and Direct Sums

Omit this section for the present

A.15 Supplementary Notes for the Lecture of Wednesday, February 8th, 2006

Release Date: Wednesday, February 8th, 2006
Subject to further revision

A.15.1 §4.11 Coordinates

Please read these notes carefully, as they differ slightly from the treatment in the textbook.

Remember that our definition of a *basis* of a vector space was a *set* of vectors with certain properties. I will be distinguishing here between a *set* and an *ordered* set, wherein the elements are written in a certain sequential order.

Definition A.37 An *ordered basis* for a vector space $\mathcal{V}$ of finite dimension n is an ordered sequence of basis elements for $\mathcal{V}$. Usually we indicate the order by subscripts 1, 2, ..., $\dim \mathcal{V}$, so an ordered basis has the form

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \qquad \text{or} \qquad \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}\,.$$

But note well: I am talking about a *sequence*, not a *set*.

We saw in Theorem A.37.1 on page 1061 of these notes that every vector is expressible in an unique way in terms of the vectors of a basis. That unique expression involved the association of a specific scalar with each of the basis vectors. If we now *order* the basis vectors, then the scalars become an *ordered n-tuple* of scalars, which we may write as either a row or column vector of $\mathbb{K}^n$. Thus the choice of an *ordered basis* creates a correspondence between $\mathcal{V}$ and the points of the vector space $\mathbb{K}^n$. We can show that this correspondence is, in fact, what algebraists call an *isomorphism* because it agrees with vector addition and scalar multiplication: if one adds vectors and then finds their coordinates, one obtains the same results as if one first finds their coordinates and then adds the coordinates as ordered n-tuples; ditto for scalar multiples. Thus the selection of an ordered basis permits one to study $\mathcal{V}$ as though it was the vector space of ordered n-tuples of field elements.

Definition A.38 The *standard basis* for $\mathbb{K}^n$ is the set of vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{e}_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

The *standard coordinate system* will be these basis vectors ordered according to their subscripts. Frequently we speak of the *standard basis* when we actually mean the *standard coordinate system*.

A.15.2 Unsolved problems

[1, Exercise 4.130(a), p. 167] Show that

$$\{t^3+t^2, t^2+t, t+1, 1\}$$

is a basis for $\mathbb{R}^3[t]$. Then, for the vector $\mathbf{v} = 2t^3+t^2-4t+2$, find the coordinate vector

$$[\mathbf{v}]_{\{t^3+t^2,t^2+t,t+1,1\}}\,.$$

Solution:

1. Suppose that the vector $\mathbf{0}$ is expressed as a linear combination of the given vectors:

   $$a(t^3+t^2)+b(t^2+t)+c(t+1)+d(1) = \mathbf{0} = 0t^3+0t^2+0t+0\,,$$

   where a, b, c, d are scalars. The meaning of the sum on the left is the polynomial obtained by collecting together all terms in the various powers of the indeterminate t, i.e.,

   $$at^3+(a+b)t^2+(b+c)t+(c+d)\,.$$

   This polynomial must be the 0 polynomial, i.e., the coefficients of the various powers of t must each be 0. This gives rise to a system of 4 linear equations,

   $$\begin{aligned} a &= 0\\ a+b &= 0\\ b+c &= 0\\ c+d &= 0\,, \end{aligned}$$

   which we can solve in the usual way to show it is equivalent to $a=b=c=d=0$. We have proved that the only linear combination of the given vectors equal to the vector $\mathbf{0}$ is the trivial linear combination: this is precisely the meaning of linear independence.

2. The dimension of the space of polynomials of degree at most 3 is known to be 4, since one basis is $\{1, t, t^2, t^3\}$. Any other set of 4 linearly independent vectors is also a basis!

3. We wish to express the given vector $\mathbf{v}$ as a linear combination of the 4 given basis vectors, i.e., to determine scalars a, b, c, d such that

$$a(t^3+t^2)+b(t^2+t)+c(t+1)+d(1)=\mathbf{v}=2t^3+1t^2-4t+2\,,$$

which is equivalent to the system of linear equations, $a=2$, $a+b=1$, $b+c=-4$, $c+d=2$, whose augmented matrix we proceed to row reduce:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -4 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right) \quad \text{reduces to} \quad \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right)$$

from which we read that

$$[\mathbf{v}]=\left(\begin{array}{r} 2 \\ -1 \\ -3 \\ 5 \end{array}\right).$$

We can verify that

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

as required. Evidently the solution given in [1, p. 169] is incorrect!

[1, Exercise 4.131(b), p. 168] Let $\mathcal{V}=\mathbb{R}_{2,2}$. Find the coordinate vector of the matrix $M=\left(\begin{array}{cc} a & b \\ c & d \end{array}\right)$ referred to the coordinate system $M_1=\left(\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right)$, $M_2=\left(\begin{array}{cc} 1 & -1 \\ 1 & 0 \end{array}\right)$, $M_3=\left(\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right)$, $M_4=\left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right)$.

Solution: The objective is to determine scalars x_1, x_2, x_3, x_4 such that $x_1M_1+x_2M_2+x_3M_3+x_4M_4=M$ where $M=\left(\begin{array}{cc} a & b \\ c & d \end{array}\right)$, i.e.,

$$\left(\begin{array}{cc} x_1+x_2+x_3+x_4 & x_1-x_2+x_3 \\ x_1+x_2 & x_1 \end{array}\right)=\left(\begin{array}{cc} a & b \\ c & d \end{array}\right),$$

equivalently the 4 scalar equations

$$\begin{array}{rcl} +1x_1+1x_2+1x_3+1x_4 & = & a \\ +1x_1-1x_2+1x_3+0a_4 & = & b \\ +1x_1+1x_2+0x_3+0x_4 & = & c \\ +1x_1+0x_2+0x_3+0x_4 & = & a \end{array}$$

whose augmented matrix is

$$\left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & a \\ 1 & -1 & 1 & 0 & b \\ 1 & 1 & 0 & 0 & c \\ 1 & 0 & 0 & 0 & d \end{array}\right)$$

and which may be row reduced to

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & d \\ 0 & 1 & 0 & 0 & c-d \\ 0 & 0 & 1 & 0 & b+c-2d \\ 0 & 0 & 0 & 1 & a-b-2c+2d \end{array}\right) .$$

(When $a = b = c = d$, this shows that the only linear combination of these matrices which equals the matrix 0 is the trivial linear combination: i.e., it shows that the 4 given matrices are linearly independent.) For general a, b, c, d these equations show how every matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be expressed (uniquely) as a linear combination; thus it shows that the given matrices span the space of all 2×2 matrices. More precisely, we have shown that

$$[M]_{\{M_1,M_2,M_3,M_4\}} = \begin{pmatrix} d \\ c-d \\ b+c-2d \\ a-b-2c+2d \end{pmatrix} .$$

A.16 Supplementary Notes for the Lecture of Friday, February 10th, 2006

Release Date: Friday, February 10th, 2006
Subject to further revision

Two specific facts from Chapter 4

1. When two vectors form a linearly dependent set, then one of them is a scalar multiple of the other.

2. [1, Theorem 4.1(iii)] If $k\mathbf{v} = \mathbf{0}$, then either $k = 0$ or $\mathbf{v} = \mathbf{0}$. This fact looks convincing enough, and you should remember it, and need not memorize the proof. However, you should keep in mind that, in certain environments, a product of two quantities can be zero even if neither of the factors is zero. One such example is square matrices with side of length at least 2. For example

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

   We call the factors on the left of this equation *zero divisors*, since they can appear as factors in a zero-product entirely composed of non-zero factors. No field can contain zero divisors, but they can exist in certain other algebraic structures.

First Written Assignment A preliminary version of this assignment has been posted in these notes, beginning on page 67. In the present form the only problems you are ready to consider are ##1, 2(a,b,c), 3(a,b), 4(a,c).

CHAPTER 5 — Linear Mappings

A.16.1 §5.1 Introduction

A.16.2 §5.2 Mappings, Functions

I will adopt the notation of your book. I begin with standard definitions and theorems.

Definition A.39 Let U and V be sets.

1. A *function* or *mapping* $f : U \to V$ or $X \xrightarrow{f} Y$ is a subset of the cartesian product $U \times V$ in which there is exactly one element (u, v) for every $u \in U$. This subset of $U \times V$ is also called the *graph* of the function f.
2. The element $v \in V$ which appears in that unique ordered pair (u, v) associated with a particular $u \in U$ is denoted by $f(u)$ and called the *image of u under the function f*. In some contexts we may use alternative notations, like uf or u^f or fu to denote the image.
3. U is called the *domain* of f, and
4. V is called either the *target* of f or the *codomain* of f.
5. For any $y \in V$, $f^{-1}(y)$ denotes the set
$$\{x : f(x) = y\}$$
and is called the *inverse image* or *preimage* of y under f.
6. Sometimes we may indication the "action" of a function $f : U \to V$ by writing $f(u) =$ a specific formula in terms of u. We may also denote this by a statement of the form $u \mapsto f(u)$, where $f(u)$ is a formula which indicates how the image is determined.
7. The set $\{f(u) : u \in U\}$ is called the *image* of f.[39]
8. If $f : U \to V$ and $g : V \to W$ are functions, then the composition $g \circ f : U \to W$ is defined[40] by $u \to g(f(u))$. This definition is applied precisely when the target of the first function applied is equal to the domain of the second.

Note that, since functions are defined as subsets of a cartesian product, functions $f : U \to V$ and $g : U' \to V'$ are *equal* (written $f = g$) precisely when the sets are equal. This means that three conditions have to hold:

1. The domains, U and U' have to be equal as sets.
2. The targets, V and V' have to be equal as sets.

[39] Some authors use the word *range* for the image; I try to avoid that word, because some authors use the word for the codomain.

[40] Note that this definition is complete: it tells you the domain, U, the codomain W, and the action of the function on any element of the domain.

3. The functions have to have precisely the same "actions" on their common domain. This means that, for every $u \in U$,

$$f(u) = g(u). \tag{52}$$

If condition (52) fails to hold for even one element $u_0 \in U$, then we say that the functions are not equal, and write $f \neq g$. In practice we often dispense with the checking of the equality of the domains or of the targets; in particular, the distinction between two functions which are equal in every other way except that one may have a larger codomain than the other is often ignored.

Definition A.40 Let A be any $m \times n$ matrix over a field $\mathbb{K}$. The function $\mathbf{F}_A : \mathbb{K}^n \to \mathbb{K}^m$ is defined by $\mathbf{F}_A(\mathbf{u}) = A\mathbf{u}$, where $\mathbf{u} \in \mathbb{K}^n$. In practice the function $\mathbf{F}_A$ may be denoted simply by A. (Note that, for this definition to make sense, $\mathbf{u}$ must be interpreted as a *column* vector, i.e., an $n \times 1$ matrix.)

Theorem A.48 *Function composition is associative. That is, if $f : U \to V$, $g : V \to W$, and $h : W \to X$ are functions, then*

$$(h \circ g) \circ f = h \circ (g \circ f)$$

<u>Proof</u>: To prove this equality we need first to observe that the domains of the two iterated compositions are both U, and the codomains are both X. We need only prove that the two functions have the same "actions". Note also that each of the functions in the compositions has a target which is identical to the "next" function in the composition, so the compositions are all defined.[41] Let u be any element of U. Then

$$\begin{aligned}
((h \circ g) \circ f)(u) &= (h \circ g)(f(u)) \\
&\qquad \text{by definition of the composition } (h \circ g)\underline{\circ}f \\
&= h(g(f(u)) \\
&\qquad \text{by definition of the composition } h\underline{\circ}g \\
&= h((g \circ f)(u)) \\
&\qquad \text{by definition of } g \circ f \\
&= (h \circ (g \circ f))(u) \\
&\qquad \text{by definition of } h\underline{\circ}(g \circ f)
\end{aligned}$$

Since we have proved the functions act in the same way on every element of their common domain, and proved that the domains and codomains are also common, we have completed the proof of equality of the functions.

[41] We can't just write one function after another and expect the composition to be meaningful.

Definition A.41 1. A function $f : U \to V$ is *injective* or *one-to-one* if distinct points of the domain are always mapped on to distinct points of the codomain. It is more practical to use, as our standard definition, the *contrapositive* of this statement:

$$f(u_1) = f(u_2) \Rightarrow u_1 = u_2$$

for all elements $u_1, u_2 \in U$. An injective function is called an *injection*.

2. A function $f : U \to V$ is *surjective* or *onto* if, for every element $v \in V$, there exists $u \in U$ such that $f(u) = v$. A surjective function is called a *surjection*.

3. A function which is both an injection and a surjection is called a *bijection* or a *one-to-one correspondence*[42], and is said to be *bijective*.

4. For any set U, the function $\mathbf{1}_U : U \to U$ is defined by $u \to u$ for all $u \in U$, and called the *identity function on* U. Where there is no danger of confusion, the subscript U may be suppressed.

5. Suppose that there exists, for a function $f : U \to V$, a function $g : V \to U$ such that $g \circ f = \mathbf{1}_U$ and $f \circ g = \mathbf{1}_V$. Then g is called *an inverse (function)* of f.

Theorem A.49 *1. Let $f : U \to V$ be any function. Then $f \circ \mathbf{1}_U = f$.*

2. Let $f : U \to V$ be any function. Then $\mathbf{1}_V \circ f = f$.

3. A function $f : U \to V$ cannot have more than one inverse function $g : V \to U$.

4. If $g : V \to U$ is the inverse function of f, then, for any $v \in V$, $f^{-1}(v) = \{g(v)\}$.

5. A function f has an inverse if and only if f is bijective.

Proof:

1. You should be able to supply an explanation for each of the following equations:

$$\text{For any } \mathbf{u} \in U, \quad \begin{aligned} (f \circ \mathbf{1}_U)(\mathbf{u}) &= f(\mathbf{1}_U(\mathbf{u})) \\ &= f(\mathbf{u}). \end{aligned}$$

Since the domains and targets of $f \circ \mathbf{1}_U$ and f are both U, this proof that they act in identical ways on an arbitrary element of the domain shows that they are the same function.

2. To be supplied by the student.

[42]It's best to avoid using this term, as the word *correspondence* must not be omitted.

3. A proof of this property can be modelled on the proof of Theorem A.10.2 given earlier for matrix inverses.

4. For any $v \in V$,

$$\begin{aligned} u \in f^{-1}(v) &\Leftrightarrow f(u) = v \\ &\Leftrightarrow g(f(u)) = g(v) \\ &\Leftrightarrow (g \circ f)(u) = g(v) \\ &\Leftrightarrow \mathbf{1}_U(u) = g(v) \\ &\Leftrightarrow u = g(v) \end{aligned}$$

   proving the equality.[43]

5. (a) Suppose that f is bijective. We need to prove the existence of a function g with the desired properties. So, given a point $v \in V$, we must define precisely what will be the image of v under this new function. Since f is surjective, there is some point in $u \in U$ which is mapped on to it; since f is injective, that point u is the only point mapped on to v. It is that point which we define to be $g(v)$. Thus this point has the property that $f(g(v)) = v$, i.e., that $(f \circ g)(v) = v$, i.e., that $(f \circ g)(v) = \mathbf{1}_V(v)$. The functions $f \circ g$ and $\mathbf{1}_V$ both have V as their domain and as their target; since we have shown that they act in exactly the same way, they are the same function, i.e., $f \circ g = \mathbf{1}_V$.

   Now take any point $u \in U$. We have just defined g so that $g(f(u))$ is the unique point mapped by f on to $f(u)$, i.e., it is u. So $g \circ f$ and $\mathbf{1}_U$ have the same action on all points in U. As the two functions have the same domain and the same codomain, they are the same function, i.e., $g \circ f = \mathbf{1}_U$.

   We have thus shown that, if f is bijective, there exists a function g with the desired properties, i.e., there exists an inverse of f.

   (b) Suppose that f has an inverse g. Let v be any point in V. Then

      i.

$$\begin{aligned} v \in V &\Rightarrow v = \mathbf{1}_V(v) \\ &\Rightarrow v = f(g(v)) \\ &\Rightarrow v \text{ is in the image of } f \end{aligned}$$

      proving that f is surjective.

[43] Two sets are equal when they have precisely the same members.

ii.

$$\begin{aligned} f(u_1) = f(u_2) &\Rightarrow g(f(u_1)) = g(f(u_2)) \\ &\Leftrightarrow (g \circ f)(u_1) = (g \circ f)(u_2) \\ &\Rightarrow \mathbf{1}_U(u_1) = \mathbf{1}_U(u_2) \\ &\Leftrightarrow u_1 = u_2 \end{aligned}$$

proving that f is injective.

Definition A.42 If $f : U \to V$ has an inverse function, that function can be denoted by f^{-1}.[44]

Theorem A.50 1. *Suppose that $f : U \to V$ is invertible. Then $f^{-1} : V \to U$ is also invertible, and $\left(f^{-1}\right)^{-1} = f$.*

2. *Suppose that $f : U \to V$ and $g : V \to W$ are both invertible. Then $g \circ f : U \to W$ is also invertible, and*
$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

3. *Any identity function is invertible, and is its own inverse.*

Supplementary Problems

[**1, Exercise 5.46(b), p. 198**] Determine the number of different mappings from $\{1, 2, \ldots, r\}$ to $\{1, 2, \ldots, s\}$

Solution: The action of a function on any point does not affect its action on any other point; the definition of *function* does not impose any constraint. Thus there are s choices for the image of point 1 in the domain; independently of that there are s choices for the action of the function on 2, s choices on 3, ..., s choices on r. The number of functions is s^r.

A.16.3 §5.3 Linear Mappings, Linear Transformations

Definition A.43 Let $\mathcal{U}$ and $\mathcal{V}$ be vector spaces over the same field $\mathbb{K}$. A function $F : \mathcal{U} \to \mathcal{V}$ is called a *linear transformation* or *linear mapping* or *homomorphism* if it has the following specific properties with respect to the operations of vector addition and scalar multiplication:

[44] Note that we have used this same symbol for another purpose above, in Definition 5. What the present definition says is that we will be identifying the unique element in $f^{-1}(v)$ with the set of which it is the unique element.

1. For any vectors $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$, $F(\mathbf{u}_1 + \mathbf{u}_2) = F(\mathbf{u}_1) + F(\mathbf{u}_2)$.
2. For any vector $\mathbf{u} \in \mathcal{U}$ and any scalar $k \in \mathbb{K}$, $F(k\mathbf{u}) = kF(\mathbf{u})$.

A linear transformation $F : \mathcal{V} \to \mathcal{V}$ is sometimes called a linear transformation *of* $\mathcal{V}$, or a *linear operator on* $\mathcal{V}$. We often suppress parentheses when writing the image of a vector under a linear transformation, writing $F\mathbf{u}$ rather than $F(\mathbf{u})$.

These two conditions look deceptively trivial. Remember that the plus signs on the two sides of the first equation are not the same addition: the same symbol is being used to denote vector addition in two different vector spaces. Similarly, the scalar products on the two sides of the second equation are different, as they are composed in different vector spaces. We have chosen to use these “confusing” conventions for notation because they cause no problems; that is, you can't go wrong if you forget where you are working.

As with earlier definitions, this definition can be simplified, in that we can combine the two conditions into a single one. This will be shown in part of the next theorem. As a result of that theorem, you may work with either version of the definition that pleases you.

A.17 Supplementary Notes for the Lecture of Monday, February 13th, 2006

Release Date: Monday, February 13th, 2006
Subject to further revision

A.17.1 Summary of Chapter 5

(subject to future revision)

Prior to taking some hard decisions about how to cope with the current "textbook", I prepared some elaborate notes on the chapter, and I don't wish to destroy them; so you will find some detailed, theoretical material below involving proofs of theorems, some of which will not be proved in the lecture room. You should understand the proofs, but the proving of theorems is not the main goal of this course. I would emphasize understanding the concepts, and solving problems; some problems will have general variables named instead of specific values. Don't be frightened by these — just treat them as numbers whose identity has not happened to have been exposed; the methods are the same.

Here is a list of things you should know from Chapter 5:

Definitions

- terminology of elementary set theory and functions (much of this should be review)
  - *function (=mapping), domain, target (=codomain)*
  - *identity* $\mathbf{1}_U$
  - *composition of functions*, *one-to-one (=injective) functions*, *onto =surjective functions*, *one-to-one correspondences =bijections*, *inverse functions (defined as pairs whose compositions are the identity)*
- *linear transformations*
- when the target has an element 0 (as is the case for linear transformations), the *constant* function (denoted by 0) maps all points on to 0
- *kernel* and *image*
- *rank*=dim(*image*), *nullity*=dim(*kernel*),

Specific Examples

- $\mathbf{1}_{\mathcal{U}}$
- (to be explained in [1, Chapter 6]) linear transformations as represented by matrices
- differentiation of polynomials ia a linear transformation

Important Results

- A function is invertible (i.e., possesses an inverse) iff it is both injective and surjective.
- A linear transformation is determined by its action on a spanning set of the domain.
- Any arbitrary assignment of the vectors in a basis of $\mathcal{U}$ to vectors in a vector space $\mathcal{V}$ can be extended to a linear transformation from $\mathcal{U}$ to $\mathcal{V}$.
- Any vector space $\mathcal{V}$ of dimension n over a field $\mathbb{K}$ is isomorphic to the space $\mathbb{K}^n$.

Types of Problems You Should Be Able to Solve In the Supplementary Problems (numbers all preceded by 5.):

- 45 – 48 drill on the concepts of *function*, *inverse*, and the *identity* function
- 49 – 53; 55; 60;
- 61 – 69 (but some aspects of *matrix mappings* will not be fully clarified until [1, Chapter 6]);70; 72; 74; 76
- Some of the later problems will become relevant as we return to the latter parts of the Chapter as the course progresses

A.17.2 §5.3 Linear Mappings, Linear Transformations (conclusion)

Theorem A.51 *1. Let $F : \mathcal{U} \to \mathcal{V}$ be a linear transformation. Then $F(\mathbf{0}) = \mathbf{0}$.*

2. Let $\mathcal{V}$ be a vector space over some field. Then the identity function $\mathbf{1}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}$ is a linear transformation.

3. Let $\mathcal{U}, \mathcal{V}$ be given vector spaces. Then the function defined by

$$\mathbf{u} \mapsto \mathbf{0} \tag{53}$$

is a linear transformation.[45]

45

4. *Let $F : \mathcal{U} \to \mathcal{V}$ be a linear transformation. Then, for any vectors $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$, and any scalars $k_1, k_2 \in \mathbb{K}$,*

$$F(k_1\mathbf{u}_1 + k_2\mathbf{u}_2) = k_1F(\mathbf{u}_1) + k_2F(\mathbf{u}_2)\,. \tag{54}$$

5. *Let $F : \mathcal{U} \to \mathcal{V}$ be a function such that (54) holds for any vectors $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$, and any scalars $k_1, k_2 \in \mathbb{K}$. Then F is a linear transformation. Thus this general condition may be used as an alternative definition.*

Another way of describing the condition(s) we have used to define linear transformations is to say that *a linear transformation is a function that "preserves" the structure of a vector space*.

Some examples of linear transformations We have already shown that $\mathbf{1}_{\mathcal{U}}$ and $\mathbf{0}$ are linear transformations. Here we give several other important examples and counterexamples.

Example A.52 1. Suppose a function $f \in \mathbf{F}(\mathbb{R})$ is a linear transformation. Then, for any real number x, $f(x) = f(x \cdot 1) = xf(1)$. Thus all linear transformations in $\mathbf{F}(\mathbb{R})$ have the form $x \mapsto cx$, where c is a constant (equal to $f(1)$). Note that this differs from the usage you have seen in your calculus courses; there we often call functions of the form $x \mapsto cx + d$ *linear* whenever $c, d \in \mathbb{R}$. Here we are requiring that $d = 0$. Algebraists sometimes call functions of the form $x \mapsto cx + d$ *affine*.

2. Consider the operation of *differentiation* of polynomials in $\mathbb{R}[t]$, which we will denote by $\mathbf{D}$. We know that

$$\mathbf{D}(t^r) = \begin{cases} rt^{r-1} & \text{if} \quad r > 0, \\ 0 & \text{if} \quad r = 0. \end{cases} \tag{55}$$

Prove that $\mathbf{D} : \mathbb{R}[t] \to \mathbb{R}[t]$ is a linear transformation. The definition makes sense even for fields other than $\mathbb{R}$, and the function remains linear in that generality.

Supplementary Problems

[1, Exercise 5.50, p. 198] Show that each of the following mappings is not linear:

1. $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $F(x, y) = (x^2, y^2)$.
2. $F : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $F(x, y, z) = (x + 1, y + z)$.

Definition A.44 The function defined by (53) is usually denoted by $\mathbf{0}$.

3. $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $F(x, y) = (xy, y)$.
4. $F : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $F(x, y, z) = (|x|, y + z)$.

Solution: To prove that a function is not linear we must provide a counterexample to demonstrate that one of the properties of linearity fails at least once. Since these are quantified properties, we need just one instance where the failure occurs. We could just grope in the dark, and eventually we would find a counterexample. But we could be a little more creative.

1. Consider the points $(1, 0)$ and $(-1, 0)$. Then

$$\begin{aligned} F(1,0) + F(-1,0) &= (1^2, 0^2) + ((-1)^2, 0^2) \\ &= (1,0) + (1,0) \\ &= (2,0) \\ &\neq (0,0) \qquad (56) \\ &= (0^2, 0^2) \\ &= ((1 + (-1))^2, 0^2) \\ &= F((1 + (-1)), (0 + 0)) \qquad (57) \\ &= F((1,0) + (-1,0)) \end{aligned}$$

It is the inequality in line (56) which shows that the function fails to commute with addition, one of the two properties we require in a linear transformation. You should be able to explain each of the equal signs. For example, line 57 is justified by the definition of F.

2. Here the "weak" part of the definition is in the first coordinate, $x + 1$. One counterexample could come from the point $(0, 0, 0)$ in the domain. According to the definition, $(0, 0, 0) \mapsto (0 + 1, 0 + 0) = (1, 0)$. But we know that linear transformations must take $\mathbf{0}$ on to $\mathbf{0}$. This proof requires the use of a theorem, rather than demonstrating the contradiction to one of the properties in the definition. We could get another example by taking $\mathbf{u} = (0, 0, 0)$ and $k = 0$. Then

$$F(0\mathbf{0}) = F(\mathbf{0}) = (0 + 1, 0 + 0) = (1, 0) \neq (0, 0) = 0(1, 0) = 0F(\mathbf{0})$$

3. $$F(2(1, 1)) = F(2, 2) = (2 \cdot 2, 2) = (4, 2) \neq (2, 2) = 2(1, 1) = 2F(1, 1)$$

4. The images all have a first coordinate that is non-negative. We can find a counterexample using a negative scalar multiplier. Take $k = -1$, $\mathbf{u} = (1, 0, 0)$. Then $F(k\mathbf{u}) = F(-1, 0, 0) = (1, 0)$, but $kF(\mathbf{u}) = (-1)(1, 0) = (-1, 0) \neq (1, 0)$.

[1, Exercise 5.55, p. 198] Give an example of a nonlinear map $F : \mathbb{R}^2 \to \mathbb{R}^2$ such that $F^{-1}(\mathbf{0}) = \{\mathbf{0}\}$, but F is not one-to-one.

Solution: This problem is referring to the original definition of F^{-1}, not to an inverse function.[46] There are infinitely many examples. One is $F(x) = (x^2, y^2)$. We couldn't use $F(x) = (x^2, 0)$, since there the inverse of $(0, 0)$ consists of the whole y-axis.

A.17.3 A linear transformation is determined by its action on a spanning set of the domain

Theorem A.53 *Suppose that a linear transformation $F : \mathcal{U} \to \mathcal{V}$ is known to exist, and the S is a spanning set of $\mathcal{U}$. Then F is determined by its action on S; that is — once we know how F acts on all the members of S, we know how it acts on all points of the domain.*

Proof: Any point $\mathbf{x} \in \mathcal{U}$ is expressible in the form $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{v}_i$, where k_i is some scalar, and $\mathbf{v}_i$ is some member of S. By the linearity of F,

$$\begin{aligned} F(\mathbf{x}) &= F\left(\sum_{i=1}^{n} x_i \mathbf{v}_i\right) \\ &= \sum_{i=1}^{n} F(x_i \mathbf{v}_i) \\ &\quad \text{since } F \text{ commutes with addition} \\ &= \sum_{i=1}^{n} x_i F(\mathbf{v}_i) \\ &\quad \text{since } F \text{ commutes with scalar multiplication} \end{aligned}$$

and so the action of F on the generators induces its action on any vector. □While the action of a linear transformation is determined on a spanning set of the domain, you don't have the freedom to map the spanning vectors arbitrarily. For example, if you mapped all the vectors in the domain on to $\mathbf{0}$, but mapped $\mathbf{0}$ somewhere else, the mapping would not be a linear transformation. But if, instead of *any* spanning set, you took a *linearly independent* spanning set, i.e., a *basis*, the situation would change. In such a case you do have the freedom to map the basis vectors arbitrarily. I will formulate this result in the language of coordinate systems, i.e., *ordered* bases, and, for convenience, will state only a version for a finite dimensional domain.

Theorem A.54 *Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an ordered basis*[47] *for a vector space $\mathcal{V}$*

[46] More precisely, $F^{-1}(\mathbf{0})$ is the set of points in the domain which are mapped on to $\mathbf{0}$.

[47] i.e., a coordinate system

over some field $\mathbb{K}$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be any *vectors of a vector space $\mathcal{V}$. We know that every vector $\mathbf{u} \in \mathcal{U}$ can be represented uniquely by an ordered n-tuple of coordinates* $[\mathbf{u}]_S = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}$*, where*

$$\mathbf{u} = \sum_{i=1}^{n} a_i \mathbf{u}_i \,. \tag{58}$$

Then we can define a function F by $\mathbf{u} \mapsto \sum_{i=1}^{n} a_i \mathbf{v}_i$; and that well defined function is a linear transformation from $\mathcal{U}$ to $\mathcal{V}$.

Proof: Since S is a basis, every vector $\mathbf{u}$ admits a *unique* decomposition (58) as a linear combination of basis vectors. As seen in the preceding theorem, this decomposition determines completely the action of a linear transformation at the given point.

Let k_1, k_2 be arbitrary scalars. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be any vectors in $\mathcal{U}$, having coordinates $[\mathbf{x}_i]_S = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \ldots \\ a_{in} \end{pmatrix}$ $(i = 1, 2)$ respectively; i.e.,

$$\mathbf{x}_i = \sum_{j=1}^{n} a_{i,j} \mathbf{u}_j \quad (i = 1, 2) \,. \tag{59}$$

Hence

$$\begin{aligned}
F(k_1\mathbf{x}_1 + k_2\mathbf{x}_2) &= F\left(k_1\sum_{j=1}^{n} a_{1j}\mathbf{u}_j + k_2\sum_{j=1}^{n} a_{2j}\mathbf{u}_j\right)\\
&\qquad \text{by (59)}\\
&= F\left(\sum_{j=1}^{n}(k_1a_{1j} + k_2a_{2j})\,\mathbf{u}_j\right)\\
&\qquad \text{by vector space properties}\\
&= \sum_{j=1}^{n}(k_1a_{1j} + k_2a_{2j})\,\mathbf{v}_j\\
&\qquad \text{by definition of the action of } F\\
&= \sum_{j=1}^{n} k_1a_{1j}\mathbf{v}_j + \sum_{j=1}^{n} k_2a_{2j}\mathbf{v}_j\\
&\qquad \text{by commutativity and associativity of addition in } \mathcal{V}\\
&= k_1\sum_{j=1}^{n} a_{1j}\mathbf{v}_j + k_2\sum_{j=1}^{n} a_{2j}\mathbf{v}_j\\
&\qquad \text{by distributivity of scalar multiplication over } +\\
&= k_1F(\mathbf{x}_1) + k_2F(\mathbf{x}_2)\\
&\qquad \text{by definition of the action of } F
\end{aligned}$$

showing that F is a linear transformation.

Corollary A.55 (to Theorem A.54) *Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an ordered basis*[48] *for a vector space $\mathcal{V}$ over some field $\mathbb{K}$. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be the standard basis vectors of the vector space $\mathbb{K}^n$. We know that every vector $\mathbf{u} \in \mathcal{U}$ can be represented uniquely by an ordered n-tuple of coordinates $[\mathbf{u}]_S = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}$, where*

$$\mathbf{u} = \sum_{i=1}^{n} a_i\mathbf{u}_i\,. \tag{60}$$

[48] i.e., a coordinate system

Then the mapping

$$\mathbf{u} \mapsto \sum_{i=1}^{n} a_i \mathbf{e}_i = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix} = [\mathbf{u}]_S$$

which carries **u** *on to its* coordinate vector relative to the coordinate system S *is a linear transformation from* $\mathcal{U}$ *to* $\mathcal{V}$*. Moreover, this transformation is an* isomorphism *i.e., it is invertible.*

Thus any problem in any vector space of finite dimension n can be transformed into a problem in the space of ordered n-tuples of field elements.

A.17.4 §5.4 Kernel and Image of a Linear Mapping; §5.5 Singular and Non-singular Linear Transformations; Isomorphisms

Definition A.45 Let $F : \mathcal{U} \to \mathcal{V}$ be a linear transformation from vector space $\mathcal{U}$ to vector space $\mathcal{V}$. The *kernel* or *null-space* of F, denoted by $\ker(F)$, is defined to be the set of vectors mapped by F on to $\mathbf{0}$.

Theorem A.56 *Let* $F : \mathcal{U} \to \mathcal{V}$ *be a linear transformation.*

1. $\mathrm{Ker}(F)$ *is a subspace of* $\mathcal{U}$*.*
2. *The image of a* F *is a subspace of* $\mathcal{V}$*.*

Definition A.46 Let $F : \mathcal{U} \to \mathcal{V}$ be a linear transformation.

1. We denote the image of F by $\mathrm{Im}(F)$ or by $F(\mathcal{U})$.
2. The dimension of the kernel is called the *nullity* of F, and denoted by $\mathrm{nullity}(F)$.
3. The dimension of $\mathrm{Im}(F)$ is called the *rank* of F, and denoted by $\mathrm{rank}(F)$.
4. A linear transformation with kernel equal to $\{\mathbf{0}\}$ is said to be *non-singular*; otherwise the linear transformation is said to be *singular*.

Theorem A.57 *Let* $F : \mathcal{U} \to \mathcal{V}$ *be a linear transformation. Then the following conditions are equivalent.*

1. F *is injective.*
2. *The kernel of* F *is* $\{\mathbf{0}\}$*.*
3. *The nullity of* F *is 0.*

Proof: Suppose that $\ker(F) = \{\mathbf{0}\}$. Then

$$\begin{aligned} F(\mathbf{x}) = F(\mathbf{y}) &\Leftrightarrow F(\mathbf{x} - \mathbf{y}) = \mathbf{0} \\ &\Leftrightarrow \mathbf{x} - \mathbf{y} \in \ker(F) = \{\mathbf{0}\} \\ &\Leftrightarrow \mathbf{x} - \mathbf{y} = \mathbf{0} \\ &\Leftrightarrow \mathbf{x} = \mathbf{y} \end{aligned}$$

so F is injective. Conversely, if F is injective, then, since any linear transformation F has the property that $F(\mathbf{0}) = \mathbf{0}$, no other vector can be mapped on to $\mathbf{0}$, i.e., the kernel of F consists entirely of $\mathbf{0}$. This proves the equivalence of the first two statements. The third is equivalent to the second, since the only space with dimension 0 is a space containing no non-$\mathbf{0}$ vector.

Theorem A.58 *Let $F : \mathcal{U} \to \mathcal{V}$ be a linear transformation. Then the following conditions are equivalent.*

1. *F is surjective.*
2. *The image of F is $\mathcal{V}$.*
3. *When $\mathcal{V}$ has finite dimension, the rank of F is equal to the dimension of $\mathcal{V}$.*

Proof: The first two conditions are equivalent by virtue of the definition of *surjective*. The third condition is equivalent by virtue of Theorem A.37.

Example A.59 A homogeneous system of linear equations

$$A\mathbf{x} = \mathbf{0} \tag{61}$$

may be interpreted as the statement $\mathbf{x} \in F_A$, where F_A is as defined in Definition A.40.

1. The nullity of F_A is the dimension of the solution space of (61).
2. The rank of F_A is $\operatorname{rank}(A)$.

Theorem A.60 *Let $F : \mathbb{K}^n \to \mathbb{K}^m$ be a linear transformation.*

1. $\operatorname{rank}(F) + \operatorname{nullity}(F) = n$
2. *F is invertible if and only if both of the following conditions hold:*
   (a) *$m = n$; and*
   (b) $\operatorname{rank}(F) = n$*; or, equivalently* $\operatorname{nullity}(F) = 0$

Proof (I DO NOT EXPECT STUDENTS IN MATH 223 TO BE ABLE TO PROVE THIS THEOREM.):

1. [1, Problem 5.23, pp. 191-192]

2. (Incomplete portion of the proof) Suppose that F is invertible. Then there exists a *function* $G : \mathbb{K}^m \to \mathbb{K}^n$ such that $G \circ F = \mathbf{1}_{\mathbb{K}^m}$ and $F \circ G = \mathbf{1}_{\mathbb{K}^n}$. First let us prove that G is a linear transformation, since all we are assuming is that it has an inverse *function*.

$$\begin{aligned}
G(\mathbf{v}_1 + \mathbf{v}_2) &= G((F \circ G)\mathbf{v}_1 + (F \circ G)\mathbf{v}_2) \\
&\quad \text{since } F \circ G = \mathbf{1} \\
&= G(F(G(\mathbf{v}_1)) + F(G(\mathbf{v}_2)) \\
&\quad \text{definition of } \circ \\
&= G(F(G(\mathbf{v}_1) + G(\mathbf{v}_2))) \\
&\quad \text{since } F \text{ is linear} \\
&= (G \circ F)(G(\mathbf{v}_1) + G(\mathbf{v}_2)) \\
&\quad \text{definition of } G \circ F \\
&= G(\mathbf{v}_1) + G(\mathbf{v}_2) \\
&\quad \text{since } \quad G \circ F = \mathbf{1} \\
G(k\mathbf{v}) &= G(k((F \circ G)\mathbf{v})) \\
&= G(k(F(G(\mathbf{v}))) \\
&= G(F(k(G(\mathbf{v}))) \\
&= (G \circ F)(kG(\mathbf{v})) \\
&= kG(\mathbf{v})
\end{aligned}$$

proving that G has both properties required of a linear transformation. Now that we have a pair of linear transformations, we may, without limiting generality, assume that $m \le n$. The image of F is spanned by the image of the m vectors in any basis of its domain; as these image vectors must span the target, it follows that $n = m$.

A.17.5 §5.6 Operations with Linear Transformations; §5.7 The Algebra $A(\mathcal{V})$ of Linear Operators of a Vector Space $\mathcal{V}$

I will introduce ideas from these sections as they are required.

Supplementary Problems

[1, Exercise 5.63(a), p. 199] Given the matrix $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{pmatrix}$ of a linear transformation F from $\mathbb{R}^4$ to $\mathbb{R}^3$, determine a basis for each of the *image* and the *kernel* of F.

Solution: The columns of the matrix are the images, respectively, of the vectors in the standard basis. They span the image, so we can find a basis for the image by *column* reducing A. That yields the matrix $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$. We don't need to fully reduce the columns: we can read from the matrix in this state that the vectors $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ span the image: they are evidently linearly independent (since neither is a linear combination of the other), so they constitute a basis of the image of F.

Now we will row reduce the same matrix A, obtaining $\begin{pmatrix} 1 & 0 & \frac{4}{5} & -\frac{1}{5} \\ 0 & 1 & -\frac{2}{5} & \frac{3}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$. The kernel is the solution space of the system of equations corresponding to this matrix, i.e., of the system

$$\begin{aligned} x_1 &= -\frac{4}{5}x_3 - \frac{1}{5}x_4 \\ x_2 &= \frac{2}{5}x_3 + \frac{3}{5}x_4 \,. \end{aligned}$$

Thus the kernel consists of all vectors of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -4x_3 - 1x_4 \\ 2x_3 + 3x_4 \\ 5x_3 \\ 5x_4 \end{pmatrix} = \frac{x_3}{5}\begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix} + \frac{x_4}{5}\begin{pmatrix} -1 \\ 3 \\ 0 \\ 5 \end{pmatrix},$$

i.e., of the subspace of $\mathbb{R}^4$ spanned by $\begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 3 \\ 0 \\ 5 \end{pmatrix}$. This last space is the *kernel*.

Suppose that we wished to obtain, not a *basis* for the image, but one or more *equations* that characterize the image. We could simply attempt to solve the system of linear equations with *augmented* matrix $\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & x \\ 2 & -1 & 2 & -1 & y \\ 1 & -3 & 2 & -2 & z \end{array}\right)$ which will give us the conditions on x, y, z in order that a point (x, y, z) lie in the image. After row reduction we obtain the matrix $\left(\begin{array}{cccc|c} 1 & 0 & \frac{4}{5} & -\frac{1}{5} & \frac{3x+2z}{5} \\ 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{x-z}{5} \\ 0 & 0 & 0 & 0 & x-y+z \end{array}\right)$. It is the last row that gives the equation; had this been a subspace requiring more than one equation to define it, we would have had more than one row at the bottom with 0's in all columns except the lase. Thus the image of the transformation is the plane $x - y + z = 0$. You can verify that the 2 basis vectors we found lie in this plane. If you found the plane first, you could solve its equation(s) in the usual way to obtain a basis: $x = y - z$, so the space consists of all vectors

$$\begin{pmatrix} x \\ y \\ x \end{pmatrix} = \begin{pmatrix} y-z \\ y \\ z \end{pmatrix} = y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} .$$

These are not the same basis vectors we found earlier, and there is no reason they should be.

[1, Exercise 5.66, p. 199] Let $\mathcal{V} = \mathbb{R}_{10}[t]$, the vector space of real polynomials of degree ≤ 10. Consider the linear transformation $\mathbf{D}^4 : \mathcal{V} \to \mathcal{V}$, where $\mathbf{D}^4$ denotes the fourth derivative, $\frac{d^4}{dt^4}$. Find a basis and the dimension of each of

1. the image of $\mathbf{D}^4$;
2. the kernel of $\mathbf{D}^4$.

Solution:

1. Differentiation 4 times reduces degrees by 4; thus the image cannot contain any polynomials of degree exceeding 6. But every polynomial $\sum_{i=0}^{6} a_i t^i$ is the image of, among other polynomials, $\displaystyle\sum_{i=0}^{6} \frac{a_i}{(i+1)(i+2)(i+3)(i+4)} t^{i+4}$. Hence the image is precisely $\mathbb{R}_6[t]$.
2. The polynomials that are annihilated by 4 differentiations are evidently all the polynomials in $\mathbb{R}_4[t]$, and no others. This is the kernel.

The problem asks for bases and dimensions for each of these spaces. You know that the dimension of $\mathbb{R}_n[t]$ is $n + 1$, and you know various bases, e.g., $1, t, t^2, \ldots, t^n$.

[1, Example 5.72(a), p. 199] Show that the mappings $F : \mathbb{R}^2 \to \mathbb{R}^3$, $G : \mathbb{R}^2 \to \mathbb{R}^3$, $H : \mathbb{R}^2 \to \mathbb{R}^2$ defined as follows are linearly independent:

$$\begin{aligned} F(x,y) &= (x,2y) \\ G(x,y) &= (y,x+y) \\ H(x,y) &= (0,x) \end{aligned}$$

Solution: I proceed in the usual way. I assume there exists real numbers a, b, c such that

$$aF + bG + cH = 0 \tag{62}$$

i.e., such that, for all $(x,y) \in \mathbb{R}^2$,

$$aF(x,y) + bG(x,y) + cH(x,y) = \mathbf{0}\,.$$

Let's substitute the formulæ for the three functions. We are assuming that, for all real numbers x and y,

$$a(x,2y) + b(y,x+y) + c(0,x) = (0,0)$$

i.e., that

$$(ax+by, 2ay + b(x+y) + cx) = (0,0)\,,$$

i.e., that the following system of linear equations is always satisfied:

$$\begin{aligned} ax + by &= 0 \\ (b+c)x + (2a+b)y &= 0\,. \end{aligned}$$

When $(x,y) = (1,0)$ the first equation tells us that $a = 0$, while the second tell us that $b + c = 0$. When $(x,y) = (0,1)$ we obtain that $b = 0$ and $a = 0$. From these equations we conclude that $c = 0$, so $a = b = c = 0$ is a consequence of assuming that equation (62) holds: this proves linear independence of the three given linear transformations.

[1, Example 5.76, p. 199] Determine whether or not each of the following linear maps is nonsingular. If not, find a nonzero vector $\mathbf{v}$ whose image is $\mathbf{0}$; otherwise find a formula for the inverse map.

1. $F : \mathbb{R}^3 \to \mathbb{R}^3$, defined by $F(x,y,z) = (x+y+z, 2x+3y+5z, x+3y+7z)$.
2. $G : \mathbb{R}^3 \to \mathbb{R}_2[t]$, defined by $G(x,y,z) = (x+y)t^2 + (x+2y+2z)t^1 + (y+z)t^0$.
3. $H : \mathbb{R}^2 \to \mathbb{R}_2[t]$, defined by $G(x,y) = (x+2y)t^2 + (x-y)t^1 + (x+y)t^0$.

Solution:

1. We solve the equations $F(x,y,z) = \mathbf{0}$. By row reduction of a 3×3 matrix we find that the solutions — the kernel of F — are the scalar multiples of the vector $\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$; since the kernel is non-zero, the mapping is singular. But what would we have done if the only solution to the equations were the trivial one? We would have needed to find the inverse function — how could we do that? We could have attempted to solve the equation $F(x,y,z) = (a,b,c)$ by row reducing the augmented matrix
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 2 & 3 & 5 & b \\ 1 & 3 & 7 & c \end{array}\right).$$
Under row reduction this matrix reduces to
$$\left(\begin{array}{ccr|c} 1 & 0 & -2 & 3a-b \\ 0 & 1 & 3 & -2a+b \\ 0 & 0 & 0 & 3a-2b+c \end{array}\right).$$
We see that there is no unique solution (x,y,z) for this system: the function can have (a,b,c) in its image if and only if $3a-2b+c=0$. In other words, the image of F is the set
$$\{(a,b,c)|3a-2b+c=0\},$$
which is a plane in $\mathbb{R}^3$. Points not on this plane do not lie in the image, so there cannot exist an inverse function. We could then have continued to find the kernel by setting $a=b=c=0$.

2. Let's proceed as we did in the last part of the preceding example. Solving for x,y,z the identity
$$(x+y)t^2 + (x+2y+2z)t^1 + (y+z)t^0 = at^2 + bt^1 + ct^0,$$
equivalently, solving the system of equations
$$\begin{aligned} 1x + 1y + 0z &= a \\ 1x + 2y + 2z &= b \\ 0x + 1y + 1z &= c \end{aligned}$$

by row reduction of the augmented matrix yields

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & b-2c \\ 0 & 1 & 0 & a-b+2c \\ 0 & 0 & 1 & -a+b-c \end{array}\right).$$

This shows that just the point $(x,y,z) = (b-2c, a-b+2c, -a+b+c)$ is mapped by G on to the polynomial $at^2+bt^1+ct^0$, i.e., that G^{-1} is given by

$$at^2+bt^1+ct^0 \mapsto (b-2c, a-b+2c, -a+b+c)\,.$$

3. Solving for x,y the identity

$$(x+2y)t^2+(x-y)t^1+(x+y)t^0 = at^2+bt^1+ct^0\,,$$

equivalently, solving the system of equations

$$\begin{array}{rcl} 1x+2y & = & a \\ 1x-1y & = & b \\ 1x+1y & = & c \end{array}$$

by row reduction of the augmented matrix yields

$$\left(\begin{array}{cc|c} 1 & 0 & -a+2c \\ 0 & 1 & a-c \\ 0 & 0 & 2a+b-3c \end{array}\right).$$

This shows that at most the point $(x,y) = (-a+2c, a-c)$ is mapped by H on to the polynomial $at^2+bt^1+ct^0$, but that such an inverse image exists only if $2a+b-3c=0$.

A.18 Supplementary Notes for the Lecture of Wednesday, February 15th, 2006

Release Date: Wednesday, February 16th, 2006
Subject to further revision

CHAPTER 6 — Linear Mappings and Matrices

A.18.1 §6.1 Introduction

Much of the preceding chapter was “coordinate free”, in that we defined and described properties of linear transformations in ways that did not use coordinate systems. In practice we often need to use coordinate systems for practical solutions to specific problems. That is the purpose of the present chapter.

A.18.2 §6.2 Matrix Representation of a Linear Operator

Definition A.47 Let a vector space $\mathcal{U}$ of dimension n over the field $\mathbb{K}$ be given. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an *ordered* basis of $\mathcal{U}$, and let $[\mathbf{u}]_S \in \mathbb{K}^n$ denote the coordinates of a vector $\mathbf{u} \in \mathcal{U}$ referred to the given coordinate system. Suppose that $T : \mathcal{U} \to \mathcal{U}$ is a linear operator (linear transformation from the vector space to itself). Suppose that

$$T\mathbf{u}_i = \sum_{j=1}^{n} a_{ij}\mathbf{u}_j \,,$$

i.e., that

$$[T\mathbf{u}_i]_S = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} .$$

We define the matrix

$$m_S(T) = [T]_S = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

to be *the matrix representation of T referred to the coordinate system S*. Thus the matrix $[T]_S$ has, as its *columns* the coordinates of the images under T of the vectors of coordinate system S.

Theorem A.61 *Let $T : \mathcal{U} \to \mathcal{U}$ be a linear transformation, and let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an* ordered *basis (=coordinate system) of $\mathcal{U}$. Suppose that $\mathbf{u} \in \mathcal{U}$ has coordinate vector*

$$[\mathbf{u}]_S = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

referred to the given coordinate system. Then the coordinate vector of $T(\mathbf{u})$ is

$$[T\mathbf{u}]_S = [T]_S[\mathbf{u}]_S = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

While the preceding appears complicated, it is actually more useful to consider even more generality. So now let's consider linear transformations between vector spaces that may be different; and coordinate systems in the domain and codomain that may be different. The results will be the same: the transformation will be represented by a matrix whose *columns* are the coordinate vectors of the images of the basis vectors of the domain; and those column vectors are expressed relative to the coordinate system selected for the *target* vector space.

Definition A.48 Let vector spaces $\mathcal{U}$, $\mathcal{V}$ of dimensions n, m over the field $\mathbb{K}$ be given. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be *ordered* bases respectively of $\mathcal{U}$ and $\mathcal{V}$. As previously defined, $[\mathbf{u}]_S \in \mathbb{K}^n$ and $[\mathbf{v}]_{S'} \in \mathbb{K}^m$ will denote the coordinates of vectors $\mathbf{u} \in \mathcal{U}$, $\mathbf{v} \in \mathcal{V}$, respectively referred to the given coordinate systems. Suppose that $T : \mathcal{U} \to \mathcal{V}$ is a linear transformation, and that

$$T\mathbf{u}_i = \sum_{j=1}^{m} a_{ji}\mathbf{v}_j\,, \quad (i = 1, \ldots, n)$$

i.e., that

$$[T\mathbf{u}_i]_{S'} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix}.$$

We define the matrix

$$m_{S,S'}(T) = [T]_{S,S'} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

to be *the matrix representation of T referred to the respective coordinate systems S, S'*. Thus the matrix $[T]_{S,S'}$ has, as its *columns* the coordinates, referred to system S', of the images of the vectors in the coordinate system S.

The preceding is the usage in the textbook: the first coordinate system in the subscript is that of the domain; the second is that of the codomain. I am going to follow a variant of the preceding. Instead of denoting the bases by two subscripts separated by commas, as

$$[T]_{\text{domain,codomain}}$$

I will often write the domain name in the *upper* right hand corner of the symbol, as

$$[T]^{\text{domain}}_{\text{codomain}}\,.$$

This variant has the advantage of a mnemonic suggestion of cancellation akin to what one sees in the Chain Rule: there will be an association of the upper index with the lower index of the next symbol, which may be either the coordinate vector of a point or another matrix of transformation.

Theorem A.62 *Let $T : \mathcal{U} \to \mathcal{V}$ be a linear transformation, and let S, S', $T^S_{S'}$ and $T^S_{S'}$ be defined as in Definition A.48. Then, for any vector $\mathbf{u} \in \mathcal{U}$, such that*

$$[\mathbf{u}]_S = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

referred to the given coordinate system, the coordinate vector of $T(\mathbf{u})$ is

$$[T\mathbf{u}]_{S'} = [T]^S_{S'}[\mathbf{u}]_S = [T]^S_{S'}[\mathbf{u}]_S \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Supplementary Problems

[1, Exercise 6.37, p. 231] Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $F(x, y) = (4x+5y, 2x-y)$.

1. Find the matrix A representing F in the standard basis. (That is, A denotes the matrix $[F]_E$.)
2. Find the matrix B representing F in the basis $S = \{\mathbf{u}_1, \mathbf{u}_2\} = \{(1,4),(2,9)\}$.
3. Find P such that $B = P^{-1}AP$.
4. For $\mathbf{v} = (a, b)$, find $[\mathbf{v}]_S$ and $[F(\mathbf{v})]_S$.
5. Verify that $B[\mathbf{v}]_S = [F(\mathbf{v})]_S$.

Solution:

1. $$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x+5y \\ 2x-y \end{pmatrix} = x\begin{pmatrix} 4 \\ 2 \end{pmatrix} + y\begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{63}$$
   The columns of the matrix $\begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix}$ are precisely the coordinates of the images under F of the standard basis vectors: this is the matrix we denote by $[F]_E$.
2. We can determine $[F]_S$ in various ways. From first principles we know that the columns of this matrix are to be the coordinates of the images under F of the vectors in the coordinate system S. We compute, using (63):
   $$\begin{aligned} F\begin{pmatrix} 1 \\ 4 \end{pmatrix} &= \begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 24 \\ -2 \end{pmatrix} \\ F\begin{pmatrix} 2 \\ 9 \end{pmatrix} &= \begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} 53 \\ -5 \end{pmatrix} \end{aligned}$$
   But these images are not expressed relative to the coordinate system S; putting it another way, we have found the columns of the matrix
   $$[F]_E^S = \begin{pmatrix} 24 & 53 \\ -2 & -5 \end{pmatrix}.$$
   (The superscript refers to the coordinate system taken for the *domain*; the subscript refers to the coordinate system taken for the *target*. When we multiply coordinate vectors *referred to S* by this matrix we obtain the coordinates of the image in the standard coordinate system.

But the problem asks that we have the images expressed relative to system S. We can find these images from first principles. We solve the equation

$$\begin{pmatrix} 24 \\ -2 \end{pmatrix} = x\begin{pmatrix} 1 \\ 4 \end{pmatrix} + y\begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

by row reduction, and find that $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 220 \\ -98 \end{pmatrix}$. Then we solve

$$\begin{pmatrix} 53 \\ -5 \end{pmatrix} = x\begin{pmatrix} 1 \\ 4 \end{pmatrix} + y\begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

by row reduction, and find that $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 487 \\ -217 \end{pmatrix}$. Note that we could have accomplished the solutions of these systems more elegantly by first finding the inverse of the matrix $\begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}$. Using the usual method, of row reducing the matrix $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 4 & 9 & 0 & 1 \end{array}\right)$ to $\left(\begin{array}{cc|cc} 1 & 0 & 9 & -2 \\ 0 & 1 & -4 & 1 \end{array}\right)$ we find that

$$\begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}^{-1} = \begin{pmatrix} 9 & -2 \\ -4 & 1 \end{pmatrix}.$$

Applying this matrix to the coordinates we had found for the images when referred to the standard coordinate system, we find that $\begin{pmatrix} 9 & -2 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} 24 \\ -2 \end{pmatrix} = \begin{pmatrix} 220 \\ -98 \end{pmatrix}$ etc. The matrix B is thus $\begin{pmatrix} 220 & 487 \\ -98 & -217 \end{pmatrix}$.

3. Tracing back through the computations of the previous part we can see that

$$B = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}^{-1}\begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix} \tag{64}$$

4. Generalizing the preceding to a general vector $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, we see that

$$[\mathbf{v}]_S = \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}^{-1}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 9 & -2 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 9a - 2b \\ -4a + b \end{pmatrix}.$$

In the same way we can show that

$$\begin{aligned}
[F(\mathbf{v})]_S &= \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\
&= \begin{pmatrix} 9 & -2 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\
&= \begin{pmatrix} 32 & 47 \\ -14 & -21 \end{pmatrix} \begin{pmatrix} a \\ a \end{pmatrix} = \begin{pmatrix} 32a + 47b \\ -14a - 21b \end{pmatrix} .
\end{aligned}$$

5. By virtue of equation (64) the preceding product is equal to

$$B \begin{pmatrix} 1 & 2 \\ 4 & 9 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \end{pmatrix} = B[\mathbf{v}]_S \,.$$

A.19 Supplementary Notes for the Lecture of Friday, February 17th, 2006

Release Date: Friday, February 17th, 2006
Subject to further revision

A.19.1 §6.3 Change of Basis

This section is misnamed: what is intended is **Change of Coordinate System**. The content will be discussed in connection with [1, §6.5].

A.19.2 §6.4 Similarity

This material will be deferred, to be discussed just before we begin Chapter 9.

A.19.3 §6.5 Matrices and General Linear Mappings

This is a very long section of my notes, and you may find that you don't need them all. What you want to learn is how to solve specific problems in which a linear transformation is referred to one coordinate system in its domain and another in its target. I will be representing the situation by "diagrams" with the top line showing the way the functions map, and the lower line showing the coordinate systems. In some cases I will be using the *identity* function, but with different coordinate systems at the two ends; this is an easy way to deal with changes of coordinate systems, but it does require some care; I believe it is simpler than the treatment in the textbook if you take the time to learn how it works. Of course, you may use whatever methods are shown in the textbook, and they will give the correct answers if used according to instructions. However, you need to be careful with what the textbook calls the *change-of-basis matrix* or the *transition matrix*. The reason for care is that the name suggests to some users a usage of the coordinate systems in the opposite order. So long as you are careful to check the meaning that the textbook uses for these terms, the method is fine.

So, insofar as these notes are concerned, you might be satisfied just with the portion that works examples, and even there there are some examples that are indicated as being difficult, and could be omitted. Let $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$ be vector spaces (over some fixed field $\mathbb{K}$), $A : \mathcal{U} \longrightarrow \mathcal{V}$, $B : \mathcal{V} \longrightarrow \mathcal{W}$ be linear transformations. We may describe the situation with a "commutative diagram" showing the domains and codomains of the various transformations:

$$\mathcal{U} \xrightarrow{A} \mathcal{V} \xrightarrow{B} \mathcal{W}; \tag{65}$$

we could even show another arrow — labelled $B \circ A$ — directed from $\mathcal{U}$ to $\mathcal{W}$ to indicate the *composition* of the two functions, as[49]

$$\begin{tikzcd} & \mathcal{V} \arrow[dr, "B"] & \\ \mathcal{U} \arrow[ur, "A"] \arrow[rr, "B \circ A"] & & \mathcal{W} \end{tikzcd}$$

. (In such a *diagram* the slopes of the various line segments are not significant; all paths along arrows from one vertex to another are to be equal — as mappings.)

Suppose now that *ordered* bases have been selected for the three vector spaces: $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m\}$ for $\mathcal{U}$, $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_n\}$ for $\mathcal{V}$, and $\{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_p\}$ for $\mathcal{W}$. These bases have the property that

1. they span the respective vector spaces;
2. they are each linearly independent; and
3. they have been *ordered* — i.e. enumerated — i.e. labelled with the *natural numbers* $\{1, 2, 3, ...\}$.

These three conditions imply, respectively that

1. the actions of A and B at any points are *determined* by their actions on the generating sets;
2. *any* assignment of action of the transformation at any of the generators may be extended to a linear transformation; and
3. we may represent the points in the various spaces by vectors respectively in $\mathbb{K}^m$, $\mathbb{K}^n$, $\mathbb{K}^p$, and may represent the linear transformations A and B respectively by $n \times m$ and $p \times n$ matrices.

Any vector $\mathbf{u} \in \mathcal{U}$ is expressible uniquely as a sum $\mathbf{u} = \sum_{i=1}^{m} x_i \mathbf{e}_i$; we may detach the coordinates and simply represent the point by its *coordinate vector* $\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{pmatrix}$, interpreted

[49]Note that the order of composition of functions is *from right to left* — the reverse of the direction of the arrows in (65). This "problem" would disappear if we were to rewrite (65) as $\mathcal{W} \xleftarrow{B} \mathcal{V} \xleftarrow{A} \mathcal{U}$. However there still remains the "problem" that we define the composition $A \circ B$ by $A \circ B(\mathbf{v}) = A(B(\mathbf{v}))$. One way in which mathematicians — mainly algebraists — resolve this problem is to write the symbol for a function as an exponent, writing, for example, u^A instead of $A(u)$. Then they can define a product AB by $u^{AB} = (u^A)^B$, which even "appears" to obey the usual rules for exponents. We shall <u>not</u> follow that convention in this course. In another very similar "solution", the name of the function can also be written "on the line", *following* the name of the point on which it acts.

as a point in $\mathbb{K}^m$. Indeed, we can show that the spaces $\mathcal{U}$ and $\mathbb{K}^{\dim \mathcal{U}}$ are *isomorphic* [1, §4.11]. The same holds for the other two vector spaces under discussion. The coordinate vector of $\mathbf{u}$ relative to the given *ordered* basis $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m\}$ could be denoted by $[\mathbf{u}]_{\mathbf{e}}$; we may also write it as $[\mathbf{u}]_{\mathbf{e}_1,\mathbf{e}_2,...,\mathbf{e}_m}$; where the coordinate system is fixed, it may be reasonable to suppress the subscript in this notation, and write simply $[\mathbf{u}]$.

A maps each basis vector $\mathbf{e}_i$ on to some vector in $\mathcal{V}$; that vector must be expressible as a linear combination of the given basis vectors: suppose that

$$A\mathbf{e}_j = \sum_{i=1}^{n} a_{ij}\mathbf{f}_i, \quad (j = 1, 2, ..., m);$$

similarly, we may suppose that

$$B\mathbf{f}_k = \sum_{k=1}^{p} b_{jk}\mathbf{g}_j, \quad (k = 1, 2, ..., n).$$

Thus the *coordinates* of the vector $A\mathbf{e}_j$ are in the jth column of the matrix

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix}, \quad (j = 1, 2, ..., n).^{50}$$

It is easy to show that the coordinates of $A\mathbf{u}$ are given by the matrix product

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_m \end{pmatrix}. \tag{66}$$

The matrix $\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix}$ is called *the matrix of the transformation A relative* (or *referred*) *to the given ordered bases*, and denoted by $[A]^{\mathbf{e}}_{\mathbf{f}}$, or, where the additional detail is required, $[A]^{\{\mathbf{e}_1,\mathbf{e}_2,...,\mathbf{e}_m\}}_{\{\mathbf{f}_1,\mathbf{f}_2,...,\mathbf{f}_n\}}$. To find the coordinates of a vector under a transformation multiply the *column* vector of coordinates by the *matrix of the transformation*. (Of course, the coordinates and the matrix must be referred to the same ordered basis!)

[50] Note that we are taking m to be the number of *columns* of the matrix, and n to be the number of *rows*.

Suppose now that

$$[B]^{\{\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n\}}_{\{\mathbf{g}_1,\mathbf{g}_2,\ldots,\mathbf{g}_p\}} = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ b_{p1} & b_{p2} & \ldots & b_{pn} \end{pmatrix}.$$

What is the matrix of the composition $B \circ A$? We can show that $[B \circ A]^{\mathbf{e}}_{\mathbf{g}} = [B]^{\mathbf{f}}_{\mathbf{g}}[A]^{\mathbf{e}}_{\mathbf{f}}$. Thus we need only multiply the corresponding matrices in the order of composition of the functions.[51] Note that there is an apparent "cancellation" of superscript and subscript naming the ordered basis chosen on the intermediate vector space $\mathcal{V}$ — reminiscent of the apparent "cancellation" of "numerator" and "denominator" in the chain rule.[52]

Of particular interest is the matrix of the *identity* transformation from a vector space to itself: $\mathbf{1} : \mathcal{U} \longrightarrow \mathcal{U}$ is defined by $\mathbf{u} \mapsto \mathbf{u}\ \forall \mathbf{u}$; since we can define such a transformation for *any* vector space $\mathcal{U}$, we may wish to identify the particular identity mapping under discussion by writing $\mathbf{1}_{\mathcal{U}}$. If the same ordered basis is selected for both the domain and the codomain, then $[\mathbf{1}]$ is simply the identity matrix. We shall consider below the general situation where, *different* bases are chosen.

Consider the case where $\mathcal{U} = \mathcal{V} = \mathcal{W}$, $B \circ A = \mathbf{1}$, and $n = m$, $\mathbf{e}_1 = \mathbf{g}_1$, $\mathbf{e}_2 = \mathbf{g}_2$, ..., $\mathbf{e}_m = \mathbf{g}_n$. Since $\mathbf{1}_{\mathcal{U}}$ fixes every element, *and is referred in both domain and codomain to the same basis,* its matrix will be the identity! Thus we have $[A]^{\mathbf{e}}_{\mathbf{f}}[B]^{\mathbf{f}}_{\mathbf{e}} = I$, so

$$[B]^{\mathbf{f}}_{\mathbf{e}} = \left([A]^{\mathbf{e}}_{\mathbf{f}}\right)^{-1}. \tag{67}$$

$$\text{i.e.} \quad [A^{-1}]^{\mathbf{f}}_{\mathbf{e}} = \left([A]^{\mathbf{e}}_{\mathbf{f}}\right)^{-1}. \tag{68}$$

We shall apply this equality in the next paragraph.

In the notation of (65), let us now consider the case when $\mathcal{V} = \mathcal{U}$ and $A = \mathbf{1}$. The product (66) then gives the coordinate vector $[\mathbf{u}]_{\mathbf{f}}$. Since terminology which refers to bases as "old" or "new" is very confusing, students are advised simply to treat a change of basis as no different from any other transformation. Typically the information which is given expresses the "new" vectors $\mathbf{f}_1$ etc. in terms of the "old" vectors $\mathbf{e}_1$, ... In order to obtain the opposite relationship we need only invert the matrix, by virtue of (68).

Problems involving matrices of transformations should be approached in the following way:

---

[51] If we had adopted one of the alternative labelling methods described in footnote 49, we would want our matrices to compose in the other direction. This could be achieved by transposing (66): allowing the *rows* of the matrix of transformation be the coordinates of the images of the vectors in the selected basis. We should then always write our coordinate vectors horizontally.

[52] This is no surprise, as the present theory could be proved as a consequence of the theory of changes of variable of functions of several variables!

1. First you should recognize that the "objects" you are dealing with are more than functions: in addition to the usual function data (domain, codomain, rules for determining function values) you must specify two sets of *ordered* basis vectors. While these may be completely prescribed for the matrix M you are looking for, you may be able to express M as a product of several matrices, where you have some freedom. More precisely — the matrix M you seek will usually represent a linear transformation T referred to prescribed coordinate systems. If you can express T as the composition of several linear transformations, you may have some freedom in some of the "factors". You may wish to make a "diagram" showing the vector spaces and linear transformations you are using involved; it is useful to note the respective ordered bases as well, for example

$$\begin{array}{ccccc} \mathcal{U} & \xrightarrow{A} & \mathcal{V} & \xrightarrow{B} & \mathcal{W}. \\ \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m\} & & \{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_n\} & & \{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_p\} \end{array} \quad (69)$$

   Some of the intermediate data may be undetermined: typically, you know the first and last coordinate systems, and the composition of all the linear transformations; but the number of factors you choose to use, their domains and codomains, and the corresponding intermediate coordinate systems, may all be available to you to fit to whatever information you have available. One type of freedom that is *always* available is to compose a function with the appropriate identity function: this gives you the chance to change a coordinate system while not altering the action of the function.

2. Write down the matrices you know, taking care that they represent the transformations relative to the intended *coordinate systems*[53]. The data may have been given relative to some coordinate system(s) other than the one(s) that interest you; two methods of approaching that problem are illustrated in the solution to part 2.(a) of the example below.

3. Express the matrices you need in terms of those you know *or their inverses.*

4. Carry out the necessary computations.

Supplementary Problems

[1, Exercise 6.63(a), p. 233] Let $G : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $G(x, y, z) = (2x + 3y - z, 4x - y + 2z)$.

1. Find the matrix $[G]^S_{S'}$, where $S = \{(1, 1, 0), (1, 2, 3), (1, 3, 5)\}$, and $S' = \{(1, 2), (2, 3)\}$

[53] i.e. ordered bases

2. For any $\mathbf{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$, find $[\mathbf{v}]_S$ and $[G(\mathbf{v})]_{S'}$.
3. Verify that $[G]^S_{S'}[\mathbf{v}]_S = [G(\mathbf{v})]_{S'}$.

Solution:

1. If $\mathbf{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$, we know that $G\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.
Hence, referred to the standard bases (which I will denote by E and E'), the matrix of G is
$$[G]^E_{E'} = \begin{pmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \end{pmatrix} .$$
(The coordinates of a vector in $\mathbb{R}^3$ referred to the standard coordinate system coincide with the point itself.)
Let's analyze the situation using the fact that the matrix of a composition of linear transformations is simply the product of the matrices for each of the transformations. The first and third mappings in my diagram don't move any vectors — they are both the identity. When we refer the identity function to the same domain for both the domain and the target, then the matrix of this mapping is simply the identity matrix. But now we are referring to different coordinate systems for the domain and the codomain, so one should not expect the matrix to be the identity matrix. The composition of linear transformations that we are considering is
$$G = \mathbf{1}_{\mathbb{R}^2} \circ G \circ \mathbf{1}_{\mathbb{R}^3} .$$
The sequence of linear transformations under consideration is the following, where I have written below each of the vector spaces the name of the coordinate system that is being used:
$$\begin{array}{ccccccc} \mathbb{R}^3 & \xrightarrow{\mathbf{1}_{\mathbb{R}^3}} & \mathbb{R}^3 & \xrightarrow{G} & \mathbb{R}^2 & \xrightarrow{\mathbf{1}_{\mathbb{R}^2}} & \mathbb{R}^2 \\ S & & E & & E' & & S' \end{array} \tag{70}$$
It is the composition of these three functions in succession that interests us; it is still the function G, but the coordinate system we wish to use for the domain is S, and for the codomain, S'. We know all the matrices for these transformations, or their inverses. The corresponding statement about matrices is
$$[G]^S_{S'} = [\mathbf{1}_{\mathbb{R}^2}]^{E'}_{S'}\,[G]^E_{E'}\,[\mathbf{1}_{\mathbb{R}^3}]^S_E .$$

We know that $[\mathbf{1}_{\mathbb{R}^3}]_E^S$ has, as its columns, the coordinates in the standard basis of the vectors of E: that is, we can construct this matrix by writing the given vectors as its columns:

$$[\mathbf{1}_{\mathbb{R}^3}]_E^S = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 3 & 5 \end{pmatrix} .$$

In the same way we know that

$$[\mathbf{1}_{\mathbb{R}^2}]_{E'}^{S'} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} .$$

However, what we want is $[\mathbf{1}_{\mathbb{R}^2}]_{S'}^{E'}$, which is the *inverse* of this matrix; thus

$$[\mathbf{1}_{\mathbb{R}^2}]_{S'}^{E'} = \left([\mathbf{1}_{\mathbb{R}^2}]_{E'}^{S'}\right)^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} .$$

Hence

$$[G]_{S'}^S = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 3 & 5 \end{pmatrix} = \begin{pmatrix} -9 & 1 & 4 \\ 7 & 2 & 1 \end{pmatrix} .$$

2. We know that $[\mathbf{v}]_E = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Thus

$$
\begin{aligned}
[\mathbf{v}]_S = [\mathbf{1}_{\mathbb{R}^3}]_S^E\,[\mathbf{v}]_E &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 3 & 5 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= \begin{pmatrix} -1 & 2 & -1 \\ 5 & -5 & 2 \\ -3 & 3 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= \begin{pmatrix} -a+2b-c \\ 5a-5b+2c \\ -3a+3b-c \end{pmatrix};
\end{aligned}
$$

$$
\begin{aligned}
[G\mathbf{v}]_{S'} = [\mathbf{1}_{\mathbb{R}^2}]_{S'}^E\,[G\mathbf{v}]_E &= \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= \begin{pmatrix} 2 & -11 & 7 \\ 0 & 7 & -4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= \begin{pmatrix} 2a-11b+7c \\ 7b-4c \end{pmatrix}.
\end{aligned}
$$

Example A.63 Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, $\mathbf{e}_4$ be the *standard*[54] *basis* of $\mathbb{C}^4$, and let $\mathbf{f}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1+i \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} 0 \\ i \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{f}_3 = \begin{pmatrix} i \\ 1 \\ i \\ 1 \end{pmatrix}$, $\mathbf{f}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ i \end{pmatrix}$.

1. Show that $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is a basis.

2. Find the matrix of each of the following linear transformations:

   (a) $T_1 : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ defined by $\mathbf{e}_i \mapsto \mathbf{f}_i$, $(i = 1, 2, 3, 4)$, referred to the ordered basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ for both the domain and codomain;

---

[54] $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

(b) $T_2 : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ defined by $\mathbf{e}_i \mapsto \mathbf{e}_i$, $(i = 1, 2, 3, 4)$, referred to the ordered basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ for the domain and $\{\mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_1, \mathbf{f}_4\}$ for the codomain;

(c) $T_3 : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ defined by $\mathbf{f}_i \mapsto \mathbf{f}_i$, $(i = 1, 2, 3, 4)$, referred to the ordered basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ for the domain and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ for the codomain;

(d) $T_4 : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ defined by $\mathbf{f}_i \mapsto \mathbf{e}_i$, $(i = 1, 2, 3, 4)$, referred to the ordered basis $\{\mathbf{e}_4, \mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2\}$ for both the domain and codomain (note the changed order).

<u>Solution</u>:

1. The space $\mathbb{C}^4$ has dimension 4: one obvious basis is the "standard" basis. Thus all we need to do is to show that the four given vectors are linearly independent. This can be shown by row reducing the matrix with the coordinates of the given vectors as its rows: $\begin{pmatrix} 1 & 0 & 0 & 1+i \\ 0 & i & 0 & 0 \\ i & 1 & i & 1 \\ 0 & 0 & 0 & i \end{pmatrix}$ row reduces to the identity matrix.[55]

2. (a) <u>First Method</u>:

$$\begin{aligned} T_1\mathbf{f}_1 &= T_1(1\mathbf{e}_1 + (1+i)\mathbf{e}_4) \\ &= 1T_1\mathbf{e}_1 + (1+i)T_1\mathbf{e}_4 \quad \text{by linearity of } T_1 \\ &= 1\mathbf{f}_1 + (1+i)\mathbf{f}_4, \text{ etc.} \end{aligned}$$

Thus

$$[T_1]^{\mathbf{f}}_{\mathbf{f}} = \begin{pmatrix} 1 & 0 & i & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & i & 0 \\ 1+i & 0 & 1 & i \end{pmatrix}.$$

<u>Second Method</u>: We require $[T_1]^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}}_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}}$. But the only convenient data give $[T_1]^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}}$. So we interpret T_1 as the composition of T_1 with the identity function $\mathbf{1}$, taking the appropriate coordinate systems:

$$\begin{matrix} \mathbb{C}^4 & \xrightarrow{\mathbf{1}} & \mathbb{C}^4 & \xrightarrow{T_1} & \mathbb{C}^4 \\ \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\} & & \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\} & & \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\} \end{matrix} . \tag{71}$$

[55] Alternatively, we could show that the only solution of $\sum_i \alpha_i \mathbf{f}_i = \mathbf{0}$ is the trivial solution; this leads to *column* reduction of the preceding matrix — again to the identity matrix.

We can first determine the matrix $[T_1]^{\mathbf{e}}_{\mathbf{f}}$ — here it is the identity matrix — and compose it with the matrix $[\mathbf{1}]^{\mathbf{f}}_{\mathbf{e}}$:

$$[T_1]^{\mathbf{f}}_{\mathbf{f}} = [T]^{\mathbf{e}}_{\mathbf{f}}[1]^{\mathbf{f}}_{\mathbf{e}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & i & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & i & 0 \\ 1+i & 0 & 1 & i \end{pmatrix}.$$

(b) $T_3 = \mathbf{1}$. But the first vector of the ordered basis associated with the domain is mapped on to the third vector of the ordered basis associated with the codomain, etc. Proceeding in this fashion we obtain a "permutation matrix":

$$[T_2]^{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4}_{\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_1,\mathbf{f}_4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(c) Again the mapping is the identity, since it is known to fix a basis (albeit a different one from the previous mapping). The matrix is the same as that found in the first part above, but for different reasons!

(d) The mapping in question could be broken down according to the following diagram:

$$\begin{array}{ccccccc} \mathbb{C}^4 & \xrightarrow{\mathbf{1}} & \mathbb{C}^4 & \xrightarrow{T_4} & \mathbb{C}^4 & \xrightarrow{\mathbf{1}} & \mathbb{C}^4 \\ \{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\} & & \{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\} & & \{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\} & & \{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\} \end{array}.$$

Matrices of each of the three transformations shown can be determined: since $\mathbf{f}_1 = (1+i)\mathbf{e}_4 + 0\mathbf{e}_3 + 1\mathbf{e}_1 + 0\mathbf{e}_2$, etc. we have sufficient data to write down the matrix of the first mapping:

$$[1]^{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}} = \left([1]^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}}_{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}\right)^{-1} = \begin{pmatrix} 1+i & 0 & 1 & i \\ 0 & 0 & i & 0 \\ 1 & 0 & i & 0 \\ i & i & 1 & 0 \end{pmatrix}^{-1}.$$

The second mapping has the identity matrix relative to the given coordinate systems. The matrix of the third mapping is

$$[1]^{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Accordingly,

$$\begin{aligned}
[T]^{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}_{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}} &= [\mathbf{1}]^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}[T]^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}}_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}[\mathbf{1}]^{\{\mathbf{e}_4,\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2\}}_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}} \\
&= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} I \begin{pmatrix} 0 & -1 & 1 & 0 \\ 0 & 2 & -1 & -i \\ 0 & -i & 0 & 0 \\ -i & 2-i & -1+i & 0 \end{pmatrix} \\
&= \begin{pmatrix} -i & 2-1 & -1+i & 0 \\ 0 & -i & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 2 & -1 & -i \end{pmatrix}
\end{aligned}$$

Change of Coordinate System Sometimes we wish to change our point of view within one vector space; that is, we start with coordinates referred to one coordinate system, and wish to view the vectors now referred to another system. The machinery of Definition A.48 and Theorem A.62 can be applied here.

Corollary A.64 (to Theorem A.62) *Let $\mathbf{1} : \mathcal{U} \to \mathcal{U}$ be the identity transformation of an n-dimensional vector space $\mathcal{U}$, having coordinate systems S, S'. Then the new coordinates of a vector $\mathbf{u} \in \mathcal{U}$, such that*

$$[\mathbf{u}]_S = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

referred to the coordinate system S, are given by the column vector

$$[\mathbf{u}]_{S'} = [\mathbf{1}\mathbf{u}]_{S'} = [\mathbf{1}]^S_{S'}[\mathbf{u}]_S = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

where the entries in the jth column are those of the jth basis vector in S when referred to the basis S'.

Warning: What the textbook calls a *change-of-basis* matrix is the *inverse* of the matrix found above. Because of possible confusion, I will avoid the term *change-of-basis matrix*. Of course, you should still be able to find new coordinates, given old ones, but this term will be out-of-bounds while we use this textbook.

Matrices of the Composition of Linear Transformations

Corollary A.65 (to Theorem A.62) *Let $T : \mathcal{U} \to \mathcal{V}$, $R : \mathcal{V} \to \mathcal{W}$ be linear transformations between vector spaces $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$, having respective coordinate systems S, S', S''. Then*

$$[R \circ T]_{S''}^{S} = [R]_{S''}^{S'}[T]_{S'}^{S}$$

We can apply this corollary to the particular case where $\mathcal{W} = \mathcal{U}$ and $S'' = S$:

Corollary A.66 (to Theorem A.62) *Let $T : \mathcal{U} \to \mathcal{V}$ be an invertible transformation between vector spaces $\mathcal{U}$, $\mathcal{V}$, having respective coordinate systems S, S'. Then*

$$[T^{-1}]_{S}^{S'} = \left([T]_{S'}^{S}\right)^{-1}$$

Example A.67 Rotation of coordinate axes. Suppose that the x- and y-axes in the plane $\mathbb{R}^2$ are rotated counterclockwise through an angle of θ to yield new x'- and y'-axes for the plane. If we denote the standard basis by E, and the new coordinate system by $S = \{\mathbf{u}_1, \mathbf{u}_2\}$, the new basis vectors will be

$$\begin{aligned} \mathbf{u}_1 &= \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2 && (72)\\ \mathbf{u}_2 &= \cos\left(\theta + \frac{\pi}{2}\right)\mathbf{e}_1 + \sin\left(\theta + \frac{\pi}{2}\right)\mathbf{e}_2 \\ &= -\sin\theta\,\mathbf{e}_1 + \cos\theta\,\mathbf{e}_2 && (73) \end{aligned}$$

Hence

$$[\mathbf{1}]_E^S = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

so

$$[\mathbf{1}]_S^E = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

If we denote the old and new coordinates of a general point by (x, y) and (x', y'), then we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{74}$$

Supplementary Problems

[1, Exercise 6.50(a), p. 232] Find...the coordinates of $\mathbf{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ relative to S, where S consists of the vectors

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

and E is the standard basis.

Solution: I will denote the linear transformation $\mathbf{1}_{\mathbb{R}^3}$ simply by $\mathbf{1}$. The matrix $[\mathbf{1}]_E^S$ is easiest to find first, since

$$\begin{aligned} \mathbf{u}_1 &= 1\mathbf{e}_1 + 1\mathbf{e}_2 + 0\mathbf{e}_3 \\ \mathbf{u}_2 &= 0\mathbf{e}_1 + 1\mathbf{e}_2 + 2\mathbf{e}_3 \\ \mathbf{u}_3 &= 0\mathbf{e}_1 + 1\mathbf{e}_2 + 1\mathbf{e}_3 \end{aligned}$$

Thus

$$[\mathbf{1}]_E^S = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}.$$

$$[\mathbf{1}]_S^E = \left([\mathbf{1}]_E^S\right)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ -2 & 2 & -1 \end{pmatrix}.$$

Hence, knowing that $[\mathbf{v}]_E = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, we can conclude that

$$[\mathbf{v}]_S = [\mathbf{1}\mathbf{v}]_S = [\mathbf{1}]_S^E[\mathbf{v}]_E = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ -2 & 2 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ a-b+c \\ -2a+2b-c \end{pmatrix}$$

[1, Exercise 6.49, p. 232] Suppose that the x- and y-axes in the plane $\mathbb{R}^2$ are rotated counterclockwise through an angle of $\frac{\pi}{6}$ to yield new x'- and y'-axes for the plane. Find

1. the unit vectors in the directions of the new axes;
2. the new coordinates of the points $A(1,3)$, $B(2,-5)$, $C(a,b)$.

Solution:

1. By (72), (73), the basis vectors forming the new coordinate system are

$$\mathbf{u}_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}.$$

2. The new coordinates of $C(a,b)$ are given by (74):

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{6} & \sin\frac{\pi}{6} \\ -\sin\frac{\pi}{6} & \cos\frac{\pi}{6} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{75}$$

Example A.68 (a) Find all linear mappings $F : \mathbb{R}^2 \to \mathbb{R}^2$ such that $F\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

(b) Find all linear mappings $F : \mathbb{R}^2 \to \mathbb{R}^2$ such that $F\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $F\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$.

Solution:

(a) We find it convenient to solve this problem by setting up a coordinate system which includes $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as one — say the first — basis vector. While we will make an arbitrary choice of a second basis vector, the solution will not depend on that choice. The dimension of $\mathbb{R}^2$ is 2, since one known basis is the "standard basis", consisting of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and all bases have the same number of elements. A second basis vector $\mathbf{u}_2 = \begin{pmatrix} a \\ b \end{pmatrix}$ will be such that if α_1 and α_2 were two scalars, not both zero, and if

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{76}$$

then $\alpha_1 = \alpha_2 = 0$. If $b = a$, then one solution to (76) is $\alpha_1 = a$, $\alpha_2 = -1$, *not both zero*. Thus $\mathbf{u}_2$ must have $a \neq b$, i.e. it cannot be a scalar multiple of $\mathbf{u}_1$. But then row reduction of the coefficient matrix of (76) yields that $\alpha_1 = \alpha_2 = 0$. We thus see that any vector may be chosen for $\mathbf{u}_2$, *excluding any vector whose coordinates are the same — in particular, excluding the zero vector*. Without limiting generality of the following, we will choose $\mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent in a space of dimension 2, they constitute a basis, which we order according to their subscripts. Any vector $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ therefore admits a decomposition in the form

$$\mathbf{v} = x\mathbf{u}_1 + y\mathbf{u}_2 \tag{77}$$

i.e. $[\mathbf{v}]_{\{\mathbf{u}_1,\mathbf{u}_2\}} = \begin{pmatrix} x \\ y \end{pmatrix}$. (Indeed, we can solve (77) in the form $a = x + y$, $b = x$, to obtain $x = b$, $y = a - b$.) By the definition of a linear mapping,

$$F(\mathbf{v}) = F(x\mathbf{u}_1 + y\mathbf{u}_2) = xF(\mathbf{u}_1) + yF(\mathbf{u}_2) \tag{78}$$

so the action of F on $\mathbf{v}$ is completely determined when we know its action on the two basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. We know its action on $\mathbf{u}_1$; suppose that $F(\mathbf{u}_2) = \begin{pmatrix} c \\ d \end{pmatrix}$. Then (78) may be rewritten as

$$\begin{aligned} F(\mathbf{v}) = x\begin{pmatrix} 1 \\ 3 \end{pmatrix} + y\begin{pmatrix} c \\ d \end{pmatrix} &= \begin{pmatrix} 1 & c \\ 3 & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1 & c \\ 3 & d \end{pmatrix}\begin{pmatrix} b \\ a-b \end{pmatrix} \\ &= \begin{pmatrix} 1 & c \\ 3 & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} c & 1-c \\ d & 3-d \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

Conversely, suppose that c and d are *any* real numbers. Then we have already seen that any mapping from $\mathbb{R}^2$ to itself defined by $\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} c & 1-c \\ d & 3-d \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$ is linear. Thus we have found the most general linear mapping with the desired property; that is, the solutions to this problem are all transformations of the form

$$\left\{ \mathbf{F}_{\begin{pmatrix} c & 1-c \\ d & 3-d \end{pmatrix}} : c, d, \in \mathbb{R} \right\}$$

(in the notation of Definition A.40), page 1089 of these notes.

(b) The given data are incompatible:

$$F\left(2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = F\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ 8 \end{pmatrix} \neq \begin{pmatrix} 2 \\ 6 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} = 2F\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$$

implies that F cannot be linear, i.e., that there exist *no* linear mappings with the prescribed property.

Example A.69 which follows is complicated: **Omit it unless you are particularly interested in the problem!** (You know my principle: if the example is a good one, I don't want to destroy it; but the following extended solution is probably too difficult for this course.)

Example A.69 An example to introduce change of basis (An unworked problem in an earlier version of your textbook).Find a linear mapping $F : \mathbb{R}^4 \to \mathbb{R}^3$ whose kernel is spanned by $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

Solution: This problem is most easily approachable if we can work within a basis for the domain containing the two given vectors.

In the first instance, we investigate the linear independence of

$$\mathbf{f}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \qquad \mathbf{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

We impose the condition that $\alpha\mathbf{f}_1 + \beta\mathbf{f}_2 = \mathbf{0}$, and deduce from the resulting 4 equations for α and β that $\alpha = \beta = 0$. Thus $\mathbf{f}_1$ and $\mathbf{f}_2$ are linearly independent. (Had these vectors been linearly *dependent* we could still solve the problem, but we could select only a maximal linearly independent subset of the given vectors, and extend it to a basis for $\mathbb{R}^4$.)

Now we extend the set $\{\mathbf{f}_1, \mathbf{f}_2\}$ to a basis. Knowing that the dimension of $\mathbb{R}^4$ is 4 (since one of its bases is the standard basis, containing 4 vectors), we require 2 additional vectors which, with $\mathbf{f}_1$ and $\mathbf{f}_2$, form a linearly independent set. There are infinitely many ways in which this extension may be carried out. This extension may be accomplished in such a way that the remaining 2 vectors in the basis will be elements of the standard basis! Even so there will exist more than one way in which this can be accomplished. We choose to take

$$\mathbf{f}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{f}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

That $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is linearly independent may be demonstrated in the usual way — by row or column reducing the matrix with these vectors as columns. It is obvious because that matrix is in *column-echelon form*.

We know that any linear transformation F is determined by its action on the elements of a basis; and, moreover, that any assignment of values for $F\mathbf{f}_i$ $(i = 1, 2, 3, 4)$ may be extended to a linear mapping — uniquely.

What freedom do we have in assigning these images $F\mathbf{f}_i$ $(i = 1, 2, 3, 4)$? Since $\mathbf{f}_1$ and $\mathbf{f}_2$ are in the kernel, they must be mapped on to the vector $\mathbf{0} \in \mathbb{R}^3$. The additional restrictions come from the requirement that the $\ker F$ be *spanned* by $\mathbf{f}_1$ and $\mathbf{f}_2$. That is, there are to be no other vectors in the kernel except linear combinations of $\mathbf{f}_1$ and $\mathbf{f}_2$. Consider the action of F on a general vector $\mathbf{v} = \sum_{i=1}^{4} x_i\mathbf{f}_i \in \mathbb{R}^4$; note that $[\mathbf{v}]_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{f}_3,\mathbf{f}_4\}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, and $[\mathbf{v}]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} =$

$$\begin{pmatrix} x_1 \\ 2x_1 + x_2 \\ 3x_1 + x_2 + x_3 \\ 4x_1 + x_2 + x_3 + x_4 \end{pmatrix}.$$

$$\begin{aligned} F\left(\sum_{i=1}^{4} x_i \mathbf{f}_i\right) &= \sum_{i=1}^{4} x_i F(\mathbf{f}_i) \\ &= x_1\mathbf{0} + x_2\mathbf{0} + x_3F(\mathbf{f}_3) + x_4F(\mathbf{f}_4) \\ &= x_3F(\mathbf{f}_3) + x_4F(\mathbf{f}_4) \end{aligned}$$

The kernel of F is to be spanned by $\mathbf{f}_1$ and $\mathbf{f}_2$. Provided either or both of x_3 and/or x_4 are non-zero, $x_3F(\mathbf{f}_3) + x_4F(\mathbf{f}_4)$ must be distinct from $\mathbf{0}$. Thus $F(\mathbf{f}_3)$ and $F(\mathbf{f}_4)$ must be linearly independent! This is the general solution to the problem: we have not merely produced one mapping, as required, but have characterized all solutions, by describing precisely the action that such linear mappings take on the elements of a convenient ordered basis, $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$. However, while this basis is convenient for the purposes of this problem, we would prefer to have a solution expressed in terms of the *standard* ordered basis. Our needs may now be stated as follows:

> Let the vectors of the standard basis in $\mathbb{R}^3$ or $\mathbb{R}^4$ be denoted by $\mathbf{e}_1$, $\mathbf{e}_2$, ... Given that[56]
>
> $$[F]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}}^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}} = \begin{pmatrix} 0 & 0 & \alpha_1 & \beta_1 \\ 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & \alpha_3 & \beta_3 \end{pmatrix}$$
>
> where the vectors $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ and $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$ are linearly independent, determine $[F]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}}^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}$.

We shall accomplish this by interpreting F as the composition of two linear mappings, namely as $F \circ 1_{\mathbb{R}^4}$. For the identity mapping $\mathbf{1}_{\mathbb{R}^4} : \mathbb{R}^4 \to \mathbb{R}^4$ we shall take the ordered basis associated with the domain to be $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$, and the ordered basis for the codomain to be $\mathbf{f}_1, \mathbf{f}_2, \mathbf{e}_3, \mathbf{e}_4$; for the linear function $F : \mathbb{R}^4 \to \mathbb{R}^3$ we shall take the ordered basis associated with the domain to be $\mathbf{f}_1, \mathbf{f}_2, \mathbf{e}_3, \mathbf{e}_4$, and that associated with the codomain to be $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. (Remember that we are using the symbols $\mathbf{e}_i$ to mean different things in different contexts!) Then it can be shown that

$$[F \circ \mathbf{1}_{\mathbb{R}^4}]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}}^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} = [F]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}}^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}} [\mathbf{1}_{\mathbb{R}^4}]_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}}^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} \tag{79}$$

We know the matrix on the left in the product. For the matrix on the right — the matrix representing the linear transformation applied first, chronologically — we need to know $[1\mathbf{e}_i]_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}}$ $(i = 1, 2, 3, 4)$. We can show that

$$[\mathbf{1}_{\mathbb{R}^4}]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}} [\mathbf{1}_{\mathbb{R}^4}]_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}}^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} = [\mathbf{1}_{\mathbb{R}^4}]_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}} = I\,,$$

[56] Remember, the lower index on $[F]$ denotes the coordinate system for the codomain, the upper index that of the domain.

and hence that

$$[\mathbf{1}_{\mathbb{R}^4}]^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}} = \left([\mathbf{1}_{\mathbb{R}^4}]^{\{\mathbf{f}_1,\mathbf{f}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}\right)^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 \\ 4 & 1 & 0 & 1 \end{pmatrix}^{-1}$$

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

from which we can conclude that

$$[F]^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}} = \begin{pmatrix} 0 & 0 & \alpha_1 & \beta_1 \\ 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & \alpha_3 & \beta_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -2 & -1 & 0 & 1 \end{pmatrix}$$

where the vectors $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ and $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$ are linearly independent. This is the most general solution. It is easy to verify that the given vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ are mapped on to $\mathbf{0}$; it is the restriction that they must span the kernel that requires that the other parameters give two linearly independent column vectors in the matrix factor on the left. The question asks only for *one* solution, so one could select, for example, $\alpha_1 = 1 = \beta_2$, and $\alpha_2 = \alpha_3 = \beta_1 = \beta_3 = 0$, giving the particular solution

$$[F]^{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}}_{\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}} = \begin{pmatrix} -1 & -1 & 1 & 0 \\ -2 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Our familiar row reduction algorithm gives us a general solution

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = w \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + z \begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix}$$

which is a 2-dimensional kernel, another of whose bases consists of the vectors $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively obtained with $(z, w) = (3, 4)$ and $(z, w) = (1, 1)$.

A.20 Supplementary Notes for the Lecture of Monday, February 27th, 2006

Release Date: Monday, February 27th, 2006
Subject to further revision

Due Dates for the WeBWorK Assignments

Release Date: February 18th, 2006

I find that the order of the topics, and their distribution on the assignments do not match the order in which I had planned to discuss the remaining material in the course. Here is the approximate distribution of problems on the remaining assignments:

Textbook Chapter	WW_7, WW_8	WW_9, WW_{10}	WW_{11}, WW_{12}
6	1-5, 8, 11, 13		
7	6,7,9,10,12,14-19	17-21	17-19
8		1-16	
9			1-16

In order to minimize disruption, I have decided to discuss [1, Chapter 7] immediately, instead of my original plan of discussing [1, Chapters 8,9] first.

CHAPTER 7 — Inner Product Spaces, Orthogonality

A.20.1 §7.1 Introduction

In this Chapter we specialize the field, and also impose additional structure on the vector spaces we are studying. You will see that the resulting *Inner Product Spaces* have properties that resemble properties you saw of Euclidean 2- and 3-dimensional spaces, and there will be other, significant applications.

Real and Complex Vector Spaces

Definition A.49 A *real* vector space is a vector space over the field $\mathbb{R}$. A *complex* vector space is a vector space over the field $\mathbb{C}$. In this chapter, until further notice, all vector spaces are real.

A.20.2 §7.2 Inner Product Spaces

You have seen the word *product* used in several ways in linear algebra:

- the *dot product*
- the *cross product* of vectors in $\mathbb{R}^3$ (more precisely, in *Euclidean 3-space*;
- the *scalar product* of a scalar and a vector, in any vector space

The *inner product* that we are about to define, will generalize the dot product. In fact, we shall see that it *is* the dot product, provided we make an appropriate choice of coordinate system.

Definition A.50 Let $\mathcal{V}$ be a real vector space. An *inner product* on $\mathcal{V}$ is a function

$$\begin{array}{rcl} \text{inner product} : \mathcal{V} \times \mathcal{V} & \to & \mathbb{R} \\ (\mathbf{v}_1, \mathbf{v}_2) & \mapsto & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \end{array}$$

with the following properties:

1. I_1 − **Linearity:** $\langle a_1\mathbf{u}_1 + a_2\mathbf{u}_2, \mathbf{v}\rangle = a_1 \langle \mathbf{u}_1, \mathbf{v}\rangle + a_2 \langle \mathbf{u}_2, \mathbf{v}\rangle$ for all vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and scalars a_1, a_2.
2. I_2 − **Symmetry:** $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$ for all vectors $\mathbf{u}$, $\mathbf{v}$.
3. I_3 − **Positive Definiteness:** For all vectors $\mathbf{v}$, $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$; $\langle \mathbf{v}, \mathbf{v}\rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

Sometimes the inner product is written with other types of parentheses; the separator between the vectors may be written as a vertical line, instead of a comma, as $\langle \mathbf{v}_1 \mid \mathbf{v}_2\rangle$.

Property I_1 is called "Linearity". But why is the property only assumed for the first argument?

Exercise A.1 Prove that an inner product has the property that $\langle \mathbf{v}, a_1\mathbf{u}_1 + a_2\mathbf{u}_2\rangle = a_1 \langle \mathbf{v}, \mathbf{u}_1\rangle + a_2 \langle \mathbf{v}, \mathbf{u}_2\rangle$ for all vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and scalars a_1, a_2.
Solution:

$$\begin{array}{rcl} \langle \mathbf{v}, a_1\mathbf{u}_1 + a_2\mathbf{u}_2\rangle & = & \langle a_1\mathbf{u}_1 + a_2\mathbf{u}_2, \mathbf{v}\rangle \qquad \text{by } I_2 \\ & = & a_1 \langle \mathbf{u}_1, \mathbf{v}\rangle + a_2 \langle \mathbf{u}_2, \mathbf{v}\rangle \qquad \text{by } I_1 \\ & = & a_1 \langle \mathbf{v}, \mathbf{u}_1\rangle + a_2 \langle \mathbf{v}, \mathbf{u}_2\rangle \\ & & \text{by } I_2 \text{ applied to each of the inner product terms} \end{array}$$

Definition A.51 Relative to a given inner product, we define the *norm* or *length* $\|\mathbf{v}\|$ of a vector $\mathbf{v}$ by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}\,,$$

where we always intend the non-negative square root. Vector $\mathbf{v}$ is said to be a *unit* vector if its length is 1. The operation of replacing a vector $\mathbf{v}$ by

$$\frac{1}{\|\mathbf{v}\|}\mathbf{v}$$

is called *normalizing* $\mathbf{v}$.

Exercise A.2 Show that a normalized vector is always a unit vector.

A.20.3 §7.3 Examples of Inner Product Spaces

In each of the following examples we need to show that the function that is claimed to be an inner product really has properties I_1, I_2, I_3.

Example A.70 The *standard inner product* on $\mathbb{R}^n$ is the usual dot product. Show that this product is an inner product. In this case the norm function is the usual length given by the Theorem of Pythagoras. When $\mathbb{R}^n$ is endowed with this particular inner product, the inner product space is sometimes called *Euclidean* n-space.

Exercise A.3 Show that this function has properties I_1, I_2, I_3. (Property I_3 may be deduced from the fact that a sum of squares of real numbers cannot be 0 unless every one of the squares is 0.)

Example A.71 $C[a, b]$ (where $a \leq b$) denotes the vector space of all functions which are continuous on the closed interval $[a, b]$. The "usual" inner product on this space is defined, for functions $f, g \in C[a, b]$, by

$$\langle f, g \rangle = \int_a^b f(t) \cdot g(t)\, dt\,.$$

In particular, the space $\mathbb{R}[t]$ may be interpreted as a subspace of the space of functions continuous on $[a, b]$, and the same inner product may be used.

Exercise A.4 Show that this function has properties I_1, I_2, I_3. (Property I_3 may be deduced from the fact that the definite integral of a positive, continuous function — here the square of any continuous function — cannot be 0 unless the function is identically zero.)

Example A.72 On the space $\mathbb{R}_{m,n}$ of $m \times n$ matrices, one inner product is given by

$$\langle M, N \rangle = \text{trace}(B^T A)$$

If you compute the trace of $A^T A$, you will find that it is simply the sum of the squares of all entries in A. Thus the proof that this function is an inner product will be similar to that for the "standard inner product" above.

A.21 Supplementary Notes for the Lecture of Wednesday, March 1st, 2006

Release Date: Wednesday, March 1st, 2006
Subject to further revision

A.21.1 §7.3 Examples of Inner Product Spaces (conclusion)

Example A.73 The set of infinite sequences $(a_1, a_2, \ldots, a_n, \ldots)$ of real numbers forms a real vector space with the analogous definitions of vector sum and scalar product as we have made for $\mathbb{R}^n$. One subspace is the set of all such sequences with the property that $\displaystyle\sum_{i=1}^{\infty} a_i^2$ converges (which we write as $\displaystyle\sum_{i=1}^{\infty} a_i^2 < \infty$). This hypothesis is sufficient to permit us to define an inner product on this space generalizing the dot product. If $\mathbf{a} = (a_1, a_2, \ldots, a_n, \ldots)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n, \ldots)$, then we may define and inner product by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=1}^{\infty} a_i b_i \,.$$

I do not plan to exploit this particular inner product space further in this course.

Supplementary Problems

[1, Exercise 7.57, p. 271] Verify that the following is an inner product on $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$:

$$f(\mathbf{u}, \mathbf{v}) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2 \,.$$

Solution:

I_1**:** We have to prove linearity in the first argument. Let $\mathbf{w} = (z_1, z_2)$, and let a and c be real numbers. Then

$$a\mathbf{u} + c\mathbf{w} = a(x_1, x_2) + c(z_1, z_2) = (ax_1 + cz_1, ax_2 + cz_2) \,.$$

$$\begin{aligned}
& f(a\mathbf{u} + c\mathbf{w}, \mathbf{v}) \\
= \; & (ax_1 + cz_1)y_1 - 2(ax_1 + cz_1)y_2 - 2(ax_2 + cz_2)y_1 + 5(ax_2 + cz_2)y_2 \\
= \; & ((ax_1)y_1 - 2(ax_1)y_2 - 2(ax_2)y_1 + 5(ax_2)y_2) \\
& \quad + ((cz_1)y_1 - 2(cz_1)y_2 - 2(cz_2)y_1 + 5(cz_2)y_2) \\
= \; & a\,(x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2) \\
& \quad + c\,(z_1y_1 - 2z_1y_2 - 2z_2y_1 + 5z_2y_2) \\
= \; & af(\mathbf{u}, \mathbf{v}) + cf(\mathbf{w}, \mathbf{v})
\end{aligned}$$

I_2**:**

$$\begin{aligned} f(\mathbf{u},\mathbf{v}) &= x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2 \\ &= y_1x_1 - 2y_1x_2 - 2y_2x_1 + 5y_2x_2 \\ &= f(\mathbf{v},\mathbf{u}) \end{aligned}$$

I_3**:** What we are trying to show is that the following quadratic expression can never be negative:

$$\begin{aligned} f(\mathbf{u},\mathbf{u}) &= x_1x_1 - 2x_1x_2 - 2x_2x_1 + 5x_2x_2 \\ &= x_1^2 - 4x_1x_2 + 5x_2^2 \\ &= (x_1 - 2x_2)^2 + x_2^2 \end{aligned}$$

which is a sum of two squares. Such a sum can never be negative. This is the first part of I_2. Moreover, if such a sum is equal to 0, then both of the squares must be 0 (since they cannot be less). That can occur only if $x_1 - 2x_2 = 0$ and $x_2 = 0$, which equations are equivalent to $x_1 = x_2 = 0$, equivalently, $\mathbf{u} = \mathbf{0}$.

I have decomposed the product into a sum of two squares in an arbitrary way. Your textbook suggests a more systematic way of attacking this problem, and we will probably return to it.

[1, Exercise 7.59, p. 271] Find the values of k so that the following is an inner product on $\mathbb{R}^2$, where $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2)$:

$$f(\mathbf{u},\mathbf{v}) = x_1y_1 - 3x_1y_2 - 3x_2y_1 + kx_2y_2 .$$

Solution: It is easy to see, using the same methods as used in the solution of the preceding problem, that I_1 and I_2 both hold for all values of k.

$$\begin{aligned} f(\mathbf{u},\mathbf{u}) &= x_1x_1 - 3x_1x_2 - 3x_2x_1 + kx_2x_2 \\ &= x_1^2 - 6x_1x_2 + kx_2^2 \\ &= (x_1 - 3x_2)^2 + (k-9)x_2^2 . \end{aligned}$$

If $k < 9$, it is possible to find values of x_1, x_2 which will make this sum negative: e.g., take $x_2 = 1$, $x_1 = 3$. Thus it is certainly necessary that $k \geq 9$. Even if $k = 9$, there are non-zero vectors for which the proposed inner product with itself will be zero, e.g., the vector $(3, 1)$. However, if we take $k > 9$, then the inner product of a vector with itself will be a sum of squares that cannot be negative; and it will be 0 precisely when both $x_2 = 0$ and $x_1 - 3x_2 = 0$, i.e., for the vector $\mathbf{0}$ alone.

[1, Exercise 7.60, p. 271] Here the object is to show that these forms do not satisfy the 3 conditions for an inner product, where $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$.

(a) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1y_1 + x_2y_2$.
Solution: Had the space been only $\mathbb{R}^2$, the proposed product would have been an inner product, in fact, it would have been the dot product. By applying it in $\mathbb{R}^3$ we are not violating either of I_1 or I_2, and even part of I_3 still remains true. The violation is in that there will be non-zero vectors whose norm is 0. For example, consider the vector $\mathbf{u} = (0, 0, 1)$.

(b) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1y_2x_3 + y_1x_2y_3$.
Solution: In this case it is clear the I_2 holds. There may be several other types of violations here. One is that there will again be non-zero vectors whose norm will be 0; in fact the same vector $\mathbf{u} = (0, 0, 1)$ used in the previous part can serve as a counterexample.

A.22 Supplementary Notes for the Lecture of Friday, March 3rd, 2006

Release Date: Friday, March 3rd, 2006
Subject to further revision

Remember the decision taken by the class on Friday, March 3rd: a second class test will be available on *either* March 8th, or March 15th. Students who write the test on March 8th may not write again on March 15th.

A.22.1 Some Geometrical Concepts From Euclidean 2-space

On **WeBWorK** assignments WW_7 and WW_8 there are a number of concepts that are not explicitly defined in the textbook. These concepts can generalized to higher numbers of dimensions, and can also be specialized to 1 dimension.

Definition A.52

Let k be a positive real number. A *dilatation* or *dilation* by a factor k is a linear transformation which maps any vector $\mathbf{x}$ of $\mathbb{R}^2$ on to $k\mathbf{x}$. Some authors would restrict this term to cases where the constant k is not less than 1.

Let k be a positive real number. A *contraction* by a factor k is a linear transformation which maps any vector $\mathbf{x}$ of $\mathbb{R}^2$ on to $\frac{1}{k}\mathbf{x}$. Some authors would restrict this term to cases where the constant k is not less than 1.

Let a line L through the origin be given in $\mathbb{R}^2$. The *reflection in* L is the linear transformation that maps any point $\mathbf{x}$ on to its mirror image in the line L.

While I may say more about these topics as the course progresses, they are not discussed in a timely way in the textbook. Accordingly, I ask you to omit the problems that refer to them, and you can be assured that the assignment will be graded out of less than the maximum possible grade. There is a discussion of reflections in [1, Problem 9.13, p. 326], but we have not reached that chapter yet.

These concepts are more appropriate to consider not as linear transformations but as *affine* transformations, where the origin itself may be moved. In one case the problem expects students to know that the product of two reflections is a rotation.

A.22.2 §7.4 The Cauchy-Schwarz Inequality, Applications

For some reason the authorship of this theorem has excited national pride in several countries. While I have shown 3 names below, there could be other mathematicians who are claimed to be involved. Depending on the origin of the book you are reading, you may see only one of these names given as the author of the theorem, or see the order of the names varied.

Theorem A.74 (Cauchy, Schwarz, Buniakovsky, et al.) *For any vectors* $\mathbf{u}, \mathbf{v}$ *in an inner product space,*

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\| \cdot \|\mathbf{v}\|\,.$$

Proof: While I haven't been proving all results, this is one that is very easy to prove.

Since the inequality is true when either $\|\mathbf{u}\| = \mathbf{0}$ or $\|\mathbf{v}\| = \mathbf{0}$, we may assume that neither of the vectors is $\mathbf{0}$.

Let t be a real variable, and consider the square of the norm of the vector $t\mathbf{u} - \mathbf{v}$, i.e., the inner product of this vector with itself. By I_3, this inner product cannot be negative, so $\langle t\mathbf{u} - \mathbf{v}, t\mathbf{u} - \mathbf{v}\rangle \ge 0$. By axioms I_1 and I_2, this implies that

$$t^2\|\mathbf{u}\|^2 - 2t\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \ge 0\,.$$

Suppose that this polynomial could be factored into factors of the form $\|\mathbf{u}\|^2(t-A)(t-B)$, where A, B were real numbers. Then, if $A \ne B$, we could take a value of t lying between A and B, and thereby make the product *negative*. From this contradiction we may infer that the polynomial has either no real roots, or equal roots; this implies that the *discriminant* of the quadratic cannot be positive, which yields the desired result.

Alternatively, let's now proceed as in the earlier problems, by completing the square.

$$\begin{aligned} & t^2\|\mathbf{u}\|^2 - 2t\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \\ = \; & \left(t\|\mathbf{u}\| - \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|}\right)^2 + \left(-\left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|}\right)^2 + \|\mathbf{v}\|^2\right) \end{aligned}$$

This sum will be non-negative if and only if

$$\langle \mathbf{u}, \mathbf{v}\rangle^2 \le \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2\,,$$

equivalently, if and only if

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\| \cdot \|\mathbf{v}\|\,.$$

Example A.75 1. For the dot product in $\mathbb{R}^n$ the inequality may be written as

$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \le \left(a_1^2 + a_2^2 + \cdots + a_n^2\right)\left(b_1^2 + b_2^2 + \cdots + b_n^2\right)$$

2. For $C[0,1]$ the inequality may be written as

$$\left(\int_0^1 f(t)g(t)\,dt\right)^2 \le \int_0^1 (f(t))^2\,dt \quad \times \quad \int_0^1 (g(t))^2\,dt$$

Corollary A.76 (to Theorem A.74) *The norm defined from an inner product satisfies the following three properties for any vectors* $\mathbf{u}$, $\mathbf{v}$, *and any scalar* k:

1. N_1**:** $\|\mathbf{v}\| \geq 0$; $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$.
2. N_2**:** $\|k\mathbf{v}\| = |k| \cdot \|\mathbf{v}\|$
3. N_3 **(Triangle Inequality):** $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Proof: (cf., [1, Exercise 7.8, p. 257].)

1. N_1**:** Suppose that $\mathbf{v} = \mathbf{0}$. Then $\|\mathbf{v}\|^2 = 0$ by I_3, so $\|\mathbf{v}\| = 0$. Conversely, if $\|\mathbf{v}\| = 0$, then the square of the norm is also 0, implying that $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, which, by I_3, implies that $\mathbf{v} = \mathbf{0}$.

2. N_2**:**
$$\begin{aligned} \|k\mathbf{v}\|^2 &= \langle k\mathbf{v}, k\mathbf{v} \rangle && \text{by definition of the norm} \\ &= k^2 \langle \mathbf{v}, \mathbf{v} \rangle && \text{by } I_1, I_2 \\ &= k^2 \|\mathbf{v}\|^2 && \text{by definition} \\ &= |k|^2 \|\mathbf{v}\|^2 . \end{aligned}$$
This implies that
$$(\|k\mathbf{v}\| - |k|\|\mathbf{v}\|)(\|k\mathbf{v}\| + |k|\|\mathbf{v}\|) = 0$$
so that one of the two factors $\|k\mathbf{v}\| - |k|\|\mathbf{v}\|$, $\|k\mathbf{v}\| + |k|\|\mathbf{v}\|$ must be zero. But the second of these is factors is a sum of two non-negative quantifies, so it can be zero only if
$$\|k\mathbf{v}\| = 0 = |k|\|\mathbf{v}\| , \tag{80}$$
while the first can be zero only if $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$, which includes (80) as a special case.

3. N_3 **(Triangle Inequality):**
$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle \quad \text{by definition} \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \quad \text{by } I_1, I_2 \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \quad \text{by definition and } I_2 \\ &\leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \\ &\quad \text{by the Cauchy-Schwarz inequality.} \end{aligned}$$
We can now move both terms to the same side of the equation and factorize the difference of squares into
$$(\|\mathbf{u} + \mathbf{v}\| - (\|\mathbf{u}\| + \|\mathbf{v}\|))(\|\mathbf{u} + \mathbf{v}\| + (\|\mathbf{u}\| + \|\mathbf{v}\|)) \leq 0 .$$

The second factor cannot be negative, but can be zero. If it is zero then $\|\mathbf{u}+\mathbf{v}\| = 0 = \|\mathbf{u}\| + \|\mathbf{v}\|$; the first factor is zero if and only if $\|\mathbf{u}+\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$, which is again more general than the earlier alternative. This is the justification of the textbook's claim about "taking the square root of both sides" of the equation.

Definition A.53 We know from the Cauchy Inequality that

$$\left|\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}\right| \le 1\,,$$

we may define this ratio to be the arccosine of the *angle between the two given vectors* if both are non-**0**.

(We have to exclude **0** from this definition, because its norm is 0; that is, we will not talk about the angle between the vector **0** and any other vector.)

A.22.3 §7.5 Orthogonality

Orthogonal means *perpendicular*. Two vectors will be said to be *orthogonal* when their inner product is 0. Thus the vector **0** is orthogonal to all vectors. Two non-**0** vectors are orthogonal if and only if the arccosine of the angle between them is $\frac{\pi}{2}$.

Definition A.54 Let S be a subset of an inner product space $\mathcal{V}$. We denote by $S^\perp$ the set

$$S^\perp = \{\mathbf{v} \in \mathcal{V} : \langle \mathbf{u}, \mathbf{v}\rangle = 0 \quad \forall \mathbf{u} \in S\}\,.$$

Theorem A.77 *$S^\perp$ as defined above is a subspace of $\mathcal{V}$. We call $S^\perp$ the* orthogonal complement *of S.*

Example A.78

1. In $\mathbb{R}^n$ the orthogonal complement of a non-zero vector is the hyperplane through the origin whose normal direction is that of the given vector. So, when $r = 2$, the orthogonal complement of a non-zero vector $\mathbf{u}$ represented by a directed line segment emanating from the origin is the line through the origin perpendicular to the line determined by the given vector; in three dimensions the orthogonal complement of a given non-zero vector is the plane through the origin to which the given vector is normal.

2. The solution space of a homogeneous linear system given by $AX = 0$, where A is an $m \times n$ matrix, and $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is the orthogonal complement of the row space of A.

Supplementary Problems

[1, Exercise 7.64, p. 272] Let $\mathcal{V}$ be the vector space of polynomials over $\mathbb{R}$ of degree ≤ 2, with inner product defined by

$$\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt\,.$$

Find a basis of the subspace $\mathcal{W}$ orthogonal to $h(t) = 2t + 1$.

Solution: Denote a general member of $\mathcal{W}$ by $a_0 + a_1 t + a_2 t^2$, where a_0, a_1, a_2 are real numbers. Then

$$\begin{aligned}
&= \int_0^1 \left(a_0 + a_1 t + a_2 t^2\right)(1+2t)\, dt \\
&= \int_0^1 \left(a_0 + (a_1 + 2a_0)t + (a_2 + 2a_1)t^2 + (2a_2)t^3\right)\, dt \\
&= \left[a_0 t + \frac{a_1 + 2a_0}{2} t^2 + \frac{a_2 + 2a_1}{3} t^3 + \frac{a_2}{2} t^4\right]_0^1 \\
&= 2a_0 + \frac{7}{6} a_1 + \frac{5}{6} a_2\,.
\end{aligned}$$

Imposing the condition that this inner product be equal to 0, equivalently, that

$$a_0 = -\frac{7}{12} a_1 - \frac{5}{12} a_2\,,$$

we find that the general member of $\mathcal{W}$ is

$$\left(-\frac{7}{12} a_1 - \frac{5}{12} a_2\right) + a_1 t + a_2 t^2 = \frac{a_1}{12}(-7 + 12t) + \frac{a_2}{12}(-5 + 12t^2)$$

and so, setting first $(a_1, a_2) = (12, 0)$, and then $(a_1, a_2) = (0, 12)$, we obtain a basis of $-7 + 12t$ and $-5 + 12t^2$. This is, of course, not unique: the textbook gives a different basis.

A.22.4 §7.6 Orthogonal Sets and Bases

Definition A.55 A set of vectors are said to be *orthogonal* if any two distinct vectors in the set are orthogonal. The set is *orthonormal* if, in addition to being orthogonal, the vectors are all of unit length.

A.23 Supplementary Notes for the Lecture of Monday, March 6th, 2006

Release Date: Monday, March 6th, 2006
Subject to further revision

A.23.1 §7.6 Orthogonal Sets and Bases (continued)

Supplementary Problems

[**1, Exercise 7.66, p. 272**] Find a basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal to the vectors $\mathbf{u}_1 = (1, 1, 3, 4, 1)$ and $\mathbf{u}_2 = (1, 2, 1, 2, 1)$.

Solution: This problem is ambiguous: we can't impose orthogonality until the inner product has been specified! I will assume that the author intends that we use the *standard* inner product.

We wish to determine a basis for the subspace $\{\mathbf{u}_1, \mathbf{u}_2\}^\perp$. Let's denote the general vector of $\mathcal{W}$ by $\mathbf{v} = (x_1, x_2, x_3, x_4, x_5)$. The conditions that we wish to impose are

$$\mathbf{u}_1 \bullet \mathbf{v} = 0 = \mathbf{u}_2 \bullet \mathbf{v},$$

which are equivalent to the matrix equation

$$\begin{pmatrix} 1 & 1 & 3 & 4 & 1 \\ 1 & 2 & 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Row reduction of the coefficient matrix yields

$$\begin{pmatrix} 1 & 0 & 5 & 6 & 1 \\ 0 & 1 & -2 & -2 & 0 \end{pmatrix}.$$

We see that the general solution is, therefore,

$$\begin{pmatrix} -5x_3 - 6x_4 - x_5 \\ 2x_3 + 2x_4 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -5 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus one basis of $\mathcal{W}$ is

$$\mathbf{v}_3 = \begin{pmatrix} -5 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

You may wonder why I have labelled these basis vectors in the reverse order. Of course, a basis is only a set, so there is no order involved (as there is in a coordinate system). The reason is that I am preparing myself for another problem, which will be worked later in the lecture.

Theorem A.79 *Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal set of non-zero vectors.*

1. *The vectors in S are linearly independent.*
2. $\|\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_n\|^2 = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \cdots + \|\mathbf{u}_n\|^2$

Proof:

1. Suppose that $a_1, a_2, \ldots, a_n$ are scalars such that
$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n = \mathbf{0}. \tag{81}$$
Taking an inner product with any vector $\mathbf{u}_i$ yields
$$\begin{aligned} & \langle a_1\mathbf{u}_1, \mathbf{u}_i\rangle + \langle a_2\mathbf{u}_2, \mathbf{u}_i\rangle + \cdots + \langle a_n\mathbf{u}_n, \mathbf{u}_i\rangle = \langle \mathbf{0}, \mathbf{u}_i\rangle \\ \Rightarrow\ & a_1\langle \mathbf{u}_1, \mathbf{u}_i\rangle + a_2\langle \mathbf{u}_2, \mathbf{u}_i\rangle + \cdots + a_n\langle \mathbf{u}_n, \mathbf{u}_i\rangle = 0 \\ \Rightarrow\ & a_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle = 0\,. \end{aligned}$$
Since $\mathbf{u}_i \neq \mathbf{0}$, $\langle \mathbf{u}_i, \mathbf{u}_i\rangle \neq 0$ (by I_3). Hence $a_i = 0$, and this proof applies to all $i = 1, 2, \ldots, n$.
2. left to the reader

Definition A.56 Let $\mathbf{v}$ and $\mathbf{w}$ be arbitrary vectors in $\mathcal{V}$, where $\mathbf{w} \neq \mathbf{0}$.

1. The *projection of* $\mathbf{v}$ *along* $\mathbf{w}$, denoted by $\mathrm{proj}(\mathbf{v}, \mathbf{w})$ is defined by
$$\mathrm{proj}(\mathbf{v}, \mathbf{w}) = \frac{\langle \mathbf{v}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{w}\rangle}\mathbf{w}$$

2. Your textbook calls the ratio $\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}$ the *Fourier coefficient of* $\mathbf{v}$ *with respect to* $\mathbf{w}$.[57] [58]

Exercise A.5 Prove that, for a fixed vector $\mathbf{w}$, the function

$$\mathbf{v} \mapsto \operatorname{proj}(\mathbf{v}, \mathbf{w})$$

is a linear transformation.

Exercise A.6 We can generalize the concept of projection, by observing that, for a fixed vector $\mathbf{w}$, the linear transformation

$$\mathbf{v} \mapsto \operatorname{proj}(\mathbf{v}, \mathbf{w})$$

has the property that it is equal to its square; we call such a function an *idempotent*. Prove that the function has this property.

Theorem A.80 *Let* $\mathbf{w}$ *be a non-zero vector. Then*

1. $\langle \mathbf{v} - \operatorname{proj}(\mathbf{v}, \mathbf{w}), \mathbf{w} \rangle = 0$; *and*
2. *The decomposition*
$$\mathbf{v} = \operatorname{proj}(\mathbf{v}, \mathbf{w}) + (\mathbf{v} - \operatorname{proj}(\mathbf{v}, \mathbf{w}))$$
expresses $\mathbf{v}$ *as a sum of orthogonal vectors.*

[57] Your textbook also calls the ratio the *component of* $\mathbf{v}$ *along* $\mathbf{w}$. I will not adopt this part of the textbook definition, since my understanding of the use of the word *component* is that it should be used for a *vector* and not for the magnitude of that vector.

[58] The "Fourier coefficient" could also be called the *scalar projection*, since it is the magnitude of the projection, which, to distinguish it from the scalar projection, is sometimes called the *vector projection*.

A.24 Supplementary Notes for the Lecture of Friday, March 10th, 2006

Release Date: Friday, March 10th, 2006
Subject to further revision

A.24.1 §7.6 Orthogonal Sets and Bases (continued)

Theorem A.81 *1. Suppose that $\mathbf{u}_i$ $(i = 1, 2, \ldots, r)$ are orthogonal vectors in $\mathcal{V}$, and $\mathbf{v}$ is any vector in $\mathcal{V}$. Then the vector*

$$\mathbf{v} - \sum_{i=1}^{r} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$$

is orthogonal to each of $\mathbf{u}_i$ $(i = 1, \ldots, r)$.

2. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an orthogonal basis for a normed vector space $\mathcal{V}$, and $\mathbf{v}$ is an arbitrary vector in the space, then

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle} \mathbf{u}_n$$

A.24.2 §7.7 Gram-Schmidt Orthogonalization Process

We know that the property of a set of vectors of being a basis is not lost when a linear combination of some of them are subtracted from another. It is possible to progressively transform a basis of an inner product space into an *orthonormal* basis, by systematically subtracting from vectors appropriate scalar multiples of other vectors in the system. This algorithm is the following:

Theorem A.82 (Gram-Schmidt Orthogonalization Process) *Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis for an inner product space $\mathcal{V}$. Then an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ for this space can be constructed as follows:*

$$\begin{array}{rcl}
\mathbf{w}_1 & = & \mathbf{v}_1 \\
\mathbf{w}_2 & = & \mathbf{v}_2 - \dfrac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 \\
\mathbf{w}_3 & = & \mathbf{v}_3 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 \\
\ldots & = & \ldots \\
\mathbf{w}_n & = & \mathbf{v}_n - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \cdots - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\langle \mathbf{w}_{n-1}, \mathbf{w}_{n-1} \rangle} \mathbf{w}_{n-1}
\end{array}$$

Note that the process given here does not normalize the basis vectors. Sometimes the normalization is built into the process; in that case the process can be called the *Gram-Schmidt Orthonormalization Process*. When you are doing hand calculations, that version of the process can be very complicated.

Example A.83 Using the data of [1, Exercise 7.66, p. 272] (solved above on page 1145 of these notes), find an *orthogonal* and an *orthonormal* basis for the subspace $\mathcal{W}$ of $\mathbb{R}^5$ orthogonal to the vectors $\mathbf{v}_1 = (1, 1, 3, 4, 1)$ and $\mathbf{v}_2 = (1, 2, 1, 2, 1)$.
Solution: We saw above that one basis for $\mathcal{W}$ is

$$\mathbf{v}_3 = \begin{pmatrix} -5 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

You may have wondered why I have labelled these basis vectors in the reverse order. Of course, a basis is only a set, so there is no order involved (as there is in a coordinate system). The reason was that I was preparing myself for the present problem, which was not part of the textbook exercise:

> Find an *orthogonal* basis for $\mathcal{W}$.

Why should the order matter when we invoke the Gram-Schmidt process? It does not matter in principal, but; here, when I am doing my calculations by hand, it is convenient to have the vector with the most 0's appear early.

We take

$$\mathbf{w}_1 = \mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then we compute $\langle \mathbf{v}_1, \mathbf{w}_1 \rangle = 6$, $\langle \mathbf{w}_1, \mathbf{w}_1 \rangle = 2$, so

$$\mathbf{w}_2 = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} - 3 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ 0 \\ 1 \\ -3 \end{pmatrix}.$$

We compute $\langle \mathbf{v}_3, \mathbf{w}_1\rangle = 5$, $\langle \mathbf{v}_3, \mathbf{w}_2\rangle = 19$, $\langle \mathbf{w}_2, \mathbf{w}_2\rangle = 23$. Thus

$$\mathbf{w}_3 = \begin{pmatrix} -5 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{5}{2}\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{19}{23}\begin{pmatrix} -3 \\ 2 \\ 0 \\ 1 \\ -3 \end{pmatrix} = \frac{1}{46}\begin{pmatrix} -1 \\ 16 \\ 1 \\ -38 \\ -1 \end{pmatrix}.$$

We may suppress the scalar factor $\frac{1}{46}$, since that will not affect orthogonality. Thus we have found the *orthogonal* basis,

$$\mathbf{w}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \mathbf{w}_2 = \begin{pmatrix} -3 \\ 2 \\ 0 \\ 1 \\ -3 \end{pmatrix}, \mathbf{w}_3 = \begin{pmatrix} -1 \\ 16 \\ 1 \\ -38 \\ -1 \end{pmatrix}.$$

And, if our ultimate goal is to find an ortho*normal* basis, we can normalize this basis by dividing the vectors respectively by $\sqrt{2}$, $\sqrt{23}$, and $\sqrt{1703}$.

Supplementary Problems

[1, Exercise 7.74(c), p. 273] Find the "Fourier coefficient" c and projection $c\mathbf{w}$ of $\mathbf{v}$ along $\mathbf{w}$, where $\mathbf{v} = t^2$, $\mathbf{w} = t+3$ in $\mathbf{R}[t]$, with inner product $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$.

Solution:

$$\begin{aligned}
\langle \mathbf{w}, \mathbf{w}\rangle &= \int_0^1 (t+3)^2\, dt \\
&= \left[\frac{(t+3)^3}{3}\right]_0^1 = \frac{4^3 - 3^3}{3} = \frac{37}{3} \\
\langle \mathbf{v}, \mathbf{w}\rangle &= \int_0^1 t^2(t+3)\, dt \\
&= \left[\frac{t^4}{4} + t^3\right]_0^1 = \frac{5}{4}
\end{aligned}$$

The "Fourier coefficient" is therefore $\dfrac{\langle \mathbf{v}, \mathbf{w}\rangle}{\langle \mathbf{w}, \mathbf{w}\rangle} = \dfrac{15}{148}$.

A.24.3 §7.8 Orthogonal and Positive Definite Matrices

Definition A.57 A square matrix A is said to be *orthogonal* if its rows form an orthonormal[59] set with respect to the usual inner product. Equivalently, A is orthogonal if $AA^T = I$.

Theorem A.84 *A is orthogonal if and only if any one of the following conditions holds:*

1. $AA^T = I$
2. $A^TA = I$
3. $A^{-1} = A^T$

Matrix Representation of an Inner Product

Definition A.58 Let $\mathcal{V}$ be an inner product space having an ordered basis (coordinate system) $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ The *matrix representation of the inner product* is obtained from the matrix $A = [a_{ij}]$, where $a_{ij} = \langle \mathbf{u}_i, \mathbf{u}_j \rangle$ $(i, j = 1, \ldots, n)$.

Theorem A.85 *In the inner product space just described, if* $\mathbf{v}$, $\mathbf{w}$ *be any two vectors, their inner product is given by*

$$\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{v}]^T A [\mathbf{w}]$$

(We may return to this section later in the course.)

Supplementary Problems

[1, Exercise7.80, p. 273] Find a 3×3 orthogonal matrix P whose first two rows are multiples of $\mathbf{u} = (1, 1, 1)$ and $\mathbf{v} = (1, -2, 3)$.

Solution: We first verify that the given vectors are orthogonal, by computing their dot product: $1(1) + 1(-2) + 1(3) = 2 \neq 0$. Thus the problem has an error: it is not possible to complete multiples of these rows to an orthogonal matrix!

Let's fix this blunder: Assume that the problem reads

> Find a 3×3 orthogonal matrix P whose first two rows are multiples of $\mathbf{u} = (1, 1, 1)$ and $\mathbf{v} = (1, +2, -3)$.

[59]Note that we require orthonormality, not simply orthogonality. It might have been better to call the matrix *orthonormal*, but the term *orthogonal* is the one in general use at this time.

Now the inner product of these vectors is $1(1) + 1(2) + 1(-3) = 0$.

Suppose the third row of the matrix is $\mathbf{w} = (a, b, c)$. Then this vector must be orthogonal (using the standard inner product) to the two given equations, giving a system of equations

$$\begin{aligned} 1a + 1b + 1c &= 0 \\ 1a + 2b - 3c &= 0 \end{aligned}$$

whose solution, after row reduction, etc., is the set of all multiples of $\left(-\frac{5}{3}, \frac{2}{3}, 1\right)$, equivalently, all multiples of $(-5, 4, 1)$. We now have multiples of the rows of the matrix; all we need to do is to normalize them, after first observing that the lengths of the rows are, respectively, $\sqrt{3}$, $\sqrt{14}$, and $\sqrt{42}$. However, in this normalization, we may choose either sign for the square roots. Thus there are 8 possible solutions to this problem; the textbook, gives only one.[60] Here they are:

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & -\frac{3}{\sqrt{14}} \\ -\frac{5}{\sqrt{42}} & \frac{4}{\sqrt{42}} & \frac{1}{\sqrt{42}} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & -\frac{3}{\sqrt{14}} \\ -\frac{5}{\sqrt{42}} & \frac{4}{\sqrt{42}} & \frac{1}{\sqrt{42}} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{-\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ -\frac{5}{\sqrt{42}} & \frac{4}{\sqrt{42}} & \frac{1}{\sqrt{42}} \end{pmatrix},$$

$$\begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ -\frac{5}{\sqrt{42}} & \frac{4}{\sqrt{42}} & \frac{1}{\sqrt{42}} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & -\frac{3}{\sqrt{14}} \\ \frac{5}{\sqrt{42}} & -\frac{4}{\sqrt{42}} & -\frac{1}{\sqrt{42}} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & -\frac{3}{\sqrt{14}} \\ \frac{5}{\sqrt{42}} & -\frac{4}{\sqrt{42}} & -\frac{1}{\sqrt{42}} \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{5}{\sqrt{42}} & -\frac{4}{\sqrt{42}} & -\frac{1}{\sqrt{42}} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{5}{\sqrt{42}} & -\frac{4}{\sqrt{42}} & -\frac{1}{\sqrt{42}} \end{pmatrix}.$$

A.24.4 §7.9 Complex Inner Product Spaces

Omit for now.

A.24.5 §7.10 Normed Vector Spaces

Omit for now.

[60]If, by "multiples", the author meant "positive multiples", there would be 2 solutions, still more than 1.

A.25 Supplementary Notes for the Lecture of Monday, March 13th, 2006

Release Date: Monday, March 13th, 2006
Subject to further revision

A.25.1 §7.8 Orthogonal and Positive Definite Matrices (amplification)

If a real inner product space $\mathcal{V}$ is being referred to an ordered basis

$$B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n, \ldots\}$$

then the inner product of vectors $\mathbf{v}$ and $\mathbf{w}$ such that

$$[\mathbf{v}]_B = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \vdots \end{pmatrix}, \qquad [\mathbf{w}]_B = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \\ \vdots \end{pmatrix}$$

is, by virtue of the linearity of the inner product, the sum

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_i \sum_j x_i y_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle \,. \tag{82}$$

When the basis is finite, this double sum can be expressed as a matrix product,

$$\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{v}]^T A [\mathbf{w}]$$

where $a_{ij} = \langle \mathbf{u}_i, \mathbf{u}_j \rangle$. Because of condition I_2, this matrix will be *symmetric*, i.e., $a_{ij} = a_{ji}$ for all i, j. But condition I_3 imposes additional conditions on the matrix, so not every symmetric matrix can be realized as the matrix of an inner product. We may return to this section later in the course.

In particular, if we use the integral inner product

$$\langle f, g \rangle = \int_k^\ell f(t)\, g(t)\, dt \qquad (k < \ell)$$

for functions (including polynomials) $f(t)$, $g(t)$, then it suffices to compute the inner products of all possible combinations of basis vectors. Where the basis of $\mathbb{R}_n[t]$ is the powers $1, t, t^2, \ldots, t^n$, we need only observe that

$$a_{r+1,s+1} = \langle t^r, t^s \rangle = \int_k^\ell f(t)\, g(t)\, dt = \left. \frac{t^{r+s+1}}{r+s+1} \right]_k^\ell = \frac{\ell^{r+s+1} - k^{r+s+1}}{r+s+1} \,.$$

If the coordinate system B is *orthonormal*, then sum (82) simplifies to

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_i x_i y_j .$$

CHAPTER 8 — Determinants

In this course we will study only parts of this Chapter, in particular the portions that we need to comfortably read [1, Chapter 9]. Some of this theory is "obsolete", in that we can accomplish the same results in other ways, mainly by row reduction of matrices. Other portions are interesting, but have to be sacrificed because of the limited time we have available. Many of you have already encountered this material in your earlier course in linear algebra.

A.25.2 §8.1 Introduction

The *determinant* function is defined on square matrices — only on *square* matrices. The function takes its values in the field over which the matrix is defined. The value of the determinant function is commonly denoted in two ways, either

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \mapsto \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{vmatrix} \tag{83}$$

or

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \mapsto \det \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{84}$$

If the matrix is denoted by a single letter, as A, then we would write $|A|$ or $\det A$ for (the value of) its determinant. Don't confuse the vertical lines in the notation with the square brackets used by some authors, e.g., those of your textbook, for matrices. Also, don't confuse the vertical lines with absolute value. For example, a square 1×1 matrix whose only entry is -6 has determinant equal to -6, not the absolute value of -6: it's just another use for the same symbol of vertical lines.

There are (at least) 2 interesting ways in which we can define $\det A$.

An explicit definition for the determinant One of these definitions is explicit: if the size of the matrix has been specified, then the determinant is defined to be a certain sum of products of its n^2 entries: we use each of the products of entries chosen 1 from every row and 1 from every column; and we associate with each of these products a sign, determined according to a complicated formula. This definition will be given below for determinants of matrices of sizes 1×1, 2×2, and 3×3. However, it is a difficult definition for students to use the first time around, so we will instead define the determinant in another way, *recursively*. This will be discussed below in Definition A.62, on page 1157 of these notes. We could also give a definition which is based on the *volume* of a parallelepiped determined by the row or column vectors of the matrix. The determinant of an $n \times n$ matrix is sometimes called a *determinant of order n*.

A.25.3 §8.2 Determinants of Orders 1 and 2

We define

Definition A.59

$$\det \begin{pmatrix} a_{11} \end{pmatrix} = |a_{11}| = a_{11} \tag{85}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21} \tag{86}$$

I will discuss in class a way of remembering this formula.

A.25.4 §8.3 Determinants of Order 3

We define

Definition A.60

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} \\ &\quad -a_{13} \cdot a_{22} \cdot a_{31} - a_{12} \cdot a_{21} \cdot a_{33} - a_{11} \cdot a_{23} \cdot a_{32} \end{aligned} \tag{87}$$

Here also I will discuss in class simple ways of remembering this definition, which is not as complicated as it looks. We can group the terms in the determinant of a matrix of side 3 in the following way, which suggest the general, recursive definition that I will be giving below:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{88}$$

This will also be explained in the lecture. We call this way of computing the determinant, *expanding the determinant by its 1st row*.

A.25.5 §8.4 Permutations

Omit this section.

A.25.6 §8.5 Determinants of Arbitrary Order

This section is concerned with the definition we will not be using. Omit it.

A.25.7 §8.6 Properties of Determinants

I will return to this section below.

A.25.8 §8.7 Minors and Cofactors

Definition A.61 The *minor* of an entry a_{ij} of a determinant $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$ is defined to be the determinant of the matrix obtained from the preceding matrix by deleting the ith row and the jth column. The *cofactor* A_{ij} of a_{ij} is defined to be $(-1)^{i+j}$ times the minor of a_{ij}.

The signs of which this definition speaks form a "checkerboard" pattern:

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

A.25.9 §8.8 Evaluation of Determinants

In the lectures I will discuss a convenient way of evaluating the determinants of 2×2 and 3×3 matrices. The method has a "cyclic" nature, which suggests that it can be used in general. *It can't!* So don't attempt to generalize it to, say 4×4 matrices.[61] This section presents as an algorithm what I will take to be the definition:

---

[61] The determinant of a 4×4 matrix has $4! = 4 \times 3 \times 2 = 24$ terms: don't even think of using any formula that has fewer terms.

Definition A.62

$$\det A = \sum_j a_{1j} A_{1j} \,. \tag{89}$$

We speak of *expanding the determinant by its 1st row*.

The following theorem will not be proved in the course, but can be derived from the properties of determinants that will be discussed in the section [1, §8.6]:

Theorem A.86

$$\det A = \sum_j a_{kj} A_{kj} = \sum_i a_{i\ell} A_{i\ell} \tag{90}$$

where k and l are each any integer between 1 and n. We speak of expanding the determinant by its kth row, *or of* expanding the determinant by its ℓth column.

A.26 Supplementary Notes for the Lecture of Friday, March 17th, 2006

Release Date: Friday, March 17th, 2006; updated April 17, 2006;
Subject to further revision

A.26.1 §8.6 Properties of Determinants

While the preceding definition can be applied to determine the determinant of any square matrix, it is possible to streamline the computation by applying a number of results that we can prove about determinants. The proofs will not be given in this course.

Theorem A.87 *1. The determinant of A and A^T are the same.*

2. *If a square matrix has a row or column entirely composed of 0's, then its determinant is 0.*
3. *If a square matrix has two rows that are the same, or two columns that are the same, then its determinant is 0.*
4. **Determinant of a diagonal matrix**: *An $n \times n$ matrix whose only non-zero entries are along the main diagonal*[62] *has, as its determinant, the product of all the main diagonal entries. More generally,*
5. *If a square matrix is either upper triangular*[63] *or lower triangular*[64]*, then its determinant is simply the product of its entries in the main diagonal.*
6. **Effect of row and column operations on the determinant**: *Since a row or column operation may be effected by multiplying by an elementary matrix, the effect on the determinant will be given by the determinant of the elementary matrix. The following may be shown:*
   (a) *When 2 rows or 2 columns of a matrix are interchanged, the determinant is multiplied by -1.*
   (b) *When a multiple of one row is added to another (or a multiple of one column is added to another) the determinant is not changed.*
   (c) *When all entries in one row (or in one column) of a matrix are multiplied by a number α, then the value of the determinant is also multiplied by α. (Note that this means that, for an $n \times n$ matrix A, $\det(\alpha A) = \alpha^n \det A$.*

---

[62] called a *diagonal* matrix
[63] i.e., all entries are either on or above the main diagonal
[64] i.e., all entries are either on or below the main diagonal

Thus we may row or column reduce a matrix while we compute its determinant. Indeed, we may mix row and column operations in any order.

7. **Determinant of a product of matrices**: $\det(AB) = \det A \cdot \det B$.

   One consequence is that, if a matrix A is invertible, it follows from $AA^{-1} = I$ that $\det A \cdot \det A^{-1} = \det I = 1$, *so* $\det A \neq 0 \neq \det A^{-1}$, *and* $\det A^{-1} = (\det A)^{-1}$.

8. *Since the values of determinants are scalars, and since multiplication of scalars is* commutative, *note that the preceding result may be written as*

$$\det(AB) = \det A \cdot \det B = \det(BA) = \det B \cdot \det A\,.$$

9. *A square matrix is invertible if and only if its determinant is not* 0.

10. *If A is a square matrix whose determinant is not 0, then*

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

11. *If square matrices A and B are such that there exists a matrix P for which $B = P^{-1}AP$, then* $\det A = \det B$.

A.26.2 The "Classical" Adjoint

The modifier *classical* is used here because the word *adjoint* may appear with another usage in connection with [1, Chapter 13].

Definition A.63 The cofactors of a matrix $A = (a_{ij})_{\substack{i=1,2,\ldots,n\\ j=1,2,\ldots,n}}$ form the entries of the *"classical" adjoint* matrix: $\operatorname{adj} A = (A_{ji})_{\substack{i=1,2,\ldots,n\\ j=1,2,\ldots,n}}$.

Note that the cofactors corresponding to the ith row appear down the ith *column* of the classical adjoint.

Theorem A.88 ([1, Theorem 8.9, p. 285])

$$A \cdot adj\, A = adj\, A \cdot A = (\det A) \quad I$$

(Note that the extreme right member of the equations is a diagonal matrix each of whose diagonal entries is equal to $\det A$.) Note, in particular, that the sum of products of entries in some row (column) of A with the entries in the corresponding row (column) of the classical adjoint is always zero.

Corollary A.89 (to Theorem A.88) Use of the determinant in determining the inverse matrix: *From the preceding theorem we see that if* $\det A \neq 0$ *then*

$$A^{-1} = \frac{1}{\det A} adj\, A$$

As seen above, if $\det A = 0$, A cannot have an inverse. Thus *a square matrix A is invertible if and only if its determinant is non-zero*. It will be this property that we apply when we define the *characteristic polynomial* of a matrix in [1, Chapter 9].

A.26.3 §8.10 Applications to Linear Equations, Cramer's Rule

Cramer's Rule is a formula for the solution of a system of n possibly inhomogeneous equations in n variables, where the determinant of coefficients is non-zero. Cramer's Rule has occasional applications in theoretical situations where a symmetric formula is useful, but is otherwise not of much use in solving systems of equations. Avoid using it if you can.

Suppose you are given a general system of n linear equations in n variables:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \ldots & & \ldots & & \ldots & & \ldots & = & \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \tag{91}$$

and that you know that the determinant of the matrix of coefficients — which matrix I will denote by A — is not equal to 0. Suppose that we multiply the first equation by the cofactor A_{1i} of a_{1i} in the matrix A, the second equation by the cofactor A_{2i} of a_{2i}, etc., and that the products are summed. It follows from Theorem A.88 that all but one of the variables appears with coefficient 0, and the variable x_i appears with coefficient $\det A$. On the right side we have the sum $\sum_{j=1}^{n} b_j A_{ji}$, which is equal to the determinant of the matrix obtained from A by replacing the ith column by the column vector $\begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix}$. Since we assumed that the determinant of A is non-zero, we may divide, and thereby obtain the value of x_i as the ratio of values of two determinants.

This "Rule" is occasionally useful in solving general systems of equations, particularly if the coefficients are not known. The "Rule" has very limited uses in numerical situations: don't use it unless there is some good reason to do so, as row reduction is likely to be faster and more reliable. When we apply Cramer's Rule to a *homogeneous* system of n equations in n variables, we obtain the following

Theorem A.90 *The homogeneous system of equations*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n} & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n} & = & 0 \\ \ldots & & \ldots & & \ldots & & \ldots & = & \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn} & = & 0 \end{array} \tag{92}$$

has a non-trivial *solution, i.e., a solution* $\mathbf{x}$ *distinct from the vector* $\mathbf{0}$, *if and only if* $\det\left(\begin{array}{c} a_{ij} \end{array}\right) = 0$.

Proof: If the matrix A of coefficients had an inverse, we could multiply both sides of the equation $A\mathbf{x} = \mathbf{0}$ by it, and obtain

$$\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}\,,$$

i.e., uniquely the trivial solution. From that contradiction we infer that A is not invertible, hence $\det A = 0$.

A.26.4 §8.12 Submatrices, Minors, Principal Minors

Omit this section for now.

A.26.5 §8.13 Block Matrices and Determinants

Omit this section for now.

A.26.6 §8.14 Determinants and Volume

Omit this section for now.

A.26.7 §8.15 Multilinearity and Determinants

Omit this section for now.

A.26.8 §6.4 Similarity

A linear transformation is a mapping between vector spaces. However, in order to work with a transformation we often need to select coordinate systems for the spaces. Different selections of coordinate systems will produce different matrices; while the matrices may look different, all matrices representing a given linear transformation will have certain intrinsic properties. In this section we are interested specifically in linear transformations that map a vector space on to itself.

Definition A.64 Let $\mathcal{V}$ be a vector space over a field $\mathbb{K}$, and let $F : \mathcal{V} \to \mathcal{V}$ be a linear transformation. We may call F a *linear operator on* $\mathcal{V}$.

While the concept of linear operator makes sense for spaces that do not have finite dimension (e.g., differentiation is a linear operator on the space of polynomials $\mathbb{R}[t]$), let's confine our attention for the present to linear operators on finite dimensional spaces. Suppose that F is a linear operator on $\mathbb{K}^n$, and that B and B' are two coordinate systems for $\mathcal{V}$. From properties of the identity functions we know that

$$F = \mathbf{1}_{\mathcal{V}} \circ F \circ \mathbf{1}_{\mathcal{V}} . \tag{93}$$

Let's consider the relationships between $[F]_B$ and $[F]_{B'}$. If we think of the middle F in the product (93) to be referred to B, and the product to be referred to B', then we have

$$[F]_{B'}^{B'} = [\mathbf{1}_{\mathcal{V}}]_{B'}^{B} \cdot [F]_B^B \cdot [\mathbf{1}_{\mathcal{V}}]_B^{B'} . \tag{94}$$

But we know that the matrices $[\mathbf{1}_{\mathcal{V}}]_B^{B'}$ and $[\mathbf{1}_{\mathcal{V}}]_{B'}^{B}$ are mutual inverses. Thus, if we define $P = [\mathbf{1}_{\mathcal{V}}]_B^{B'}$, then the matrices of F referred to the different bases have a relation of the form

$$[F]_{B'}^{B'} = P^{-1}[F]_B^B P . \tag{95}$$

This motivates the following definition:

Definition A.65 We say that a square matrix B is *similar* to a square matrix A if there exists an invertible matrix P of the same size such that $B = P^{-1}AP$.

Theorem A.91 *Square matrices which are similar have equal determinants.*

Proof: Suppose that $B = P^{-1}AP$, where A, B, P are $n \times n$ matrices, and P is invertible. Then

$$\det B = \det\left(P^{-1}\right) \cdot \det A \cdot \det P = (\det P)^{-1} \cdot \det A \cdot \det P = \det A$$

since the multiplication of determinants is commutative (since they are real numbers). □

Just as we found "canonical" matrices to which every matrix is row-equivalent, we now wish to investigate canonical matrices to which every matrix is similar. This problem is much more difficult than row reduction, and we will not be completing the investigation in this course. For reasons that will become apparent later, we will be particularly interested in matrices which are similar to diagonal matrices, and to matrices that are "nearly" diagonal.

A.27 Supplementary Notes for the Lecture of Monday, March 20th, 2006

Release Date: Monday, March 20th, 2006
Subject to further revision

Chapter 8: Supplementary Problem

[1, Exercise 8.43(b), p. 302] Evaluate

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

Solution: This problem was probably intended to illustrate a result in [1, §8.11], which I will not be discussing. No matter, we can still evaluate it easily. Basically we will apply row *or column* operations and make the appropriate adjustments to the value. Recall that elementary operations of type E_1 change the sign of the determinant; operations of type E_2 multiply the value by the constant used to scale the row (or column), and E_3 has no effect on the value.

$$\begin{vmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 6 & 2 \\ 0 & 0 & 2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 & 7 & 9 \\ 0 & -2 & -8 & -10 & -16 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 6 & 2 \\ 0 & 0 & 2 & 3 & 1 \end{vmatrix},$$

by subtraction of twice the 1st row from the second. Now expand this determinant by the first column: we obtain the determinant of the minor of the entry in position $(1,1)$ multiplied by that entry, i.e., 1, and by the sign associated with that position, i.e., $+1$:

$$\begin{vmatrix} -2 & -8 & -10 & -16 \\ 0 & 1 & 2 & 3 \\ 0 & 5 & 6 & 2 \\ 0 & 2 & 3 & 1 \end{vmatrix}$$

and we may again expand by the first column, this time obtaining

$$-2\begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 2 \\ 2 & 3 & 1 \end{vmatrix}$$

which we could expand using the formula we have for determinants of order 3. Or we can continue with row operations:

$$-2\begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 2 \\ 2 & 3 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 3 \\ 0 & -4 & -13 \\ 0 & -1 & -5 \end{vmatrix} == -2\begin{vmatrix} -4 & -13 \\ -1 & -5 \end{vmatrix} = (-2)(20-13) = -14\,.$$

CHAPTER 9 — Diagonalization: Eigenvalues and Eigenvectors

A.27.1 §9.1 Introduction

This chapter explores first steps to associating with a linear operator on a vector space $\mathcal{V}$ a particular basis, relative to which the operator can be seen to have "simple" properties. In this chapter we will be using, mainly, the language of matrices.

A.28 Supplementary Notes for the Lecture of Wednesday, March 22nd, 2006

Release Date: Wednesday, March 22nd, 2006
Subject to further revision

A.28.1 §9.1 Introduction (continued)

Definition A.66 1. A linear operator T on a vector space $\mathcal{V}$ of dimension n over a field $\mathbb{K}$ is *diagonalizable* if there exists a coordinate system B for $\mathcal{V}$, referred to which $[T]_B$ is a diagonal matrix.

2. An $n \times n$ matrix M over a field $\mathbb{K}$ is *diagonalizable* if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Theorem A.92 *A linear operator T is diagonalizable if and only if the matrix $[T]$ is diagonalizable.*

The Geometric Significance of Diagonalizability When a matrix is diagonalizable, it is the product of diagonal matrices having all entries equal to 1, and the remaining entry equal to one of its diagonal entries: geometrically, a positive entry in that location represents a scale change in the direction of that vector — what we called a *dilation* or *dilatation* (or, to some authors, a *contraction* if the entry is a positive scalar less than 1); a negative entry represents a *reflection* in the hyperplane perpendicular to the vector corresponding to this entry (provided the appropriate inner product is adopted). We don't have time in this course to investigate these interesting geometric issues; for example, the product of two reflections corresponds to a *rotation*.

The Underlying Field Until now we have not seen any serious issues that depended on the field associated with a given vector space. That is going to change in this chapter. Our main interest here will be in vector spaces and matrices over $\mathbb{R}$ and $\mathbb{C}$, but the concepts make sense for any vector spaces.

A.28.2 §9.2 Polynomials of Matrices

We have already defined, in [1, §2.8] powers and polynomials in a square matrix A. Analogous concepts can be introduced for linear operators.

Unsolved Problems

[1, Exercise 9.41(a)] Find a polynomial having the following matrix as a root:

$$A = \begin{pmatrix} 2 & 5 \\ 1 & -3 \end{pmatrix}.$$

Solution: **This is an exploratory solution, not a model solution to be applied to solving problems of this type.** The first question to ask is why there should exist a polynomial. We know that 2×2 matrices form a real vector space of dimension 4 Hence, if we take the 5 matrices $A^0 = I$, A, A^2, A^3, A^4, they cannot be linearly independent. The implies the existence of a polynomial of degree not more than 4 having A as a root. We shall see, in the next section, that there will always exist a polynomial whose degree is no more than the number of rows or columns of the matrix — here 2. We will eventually have better ways of answering a question like this; but, for now, I will solve it "by brute force". So let's pretend we don't know that there can be a polynomial of degree at most 2, and look for one of degree at most 4. I will calculate the powers of A.

$$\begin{aligned}
A^0 &= I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
A^1 &= \begin{pmatrix} 2 & 5 \\ 1 & -3 \end{pmatrix} \\
A^2 &= \begin{pmatrix} 9 & -5 \\ -1 & 14 \end{pmatrix} \\
A^3 &= \begin{pmatrix} 13 & 60 \\ 12 & -47 \end{pmatrix} \\
A^4 &= \begin{pmatrix} 86 & -115 \\ -23 & 201 \end{pmatrix}
\end{aligned}$$

Now let's look for a dependence relation between these 5 matrices. I will write them linearly as the *columns* of the matrix of coefficients of a system of linear equations, and row reduce:

$$\begin{aligned}
&\begin{pmatrix} 1 & 2 & 9 & 13 & 86 \\ 0 & 5 & -5 & 60 & -115 \\ 0 & 1 & -1 & 12 & -23 \\ 1 & -3 & 14 & -47 & 201 \end{pmatrix} \\
\sim &\begin{pmatrix} 1 & 2 & 9 & 13 & 86 \\ 0 & 1 & -1 & 12 & -23 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
\sim &\begin{pmatrix} 1 & 2 & 9 & 13 & 86 \\ 0 & 1 & -1 & 12 & -23 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
\sim &\begin{pmatrix} 1 & 0 & 11 & -11 & 132 \\ 0 & 1 & -1 & 12 & -23 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

which tells us that the real numbers $x_0, x_1, \ldots, x_4$ with the property that

$$x_0 I + x_1 A + x_2 A^2 + x_3 A^3 + x_4 A^4 = 0$$

are characterized by the properties that

$$\begin{aligned} x_0 &= 11(-x_2 + x_3 - 12x_4) \\ x_1 &= x_2 - 12x_3 + 23x_4 \end{aligned}$$

We are not, at this time, asking for *all* polynomials of degree not exceeding 4 satisfied by the given matrix: we want only one non-trivial polynomial. We can find one by setting one of the independent variables x_2, x_3, x_4 equal to 1, and the other 2 equal to 0. So, with $x_2 = 1$, $x_3 = x_4 = 0$, we obtain the polynomial represented by

$$\begin{aligned} x_0 &= -11 \\ x_1 &= 1 \end{aligned}$$

i.e., $-11 + t + t^2$, which is indeed satisfied by the given matrix.

After reading further in the chapter, we will have much more efficient ways of finding a polynomial satisfied by a given matrix.

Any multiple of this polynomial by any other polynomial would also be satisfied by the given matrix. The polynomial we found has degree 2. Could there be another polynomial with this property having degree 1? Such a polynomial would have the form $a_0 + a_1 t$, where $a_1 \neq 0$. If A were a root, then A would be expressible in the form $A = -\frac{a_0}{a_1} I$, i.e., it would be a diagonal matrix. Since A is not diagonal, we know that the lowest degree of polynomial it satisfies has degree at least 2. Could A satisfy two distinct polynomials of degree 2, where neither is a constant multiple of the other? If that were the case, we could subtract a multiple of one from the other, and show that A satisfies a polynomial of degree at most 1, which we know is not true. Hence there is essentially only one quadratic polynomial with the property we have found. If we make the highest degree term of this polynomial have coefficient 1, we call it a *monic* polynomial. In this case $t^2 + t - 11$ is the *minimal* polynomial of A.

A.28.3 §9.3 Characteristic Polynomial, Cayley-Hamilton Theorem

We saw in the preceding section that every $n \times n$ matrix A must satisfy a polynomial equation of degree not more than n^2, since $I, A, A^2, \ldots, A^{n^2}$ are $n^2 + 1$ elements of the space $\mathbb{R}_{n,n}$ which has dimension n^2, so they must be linearly dependent. However, when $n > 1$, we can find a polynomial of lower degree satisfied by A; moreover, the following theorem will provide a *constructive* proof of that theorem — we don't simply have to work from the known *existence* of such a polynomial.

Definition A.67 Let A be a given $n \times n$ matrix. Let t be a real variable. The matrix

$$\Delta(t) = \det(tI_n - A) = (-1)^n \det(A - tI)$$

is called the *characteristic polynomial* of A.[65]

[65]Some authors adopt a slightly different definition, taking the characteristic polynomial to be $\det(tI_n - A)$; thus they take the characteristic polynomial to be *monic*. This leads to slightly different results from those appearing in our textbook.

A.29 Supplementary Notes for the Lecture of Friday, March 24th, 2006

Release Date: Friday, March 24th, 2006
Subject to further revision

A.29.1 §9.3 Characteristic Polynomial, Cayley-Hamilton Theorem (continued)

We saw, through an example, that every matrix satisfies a polynomial. That approach, could, in theory, yield for an $n \times n$ matrix a polynomial of degree as high as n^2. But fortunately we have a result which ensures that there exists a polynomial of degree n satisfied by such a matrix:

Theorem A.93 (Cayley-Hamilton) *Any matrix A satisfies its own characteristic polynomial.*

It may happen that there exists a matrix of even lower degree satisfied by the matrix; I will return to this question later.

Example A.94 Consider the matrix of [1, Exercise 9.41(a), p. 336], considered above on page 1166 of these notes. Its characteristic polynomial is

$$\begin{vmatrix} t-2 & -5 \\ -1 & t+3 \end{vmatrix} = t^2 + t - 11$$

which happens to be the minimal polynomial that we calculated. Note that the coefficient of the term in t^0 is -11, which is the determinant of the given matrix. A; this is no accident.

A.29.2 §9.4 Diagonalization, Eigenvalues and Eigenvectors

Definition A.68 Let A be a real $n \times n$ matrix. Suppose that λ is a scalar, and $\mathbf{v}$ is a non-zero column vector such that

$$A\mathbf{v} = \lambda\mathbf{v}\,. \tag{96}$$

Then λ is called an *eigenvalue* or *proper value* of A, and $\mathbf{v}$ is called an *eigenvector* or *proper vector* of A. We say that $\mathbf{v}$ is an eigenvector *corresponding to* or *belonging to* the eigenvalue λ.

Since equation (96) is equivalent to

$$(\lambda I - A)\mathbf{v} = \mathbf{0}\,,$$

there can exist an eigenvector belonging to λ only if the matrix $\lambda I - A$ is singular, equivalently, if its determinant is equal to 0, equivalently if λ is a root of the characteristic polynomial of A.

Definition A.69 Let λ be an eigenvalue of A. The kernel of the matrix $\lambda I - A$ is called the *eigenspace* of A corresponding to eigenvalue λ.

Theorem A.95 *Let* $\mathbf{x}$ *be an eigenvector corresponding to eigenvalue* λ *of an* $n \times n$ *matrix* A *which satisfies a polynomial* $p(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_r x^r$, *i.e., such that*

$$p(A) = p_0 I + p_1 A + p_2 A^2 + \cdots + p_r A^r = 0\,.$$

Then λ *is a root of* $p(x)$, *i.e.,* $p(\lambda) = 0$.

Proof: For any non-negative integer i, $A^i\mathbf{x} = A^{i-1}\lambda\mathbf{x} = \lambda A^{i-1}\mathbf{x} = \cdots = \lambda^i\mathbf{x}$; hence

$$\mathbf{0} = 0\mathbf{x} = p(A)\mathbf{x} = p(\lambda)\mathbf{x}\,.$$

But $\mathbf{x} \neq \mathbf{0}$, so the scalar $p(\lambda)$ must equal 0. □Note that, the preceding applies to any polynomial satisfied by the matrix, not only to the characteristic polynomial. For the characteristic polynomial we have a converse:

Theorem A.96 *Let* λ *be a root of the characteristic polynomial of* A, *i.e., let* $\det(\lambda I - A) = 0$. *Then* λ *is an eigenvalue of* A.

Proof: If $|\lambda I - A| = 0$, then the square matrix $\lambda I - A$ is not invertible; hence there exists a non-$\mathbf{0}$ solution $\mathbf{x}$ to the equation $(\lambda I - A)\mathbf{x} = \mathbf{0}$; that solution has the property that $\lambda\mathbf{x} = A\mathbf{x}$, i.e., $\mathbf{x}$ is an eigenvector corresponding to λ. □

Supplementary Problems

[**1, Exercise 9.41(c)**] Find a polynomial having the following matrix as a root:

$$C = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix}.$$

Solution: While the solution given to part (a) above on 1166 of these notes was exploratory; this one is given as a routine application, by determining the characteristic polynomial of C, which out theory (the Cayley-Hamilton Theorem) tells us will be a polynomial satisfied by C. The characteristic polynomial is

$$\det(tI - C) = \begin{vmatrix} t-1 & -1 & -2 \\ -1 & t-2 & -3 \\ -2 & -1 & t-4 \end{vmatrix}.$$

$$= (t-1)(t^2 - 6t + 5) + (-t + 4 - 6) - 2(1 - 2t - 4) = t^3 - 7t^2 + 6t - 1\,.$$

[**1, Exercise 9.48(a), p. 336**] For the following matrix, find all eigenvalues, and a maximum set S of linearly independent eigenvectors.

1. $A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$

<u>Solution</u>:

1. The characteristic polynomial of A is $\begin{vmatrix} t-1 & -1 & -2 \\ -1 & t-2 & -3 \\ -2 & -1 & t-4 \end{vmatrix} = (t+2)^2(t-4)$.

   The eigenvalues are -2, of multiplicity 2; and 4, of multiplicity 1.

   Eigenvalue $t = 4$**:** We solve the system of equations $(4I - A)\mathbf{x} = \mathbf{0}$. The matrix row reduces to $\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$, so the general solution is

   $$\mathbf{x} = k\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

   or, equivalently,

   $$\mathbf{x} = \ell\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

   so $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ is one eigenvector corresponding to eigenvalue 4.

   Eigenvalue $t = -2$**:** This time we solve the system $(4I - A)\mathbf{x} = \mathbf{0}$ by row reduction of the coefficient matrix, finding that matrix to reduce to $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. The general solution is

   $$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Two linearly independent eigenvectors corresponding to eigenvalue -2 are, therefore, $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. These eigenvectors and the vector given earlier for $t = -4$ form a linearly independent basis of $\mathbf{R}^3$: this is a maximal set of linearly independent eigenvectors.

You are advised to verify that the vectors you have found are, indeed, eigenvectors. One way to do this is simply to compute $A\mathbf{x}$ and to check whether it is a scalar multiple of $\mathbf{x}$. In cases where we have found a full set of linearly independent eigenvectors, those vectors form the columns of a matrix which *diagonalizes* A. More precisely, let's calculate

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 2 \\ 4 & -2 & 0 \\ 8 & 0 & -2 \end{pmatrix}.$$

This last matrix has columns which are scalar multiples of the eigenvectors. We can "factorize" the matrix into a product

$$\begin{pmatrix} 4 & -2 & 2 \\ 4 & -2 & 0 \\ 8 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

the product of the matrix of eigenvectors followed by a diagonal matrix whose entries are the corresponding eigenvalues.

A.30 Supplementary Notes for the Lecture of Monday, March 27th, 2006

Release Date: Monday, March 27th, 2006
Subject to further revision

A.30.1 §9.4 Diagonalization, Eigenvalues and Eigenvectors (continued)

Supplementary Problems

[1, Exercise 9.48, p. 336 (continuation)] For each of the following matrices, find all eigenvalues, and a maximum set S of linearly independent eigenvectors.

2. $B = \begin{pmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{pmatrix}$

Solution:

2. I will sketch this solution briefly, as it parallels the preceding. The characteristic polynomial is $\begin{vmatrix} t-3 & 1 & -1 \\ -7 & t+5 & -1 \\ -6 & 6 & t-2 \end{vmatrix}$, and this polynomial factorizes as $(t-2)^2(t+4)$. When we set $t=2$ in the matrix $tI-B$, we find that it row reduces to $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, whose null space is spanned by the eigenvector $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$; even though the eigenvalue has multiplicity 2, the eigenspace has dimension 1! The null space of the matrix $-4I-B$ is spanned by $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Thus a maximal set of linearly independent eigenvectors of B contains only 2 members. This matrix cannot be diagonalized.

Suppose that the *columns* of a matrix P are eigenvectors of an $n \times n$ matrix A, respectively corresponding to eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then the product AP is a matrix whose ith column is λ_i times the ith column of P. If P is invertible, then the product $P^{-1}AP$ is a diagonal matrix whose ith diagonal entry is simply λ_i. In this case the matrix A has been shown to be *diagonalizable*. Conversely, if there exists an invertible matrix P such that

$P^{-1}AP$ is a diagonal matrix $D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$, then $AP = PD$, and so each entry λ_i is an eigenvalue of P, and the ith column of P is an eigenvector corresponding to the eigenvalue λ_i $(i = 1, 2, \ldots, n)$. We have seen that $\det(P^{-1}AP) = \det A$; hence the determinant of a diagonalizable matrix A is just the product of the entries of the similar, diagonal matrix, which are equal to the eigenvalues of A in the appropriate multiplicities.

A.31 Supplementary Notes for the Lecture of Wednesday, March 29th, 2006

Release Date: Wednesday, March 29th, 2006
Subject to further revision

A.31.1 §9.4 Diagonalization, Eigenvalues and Eigenvectors (continued)

Supplementary Problems

[**1, Exercise 9.48, p. 336 (conclusion)**] For each of the following matrices, find all eigenvalues, and a maximum set S of linearly independent eigenvectors.

3. $C = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}$

Solution:

3. The characteristic polynomial is $\begin{vmatrix} t-1 & -2 & -2 \\ -1 & t-2 & 1 \\ 1 & -1 & t-4 \end{vmatrix}$, and this polynomial factorizes as $(t-3)^2(t-1)$. When we set $t = 3$ in the matrix $tI - C$, we find that it row reduces to $\begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, corresponding to the equation $x = y + z$, whose null space consists of all vectors

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y+z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

spanned by the eigenvectors $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. The null space of the matrix $1I - B$ is spanned by $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$. For this matrix a maximal set of linearly independent eigenvectors of B contains only 3 members, and the matrix is diagonalizable. (We shall see later that eigenvectors corresponding to distinct eigenvalues must be linearly independent.)

I summarize some theorems in the textbook — theorems that we will not have time to derive completely, but which you will be expected to understand and to be able to apply:

Theorem A.97 *Let A be an $n \times n$ matrix.*

1. *The following statements are equivalent:*
   (a) *The scalar λ is an eigenvalue of A.*
   (b) *The matrix $\lambda I - A$ is singular.*
   (c) *λ is a root of the characteristic polynomial $\Delta(t) = tI - A$ of A.*
2. *Eigenvectors of A corresponding to distinct eigenvalues are linearly independent.*
3. *When the characteristic polynomial $\Delta(t) = tI - A$ of A admits a factorization into n distinct linear (i.e., 1st degree) factors $t - \lambda_i$ $(i = 1, 2, \ldots, n)$, then A is similar to a diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$ (in any order).*
4. *If λ is a root of the characteristic polynomial of multiplicity r, then there will exist at least 1 eigenvector, and not more than r corresponding linearly independent eigenvectors.*
5. *A is diagonalizable by a matrix P, i.e., $P^{-1}AP = D$ iff the columns of P are n linearly independent eigenvectors.*

A.31.2 The coefficients of the characteristic polynomial

Suppose that the characteristic polynomial of A, i.e., the polynomial $|tI - A|$ factorizes completely into 1st degree factors, as

$$|tI - A| = (t - \lambda_1)(t - \lambda_2)\ldots(T - \lambda_n)\,.$$

Then all of the coefficients of the polynomial can be expressed in terms of these eigenvalues, $\lambda_1, \ldots, \lambda_n$. In particular, the coefficient of t^{n-1} is $-\sum_{i=1}^{n} \lambda_i$, and the coefficient of t^0 is $\lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n$. These two statements can be paraphrased as follows:

Theorem A.98 *Suppose that the characteristic polynomial of A, i.e., the polynomial $|tI - A|$ factorizes completely into 1st degree factors, as*

$$|tI - A| = (t - \lambda_1)(t - \lambda_2)\ldots(T - \lambda_n)\,.$$

Then

1. $\text{trace}(A) = -\sum_{i=1}^{n} \lambda_i$

2. $\det(A) = (-1)^n \prod_{i=1}^{n} \lambda_i$

This provides a convenient way of verifying calculations of eigenvalues.

The other coefficients can also be expressed as *symmetric functions* of the eigenvalues.

A.31.3 Diagonalization of powers of a matrix

Suppose that an $n \times n$ matrix A admits diagonalization as $P^{-1}AP = D$. Then, if n is a non-negative integer, we can show that

$$P^{-1}A^nP = \left(P^{-1}AP\right)^n = D^n$$

so A^n admits the same eigenvectors as A, and the corresponding eigenvalues are the nth powers of the eigenvalues of A.

Square roots of a diagonalizable matrix Must a diagonalizable matrix A always admit a square root B, i.e., a matrix B such that $B^2 = A$? Obviously not, since, for example, the (diagonal) 1×1 matrix $\left(\begin{array}{c} -1 \end{array}\right)$ clearly cannot have a square root. Obviously we will need to require that the eigenvalues of A be non-negative. In that case it can be shown that the matrix will have a diagonalizable square root. Once we know that, we also know that the same matrix P can diagonalize any square root matrix B. Indeed, if μ_i is either square root of λ_i $(i = 1, 2, \ldots, n)$, then any matrix

$$P \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ 0 & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & \mu_n \end{pmatrix} P^{-1}$$

has the property that its square is PDP^{-1}, which is A.

A.32 Supplementary Notes for the Lecture of Friday, March 31st, 2006

Release Date: Friday, March 31st, 2006; updated April 19th, 2006;
Subject to further revision

Supplementary Problems

[**1, Exercise 9.51, p. 337**] Show that the matrices A and A^T have the same eigenvalues. Give an example of a 2×2 matrix A which has different eigenvectors from A^T.

Solution: Observe that the transpose of $tI - A$ is $(tI)^T - A^T = tI - A^T$. We know that a matrix and its transpose have the same determinant. Hence

$$\begin{aligned} \det(tI - A^T) &= \det(tI^T - A^T) \\ &= \det\left((tI - A)^T\right) \\ &= \det(tI - A)\,. \end{aligned}$$

Since matrices A, A^T have the same characteristic polynomials, the roots of those polynomials will be the same. Those roots are the eigenvalues of the respective matrices.

But having the same eigenvalues does not mean necessarily having the same eigenvectors. For example, let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then the characteristic polynomial of A is $\det \begin{pmatrix} t-1 & -1 \\ 0 & t-1 \end{pmatrix} = (t-1)^2$, so the only eigenvalue of A is 1, of multiplicity 2. When we substitute $t = 1$ in $tI - A$, we find its null space is generated by the vector $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The characteristic polynomial of A^T is again $(t-1)^2$, but the null space of $1I - A^T$ is generated by $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

[**1, Exercise 9.55, p. 337**] Let $E : \mathcal{V} \to \mathcal{V}$ be a projection mapping, that is $E^2 = E$. Show that E is diagonalizable, and, in fact, can be represented by a diagonal matrix whose diagonal entries consist of 0's and 1's, where the number of 1's is exactly the rank of E.

Solution: Since $E^2 = E$, we have a polynomial of degree 2 satisfied by the matrix. We saw above that this implies that the eigenvalues are roots of the polynomial $t^2 - t = t(t-1)$. However, it could happen that the *minimal* polynomial of the matrix is of lower degree: it could be t or $t - 1$ — since this polynomial can be shown to always be a divisor of all polynomials satisfied by the given matrix/transformation. Let's consider the possible cases.

Minimal polynomial is t: Here E has the entire space as its kernel, and so all non-zero vectors in $\mathbb{R}^n$ are eigenvectors, and any basis for $\mathbb{R}^n$ can form the columns of a diagonalizing matrix P, where the diagonalized matrix is 0. The eigenvalues are all equal to 0.

Minimal polynomial is $t - 1$: In this case $E = I$; again all non-zero vectors in $\mathbb{R}^n$ are eigenvectors, and any basis for $\mathbb{R}^n$ can form the columns of a diagonalizing matrix P, where the diagonalized matrix this time is I.

In all other cases the minimal polynomial of E is $t^2 - t = t(t-1)$. There will be eigenvectors corresponding to each of the eigenvalues 0 and 1, and no other eigenvalues. The eigenspace corresponding to eigenvalue 0 is the kernel of the function E, and the maximum number of linearly independent eigenvectors corresponding to eigenvalue 0 is the nullity of E. The number of eigenvectors corresponding to 1 will be the excess of the dimension of $\mathcal{V}$ over the nullity, i.e., the rank of E.

A.32.1 §9.7 Minimum Polynomial

If $\mathbf{x}$ is an eigenvector of a square matrix A, corresponding to eigenvalue λ, and if

$$p(t) = p_0 t^0 + p_1 t^1 + \cdots + p_r t^r$$

is any polynomial, then

$$p(A)\mathbf{x} = p(\lambda)\mathbf{x}\,.$$

Thus, if A satisfies a polynomial $p(t)$, then λ is a root of the polynomial. We know by the Cayley-Hamilton Theorem that A satisfies its own characteristic polynomial. There may possibly be polynomials of lower degree also satisfied by A. Among the roots of such polynomials will have to be the eigenvalues of A. In investigating polynomials satisfied by a matrix it is convenient to insist that the polynomials be *monic*.

Suppose that $p(x)$ is *any* polynomial satisfied by A, and $m(x)$ is a monic polynomial of minimum degree satisfied by A. Then we can divide $m(x)$ into $p(x)$, obtaining a *quotient* polynomial $q(x)$ and a *remainder* polynomial $r(x)$. The degree of $r(x)$ will be strictly less than the degree of the divisor, $m(x)$. This leads to an equation

$$p(x) = m(x) \cdot q(x) + r(x)\,.$$

If we evaluate these polynomials at A, then both $p(A)$ and $m(A)$ are known to be equal to the matrix 0, hence $r(A) = 0$. But m was a polynomial of lowest degree satisfied by A. From this contradiction we can conclude that $r(x)$ is the 0 polynomial, i.e., that

$m(x)$ is a factor of $p(x)$. So the minimum polynomial will always be a divisor of the characteristic polynomial, and it will have all the roots of that characteristic polynomial among its roots. In particular, if the roots of the characteristic polynomial are distinct, then the minimum polynomial must be the same as the characteristic polynomial. But, if some of the roots of the characteristic polynomial appear with multiplicities higher than 1, it is possible that the minimum polynomial may be of lower degree than the characteristic. We will not have time in this course to investigate this situation fully. (To do so it would be desirable to work over the field of complex numbers.)

Much more is known about the minimum polynomial, but we will not have time to go beyond this point in this course this term.

A.32.2 Final Examination, MATH 223 2004 09

1. Let $A = \begin{pmatrix} 1 & -3 & 0 & 1 \\ -1 & 2 & 1 & 2 \end{pmatrix}$

   (a) [5 MARKS] Find a basis for the null space, the column space, and the row space of A.

   (b) [2 MARKS] Find the nullity and rank of A.

   (c) [2 MARKS] Compute AA^T.

   (d) [1 MARK] Is AA^T invertible?

   Solution:

   (a) **Row Space:** Under row reduction, $A = \begin{pmatrix} 1 & -3 & 0 & 1 \\ -1 & 2 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -3 & 0 & 1 \\ 0 & -1 & 1 & 3 \end{pmatrix} \to$ $\begin{pmatrix} 1 & 0 & -3 & -8 \\ 0 & -1 & 1 & 3 \end{pmatrix}$. Hence one basis for the row space is

   $$\begin{pmatrix} 1 & 0 & -3 & -8 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 & 1 & 3 \end{pmatrix}.$$

   Null Space: The row reduced matrix we obtained corresponds to the system of equations

   $$\begin{aligned} x &= 3z + 8t \\ y &= z + 3t\,. \end{aligned}$$

   Hence the general solution of the system of equations

   $$A \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 3z+8t \\ z+3t \\ z \\ t \end{pmatrix} = z\begin{pmatrix} 3 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 8 \\ 3 \\ 0 \\ 1 \end{pmatrix},$$

and a basis for the null space is the 2 vectors $\begin{pmatrix} 3 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 8 \\ 3 \\ 0 \\ 1 \end{pmatrix}$.

Column Space: We can find a basis for the column space by naively column reducing A, which will lead to columns $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. A more sophisticated solution could be based on a theorem that tells us that the columns of A in which the pivots for our row space are located will constitute a basis for the column space. This means, in this case the columns $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} -3 \\ 2 \end{pmatrix}$.

(b) The *rank* is the dimension of the row space, which we have found to be 2. The *nullity* is the dimension of the kernel, equivalently, the excess of the number of variables over the rank, i.e., 4-2=2.

(c) $AA^T = \begin{pmatrix} 11 & -5 \\ -5 & 10 \end{pmatrix}$.

(d) Since $\det AA^T = 110 - 25 = 85 \neq 0$, AA^T is invertible. Its inverse can be found by either row reduction of a 2×4-matrix, or as $\frac{1}{|A|}$ times the classical adjoint: $A^{-1} = \frac{1}{85}\begin{pmatrix} 10 & 5 \\ 5 & 11 \end{pmatrix}$.

2. Let $\mathcal{V} = \operatorname{span}\{e^x, \sin x, \cos x\}$, where these functions are defined on $\mathbb{R}$.

   (a) [3 MARKS] Show that $B = \{e^x, \sin x, \cos x\}$ is linearly independent.

   (b) [3 MARKS] Let $T : \mathcal{V} \to \mathcal{V}$ be defined by $T(f) = f' - f$. Show that T is a linear transformation.

   (c) [3 MARKS] Find the kernel and range of T.

   (d) [3 MARKS] Find $[T]_B$.

   (e) [3 MARKS] Is T a linear isomorphism? Is it onto? Is it one-one?

   Solution:

(a) Let a, b, c be scalars.

$$\begin{aligned} & ae^x + b\sin x + c\cos x = 0 \quad \text{for all real numbers } x \\ \Rightarrow\ & ae^x + b\sin x + c\cos x = 0 \quad \text{for } x = 0, \tfrac{\pi}{2}, \pi \\ \Rightarrow\ & \begin{pmatrix} 1 & 0 & 1 \\ e^{\frac{\pi}{2}} & 1 & 0 \\ e^{\pi} & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

As the matrix of coefficients has determinant $\pi - e^{\pi} \neq 0$, the only solution is the trivial one, i.e., $a = b = c = 0$. Hence the functions are linearly independent.

A.33 Supplementary Notes for the Lecture of Monday, April 3rd, 2006

Release Date: Monday, April 3rd, 2006
Subject to further revision

A.33.1 Final Examination, MATH 223 2004 09 (continued)

1.

2. Let $\mathcal{V} = \text{span}\{e^x, \sin x, \cos x\}$, where these functions are defined on $\mathbb{R}$.

   (a) [3 MARKS] Show that $B = \{e^x, \sin x, \cos x\}$ is linearly independent.

   (b) [3 MARKS] Let $T : \mathcal{V} \to \mathcal{V}$ be defined by $T(f) = f' - f$. Show that T is a linear transformation.

   (c) [3 MARKS] Find the kernel and range of T.

   (d) [3 MARKS] Find $[T]_B$.

   (e) [3 MARKS] Is T a linear isomorphism? Is it onto? Is it one-one?

   Solution:

   (a) Let a, b, c be scalars.

$$\begin{array}{rll} & ae^x + b\sin x + c\cos x = 0 & \text{for all real numbers } x \\ \Rightarrow & ae^x + b\sin x + c\cos x = 0 & \text{for } x = 0, \frac{\pi}{2}, \pi \\ \Rightarrow & \begin{pmatrix} 1 & 0 & 1 \\ e^{\frac{\pi}{2}} & 1 & 0 \\ e^{\pi} & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \end{array}$$

   As the matrix of coefficients has determinant $\pi - e^{\pi} \neq 0$, the only solution is the trivial one, i.e., $a = b = c = 0$. Hence the functions are linearly independent.

   (b) Let a, b be scalars, and f, g be functions in $\mathcal{V}$. Then we need to investigate properties of $af + bg$.

$$\begin{aligned}
T(af+bg) &= (af+bg)' - (af+bg) \\
&= (af)' + (bg)' - (af+bg) \\
&\quad \text{by the Sum Rule for Differentiation} \\
&= af' + bg' - af - bg \\
&\quad \text{by the “Constant Multiple” Rule for Differentiation} \\
&= a(f'-f) + b(g'-g) \\
&= aTf + bTg\,.,
\end{aligned}$$

proving that T is linear.

(c) The kernel of T is the set of functions $f \in \mathcal{V}$ such that $Tf = 0$, i.e., such that $f' - f = 0$, i.e., such that $e^{-x}f'(x) - e^{-x}f(x) = 0$ for all x. But this last equation is equivalent to $\frac{d}{dx}\left(e^{-x}f(x)\right) = 0$, hence to $e^{-x}f(x) = C$, some constant, hence to $f(x) = Ce^x$. Thus the kernel is generated by e^x. (This could have been proved in a simpler way: consider the action of T on a sum $ae^x + b\sin x + c\cos x$, set that action equal to $\mathbf{0}$, and solve for a, b, c.)
$T\sin x = \cos x - \sin x$; $T\cos x = -\sin x - \cos x$. Hence $\sin x = T\left(\frac{-\sin x - \cos x}{2}\right)$, $\cos x = T\left(\frac{\sin x - \cos x}{2}\right)$. Thus the Image (Range) of T is spanned by $\sin x$ and $\cos x$.

(d) This is a straightforward computation: we determine the images of the three vectors in the coordinate system, expressed as linear combinations of themselves, in the prescribed orders, and detach the coefficients.

$$[T]_B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & -1 \end{pmatrix}.$$

(e) Omit the reference to a “linear isomorphism” we have not used this term.
T is not surjective(=onto): the vector e^t is not the image of any function in $\mathcal{V}$. Nor is T injective(=one-to-one), since there are distinct vectors with the same image, for example the vectors e^x, $2e^x$, and 0 all map to 0.

3. The next question referred to a *stochastic matrix*, a topic not covered in our textbook.

4. A volume question in $\mathbb{R}^4$. We have not studied volumes, (which can be defined using the determinant).

5. A question on partitioned matrices, which we will not be studying.

6. A question of differential equations, which we will not be studying.

7. On the space $\mathbb{R}_2[t]$, an inner product is defined by

$$\langle a + bt + ct^2, a' + b't + c't^2 \rangle = aa' + bb' + cc' .$$

$B_0 = \{1, t, t^2\}$.

(a) [4 MARKS] Show that $\langle , \rangle$ actually defines an inner product on $\mathbb{R}_2[t]$.

(b) [4 MARKS] Let $T(a+bt+ct^2) = (8a-2b+2c)+(-2a+5b+4c)t+(2a+4b+5c)t^2$. Find $[T]_{B_0}$, and show that it is a symmetric matrix.

(c) [2 MARKS] Briefly explain why T is orthogonally diagonalizable, without doing any calculation.

(d) [4 MARKS] Find $\ker T$ and range T.

(e) [2 MARKS] Is T invertible?

Solution:

(a) The product is just the dot product — find a reference in the textbook.

(b) To find the matrix we need to determine the action of T on the given ordered basis.

$$\begin{aligned}
T1 &= 8 - 2t + 2t^2 \\
Tt &= -2 + 5t + 4t^2 \\
Tt^2 &= 2 + 4t + 5t^2 \\
\Rightarrow [T]_{B_0} &= \begin{pmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{pmatrix} .
\end{aligned}$$

Evidently the matrix has reflective symmetry in its main diagonal.

(c) The matrix we found for T is symmetric. A symmetric matrix is known to have real eigenvalues; it is known that eigenvectors corresponding to distinct eigenvalues are orthogonal; and that, for any eigenvalue whose multiplicity as a root of the characteristic polynomial is r, there exists a set of r linearly independent eigenvectors. To those r eigenvectors we may apply the Gram-Schmidt process to obtain a mutually orthogonal set of eigenvectors. Hence a symmetric $n \times n$ matrix A possesses a set of n mutually orthogonal eigenvectors, which can form the columns of the *orthogonal* matrix P with the property that $P^{-1}AP$ is diagonal. (It wasn't necessary to do any more calculations here. It turns out that the characteristic polynomial is $(t-9)^2t$. While one eigenvalue, 9, appears with multiplicity 2, we know that there will exist two mutually orthogonal eigenvectors for this eigenvalue. The 0 eigenvalue is not surprising, since we know that the determinant of the given matrix is 0.)

(d) We row reduce the matrix of T, obtaining $\begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Hence the kernel is generated by $\begin{pmatrix} -\frac{1}{2} \\ -1 \\ 1 \end{pmatrix}$. By column reduction we find that the image is generated by $\begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

(e) T is not invertible: since its kernel is not empty, there are two distinct vectors which are mapped on to $\mathbf{0}$: an invertible transformation must be injective.

A.34 Supplementary Notes for the Lecture of Wednesday, April 5th, 2006

Release Date: Monday, April 3rd, 2006
Subject to further revision

A.34.1 Final Examination, MATH 223 2004 09 (continued)

8. (a) [3 MARKS] Find an orthogonal basis for

$$\mathcal{W} = \text{span}\left\{(1,-1,0,1),(2,1,0,0),(1,1,0,1)\right\}\ .$$

   (b) [1 MARKS] Find an orthonormal basis for $\mathcal{W}$.

   (c) (This question involved a "best approximation", a topic we will not be studying this term.)

   Solution:

   (a) Define

$$\begin{aligned}\mathbf{u}_1 &= (1,-1,0,1)\\ \mathbf{u}_2 &= (2,1,0,0)\\ \mathbf{u}_3 &= (1,1,0,1).\end{aligned}$$

   In a naive solution, without prior planning (see below), we can apply the Gram-Schmidt process to obtain an orthogonal basis, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$.

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{u}_1 = (1,-1,0,1)\\ \langle\mathbf{u}_2,\mathbf{v}_1\rangle &= 2-1+0+0=1\\ \langle\mathbf{v}_1,\mathbf{v}_1\rangle &= 1+1+0+1=3\\ \mathbf{v}_2 &= \mathbf{u}_2-\frac{1}{3}\mathbf{v}_1=\left(\frac{5}{3},\frac{4}{3},0,-\frac{1}{3}\right)\\ \langle\mathbf{u}_3,\mathbf{v}_1\rangle &= 1-1+0+1=1\\ \langle\mathbf{u}_3,\mathbf{v}_2\rangle &= \frac{5}{3}+\frac{4}{3}+0-\frac{1}{3}=\frac{8}{3}\\ \langle\mathbf{v}_2,\mathbf{v}_2\rangle &= \frac{25}{9}+\frac{16}{9}+0+\frac{1}{9}=\frac{14}{3}\\ \mathbf{v}_3 &= \mathbf{u}_3-\frac{1}{3}\mathbf{v}_1-\frac{8}{14}\mathbf{v}_2=\frac{2}{7}(-1,2,0,3).\end{aligned}$$

   This problem could have been made much simpler by replacing the given basis by one which would have been simpler to apply the Gram-Schmidt

Process. For example, row reduction would yield the alternative spanning set $\{(1,0,0,0),(0,1,0,0),(0,0,0,1)\}$ to which the application of the process would have been trivial!

Another simplification, even if we followed the naive solution above, would have been to scale the vectors so as to eliminate denominators.

(b) Continuing with the naive solution we gave originally, the lengths of the orthogonal vectors found are respectively $\sqrt{3}$, $\frac{\sqrt{14}}{3}$, $\frac{2}{7}\sqrt{14}$. The vectors must be multiplied by the reciprocals of their respective lengths.

9. Let $A = \begin{pmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{pmatrix}$.

(a) [4 MARKS] Find the eigenvalues and eigenspaces of A.

(b) [2 MARKS] Find an orthonormal basis for $\mathbb{R}^3$ consisting of eigenvectors of A.

(c) [4 MARKS] Find an orthogonal matrix Q and a diagonal matrix D such that $D = Q^T AQ$.

(d) [2 MARKS] Find $\min\limits_{\|X\|=1, X\in\mathbb{R}^3} X^T AX$ and $\max\limits_{\|X\|=1, X\in\mathbb{R}^3} X^T AX$

Solution:

(a) The characteristic polynomial is

$$|tI - A| = \begin{vmatrix} t-8 & 2 & -2 \\ 2 & t-5 & -4 \\ -2 & -4 & t-5 \end{vmatrix} = (t-9)^2 t\,,$$

having roots 9 (of multiplicity 2) and 0 (of multiplicity 1).

Eigenvalue $t = 9$: The matrix $9I - A$ row reduces to $\begin{pmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, whose kernel is the set of all vectors of the form

$$\begin{pmatrix} -2y+2z \\ y \\ z \end{pmatrix} = y\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

Thus the 2-dimensional eigenspace is generated by the two vectors $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$.

Eigenvalue $t = 0$: The matrix $0I - A$ row reduces to $\begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, whose kernel is the set of all vectors of the form

$$\begin{pmatrix} -\frac{1}{2}z \\ -z \\ z \end{pmatrix} = z\begin{pmatrix} -\frac{1}{2} \\ -1 \\ 1 \end{pmatrix}.$$

Thus the 1-dimensional eigenspace is generated by the vector $\begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}$.

(b) We know that the eigenvector found for eigenvalue 0 has to be — and is — orthogonal to both eigenvectors found for eigenvalue 9. We apply the Gram-Schmidt process to the latter two vectors in the order in which they are listed, obtaining a second eigenvector for eigenvalue 9:

$$\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\rangle} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}.$$

We now have three *orthogonal* eigenvectors:

$$\begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix},$$

whose lengths are, respectively $\sqrt{9}$, $\sqrt{5}$, and $\sqrt{45}$; hence the three eigenvectors can be normalized to yield an *orthonormal* set

$$\frac{1}{3}\begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}, \quad \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{3\sqrt{5}}\begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}.$$

(c) Any matrix Q with the preceding three vectors as its columns will orthogonally diagonalize A, for example $Q = \begin{pmatrix} -\frac{1}{3} & -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}$. The diagonal matrix obtained has the eigenvalues down the main diagonal in the order corresponding to the order of the eigenvectors in Q: i.e., 0, 9, 9.

(d) $X^TAX = \left(Q^TX\right)^T D \left(Q^TX\right)$ The vector X has length 1. When it is mapped by the orthogonal matrix Q, the resulting vector Q^TX also has length 1. Suppose that the vector Q^TX has components (a, b, c). Then the product X^TAX is equal to $0a^2 + 9(b^2 + c^2)$. This vector can be shown to be maximized when $a = 0$ and $b = c = \frac{1}{\sqrt{2}}$, and minimized when $a = 0$ and $b = c = -\frac{1}{\sqrt{2}}$; the extremal values are ± 9.

10. This problem is concerned with two topics that will not be studied this year (LU-factorization, QR-factorization).

11. This problem is concerned with a topic that will not be studied this year (least squares solutions of a linear system)

12. This problem is concerned with a topic that will not be studied this year (quadratic forms).

A.35 Supplementary Notes for the Lecture of Friday, April 7th, 2006

Release Date: Friday, April 7rd, 2006
Subject to further revision

A.35.1 Final Examination, MATH 223 1997 09

(The examiner was Professor S. Zlobec, with Associate Examiner A. Sathaye. The examination book contained the following statement on the cover: **Be sure to provide supporting calculations or arguments for all your conclusions; random guesses will not be awarded marks!**)

1. [11 MARKS] Let $\mathcal{V}$ be the subspace of $\mathbb{R}^4$ spanned by the vectors

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -2 \\ 2 \\ -2 \end{pmatrix}.$$

   Answer the following questions

   (a) [4 MARKS] What is the dimension of $\mathcal{V}$? Determine a basis for $\mathcal{V}$.

   (b) [3 MARKS] Let $\mathcal{W}$ be the set of all vectors in $\mathbb{R}^4$ which are perpendicular to every vector in $\mathcal{V}$. Explain why $\mathcal{W}$ is a subspace of $\mathbb{R}^4$. You should either prove the correct axioms or quote appropriate known theorems.

   (c) [3 MARKS] Determine the dimension of $\mathcal{W}$ and a basis for it.

   (d) [2 MARKS] Use the above to find a matrix M such that $\mathcal{V}$ is the null space of M.

   Solution:

   (a) Before solving this problem naively, I must observe that the last vector is obviously linearly dependent on the 2nd vector; hence no basis could contain both of them. If we consider the first and third vectors, we see from the presence of the 0's that they have to be linearly independent, since no scalar multiple of the one could equal the other. Now let's solve the problem naively.

Let's row reduce the matrix having the three vectors as rows:

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & -2 & 2 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (R_3 \rightarrow R_3 + 2R_2)$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 5 & -2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (R_1 \rightarrow R_1 - 2R_2)$$

The first 2 rows, as they contain pivot entries, constitute a basis for the row space of the matrix, which is the space spanned by the 3 given vectors.

As we have exhibited a basis with 2 elements, the dimension of $\mathcal{V}$ is 2.

(b) We are considering vectors $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ which are orthogonal to the three given vectors, hence to all linear combinations of those 3 vectors, hence to the space spanned by those vectors. We have seen in class (see notes) that $\mathcal{V}^\perp$ is a subspace: we have to prove

i. that the vector $\mathbf{0}$ is perpendicular to all three given vectors;[66]

ii. that the set is closed under addition;[67]

iii. that the set is closed under multiplication by a scalar.[68]

(c) $\mathcal{W}$ is the null space of the 3×4 matrix we set up earlier. The null space of that matrix is

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} -5z + 2t \\ z - t \\ z \\ t \end{pmatrix} = z \begin{pmatrix} -5 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

so one basis consists of the 2 vectors $\begin{pmatrix} -5 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$; the dimension is the size of any basis — hence 2.

[66]This is immediate, since the dot product of $\mathbf{0}$ with any vector is equal to 0.

[67]If $\langle \mathbf{v}, \mathbf{w}_1 \rangle = 0$ and $\langle \mathbf{v}, \mathbf{w}_2 \rangle = 0$, then $\langle \mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2 \rangle = \langle \mathbf{v}, \mathbf{w}_1 \rangle + \langle \mathbf{v}, \mathbf{w}_2 \rangle = 0 + 0 = 0$.

[68]This follows from the property that, for any scalar k, $\langle \mathbf{v}, \mathbf{w} \rangle = 0 \Rightarrow \langle \mathbf{v}, k\mathbf{w} \rangle = k \langle \mathbf{v}, \mathbf{w} \rangle = k0 = 0$.

(d) We have found the space $\mathcal{W} = \mathcal{V}^\perp$; it follows that $\mathcal{V} = \mathcal{W}^\perp$. (We didn't discuss this in the lectures: it depends on the finiteness of the dimensions.) One matrix can be formed from the basis vectors we found for $\mathcal{W}$, e.g., the matrix $\begin{pmatrix} -5 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{pmatrix}$.

2. Consider a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ which satisfies $T(\mathbf{v}_1) = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, $T(\mathbf{v}_2) = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$, where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Answer the following questions:

(a) [5 MARKS] Calculate the matrix representation for T relative to the (usual) standard bases in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

(b) [2 MARKS] Let $\mathbf{w} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Determine the coordinate vector $[\mathbf{w}]_B$ with respect to the basis $B = [\mathbf{v}_1, \mathbf{v}_2]$ for the domain.

(c) [2 MARKS] Using the above or otherwise, determine $T\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

(d) [2 MARKS] Determine the dimensions of the kernel and the range of the transformation T.

<u>Solution</u>:

(a) We are given that $[T]_{S'}^B = \begin{pmatrix} 2 & 4 \\ 1 & 1 \\ 3 & 3 \end{pmatrix}$, and wish to determine $[T]_{S'}^S$, where S, S' are the standard bases for $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively. We need to express the vectors of the standard basis as linear combinations of the given basis vectors. In general, suppose a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is given. We need to determine scalars a, b such that

$$a\begin{pmatrix} 1 \\ 2 \end{pmatrix} + b\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

i.e., such that

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

We can do this by inverting the 2×2 matrix, obtaining

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix};$$

that is $\left[\begin{pmatrix} x \\ y \end{pmatrix}\right]_B = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x+2y \\ 2x-y \end{pmatrix}$. Hence $\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_B = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$, $\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. It follows that the images under T of these basis vectors from the standard basis are the matrix products $[T]^B_{S'} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}$, $[T]^B_{S'} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, so

$$[T]^S_{S'} = \begin{pmatrix} 2 & 4 \\ 1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ -3 & 3 \end{pmatrix}.$$

We could also have consider the composition of functions given by the diagram

$$\begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{\mathbf{1}_{\mathbb{R}^2}} & \mathbb{R}^2 & \xrightarrow{T} & \mathbb{R}^3 \\ S & & B & & S' \end{array} \tag{97}$$

We know that $[\mathbf{1}_{\mathbb{R}^2}]^B_S = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, so $[\mathbf{1}_{\mathbb{R}^2}]^S_B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$. Then

$$[T]^S_{S'} = [T]^B_{S'} \, [\mathbf{1}_{\mathbb{R}^2}]^S_B = \begin{pmatrix} 2 & 4 \\ 1 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ -3 & 3 \end{pmatrix}$$

as before.

(b) Applying the preceding formula, we have $\left[\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right]_B = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

(c) $[T]^S_{S'} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ -3 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 3 \end{pmatrix}$.

(d) By the usual methods of row reduction we can show that the kernel consists only of the vector $\mathbf{0}$, having dimension 0. The rank is, therefore 2; the image

(=range) can be seen — by column reduction of the matrix $[T]_{S'}^{S}$ — to be generated by the two vectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$.

3. [6 MARKS] State without proof which of the following claims are true.

   (a) [2 MARKS] An $n \times n$ real matrix A and its transpose have the same eigenvalues.

   (b) [2 MARKS] If an $n \times n$ matrix A satisfies the equation $A^2 = 0$, then $\lambda = 0$ is its only eigenvalue with multiplicity n.

   (c) [2 MARKS] Assume that a 5×5 matrix A has three distinct eigenvalues, p, q, r, such that the eigenspaces of p, q are each 2-dimensional. The matrix cannot be diagonalized.

   Solution:

   (a) This is true, since A and its transpose have the same characteristic polynomial.

   (b) The statement is not well formulated: there should be a comma after the word "eigenvalue". The eigenvalues can only be among the roots of any equation satisfied by the matrix: as t^2 has only 0 as its roots, the only possible eigenvalue for A is 0, and it must be repeated as often as needed. (This does not, however, say that A can be diagonalized.)

   (c) If p and q each have eigenspaces which are 2 dimensional, then r must have a 1-dimensional eigenspace. Eigenvectors corresponding to distinct eigenvalues are known to be linearly independent. Thus, if we choose 2 linearly independent eigenvectors for each of p and q, and one eigenvector independent of all of them corresponding to eigenvalue r, we have the columns of a 5×5 matrix which will diagonalize A. The statement is FALSE.

A.36 Supplementary Notes for the Lecture of Monday, April 10th, 2006

Release Date: Friday, April 10th, 2006
Subject to further revision

A.36.1 Final Examination, MATH 223 1997 09 (conclusion)

4. [12 MARKS] A linear transformation $T : \mathbb{R}_2[t] \to \mathbb{R}_2[t]$ is represented in the coordinate system $B = [1, t, t^2]$ by $[T]_B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix}$.

   (a) [10 MARKS] Find the representation of T relative to the coordinate system $[1 + t, t, 1 + t^2]$.

   (b) [2 MARKS] Calculate $T(1 + t + t^2)$.

   Solution:

   (a)

$$\begin{aligned}
[T(1+t)]_B &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \\
[T(t)]_B &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\
[T(1+t^2)]_B &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \\
\Rightarrow [T]_B^{1+t,t,1+t^2} &= \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ 1 & 1 & 0 \end{pmatrix}.
\end{aligned}$$

   What we have shown is that

$$[T]_B^{1+t,t,1+t^2} = [T]_B^B \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By the same reasoning as seen in an earlier problem, we can prove that

$$\begin{aligned}
[T]_{1+t,t,1+t^2} &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} [T]_B^B \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 & 2 \\ 2 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}.
\end{aligned}$$

(b) $[T(1+t+t^2)]_B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$, so $T(1+t+t^2) = 2+4t+t^2$.

5. This problem was concerned with a stochastic matrix for a Markov chain, a topic we have not studied this term.

6. [12 MARKS] Suppose that an inner product is to be defined on the vector space $\mathbb{R}_2[t]$ by the matrix $M = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 5 \end{pmatrix}$. This means that the inner product of t^i and t^j is computed by looking up the $(i+1, j+1)$th entry of M. This works for all $0 \le i, j \le 2$. More explicitly, the rows and columns correspond to the basis $1, t, t^2$ respectively, and thus, to find the inner product $\langle 1, t\rangle = \langle t^0, t^1\rangle$ you look up the $(1, 2)$-th entry, which is 2. Similarly, the product $\langle t^2, t^2\rangle$ is the $(3, 3)$-th entry, 5.

   (a) [9 MARKS] Use this definition of the inner product to convert the basis $1, t, t^2$ of $\mathbb{R}_2[t]$ to an orthonormal basis. Use the Gram-Schmidt process, either as in the book, or use the direct matrix method as described in the handout.

   (b) [3 MARKS] Use the above definition of the inner product to calculate $\langle 1 + t + t^2, 2 + t^2\rangle$.

   Solution:

   (a) Linearity of the inner product implies that

$$\langle at^0 + bt^1 + ct^2, dt^0 + et^1 + ft^2\rangle = \begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 5 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix}.$$

We proceed to apply the Gram-Schmidt process to the vectors 1, t, t^2, in that order.

$$\begin{aligned}
\mathbf{v}_1 &= 1\\
\mathbf{v}_2 &= t - \frac{\langle t, 1\rangle}{\langle 1, 1\rangle}1\\
&= t - \frac{2}{1}1 = t-2\\
\mathbf{v}_3 &= t^2 - -\frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle t^2, t-2\rangle}{\langle t-2, t-2\rangle}t-2\\
&= t^2 - \frac{0}{1}1 - \frac{1}{1}(t-2) = t^2 - t + 2\,.
\end{aligned}$$

In these calculations we have already shown that $\|1\| = 1$ and $\|t-2\| = 1$. To complete normalization, we need to compute

$$\langle 2-t+t^2, 2-t+t^2\rangle = \begin{pmatrix} 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0\\ 2 & 5 & 1\\ 0 & 1 & 5 \end{pmatrix}\begin{pmatrix} 2\\ -1\\ 1 \end{pmatrix} = 4$$

implying that $\|2-t+t^2\| = 2$. Hence one orthonormal basis is

$$1, -2+t, 1 - \frac{1}{2}t + \frac{1}{2}t^2\,.$$

(b)

$$\langle 1t^0 + 1t^1 + 1t^2, 2t^0 + 0t^1 + 1t^2\rangle = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0\\ 2 & 5 & 1\\ 0 & 1 & 5 \end{pmatrix}\begin{pmatrix} 2\\ 0\\ 1 \end{pmatrix} = 12.$$

7. [13 MARKS] A linear transformation T is defined in the space $\mathcal{V}$ of 2×2 real matrices by $T(X) = X\begin{pmatrix} 1 & 1\\ 0 & 1 \end{pmatrix}$.

   (a) [4 MARKS] Use the following standard basis for $\mathcal{V}$:

   $$E_1 = \begin{pmatrix} 1 & 0\\ 0 & 0 \end{pmatrix},\quad E_2 = \begin{pmatrix} 0 & 1\\ 0 & 0 \end{pmatrix},\quad E_3 = \begin{pmatrix} 0 & 0\\ 1 & 0 \end{pmatrix},\quad E_4 = \begin{pmatrix} 0 & 0\\ 0 & 1 \end{pmatrix}.$$

   Determine the matrix representation of T with respect to this basis.

(b) [4 MARKS] Find the eigenvalues of T, and hence verify that T is nonsingular[69].

(c) [2 MARKS] Find a 2×2 matrix X such that $T(X) = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$.

(d) [3 MARKS] Does $\mathcal{V}$ have a basis consisting of eigenvalues of T? Why? If it exists, find such a basis.

Solution:

(a)

$$\begin{aligned}
T(E_1) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1E_1 + 1E_2 + 0E_3 + 0E_4 \\
T(E_2) &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0E_1 + 1E_2 + 0E_3 + 0E_4 \\
T(E_3) &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = 0E_1 + 0E_2 + 1E_3 + 1E_4 \\
T(E_4) &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0E_1 + 0E_2 + 0E_3 + 1E_4 \\
\Rightarrow [T]_{E_1,E_2,E_3,E_4} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.
\end{aligned}$$

(b) The characteristic polynomial is $\det \begin{pmatrix} t-1 & 0 & 0 & 0 \\ -1 & t-1 & 0 & 0 \\ 0 & 0 & t-1 & 0 \\ 0 & 0 & -1 & t-1 \end{pmatrix} = (t-1)^4$.
The eigenvalues are all equal to 1; for a transformation or matrix to be singular there needs to be a 0 eigenvalue.

(c) Since the transformation is non-singular, it has an inverse. The inverse matrix is

$$\left[T^{-1}\right]_{E_1,E_2,E_3,E_4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

[69] i.e., that T is injective

The coordinates of the inverse matrix sought are, therefore,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

representing the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. We could have found this also simply by solving the equation $T(X) = X \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$:

$$X = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

(d) Note that there is a typographical error in this question: it asks for a basis of eigen*values*, but surely means eigen*vectors*.

The transformation T does not possess 4 linearly independent eigenvectors. The system of equations that would have to be solved for eigenvalue 1 can be seen to have rank 2. We can show that the eigenspace is generated by the matrices E_2 and E_4.

8. [12 MARKS] Consider the matrix $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

(a) [4 MARKS] Find all the eigenvalues of A.

(b) [6 MARKS] Find the bases for the eigenspaces for each of the eigenvalues found above.

(c) [2 MARKS] Orthogonally diagonalize A, or prove that this is not possible.

<u>Solution</u>:

(a) The characteristic polynomial is $\det \begin{pmatrix} t-2 & 0 & 0 & 0 \\ 0 & t-2 & 0 & 0 \\ 0 & 0 & t-1 & -1 \\ 0 & 0 & -1 & t-1 \end{pmatrix} = t(t-2)^3$. The eigenvalues are thus 0 (once) and 2 (three times).

(b) There is a small error in this problem. One should not speak of *the* bases, as bases are usually not unique.

For eigenvalue 0 the eigenspace is the null space of the matrix $-A$, which matrix reduces to $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. So the eigenspace is generated by the vector $\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$.

For eigenvalue 2 the eigenspace is the null space of the matrix $\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

One basis for that eigenspace is the set of 3 vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

(c) The 4 eigenvectors given are mutually orthogonal; they can be normalized by multiplying them respectively by $2^{-\frac{1}{2}}$, 1, 1, $2^{-\frac{1}{2}}$. The matrix will diagonalize to one having its eigenvalues along the main diagonal: 0, 2, 2, 2 respectively.

9. [10 MARKS] Consider the matrix $A = \begin{pmatrix} \alpha & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

(a) [6 MARKS] What are the conditions on α for which the quadratic form $Q(X) = X^T AX$ is positive definite?

(b) [4 MARKS] What are the conditions on α such that the function $f(X, Y) = X^T AY$ defines an inner product on $\mathbb{R}^3$? Explicitly find a value of α for which **such a function does not define an inner product** on $\mathbb{R}^3$. Explain what goes wrong.

<u>Solution</u>: We have not covered this topic this term.

(a) There are a number of equivalent criteria for a quadratic form to be positive definite. For this problem the most convenient is probably that all of the upper left submatrices should have positive determinants. This means that we need to determine the values of the following 3 determinants:

$$\det\left(\begin{array}{c}\alpha\end{array}\right) = \alpha \det\left(\begin{array}{cc}\alpha & 1\\ 1 & 2\end{array}\right) = 2\alpha - 1$$

$$\det\left(\begin{array}{ccc}\alpha & 1 & 1\\ 1 & 2 & 0\\ 1 & 0 & 1\end{array}\right) = 2\alpha - 2$$

The condition reduces to $\alpha > 1$.

B Assignments and Tests from Previous Years

The only past year for which I can find notes in my files is the following. THIS MATERIAL HAD TO BE INCORPORATED WITHOUT FULL EDITING — BE ALERT TO POSSIBLE MISPRINTS AND OTHER ERRORS. The textbook used at the time was an earlier version of your textbook, in which the materials had a different order.

B.1 Fall 1993 Problem Assignments

B.1.1 First 1993 Problem Assignment, with Solutions

Release Date: Friday, September 24th, 1993
(Solutions were to be submitted by Monday, September 20th, 1993)

In all of the following problems you are expected to *show all of your work.* Some of the problems are computationally challenging, and are not typical of examination questions. The assignment is intended as a review and learning experience: you may not have been shown in the lectures how to solve every problem, although there are no tricks needed for any of them. In most of these problems there will be more than one method that can be used for a solution. You will be provided with solutions showing one such method, not necessarily the "best". Whatever method you use, try to *verify* your work in some way; if you can't find any other way of verifying, you can always solve the problem a second time using a different method.

1. *By row-reducing a matrix,* test the following systems of equations for consistency and — if they are consistent — determine the general solution, and verify. *Show all your work, carefully indicating what operations you are applying at any given stage.* Where the coefficients depend on parameters, your analysis may also depend on those parameters.

   (a)

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 &= 3 && (98)\\ x_1 + 4x_2 + 5x_3 + 2x_4 &= 2 && (99)\\ 2x_1 + 9x_2 + 8x_3 + 3x_4 &= 7 && (100)\\ 3x_1 + 7x_2 + 7x_3 + 2x_4 &= 12 && (101)\\ 5x_1 + 7x_2 + 9x_3 + 2x_4 &= 20 && (102) \end{aligned}$$

   <u>Solution</u>: The "augmented" matrix of the system is $\left(\begin{array}{cccc|c} \mathbf{1} & 2 & 3 & 1 & 3\\ 1 & 4 & 5 & 2 & 2\\ 2 & 9 & 8 & 3 & 7\\ 3 & 7 & 7 & 2 & 12\\ 5 & 7 & 9 & 2 & 20 \end{array}\right)$.

In row reduction we select one row with a non-zero entry as far left as possible: it is convenient — but not mandatory — to use the first row in this example. Subtracting multiples of that row from the others yields $\left(\begin{array}{cccc|c} 1 & 2 & 3 & 1 & 3 \\ 0 & 2 & 2 & 1 & -1 \\ 0 & 5 & 2 & 1 & 1 \\ 0 & \mathbf{1} & -2 & -1 & 3 \\ 0 & -3 & -6 & -3 & 5 \end{array}\right)$. Here we must select a row having a non-zero entry as far left as possible excluding the column containing the 1 we have used as our first *pivot*. We select the fourth row, although any of rows ##2, 3, 4, or 5 could be used, and subtract multiples from the other rows. (Note that in these first two stages of the row-reduction algorithm the pivot entry has been 1, so we have not had to scale the row by multiplying by a non-zero scalar.) We obtain $\left(\begin{array}{cccc|c} 1 & 0 & 7 & 3 & -3 \\ 0 & 0 & \mathbf{6} & 3 & -7 \\ 0 & 0 & 12 & 6 & -14 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 0 & -12 & -6 & 14 \end{array}\right)$. At this point a convenient pivot is found in the first non-zero entry in the second row. This entry is not 1 at present, so we could scale the row by multiplying by 1/6. (An alternative method is simply to multiply the other rows by appropriate factors prior to addition or subtraction.) We obtain $\left(\begin{array}{cccc|c} 1 & 0 & 0 & -\frac{1}{2} & \frac{31}{6} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{7}{6} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$. Except for rearrangement of the rows, which yields $\left(\begin{array}{cccc|c} 1 & 0 & 0 & -\frac{1}{2} & \frac{31}{6} \\ 0 & 1 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{7}{6} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$, the matrix is in *row-canonical* form.

It corresponds to the system of equations

$$x_1 = \frac{1}{2}x_4 + \frac{31}{6} \tag{103}$$

$$x_2 = 0x_4 + \frac{2}{3} \tag{104}$$

$$x_3 = -\frac{1}{2}x_4 - \frac{7}{6} \tag{105}$$

which show that any value of x_4 can be extended to values for the other three variables which satisfy all of the five original equations. Thus the original system is consistent, and the "general solution" is given by the last three equations. (Actually, there are two other equations in the last mentioned equivalent system, specifically $0 = 0$, $0 = 0$.) These equations give the solutions in *parametric* form. We could introduce a new variable as parameter, and rewrite the system as

$$\begin{aligned} x_1 &= \frac{1}{2}t + \frac{31}{6} &(106)\\ x_2 &= 0t + \frac{2}{3} &(107)\\ x_3 &= -\frac{1}{2}t - \frac{7}{6} &(108)\\ x_4 &= 1t + 0, &(109) \end{aligned}$$

which you should recognize as equations of a line in 4-dimensional Euclidean space.

(b)

$$\begin{aligned} \lambda x_1 + x_2 + x_3 &= 2 &(110)\\ x_1 + \lambda x_2 + x_3 &= 3 &(111)\\ x_1 + x_2 + \lambda x_3 &= 4 &(112) \end{aligned}$$

<u>Solution</u>: This system can be attacked from first principles, and we shall proceed in that way below. We observe, though, that a more pleasing solution could be found by making use of the symmetry of the coefficients. For example, we could first add the three equations together (i.e. replace the first by the sum of the three); from such an equation it is clear that there cannot be a solution unless $\lambda + 2 \neq 0$.

While we shall proceed from first principles, there is no need to make the computations unnecessarily difficult. Thus, were we to take the first equation to contain the pivot entry in the first column, we would have to assume that $\lambda \neq 0$. We delay any such problems by taking the *second* row as containing a pivot in position (2,1) (i.e. row #2, column #1). The original augmented matrix, $\left(\begin{array}{ccc|c} \lambda & 1 & 1 & 2\\ \mathbf{1} & \lambda & 1 & 3\\ 1 & 1 & \lambda & 4 \end{array}\right)$, transforms to $\left(\begin{array}{ccc|c} 0 & 1-\lambda^2 & 1-\lambda & 2-3\lambda\\ 1 & \lambda & 1 & 3\\ 0 & 1-\lambda & \lambda-1 & 1 \end{array}\right)$. The second column contains, outside of the row in which we have already selected a pivot, entries $1-\lambda^2 = (1-\lambda)(1+\lambda)$ and $1-\lambda$. We recognize two cases:

<u>**Case $\lambda = 1$**</u>**:** The third row is then $\left(\begin{array}{ccc|c} 0 & 0 & 0 & 1 \end{array}\right)$, corresponding to a contradictory equation $0 = 1$. We conclude that there is *no solution* when $\lambda = 1$.

<u>**Case $\lambda \neq 1$**</u>**:** We may scale the third row to obtain a pivot there. Subtracting appropriate multiples from the other two rows yields, after reduction, $\left(\begin{array}{ccc|c} 0 & 0 & (2+\lambda)(1-\lambda) & 1-4\lambda \\ 1 & 0 & 1+\lambda & \frac{3-4\lambda}{1-\lambda} \\ 0 & 1 & -1 & \frac{1}{1-\lambda} \end{array}\right)$. As we have assumed that $\lambda \neq 1$ we may scale the first row, which becomes the third after rearrangement: $\left(\begin{array}{ccc|c} 1 & 0 & 1+\lambda & \frac{3-4\lambda}{1-\lambda} \\ 0 & 1 & -1 & \frac{1}{1-\lambda} \\ 0 & 0 & 2+\lambda & \frac{1-4\lambda}{1-\lambda} \end{array}\right)$. There are now two sub-cases to consider:

<u>**Subcase $\lambda = -2$**</u>**:** Here the last row becomes $\left(\begin{array}{ccc|c} 0 & 0 & 0 & 3 \end{array}\right)$, a contradiction. Thus there is *no solution* in this case.

<u>**Subcase $\lambda \neq -2, 1$**</u>**:** $\left(\begin{array}{ccc|c} 1 & 0 & 1+\lambda & \frac{3-4\lambda}{1-\lambda} \\ 0 & 1 & -1 & \frac{1}{1-\lambda} \\ 0 & 0 & 1 & \frac{1-4\lambda}{(2+\lambda)(1-\lambda)} \end{array}\right)$ further transforms to $\left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{5-2\lambda}{(2+\lambda)(1-\lambda)} \\ 0 & 1 & 0 & \frac{3}{2+\lambda} \\ 0 & 0 & 1 & \frac{1-4\lambda}{(2+\lambda)(1-\lambda)} \end{array}\right)$. There is only one particular solution to this system, given by the fourth column of this matrix. We *verify* by substitution in the original equations.

(c) (Optional: This problem was described as optional, and not to be graded, but it was announced that a solution would be circulated.)

$$\begin{array}{rcl} 6x_1 - 2x_2 + 3x_3 + 4x_4 + 9x_5 & = & \lambda \\ 3x_1 - 1x_2 + 2x_3 + 6x_4 + 3x_5 & = & \mu \\ 6x_1 - 2x_2 + 5x_3 + 20x_4 + 3x_5 & = & 0 \\ 9x_1 - 3x_2 + 4x_3 + 2x_4 + 15x_5 & = & 0 \end{array}$$

<u>Solution</u>: The augmented matrix in this case is $\left(\begin{array}{ccccc|c} 6 & -2 & 3 & 4 & 9 & \lambda \\ \mathbf{3} & -1 & 2 & 6 & 3 & \mu \\ 6 & -2 & 5 & 20 & 3 & 0 \\ 9 & -3 & 4 & 2 & 15 & 0 \end{array}\right)$.

We may select the entry in position (2,1) as the first pivot, and subtract multiples of the second row from the others; we could scale the row first, but procrastinate in order to delay the appearance of fractions in our computations: $\left(\begin{array}{ccccc|c} 0 & 0 & -1 & -8 & 3 & \lambda-2\mu \\ 3 & -1 & 2 & 6 & 3 & \mu \\ 0 & 0 & \mathbf{1} & 8 & -3 & -2\mu \\ 0 & 0 & -2 & -16 & 6 & -3\mu \end{array}\right)$. Now we select the entry in position (3,3) as the second pivot; (note that there is no pivot in the second column of the matrix): $\left(\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 0 & \lambda-4\mu \\ 3 & -1 & 0 & -10 & 9 & 5\mu \\ 0 & 0 & 1 & 8 & -3 & -2\mu \\ 0 & 0 & 0 & 0 & 0 & -7\mu \end{array}\right)$. The corresponding system of equations will be contradictory unless $\lambda - 4\mu = 0 = -7\mu$, i.e. unless $\lambda = 0 = \mu$. Where this is the case the system becomes, after scaling and rearrangement,

$$x_1 = \frac{1}{3}x_2 + \frac{10}{3}x_4 - 3x_5 \tag{113}$$
$$x_3 = 0x_2 - 8x_4 + 3x_5 \tag{114}$$
$$0 = 0 \tag{115}$$
$$0 = 0. \tag{116}$$

As in the first system above, this could be rephrased in terms of three parameters: $x_1 = \frac{1}{3}t + \frac{10}{3}u - 3v$, $x_2 = t$, $x_3 = -8u + 3v$, $x_4 = u$, $x_5 = v$.

2. The following simple problems are intended as a review of geometric material in Mathematics 189–221 or CEGEP –105.

   (a) Find an equation of the hyperplane[70] in $\mathbb{R}^3$ which

      i. passes through $(-7, 8, -9)$ and is normal to the direction of the line joining $(1, 2, -3)$ to $(4, 5, -6)$;
      Solution: The normal direction is that of $(4-1)\mathbf{i}+(5-2)\mathbf{j}+(-6+3)\mathbf{k}$, i.e. of $\mathbf{i}+\mathbf{j}-\mathbf{k}$. Thus the plane has equation of the form $x+y-z = D$, where D is a constant. Imposing the condition that the coordinates of $(-7, 8, -9)$ satisfy the equation, we find the constant D to be $-7 + 8 - (-9) = 10$, so the equation is $x - y + z = 10$.

      ii. contains the points $(0, -2, -1)$, $(1, 4, 0)$, $(-2, -2, 1/2)$;
      Solution: There are several effective methods of solving a problem of this type. One method would be to assume the equation to be $\ell x+my+nz = d$, then impose the condition that the three sets of coordinates satisfy the equation, to obtain three linear equations in the four undetermined

[70] A *hyperplane* of a space of dimension n is a "flat" of dimension $n - 1$ — i.e. the subset obtained by adding a constant vector to all points in a subspace of dimension $n - 1$; a hyperplane of $\mathbb{R}^3$ is a plane.

constants. These could be solved to express the constants as multiples of a parameter, which would yield the equation.
Another method is to determine vectors joining one of the points to the other two, viz. $\mathbf{i}+6\mathbf{j}+\mathbf{k}$ and $-2\mathbf{i}-(3/2)\mathbf{k}$; their cross product $-9\mathbf{i}-(1/2)\mathbf{j}+12\mathbf{k}$ yields the normal direction of the plane. The equation could then be determined as in the preceding part: $18x+y-24z = 18(0)+1(-2)-24(-1) = 22$.

iii. contains the point $(-1,-4,8)$, and is parallel to the plane $3y-7z=16$.
<u>Solution</u>: All planes parallel to $3y-7z=16$ have equation of the form $3y-7z=$ constant. Imposing the condition that $(-1,-4,8)$ lie on the plane determines the constant to be $0(-1)+3(-4)-7(8)=-68$.

(b) Determine the value of k such that the plane $-2x+ky-(3k+2)z = 1-k^2$ is perpendicular to $x+2y-3z=6$. ("Two hyperplanes are perpendicular iff[71] corresponding normal vectors are orthogonal.")

<u>Solution</u>: The respective normal vectors are $-2\mathbf{i}+k\mathbf{j}-(3k+2)\mathbf{k}$, $\mathbf{i}+2\mathbf{j}-3\mathbf{k}$, which will be orthogonal iff their dot product, $-2+2k+(9k+6)$ is zero, i.e. iff $k=-4/11$.

(c) Find a parametric representation of the line which

i. passes through $(6,-1,0)$ and $(1,1,\pi)$;
<u>Solution</u>: The direction of the line is that of the vector $(6-1)\mathbf{i}+(-1-1)\mathbf{j}+(0-\pi)\mathbf{k}$. One set of parametric equations is $x=6+5t$, $y=-1-2t$, $z=0-\pi t$. (Question: How can you determine if two sets of parametric equations represent the same line?)

ii. passes through $(6,-1,0)$ and has the direction of the line of intersection of the planes $x-2y+3z=1$, $y-4z=-2$;
<u>Solution</u>: The direction is that of the cross product of $\mathbf{i}-2\mathbf{j}+3\mathbf{k}$ and $0\mathbf{i}+1\mathbf{j}-4\mathbf{k}$, i.e. $5\mathbf{i}+4\mathbf{j}+\mathbf{k}$. One set of parametric equations is $x=6+5u$, $y=-1=4u$, $z=0+u$. (Note that we may use any symbol for the parameter; indeed, in problems involving several curves or surfaces we are obliged to use different symbols.)

iii. passes through $(0,3,2)$ and is perpendicular to the plane $7y-4x+z=11$.
<u>Solution</u>: The line will be in the normal direction to the plane, viz. $-4\mathbf{i}+7\mathbf{j}+\mathbf{k}$. One set of equations is $x=0-4v$, $y=3+7v$, $z=2+v$.

3. (a) $z_1 = x_1+iy_1$ and $z_2 = x_2+iy_2$ are two complex numbers. Show that $\overline{z_1+z_2} = \overline{z_1}+\overline{z_2}$ and $\overline{z_1\cdot z_2} = \overline{z_1}\cdot\overline{z_2}$.

[71] if and only if

(b) (This problem was described as optional, and not to be graded, but it was announced that a solution would be circulated.)

If $z = a + ib$ is a complex number, and a and b are real, we call a the *real part* of z, and b the "*imaginary*" part, and write $a = \Re z$, $b = \Im z$. Since

$$z = |z|\left(\frac{a}{|z|} + i\frac{b}{|z|}\right),$$

and

$$\left(\frac{a}{|z|}\right)^2 + \left(\frac{b}{|z|}\right)^2 = 1,$$

there exists a real number θ (called an *argument* of z) such that $\Re z = |z|\cos\theta$ and $\Im z = |z|\sin\theta$; evidently, if θ is an argument of z, then $\theta + 2\pi n$ is also an argument, for any integer n.

Let $z_1 = 2 - 3i$, $z_2 = -1 + 5i$, $z_3 = 2i$, and $z_4 = -7$. Evaluate the following, writing each answer first in the standard form $a + bi$ with a and b real, *and then in the form* $|z|(\cos\theta + i\sin\theta)$. Arguments should be correct to 3 decimal places.[72]

i. $(1 + i)\overline{z_1}$

ii. $z_3 z_4$

iii. z_2^4

iv. $|z_1|$

v. $|i\overline{z_2}|$

vi. z_2/z_1 (Hint: $\frac{1}{z} = \frac{1}{z}\frac{\overline{z}}{\overline{z}} = \frac{\overline{z}}{|z|^2}$.)

<u>Solution</u>:

(a)

$$\begin{aligned}\overline{z_1 + z_2} &= \overline{(x_1 + x_2) + i(y_1 + y_2)}\\ &= (x_1 + x_2) - i(y_1 + y_2) = (x_1 - iy_1) + (x_2 - iy_2)\\ &= \overline{z_1} + \overline{z_2}\end{aligned}$$

$$\begin{aligned}\overline{z_1 z_2} &= \overline{(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)}\\ &= (x_1x_2 - y_1y_2) - i(x_1y_2 + x_2y_1)\\ &= (x_1 - iy_1)(x_2 - iy_2) = \overline{z_1}\,\overline{z_2}\end{aligned}$$

[72]Note that, although calculators are not permitted on the test or examination in this course, they may be needed for this particular homework problem.

(b) i. $(1+i)\overline{z_1} = (1+i)(2+3i) = -1+5i$. $|-1+5i| = \sqrt{1+25} = \sqrt{26}$. The argument is a real number θ such that $\cos\theta = -\frac{1}{\sqrt{26}}$, $\sin\theta = \frac{5}{\sqrt{26}}$. One approximate value is $\theta = 1.7682$ (here measured in radians). The argument may also be expressed as $\pi - \arctan 5$. (Note that, since the image of the arctangent function is usually taken to be the open interval $(-\frac{\pi}{2}, \frac{\pi}{2})$, there is no way to express this argument purely as an arctangent.)

ii. $z_3z_4 = (2i)(-7) = -14i$; the *modulus* (another word for *magnitude*) is 14; the argument is $\frac{3\pi}{2}$, since this complex number is located on the negative imaginary axis.

iii. $(-1+5i)^4 = 1-4(5i)+6(5i)^2-4(5i)^3+(5i)^4 = (1-150+625)+i(-20+500) = 476+i480$. The modulus is $\sqrt{476^2+480^2} = 676$. The argument is a real number θ such that $\cos\theta = \frac{119}{169}$, $\sin\theta = \frac{120}{169}$, approximately $0.7896+2n\pi$, where n is any integer. The argument may also be expressed as $\arctan\frac{120}{119}$.

iv. $|z_1| = \sqrt{2^2+(-3)^2} = \sqrt{13}+i0$. The argument is $0+2n\pi$ for any integer n, since this complex number is located on the *real axis*.

v. $|i\overline{z_2}| = |i\cdot\overline{-1+5i}| = |i(-1-5i)| = |5-i| = \sqrt{25+1} = \sqrt{26}+0i$ in *standard form*. The argument is again 0 etc., since a modulus, being real, is located on the real axis (indeed, only on the *nonnegative* real axis).

vi.

$$\begin{aligned}\frac{z_2}{z_1} &= \frac{-1+5i}{2-3i} = \frac{-1+5i}{2-3i}\cdot\frac{2+3i}{2+3i}\\ &= \frac{(-1+5i)(2+3i)}{|2-3i|^2} = \frac{-17+7i}{13}\\ &= -\frac{17}{13}+i\frac{7}{13}\end{aligned}$$

The modulus is $\sqrt{\frac{17^2+7^2}{13^2}} = \frac{\sqrt{338}}{13}$. The argument is a real number θ such that $\cos\theta = -\frac{17}{\sqrt{338}}$, $\sin\theta = \frac{7}{\sqrt{338}}$, approximately 2.7510 (or any real number obtained from this by adding an even integer multiple of π.) The argument may also be expressed as $\pi - \arctan\frac{7}{17}$.

B.1.2 Second 1993 Problem Assignment, with Solutions

Distribution Date: Friday, October 8th, 1993
Solutions were to be submitted by Monday, October 4th, 1993

1. Calculate the value of each of the polynomials at each of the matrices that follow:

(i) $x^2 - 3x + 2$ $(= (x-1)(x-2))$ (ii)$x^3 + 2x^2 + x + 2$ $(= (x^2+1)(x+2))$

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 5 & -2 \\ 6 & -2 \end{pmatrix}$ (d) $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$

Solution: Let f and g stand for the polynomials in (i) and (ii), respectively and let A, B, C and D stand for the matrices in parts a, b, c and d, respectively. We calculate that A^2, B^2, C^2 and D^2 are the four matrices

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$C^2 = \begin{pmatrix} 13 & -6 \\ 18 & -8 \end{pmatrix}, \quad D^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

and their cubes are:

$$A^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$C^3 = \begin{pmatrix} 29 & -14 \\ 42 & -20 \end{pmatrix}, \quad D^3 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

From these computations, we can work out that

$$\begin{aligned} f(A) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} + \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -6 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

In a similar way, we can calculate that

$$f(B) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} - 3\begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 3 & 1 \end{pmatrix}$$

$$f(C) = \begin{pmatrix} 13 & -6 \\ 18 & -8 \end{pmatrix} - 3\begin{pmatrix} 5 & -2 \\ 6 & -2 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$f(D) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} - 3\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-3i & 0 \\ 0 & 1+3i \end{pmatrix}$$

The computations with g are similar; here we just record the results.

$$g(A) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 20 \end{pmatrix},\ g(B) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$g(C) = \begin{pmatrix} 62 & -28 \\ 84 & -36 \end{pmatrix},\ g(D) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

2. Determine if each of the following matrices is invertible, and, if it is, invert it.

(a) $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 6 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & -3 & 1 \\ 2 & 1 & 4 \\ 1 & 11 & 5 \end{pmatrix}$

Solution: Although there are various ways of determining if a matrix is invertible and of inverting it, the best way to invert the matrix A by hand (and even by computer until you get to very large sizes where some very sophisticated methods have been developed), for a 3×3 or larger matrix is to row reduce the block matrix $(A|I)$. If it reduces to the block matrix $(I|B)$, then $B = A^{-1}$ as shown in class. Otherwise, the reduction process shows that A can be row reduced to a matrix with a row of 0's, in which case it cannot be inverted. This process thus answers both halves of the question. It also works with 2×2 matrices, although in that case, it is probably easier to just learn the answer.

(a)

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

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

so that $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$.

(b)

$$\begin{pmatrix} 1 & 6 & 1 & 1 & 0 & 0 \\ 0 & 2 & 4 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 6 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & -11 & 1 & -3 & 0 \\ 0 & 1 & 2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 & -3 & -11 \\ 0 & 1 & 0 & 0 & 1/2 & 2 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}$$

so that the inverse is $\begin{pmatrix} 1 & -3 & -11 \\ 0 & 1/2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$.

(c)
$$\begin{pmatrix} 1 & -3 & 1 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 1 & 11 & 5 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -3 & 1 & 1 & 0 & 0 \\ 0 & 7 & 2 & -2 & 1 & 0 \\ 0 & 14 & 4 & -1 & 0 & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & -3 & 1 & 1 & 0 & 0 \\ 0 & 7 & 2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 3 & -2 & 1 \end{pmatrix}$$
which shows that the original matrix is not invertible.

3. Which of the following subsets of the indicated coordinate spaces is a vector space over the real numbers. In each case the addition and scalar multiplication are coordinate-wise. You should explain your reasoning carefully.

1. The set of all $(x, y, z) \in \mathbb{R}^3$ such that $x + y = z$.

   Solution: As explained in class, once you have a vector space, say V, a subset W is a vector space if and only if (i) $0 \in W$, (ii) $\vec{u}$, $\vec{v} \in W$ implies $\vec{u} + \vec{v} \in W$ and (iii) for any scalar a and any $\vec{u} \in W$, $a\vec{u} \in W$. So let W be the set in question. Since $0 + 0 = 0$, it follows that $(0, 0, 0) \in W$. If (x, y, z) and (x', y', z') and in W, then $x+y = z$ and $x'+y' = z'$, so that $(x+x')+(y+y') = (x+y)+(x'+y') = z+z'$ so that $(x, y, z)+(x', y', z') \in W$. If $x+y = z$, then for any scalar a, $ax + ay = a(x + y) = az$ so that $(ax, ay, az) \in W$. Thus the three conditions are satisfied and we have a subspace.

2. The set of all $(x, y, z) \in \mathbb{R}^3$ such that $xy = z$.

   Solution: Although the first condition is satisfied, neither of the other two is. For example, the vector $(1, 1, 1)$ is in the set, but $(2, 2, 2)$ is not. Thus it is not a subspace.

3. The set of all $(x, y, z) \in \mathbb{R}^3$ such that $x^2 + y^2 + z^2 = 0$.

   Solution: This actually is a subspace, even though the condition is non-linear. The reason is that the only vector that satisfies the condition is $(0, 0, 0)$ and the set consisting of the 0 vector alone is always a vector space.

4. The set of all $(\alpha, \beta, \gamma) \in \mathbb{C}^3$ such that $\alpha = \overline{\beta} - \gamma$.

   Solution: This is a subspace. Since $\overline{0} = 0$, it follows that $(0, 0, 0)$ is in the set we will call W. If $\alpha = \overline{\beta} - \gamma$ and $\alpha' = \overline{\beta}' - \gamma'$, then from the fact that $\overline{\beta + \beta'} = \overline{\beta} + \overline{\beta}'$, it is easy to see that $\alpha + \alpha' = \overline{(\beta + \beta')} - (\gamma + \gamma')$. Finally, for real scalars a, we have $\overline{a\beta} = a\overline{\beta}$ so that if $\alpha = \overline{\beta} - \gamma$, then $a\alpha = \overline{a\alpha} - a\beta$.

4. Which of the following sets is a vector space over the complex numbers. In the first one, the operations are coordinate-wise and in the other two, entry-wise. Explain your answers carefully.

   (a) The set of all $(\alpha, \beta, \gamma) \in \mathbb{C}^3$ such that $\alpha = \overline{\beta} - \gamma$.
   Solution: Unlike the last part of the previous question, this is not a subspace. The reason is that the set is not closed under scalar multiplication. For example, $(0, 1, 1)$ satisfies the condition, since $0 = \overline{1} - 1$. The scalar multiple $(0, i, i)$ does not satisfy the condition since $\overline{i} - i = -i - i = -2i \neq 0$.

   (b) The set of all symmetric 2×2 matrices of complex numbers.
   Solution: This is the set of all matrices $\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$. This clearly includes the 0 matrix. Since
   $$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} + \begin{pmatrix} \alpha' & \beta' \\ \beta' & \gamma' \end{pmatrix} = \begin{pmatrix} \alpha + \alpha' & \beta + \beta' \\ \beta + \beta' & \gamma + \gamma' \end{pmatrix}$$
   the set is closed under addition. Similarly, it is closed under scalar multiplication and is thus a subspace.

   (c) The set of all Hermitian 2×2 matrices of complex numbers.
   Solution: This is not closed under scalar multiplication. For example, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is Hermitian, since $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, while for the scalar multiple $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$, we have $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}^H = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$.

B.1.3 Third 1993 Problem Assignment, with Solutions

Distribution Date: Monday, October 25th, 1993
Solutions were to be submitted by Friday, October 22nd, 1993
Caveat lector! There may be misprints and/or other errors.

REMINDER: MIDTERM TEST ON TUESDAY, 26 OCTOBER, 1993, 19 TO 21 HOURS, LEACOCK 26

1. Find all values of the scalar μ such that the vector $v = \begin{pmatrix} -4 \\ -12 \\ -20 \end{pmatrix}$ is a linear combination of the vectors $u_1 = \begin{pmatrix} 5 \\ 6 \\ \mu \end{pmatrix}$, $u_2 = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$, $u_3 = \begin{pmatrix} 4 \\ 8 \\ 14 \end{pmatrix}$.

Solution: We seek all values of μ for which there exist scalars α, β, γ such that $\alpha u_1 + \beta u_2 + \gamma u_3 = v$, i.e. such that the equation

$$\begin{pmatrix} 5 & 3 & 4 \\ 6 & 2 & 8 \\ \mu & 5 & 14 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} -4 \\ -12 \\ -20 \end{pmatrix} \tag{117}$$

admits a solution. We row reduce the augmented matrix in equation (117):

$$\begin{aligned}
\left(\begin{array}{ccc|c} 5 & 3 & 4 & -4 \\ 6 & 2 & 8 & -12 \\ \mu & 5 & 14 & -20 \end{array}\right) &\to \left(\begin{array}{ccc|c} 1 & 3/5 & 4/5 & -4/5 \\ 6 & 2 & 8 & -12 \\ \mu & 5 & 14 & -20 \end{array}\right) \\
&\to \left(\begin{array}{ccc|c} 1 & 3/5 & 4/5 & -4/5 \\ 0 & -8/5 & 16/5 & -36/5 \\ 0 & 5-3\mu/5 & 14-4\mu/5 & -20-4\mu/5 \end{array}\right) \\
&\to \left(\begin{array}{ccc|c} 1 & 3/5 & 4/5 & -4/5 \\ 0 & 1 & -2 & 9/2 \\ 0 & 5-3\mu/5 & 14-4\mu/5 & -20-4\mu/5 \end{array}\right) \\
&\to \left(\begin{array}{ccc|c} 1 & 3/5 & 4/5 & -4/5 \\ 0 & 1 & -2 & 9/2 \\ 0 & 0 & 24-2\mu & (-85+7\mu)/2 \end{array}\right)
\end{aligned}$$

The system will certainly have a solution when the entry in position $(3,3)$ is non-zero, i.e. when $\mu \neq 12$, since the coefficient matrix and the augmented matrix both have rank 3. When, however, that entry is 0, the last row has its only non-zero entry in the last column, corresponding to an equation $0 = -\frac{1}{2}$. From this contradiction we deduce that there can be no solution when $\mu = 12$. (It is not sufficient to consider only the entry in position $(3,3)$: suppose that the entry in position $(3,4)$ had been $(-84+7\mu)/2$.)

2. Consider the real vector space V of polynomials in an indeterminate x of degree at most 3.

   (a) Show that the polynomials $e_i = x^i$ $(i = 0,1,2,3)$ form a basis $\mathcal{B}_1$ for V. (This is easy.)

   (b) Show that the polynomials $f_i = (x+2)^i$ $(i = 0,1,2,3)$ form a basis $\mathcal{B}_2$ for V.

   (c) Show that each of the following sets of vectors *does not* constitute a basis for V:

      i. $\{1+x-x^2, x-x^3, x+x^2, 1+4x+2x^2\}$

      ii. $\{1, x^2, x^3\}$

(d) Express a polynomial $a_0 + a_1x + a_2x^2 + a_3x^3$ as a linear combination of the basis vectors in each of the bases $\mathcal{B}_1$, $\mathcal{B}_2$.

Solution:

(a) Every polynomial is of the form

$$a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 = a_0e_0 + a_1e_1 + a_2e_2 + a_3e_3\,, \tag{118}$$

a linear combination of the given 4 vectors. Thus the given polynomials constitute a *generating set* for the vector space. And, $a_0e_0 + a_1e_1 + a_2e_2 + a_3e_3 = 0 \Leftrightarrow a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 = 0x^0 + 0x^1 + 0x^2 + 0x^3$, which two polynomials are equal iff their corresponding coefficients are equal, i.e. iff $a_i = 0$ $(i = 0, 1, 2, 3)$. Thus a linear combination of the given polynomials vanishes iff it is the "trivial" linear combination, where all coefficients are zero: this is the definition of linear independence. A linearly independent generating set is a basis.

(b)

$$\begin{aligned}
& a_0(x+2)^0 + a_1(x+2)^1 + a_2(x+2)^2 + a_3(x+2)^3 = 0 \\
\Leftrightarrow\ & (a_0 + 2a_1 + 4a_2 + 8a_3)x^0 + (a_1 + 4a_2 + 12a_3)x^1 + (a_2 + 6a_3)x^2 + a_3x^3 \\
& \quad = 0x^0 + 0x^1 + 0x^2 + 0x^3 \\
\Leftrightarrow\ & \left\{\begin{array}{rcl} a_0 + 2a_1 + 4a_2 + 8a_3 & = & 0 \\ a_1 + 4a_2 + 12a_3 & = & 0 \\ a_2 + 6a_3 & = & 0 \\ a_3 & = & 0 \end{array}\right. \\
\Leftrightarrow\ & \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

which may be shown, by row reduction, to imply that $a_0 = a_1 = a_2 = a_3 = 0$. Thus vectors f_i $(i0, 1, 2, 3)$ are linearly independent. As we have shown above that the vector space of polynomials of degree at most 3 has a basis consisting of 4 vectors, its dimension is 4, and *all* basis consist of 4 vectors; moreover, any linearly independent set of 4 vectors in the space must be a basis (Theorem 5.14(ii)).

(c) i. We investigate whether the given 4 vectors are linearly independent, by determining all linear combinations of them which vanish.

$$
\begin{aligned}
& a_0(1+x-x^2)+a_2(x-x^3)+a_3(x+x^2)+a_4(1+4x+2x^2)=0 \\
\Leftrightarrow\ & (a_1+a_4)x^0+(a_1+a_2+a_3+4a_4)x^1+(-a_1+a_3+2a_4)x^2+(-a_2)x^3 \\
& =0x^0+0x^1+0x^2+0x^3 \\
\Leftrightarrow\ & \left\{\begin{array}{rcl} a_1+1a_4 &=& 0 \\ a_1+a_2+a_3+4a_4 &=& 0 \\ -a_1+a_3+2a_4 &=& 0 \\ -a_2 &=& 0 \end{array}\right. \\
\Leftrightarrow\ & \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
\Leftrightarrow\ & \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}=a_4\begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \end{pmatrix}.
\end{aligned}
$$

Thus there do exist nontrivial linear combinations of these vectors which are equal to zero, implying that the vectors are not linearly independent, so they cannot form a basis. (Alternatively we could have demonstrated the existence of vectors that could not be expressed as a linear combination of these 4 polynomials: for example, x^3.)

ii. The given set contains only three vectors, but we have shown above that the dimension of the space is 4. Thus the set cannot be a basis.

(d) i. cf. equation (118)

ii. We know that

$$
\begin{array}{rcccl}
f_0 &=& (x+2)^0 &=& 1=1e_0 \\
f_1 &=& (x+2)^1 &=& x+2=2e_0+1e_1 \\
f_2 &=& (x+2)^2 &=& x^2+4x^1+4=4e_0+4e_1+1e_2 \\
f_3 &=& (x+2)^3 &=& x^3+6x^2+12x^1+8x^0=8e_0+12e_1+6e_2+1e_3
\end{array} \tag{119}
$$

But what we want is a series of equations expressing the basis vectors in $\mathcal{B}_1$ in terms of the basis vectors in $\mathcal{B}_2$. We can solve equations (119) by *forward* substitution (subtracting multiples of equations from other equations lower in the list) to obtain

$$
\begin{array}{rcl}
1f_0 &=& e_0 \\
-2f_0+1f_1 &=& e_1 \\
4f_0-4f_1+1f_2 &=& e_2 \\
-8f_0+12f_1-6f_2+1f_3 &=& e_3
\end{array}
$$

from which it follows that

$$\begin{aligned} a_0 + a_1x + a_2x^2 + a_3x^3 &= a_0f_0 + a_1(-2f_0 + f_1) \\ &\quad +a_2(4f_0 - 4f_1 + 1f_2) \\ &\quad +a_3(-8f_0 + 12f_1 - 6f_2 + 1f_3) \\ &= (a_0 - 2a_1 + 4a_2 - 8a_3)f_0 \\ &\quad +(a_1 - 4a_2 + 12a_3)f_1 \\ &\quad +(a_2 - 6a_3)f_2 + a_3f_3 \end{aligned}$$

3. Determine which sets of columns of the matrix

$$A = \begin{pmatrix} 6 & 5 & 8 & 4 & 1 & 4 \\ 2 & 3 & 0 & 2 & 1 & 2 \\ 3 & 7 & -5 & -7 & 1 & 3 \\ -2 & -6 & 6 & 5 & 0 & 1 \\ -7 & -4 & -13 & -7 & 0 & -1 \end{pmatrix}$$

constitute a basis for the space of solutions of the system of equations

$$\begin{aligned} 2x_1 - 5x_2 + 3x_3 + 2x_4 + x_5 &= 0 \\ 5x_1 - 8x_2 + 5x_3 + 4x_4 + 3x_5 &= 0 \\ x_1 - 7x_2 + 4x_3 + 2x_4 &= 0 \\ 4x_1 - x_2 + x_3 + 2x_4 + 3x_5 &= 0 \end{aligned}$$

<u>Solution</u>: We row reduce the augmented matrix of the system:

$$\left(\begin{array}{ccccc|c} 2 & -5 & 3 & 2 & 1 & 0 \\ 5 & -8 & 5 & 4 & 3 & 0 \\ 1 & -7 & 4 & 2 & 0 & 0 \\ 4 & -1 & 1 & 2 & 3 & 0 \end{array}\right) \to \left(\begin{array}{ccccc|c} 1 & -5/2 & 3/2 & 1 & 1/2 & 0 \\ 5 & -8 & 5 & 4 & 3 & 0 \\ 1 & -7 & 4 & 2 & 0 & 0 \\ 4 & -1 & 1 & 2 & 3 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|c} 1 & -5/2 & 3/2 & 1 & 1/2 & 0 \\ 0 & 9/2 & -5/2 & -1 & 1/2 & 0 \\ 0 & -9/2 & 5/2 & 1 & -1/2 & 0 \\ 0 & 9 & -5 & -2 & 1 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|c} 1 & -5/2 & 3/2 & 1 & 1/2 & 0 \\ 0 & 9/2 & -5/2 & -1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|c} 1 & -5/2 & 3/2 & 1 & 1/2 & 0 \\ 0 & 1 & -5/9 & -2/9 & 1/9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The problem is to determine which columns of matrix A form a basis for the solution space of the system of homogeneous equations whose coefficient matrix is

$$B = \begin{pmatrix} 1 & -5/2 & 3/2 & 1 & 1/2 \\ 0 & 1 & -5/9 & -2/9 & 1/9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

There are several stages in this determination. Firstly we should determine which columns of A are *solutions* of the system. We can do this by computing the product BA and looking for product columns which consist entirely of zeros.

$$BA = \begin{pmatrix} 0 & 0 & 0 & 10 & 0 & 4 \\ 0 & 0 & 0 & 4 & 4/9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We may conclude that only columns ##1–3 of A are solutions of the system. Which subsets of these three columns constitute bases for the solution space?

There are several ways in which this problem can be approached. We give only one method in these notes.

Since the rank of the coefficient matrix of the system of equations is 2, the dimension of the solution space must be $5 - 2 = 3$ (cf. Theorem 5.20 in the text-book); thus all bases for the solution space will contain precisely 3 vectors. The problem is then to determine which sets of three columns chosen from the first three columns of A constitute such a basis. There is precisely 1 subset of 3 vectors chosen from a set of 3. We apply Theorem 5.14(ii): it suffices to determine whether such a set of 3 vectors are linearly independent to conclude that they constitute a basis for a space of dimension 3. For each of these sets of 3 columns we need to determine whether the set is linearly independent. We could test each of the sets separately (which is simplest in this case since there is only one such set), but we find it convenient to test all simultaneously. We simply examine whether there is a non-trivial solution of the system

$$\begin{pmatrix} 6 & 5 & 8 & 4 \\ 2 & 3 & 0 & 2 \\ 3 & 7 & -5 & -7 \\ -2 & -6 & 6 & 5 \\ -7 & -4 & -13 & -7 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Row reducing the augmented matrix of this system yields

$$\left(\begin{array}{rrr|r} 6 & 5 & 8 & 0 \\ 2 & 3 & 0 & 0 \\ 3 & 7 & -5 & 0 \\ -2 & -6 & 6 & 0 \\ -7 & -4 & -13 & 0 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 5/6 & 4/3 & 0 \\ 2 & 3 & 0 & 0 \\ 3 & 7 & -5 & 0 \\ -2 & -6 & 6 & 0 \\ -7 & -4 & -13 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccr|c} 1 & 5/6 & 4/3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccr|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

From this matrix in row-echelon form we may read off, in particular, that

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_3 \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$$

Thus any linearly dependent set of vectors must include columns ##1, 2, 3; and these are linearly dependent because 3 times the first column is equal to twice the second column, plus the third column. Thus, no set of 3 columns is independent, so there is no basis for the solution space of the given system formed from the columns of A.

(Try the problem again with $A = \begin{pmatrix} 6 & 5 & 8 & 4 & 1 & 4 \\ 2 & 3 & 0 & -2 & 1 & 2 \\ 3 & 7 & -5 & -7 & 1 & 3 \\ -2 & -6 & 6 & 5 & 0 & 1 \\ -7 & -4 & -13 & -7 & 0 & -1 \end{pmatrix}$.)

B.1.4 Fourth 1993 Problem Assignment, with Solutions

(Solutions were to be submitted by 10 November, 1993)

Distribution Date: Monday, November 15th, 1993

1. Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be the linear transformation with matrix

$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 5 & -2 & 0 \\ 4 & -3 & -2 & 1 \end{pmatrix}$$

in the standard bases. Find $[T]_{\mathcal{C}}^{\mathcal{B}}$ where $\mathcal{B} = \{(1,2,3,4)^T, (-1,1,2,3)^T, (0,-1,-1,2)^T, (1,1,1,1)^T\}$ and $\mathcal{C} = \{(1,0,0)^T, (1,1,0)^T, (1,1,1)^T\}$.

Solution: As explained in class and also in the text, when

$$P = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 1 & -1 & 1 \\ 3 & 2 & -1 & 1 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

then $[T]_{\mathcal{C}}^{\mathcal{B}} = Q^{-1}AP$. We can readily calculate

$$Q^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and then} \quad Q^{-1}AP = \begin{pmatrix} -3 & 2 & 6 & -1 \\ 10 & 7 & -10 & 5 \\ -4 & -8 & 7 & 0 \end{pmatrix}$$

2. Let V be the space of those functions that are linear combinations of $\mathcal{B} = \{\sin x, \cos x, e^x, e^{-x}\}$.

   (a) Show that the elements of $\mathcal{B}$ are linearly independent and therefore a basis of V)

   (b) Find the matrix $[D]_{\mathcal{B}}^{\mathcal{B}}$, where D is differentation.

   (c) What is the determinant of D?

   (d) Find the characteristic polynomial of D.

   Solution:

   (a) Suppose there were a dependence among them. That means that $a\sin x + b\cos x + ce^x + de^{-x} = 0$ identically. There are several ways to go from here. One way is to replace x by four different values of x, say $x = -\pi/2$, 0, $\pi/2$ and π. This would give four equations equivalent to

$$\begin{pmatrix} -1 & 0 & e^{-\pi/2} & e^{\pi/2} \\ 0 & 1 & 1 & 1 \\ 1 & 0 & e^{\pi/2} & e^{-\pi/2} \\ 0 & 0 & e^{\pi} & e^{-\pi} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

   The matrix row reduces to

$$\begin{pmatrix} -1 & 0 & e^{-\pi/2} & e^{\pi/2} \\ 0 & 1 & 1 & 1 \\ 0 & 0 & e^{\pi/2} + e^{-\pi/2} & e^{-\pi/2} + e^{\pi/2} \\ 0 & 0 & e^{\pi} + 1 & e^{-\pi} + 1 \end{pmatrix}$$

and the latter reduces to an echelon matrix since the two non-zero elements of the third row are equal and of the fourth are unequal. Thus there is no non-zero choice for a, b, c and d. A somewhat simpler way is to observe that if the function is identically 0, so are all its derivatives. By evaluating the function and its first four derivatives at 0, we are led to the system of equations

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and we have the reduction

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 2 & -2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

whose kernel is 0. In any case, independence follows.

(b)

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

(c) This is most easily calculated by row reduction. Three operations are needed to reduce the matrix to the identity: interchange the first two rows and multiply the (new) second row and fourth by -1. Each operation multiples the determinant by -1 and since there are three of them, the determinant must be -1. It can also be easily calculated by a minors expansion along any row or column.

(d) $(t^2 + 1)(t - 1)(t + 1)$.

3. Find the characteristic polynomial, eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

Solution: The characteristic polynomial is computed

$$\begin{aligned}
&\det\begin{pmatrix}1-t & -3 & 3\\ 3 & -5-t & 3\\ 6 & -6 & 4-t\end{pmatrix}\\
&= (1-t)\det\begin{pmatrix}-5-t & 3\\ -6 & -4-t\end{pmatrix} + 3\det\begin{pmatrix}3 & 3\\ 6 & 4-t\end{pmatrix} + 3\det\begin{pmatrix}3 & -5-t\\ 6 & -6\end{pmatrix}\\
&= (1-t)((t+5)(t-4)+18) + 3(12-t-18) + 3(-18+6t+12\\
&= -t^3 + 2t^2 + 4t + 16
\end{aligned}$$

By trial, we see that one root is -2 and then the quadratic formula (or factoring) shows that the other two are -2 and 4. The eigenvectors for -2 are the kernel of $\begin{pmatrix}3 & -3 & 3\\ 3 & -3 & 3\\ 6 & -6 & 6\end{pmatrix}$ which has basis $\begin{pmatrix}-1 & 0 & 1\end{pmatrix}^T$ and $\begin{pmatrix}0 & 1 & 1\end{pmatrix}^T$. The eigenvectors for 4 generate the kernel of the matrix $\begin{pmatrix}-3 & -3 & 3\\ 3 & -9 & 3\\ 6 & -6 & 0\end{pmatrix}$ which row reduces to $\begin{pmatrix}1 & 1 & -1\\ 0 & -2 & 1\\ 0 & 0 & 0\end{pmatrix}$ which is spanned by $\begin{pmatrix}-1 & 1 & 2\end{pmatrix}^T$.

B.1.5 Fifth 1993 Problem Assignment, with Solutions

Distribution Date: Wednesday, November 24th, 1993
Solutions were to be submitted by Friday, November 19th, 1993

For each of the matrices $M = A, B, C, D$

1. Determine all eigenvalues and all eigenspaces.
2. Verify that the sum of the eigenvalues is equal to the trace, and that the product of the eigenvalues is equal to the determinant.
3. Give the algebraic and geometric multiplicity of each of the eigenvalues.
4. Determine whether or not M is diagonalizable.
5. If the matrix M is diagonalizable, find a matrix P which diagonalizes it, i.e. such that $P^{-1}MP$ is diagonal.

6. If M is diagonalizable, find a matrix Q whose square is equal to M. (You may express this "square root" as a product of matrices, which need not be reduced.)

7. Determine the minimum polynomial. (Since you will not have been shown direct methods for finding this polynomial, you may have to use what you know about the divisors of the minimum polynomial to determine a list of candidates for the minimum polynomial, and then test each of these candidates to determine which is the correct minimum polynomial.)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}$$

$$C = \begin{pmatrix} 3 & -1 & 0 \\ 6 & -3 & 2 \\ 8 & -6 & 5 \end{pmatrix} \qquad D = \begin{pmatrix} i & 1 \\ 0 & -i \end{pmatrix}$$

All computations should be performed in the *complex* field, although, in some cases, no non-real scalars will appear.

Solution:

1. A: $\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 & 0 \\ 4 & \lambda - 4 & 0 \\ 2 & -1 & \lambda - 2 \end{vmatrix} = (\lambda - 2)(\lambda(\lambda - 4) + 4) = (\lambda - 2)^3$.

   To determine the eigenspace we solve the equation $(2I - A)x = 0$. The matrix of coefficients row reduces as follows: $\begin{pmatrix} 2 & -1 & 0 \\ 4 & -2 & 0 \\ 2 & -1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & -1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

   Thus the general solution is $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ so that the dimension of the eigenspace is 2, and one basis is $\left\{ \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

   B: The characteristic polynomial is $\lambda^3 - 6\lambda^2 + 11\lambda - 6$. We test divisors of 6, and find that 1 is a root. Dividing yields $\lambda^3 - 6\lambda^2 + 11\lambda - 6 = (\lambda - 1)(\lambda^2 - 5\lambda + 6)$ and the quadratic factor may be factorized by observation, by completing the square, or by using the quadratic formula; the characteristic polynomial may then be seen to be $(\lambda - 1)(\lambda - 2)(\lambda - 3)$, so the eigenvalues are 1, 2, 3.

Solving the equation $(I-B)x = 0$ yields a solution space generated by $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$: every eigenvector is a multiple of this vector. Similarly, the eigenspaces for $\lambda = 2$ and $\lambda = 3$ are generated by the vectors $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1/2 \\ 1 \\ 1 \end{pmatrix}$ respectively.

***C*:** The characteristic polynomial is $\det(\lambda I - C) = \begin{vmatrix} \lambda-3 & 1 & 0 \\ -6 & \lambda+3 & -2 \\ -8 & 6 & \lambda-5 \end{vmatrix} = (\lambda-3)(\lambda^2-2\lambda-3)-(-6\lambda+14) = \lambda^3-5\lambda^2+9\lambda-5$. Testing the divisors of 5, we find that 1 is a root, and the characteristic polynomial factorizes to $(\lambda-1)(\lambda^2-4\lambda+5)$. The quadratic factor may be factorized by observation, by using the quadratic formula, or by completing the square, to yield $(\lambda-1)(\lambda-2+i)(\lambda-2-i)$.

For the eigenvalue $\lambda = 1$ we obtain the vector $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ as a generator of the eigenspace. For the eigenvalue $\lambda = 2+i$ we row reduce the coefficient matrix as follows:

$$\begin{pmatrix} -1+i & 1 & 0 \\ -6 & 5+i & -2 \\ -8 & 6 & -3+i \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\frac{i}{2} \\ 0 & 1 & -\frac{1}{2}-\frac{i}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

so that an eigenvector is $\begin{pmatrix} \frac{i}{2} \\ \frac{1+i}{2} \\ 1 \end{pmatrix}$. Similarly, the eigenspace for eigenvalue $2-i$ is generated by $\begin{pmatrix} -\frac{i}{2} \\ \frac{1-i}{2} \\ 1 \end{pmatrix}$

***D*:** The characteristic polynomial is $\begin{vmatrix} \lambda-i & -1 \\ 0 & \lambda+i \end{vmatrix}$, so the eigenvalues are the diagonal entries of the (upper triangular) matrix, i.e. $\pm i$. For eigenvalue i we find an eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, while, for eigenvalue $-i$, we find an eigenvector $\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}$.

2. ***A*:** The trace of A is $0+4+2 = 6$, which is equal to the sum of the eigenvalues,

$2+2+2$. $|A| = -\begin{vmatrix} -4 & 0 \\ -2 & 2 \end{vmatrix} = 8$, which is the cube of the thrice-repeated eigenvalue 2.

B: The trace is $5-4+5=6$ which is evidently equal to the sum of the eigenvalues, $1+2+3$. $|B| = \begin{vmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{vmatrix} = \begin{vmatrix} 2 & -3 & 2 \\ 2 & -4 & 4 \\ 0 & -4 & 5 \end{vmatrix} = \begin{vmatrix} 2 & -3 & 2 \\ 0 & -1 & 2 \\ 0 & -4 & 5 \end{vmatrix} = 2\begin{vmatrix} -1 & 2 \\ -4 & 5 \end{vmatrix} = 6 = 1\times 2\times 3$, the product of the eigenvalues.

C: The trace is $3-3+5=5$ which is equal to the sum of the eigenvalues, $1+(2+i)+(2-i)$.
$|C| = \begin{vmatrix} 3 & -1 & 0 \\ 6 & -3 & 2 \\ 8 & -6 & 5 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 0 \\ -3 & -3 & 2 \\ -10 & -6 & 5 \end{vmatrix} = -(-1)\begin{vmatrix} -3 & 2 \\ -10 & 5 \end{vmatrix} = 5$ which is equal to the product of the eigenvalues, $1\times(2+i)\times(2-i) = 2^2 - i^2 = 5$.

D: The trace of D is 0, which is equal to the sum of the eigenvalues, $i-i$. The determinant of D is $-i^2 = 1$, equal to the product of the eigenvalues $(i)(-i)$.

3. *A*: The algebraic multiplicity of eigenvalue 2 is its multiplicity as a root of the characteristic polynomial, i.e. 3. The geometric multiplicity is the dimension of the eigenspace, i.e. 2.

   B: Each of the eigenvalues has algebraic multiplicity equal to 1; the geometric multiplicity of each must be positive, hence also equal to 1.

   C: as in the preceding case.

   D: as in the preceding case.

4. *A*: Since there is an eigenvalue whose algebraic multiplicity exceeds its geometric multiplicity, the matrix is not diagonalizable.

   B: This matrix is diagonalizable, since the sum of the geometric multiplicities is equal to the sum of the algebraic multiplicities.

   C: as in the preceding case

   D: as in the preceding case

5. *A*: Not applicable, since A is not diagonalizable.

   B: We may form the matrix P from the eigenvectors, taken as columns in any order: $P = \begin{pmatrix} 1 & 1 & 1/2 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$; we may compute $P^{-1} = \begin{pmatrix} -2 & 2 & -1 \\ 2 & -1 & 0 \\ 2 & -2 & 2 \end{pmatrix}$.

Then $P^{-1}BP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

C:

$$\begin{pmatrix} -2 & 2 & -1 \\ 1-3i & -1+2i & 1-i \\ 1+3i & -1-21 & 1+i \end{pmatrix} \begin{pmatrix} 3 & -1 & 0 \\ 6 & -3 & 2 \\ 8 & -6 & 5 \end{pmatrix} \begin{pmatrix} 1 & \frac{i}{2} & -\frac{i}{2} \\ 2 & \frac{1+i}{2} & \frac{1-i}{2} \\ 1 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2+i & 0 \\ 0 & 0 & 2-i \end{pmatrix}$$

D:

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

6. A: A is not diagonalizable.

B: $B = PDP^{-1}$, where D is the diagonal matrix formed by the eigenvalues 1, 2, 3 in that order. A square root of D is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix}$. It follows that

$$B = \left(\begin{pmatrix} 1 & 1 & 1/2 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} -2 & 2 & -1 \\ 2 & -1 & 0 \\ 2 & -2 & 2 \end{pmatrix} \right)^2$$

C: Students were asked to omit this part, because of computational difficulties.

D: The square roots of i are points on the unit circle "half-way" to i from the positive real axis — i.e. points whose position vectors are inclined by $\pm\frac{\pi}{4}$.

$$\left(\begin{pmatrix} 1 & \frac{i}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+i}{\sqrt{2}} & 0 \\ 0 & \frac{1-i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -\frac{i}{2} \\ 0 & 1 \end{pmatrix} \right)^2 = \begin{pmatrix} i & 1 \\ 0 & -i \end{pmatrix}$$

7. A: The minimum polynomial is a product of the linear factors that appear in the characteristic polynomial, each with some multiplicity between 1 and its algebraic multiplicity. In this case there is only the factor $\lambda - 2$ to consider, so the only candidates for minimum polynomial are $\lambda - 2$, $(\lambda - 2)^2$, and the characteristic polynomial itself, $(\lambda - 2)^3$. Since A is not equal to $2I$, $\lambda - 2$ is not satisfied by A. It can be seen that $(A - 2I)^2 = 0$, so $(\lambda - 2)^2$ is the minimum polynomial.

B: In the minimum polynomial each of the factors $\lambda-1$, $\lambda-2$, $\lambda-3$ must appear with multiplicity at least 1; but their product is of degree 3 and is equal to the characteristic polynomial. Thus the minimum polynomial is equal to the characteristic polynomial.

C: As in the preceding case, since the matrix is diagonalizable, its minimum polynomial is equal to its characteristic polynomial, $\lambda^3-5\lambda^2+9\lambda-5$.

D: As in the preceding case, since the matrix is diagonalizable, its minimum polynomial is equal to its characteristic polynomial, λ^2+1.

B.1.6 Sixth 1993 Problem Assignment, with Solutions

Distribution Date:
Solutions were to be submitted by Monday, December 6th, 1993

1. [1, Problem 2.77, p. 55] Show that, for any square complex matrix A, $A+A^H$ is hermitian, and $A-A^H$ is skew-hermitian. (For definitions cf. [1, p. 40].)

   Solution: Let $\alpha=\pm 1$.

$$\begin{aligned}
\left(A+\alpha A^H\right)^H &= \overline{\left(A+\alpha\overline{A}^T\right)^T} \\
&= \overline{A^T+\alpha\left(\overline{A}^T\right)^T} \\
&= \overline{A^T+\alpha\overline{A}} \\
&= \overline{A^T}+\overline{\alpha\overline{A}} \\
&= \overline{A^T}+\alpha\overline{\left(\overline{A}\right)} \\
&= \overline{A^T}+\alpha A \\
&= A^H+\alpha A = \begin{cases} A+A^H & \text{if } \alpha=1 \\ -(A^H-A) & \text{if } \alpha=-1 \end{cases}
\end{aligned}$$

2. (a) Show that the following is not an inner product on $\mathbb{R}^3$, where $u=\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}$ and $v=\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix}$:

$$\langle u,v\rangle = y_1x_2y_3$$

(b) On the real vector space of continuous functions $f : [0, 2] \to \mathbb{R}$, show that the function

$$\langle f, g\rangle = \int_0^2 (1 - x)\, f(x)g(x)\, dx$$

does not define an inner product.

Solution:

(a) Let $u = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Then

$$\langle u, v_1\rangle + \langle u, v_2\rangle = 1 + 1 = 2 \neq 4 = \langle u, v_1 + v_2\rangle$$

One counterexample suffices to show that property (I_1) fails to hold universally.

(b) In this case the product has properties (I_1) and (I_2). It will fail to have property (I_3) because the weight function $1 - x$ is not always positive. For example, if we take $f(x) = 1 = g(x)$ for all x,

$$\|f\|^2 = \langle f, f\rangle = \int_0^2 (1 - x)\, dx = 0$$

but f is not the zero function.

3. Find an orthonormal basis for the subspace of $\mathbb{R}^4$ spanned by the vectors

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 2 \end{pmatrix} \quad v_3 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ -2 \end{pmatrix} \quad v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

by applying the Gram-Schmidt orthonormalization process to these vectors *in the given order*.

Solution: We shall first determine an orthogonal basis, and then normalize its members.

$$\begin{aligned}
w_1 &= v_2 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} \\
\|w_1\|^2 &= 1^2 + 2^2 + (-3)^2 + (-4)^2 = 30 \\
w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 \\
&= \begin{pmatrix} 1 \\ -1 \\ 2 \\ 2 \end{pmatrix} - \frac{-15}{30} \begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}; \quad \|w_2\|^2 = \frac{10}{4} \\
w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2 \\
&= \begin{pmatrix} 2 \\ 1 \\ -1 \\ -2 \end{pmatrix} - \frac{15}{30} \begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} - \frac{\frac{5}{2}}{\frac{10}{4}} \begin{pmatrix} \frac{3}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

from which it follows that v_3 is linearly dependent on w_1 and w_2. Thus w_3 can be suppressed: the subspace generated by v_1, v_2, and v_3 is generated by the orthonormal vectors w_1 and w_2. We continue:

$$\begin{aligned}
w_4 &= v_4 - \frac{\langle v_4, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_4, w_2 \rangle}{\|w_2\|^2} w_2 \\
&= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{-4}{30} \begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} - \frac{2}{\frac{10}{4}} \begin{pmatrix} \frac{3}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix} \\
&= \frac{1}{15} \begin{pmatrix} -1 \\ 19 \\ 3 \\ 7 \end{pmatrix}; \quad \|w_4\|^2 = 2\sqrt{\frac{7}{15}}
\end{aligned}$$

We may thus normalize to produce an ortho*normal* basis:

$$\begin{aligned} u_1 &= \frac{1}{\|w_1\|}w_1 = \frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ 2 \\ -3 \\ -4 \end{pmatrix} \\ u_2 &= \frac{1}{\|w_2\|}w_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ u_4 &= \frac{1}{\|w_4\|}w_4 = \frac{2}{\sqrt{105}}\begin{pmatrix} -1 \\ 19 \\ 3 \\ 7 \end{pmatrix} \end{aligned}$$

4. Let $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$.

   (a) Determine an invertible matrix P such that $P^{-1}AP$ is diagonal.

   (b) Working from the matrix P determined above, determine an orthogonal matrix Q such that Q^TAQ is diagonal.

   (c) By applying row operations and then the corresponding column operations to A, as in ??determine an invertible matrix R such that R^TAR is diagonal. From this matrix and the associated change of variables, discuss which of the following designations apply to the quadratic form x^TAx: *positive definite*, *positive semidefinite*, *negative definite*, *negative semidefinite*.

   (d) Let R is as determined in the preceding part. Let f_1, f_2, ..., f_4 be a basis for $\mathbb{R}^4$, and let x and y be any vectors in this space, with coordinate vectors $[x]$ and $[y]$ referred to the given ordered basis. Discuss whether or not

   $$\langle x, y\rangle = [x]^TA[y]$$

   defines an inner product on $\mathbb{R}^4$.

Solution:

(a) $\det(\lambda I - A) = (\lambda - 2)^3(\lambda + 2)^1$. We determine the eigenspaces.

$\lambda = 2:$ Solving the system

$$\begin{pmatrix} 1 & -1 & -1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

we find that the matrix of coefficients now reduces to $\begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,

so that the general solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_2 + x_3 \\ x_2 \\ x_3 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

and a basis of the eigenspace is the set of vectors

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

$\lambda = -2:$ The eigenspace is generated by $\begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

The matrix P whose columns are the vectors of the bases of the two eigenspaces, has the desired properties: $P = \begin{pmatrix} 1 & 1 & 0 & -3 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$; $AP = PD$ where D is the diagonal matrix with entries 2, 2, 2, -2.

(b) Matrix A is symmetric, hence it admits a set of 4 orthogonal eigenvectors. Eigenvectors corresponding to distinct eigenvalues will be orthogonal — indeed, the eigenspaces for $\lambda = 2$ and for $\lambda = -2$ will be mutually orthogonal. Thus the basis vector we found for the eigenspace $\lambda = -2$ will be orthogonal to the three basis eigenvectors we found for $\lambda = 2$. However, those three vectors are not mutually orthogonal. We apply the Gram-Schmidt process to

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

to obtain

$$\begin{aligned}
w_1 &= v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\
w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \\
&= \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix} \\
w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{\frac{1}{2}}{\frac{3}{2}} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix}
\end{aligned}$$

Normalizing these three vectors and the basis vector for the $\lambda = -2$ eigenspace yields the matrix

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{\sqrt{3}}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} \end{pmatrix}$$

Since Q is a matrix whose columns are eigenvectors, $AQ = QD$, where D is a diagonal matrix whose diagonal entries are the corresponding eigenvalues. Since the columns of Q are mutually orthogonal, the off diagonal entries in the product Q^TQ are zero. Since the columns of Q have been normalized, the diagonal entries are each 1. Thus $Q^TQ = I$, so $Q^T = Q^{-1}$, i.e. Q is orthogonal.

(c) We transform the 4×8 matrix $\left(\; A \mid I \; \right)$ by row operations followed immediately by the corresponding operations, as described in the algorithm in the text-book.

$$\left(\begin{array}{rrrr|rrrr} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & -1 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 0 & -2 & -1 & 0 & 1 & 0 \\ 1 & -2 & -2 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 0 & -2 & -1 & 0 & 1 & 0 \\ 0 & -2 & -2 & 0 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & -4 & -2 & -2 & 1 & 0 & 1 \\ 0 & -4 & 0 & -2 & -1 & 0 & 1 & 0 \\ 0 & -2 & -2 & 0 & -1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & -2 & 1 & 0 & 1 \\ 0 & 0 & 4 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \end{array}\right)$$

From which we may conclude that the matrix R given by

$$R^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

has the property that the transformation $[x] = R[y]$ yields that

$$[x]^T A[x] = [y]^T (R^T AR)[y] = y_1^2 - 4y_2^2 + 4y_3^2 + y_4^2$$

This quadratic form can have positive values — as when $y_1 = 1$, $y_2 = y_3 = y_4 = 0$; it can also attain negative values, — as when $y_2 = 1$, $y_1 = y_3 = y_4 = 0$. Thus it is neither positive semidefinite nor negative semidefinite — which require respectively that its values all be nonnegative or nonpositive; *a fortiori* it cannot be positive definite or negative definite — which require, in addition to the *semi-* conditions, that the function assume the value 0 only at the vector 0.

(d) If we attempted to define an inner product using the quadratic form with matrix R, we would find vectors whose norm is negative. Without limiting generality, take f_1, f_2, f_3, f_4 to be the standard basis, so, for any vector x, $x = [x]$. Consider the specific vector $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, which corresponds to

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = R \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$ For this vector $x^T Ax = -4$, so condition I_3 would not be satisfied were we to take $\langle x, y\rangle = x^T Ay$.

5. Let $G = \begin{pmatrix} 2 & 1+i & 0 \\ 1-i & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

   (a) Show that G is Hermitian.

   (b) Determine a unitary matrix U such that $U^H GU$ is diagonal.

<u>Solution</u>:

(a) $\overline{G} = \begin{pmatrix} 2 & 1-i & 0 \\ 1+i & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Hence $\overline{G} = G^T$.

(b) As G is Hermitian, we know that it has real eigenvalues, and that it admits an orthonormal set of 3 eigenvectors. The characteristic polynomial is

$$\begin{aligned} \det \begin{pmatrix} \lambda-2 & -1-i & 0 \\ -1+i & \lambda-3 & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix} &= (\lambda-1) \begin{vmatrix} \lambda-2 & -1-i \\ -1+i & \lambda-3 \end{vmatrix} \\ &= (\lambda-1)(\lambda^2-5\lambda+4)\,, \end{aligned}$$

whose roots are 1 (twice), and 4 (once).

$\lambda = 1$: The matrix of the system of equations we have to solve is $\begin{pmatrix} -1 & -1+i & 0 \\ -1+i & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, which row reduces to $\begin{pmatrix} 1 & 1-i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Thus the general solution for the system is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (-1+i)y \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -1+i \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and a basis for the eigenspace is $v_1 = \begin{pmatrix} -1+i \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. As these vectors happen to be orthogonal, there is no need to apply the Gram-Schmidt process.

$\lambda = 4$**:** The matrix of the system of equations we have to solve is $\begin{pmatrix} 2 & -1+i & 0 \\ -1+i & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, which row reduces to $\begin{pmatrix} 1 & -\frac{1-i}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Thus the general solution for the system is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1-i}{2}y \\ y \\ 0 \end{pmatrix} = y \begin{pmatrix} \frac{1-i}{2} \\ 1 \\ 0 \end{pmatrix},$$

and a basis for the eigenspace is the vector $v_3 = \begin{pmatrix} \frac{1-i}{2} \\ 1 \\ 0 \end{pmatrix}$.

A matrix with columns v_1, v_2, v_3 will have orthogonal columns. Normalizing to make all columns have norm equal to 1 yields the matrix $U = \begin{pmatrix} 3^{-1/2}(-1+i) & 0 & 6^{-1/2}(1-i) \\ 3^{-1/2} & 0 & 2^{1/2}3^{-1/2} \\ 0 & 1 & 0 \end{pmatrix}$ which has the property that $U^H U = I$, i.e. is unitary. Since $GU = UD$, where D is the diagonal matrix with entries 1, 1, 4, so $U^{-1}GU = D$. But $U^{-1} = U^H$.

B.2 Fall 1993 Class Tests

B.2.1 1993 Class Test, with Solutions

which was administered on Monday, October 25th, 1993

Distribution Date: Wednesday, November 3rd, 1993

The instructions stated that "All questions have equal weight." Each of questions ##1,2,3 was graded out of 25 marks.

1. Which of the following subsets of the indicated vector spaces is a subspace?

   (a) The set $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + 2x_1x_2 = 0\}$ in the vector space of real two dimensional row vectors.

   (b) The set of all 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $a + \overline{d} = c + d$ in the complex vector space of 2×2 matrices with complex number entries.

   (c) The set of all polynomials $p(x)$ such that the derivative $p'(1) = 0$.

   Solution:

(a) This is a not a subspace because the vector $(1, -1/2)$ satisfies the condition, while the scalar multiple $(2, -1)$ does not.

(b) This is a not a subspace for the matrix $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ satisfies the condition $0 + \overline{1} = 0 + 1$ since $\overline{1} = 1$ while the scalar multiple $i \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix}$ does not (for $\overline{i} = -i$).

(c) This is a subspace. The zero polynomial has 0 derivative everywhere. If $p_1'(1) = 0$ and $p_2'(1) = 0$ and $p = p_1 + p_2$, then it is shown in calculus that $p' = p_1' + p_2'$ so that $p'(1) = p_1'(1) + p_2'(1) = 0 + 0 = 0$. Similarly, for a scalar a, if $q = ap$, then $q' = ap'$ so that $p'(1) = 0$ implies that $q'(1) = a0 = 0$. Thus the three conditions are satisfied and this is a subspace.

2. Decide, for each of the following subsets of three dimensional row space if it generates the space, if it is linearly independent and if it is a basis. If it is a basis (exactly one of the sets is), find the coordinates of the standard basis in terms of it.

(a) $\{(1, 3, 2), (1, 1, 1), (1, 5, 3)\}$

(b) $\{(1, 3, 2), (1, 1, 1), (2, 5, 3)\}$

(c) $\{(1, 3, 2), (1, 1, 1)\}$

(d) $\{(1, 3, 2), (1, 1, 1), (1, 5, 3), (2, 5, 3)\}$

Solution:

(a) This set is not linearly independent since they are the rows of the matrix $\begin{pmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 1 & 5 & 3 \end{pmatrix}$, which row reduces as follows

$$\begin{pmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 1 & 5 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & -2 & -1 \\ 0 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

which has a row of 0's and so the rows were not linearly independent.

(b) A similar reduction as above would show that these vectors are linearly independent, since $A = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 2 & 5 & 3 \end{pmatrix}$ can be row reduced to the identity. If we denote the set of vectors by B and the vectors in the standard basis by v_1, v_2

and v_3, then the vectors $[v_1]_{\mathcal{B}}$, $[v_2]_{\mathcal{B}}$ and $[v_3]_{\mathcal{B}}$ are the rows of A^{-1}. Thus we can show that A is row equivalent to the identity and find its inverse at the same time, by computing

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 5 & 3 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & -2 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & -2 & -1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 & -1 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & -1 \\ 0 & -2 & -1 & -1 & 1 & 0 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -5 & 0 & 3 \\ 0 & 1 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & 1 & 1 \\ 0 & 1 & 0 & 2 & -1 & 1 \\ 0 & 0 & 1 & 0 & 1 & -2 \end{array}\right)$$

so that the rows of A^{-1}, and the coordinates that are sought are $(-2, 1, 1)$, $(-1, -1, 1)$ and $(3, 1, -2)$.

(c) Two vectors are linearly dependent if and only at least one is a multiple (possibly 0) of the other. This is clearly not the case with these two vectors, so they are linearly independent. Being only two vectors, they cannot span a 3 dimensional space.

(d) Four vectors in 3 dimensional space cannot be linearly independent, hence cannot be a basis. On the other hand, this set includes the basis in (b) above and so spans the space.

3. (a) [5 MARKS] Let V denote the vector space of $n \times n$ matrices over some field $\mathbf{F}$. Let P be an invertible matrix in V. Show that the mapping $T : V \to V$ defined by $T(A) = PAP^{-1} - A$ is linear.

(b) [6+6 MARKS] Take $n = 2$, $\mathbf{F} = \mathbb{R}$ and $P = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ in (a). Find a basis for the image of T and the kernel of T.

(c) [8 MARKS] Find a matrix for the linear transformation T of part (b) with respect to the ordered basis

$$\begin{pmatrix}1&0\\0&0\end{pmatrix}, \quad \begin{pmatrix}0&1\\0&0\end{pmatrix}, \quad \begin{pmatrix}0&0\\1&0\end{pmatrix}, \quad \begin{pmatrix}0&0\\0&1\end{pmatrix}$$

(The same basis is to be used for the source (domain) and target (codomain) space; that is, if the ordered basis is denoted $\mathcal{B}$, then find $[T]_{\mathcal{B}}^{\mathcal{B}}$.)

Solution:

(a) $T(A_1+A_2) = P(A_1+A_2)P^{-1} - (A_1+A_2) = (PA_1+PA_2)P^{-1} - A_1 - A_2 = PA_1P^{-1} + PA_2P^{-1} - A_1 - A_2 = PA_1P^{-1} - A_1 + PA_2P^{-1} - A_2 = T(A_1) + T(A_2)$. Also, for any scalar a, $T(aA) = P(aA)P^{-1} - aA = aPAP^{-1} - aA = a(PAP^{-1} - A) = aT(A)$. Thus T preserves both addition and scalar multiplication and is linear.

(b) We begin by inverting P.

$$\left(\begin{array}{cc|cc}1&2&1&0\\2&1&0&1\end{array}\right) \rightarrow \left(\begin{array}{cc|cc}1&2&1&0\\0&-3&-2&1\end{array}\right)$$
$$\rightarrow \left(\begin{array}{cc|cc}1&2&1&0\\0&1&2/3&-1/3\end{array}\right)$$
$$\rightarrow \left(\begin{array}{cc|cc}1&0&-1/3&2/3\\0&1&2/3&-1/3\end{array}\right)$$

so that $P^{-1} = \frac{1}{3}\begin{pmatrix}-1&2\\2&-1\end{pmatrix}$. Then we have that for a matrix $A = \begin{pmatrix}a&b\\c&d\end{pmatrix}$,

$$\begin{aligned} T(A) &= \frac{1}{3}\begin{pmatrix}1&2\\2&1\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}-1&2\\2&-1\end{pmatrix} - \begin{pmatrix}a&b\\c&d\end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix}-4a-2c+2b+4d & 2a+4c-4b-2d\\ -2a-c+4b+2d & 4a+2c-2b-4d\end{pmatrix}\end{aligned}$$

To find the kernel, we set this to 0. This gives four equations, but the third and fourth are just the negatives of the second and first, respectively. Solving them leads to the matrix $\begin{pmatrix}-1&-2&2&1\\2&1&-4&-2\end{pmatrix}$, which row reduces to $\begin{pmatrix}1&0&0&-1\\0&1&-1&0\end{pmatrix}$ corresponding to the equations $a=d$ and $c=b$ whose solution space is clearly generated by the two matrices $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ and $\begin{pmatrix}0&1\\1&0\end{pmatrix}$.

As for the image, we see that it consists of all matrices of the form

$$\begin{pmatrix} -4a-2c+2b+4d & 2a+4c-4b-2d \\ -2a-c+4b+2d & 4a+2c-2b-4d \end{pmatrix} = \begin{pmatrix} e & f \\ -f & -e \end{pmatrix}$$

for $e = -2a - c + b + 2d$ and $f = a + 2c - 2b - d$. To see that this is exactly the image, we could say that the matrices of the form $\begin{pmatrix} e & f \\ -f & -e \end{pmatrix}$ are a two dimensional space and the since the dimension of the kernel is known to be two, the dimension of the image is also two and thus consists of all matrices of the form $\begin{pmatrix} e & f \\ -f & -e \end{pmatrix}$. Alternately, one can solve directly the equations $T(A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $T(A) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Solutions can be found by any suitable method; one set is that $T\begin{pmatrix} -2/3 & 0 \\ 1/3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $T\begin{pmatrix} -1/3 & 0 \\ 2/3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(c) We have already calculated T. Let the four matrices in $\mathcal{B}$ be denoted v_1, v_2, v_3, and v_4 respectively. Then $T(v_1) = \begin{pmatrix} -2 & 1 \\ -1 & 2 \end{pmatrix} = -2v_1 + v_2 + (-1)v_3 + 2v_4$ so that the first column of the matrix is $\begin{pmatrix} -2 \\ 1 \\ -1 \\ 2 \end{pmatrix}$. In a similar way we find the remaining columns and the final result is that

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} -2 & 1 & -1 & 2 \\ 1 & -2 & 2 & -1 \\ -1 & 2 & -2 & 1 \\ 2 & -1 & 1 & -2 \end{pmatrix}$$